applied sciences

MDPI

Article

Detection of Nutrients and Contaminants in the Agri-Food Industry Evaluating the Probabilities of False Compliance and False Non-Compliance Through PLS Models and NIR Spectroscopy

David Castro-Reigía [1,2], Iker García [2], Silvia Sanllorente [1], María Cruz Ortiz [1] and Luis A. Sarabia [3,*]

[1] Departamento de Química, Facultad de Ciencias, Universidad de Burgos, Plaza Misael Bañuelos s/n, 09001 Burgos, Spain; dcr1002@alu.ubu.es (D.C.-R.); silsan@ubu.es (S.S.); mcortiz@ubu.es (M.C.O.)
[2] AOTECH (Advanced Optical Technologies S.L.), Edificio Ilgner-BIC Bizkaia Ezkerraldea, Módulo C-7, Avda. Altos Hornos de Vizcaya, 33, 48901 Barakaldo, Spain; igarcia@aotech.es
[3] Departamento de Matemáticas y Computación, Facultad de Ciencias, Universidad de Burgos, Plaza Misael Bañuelos s/n, 09001 Burgos, Spain
* Correspondence: lsarabia@ubu.es

Abstract: NIR spectroscopy has become one of the most prominent techniques in the food industry due to its easy and fast use. Coupled with PLS, it is a well-established method for determining nutrients, contaminants, or adulterants in foods. Nevertheless, it is not common when calculating the capability of detection or discrimination given a target/permitted value, providing probabilities of false non-compliance (α) or false compliance (β). That is exactly the main purpose of this work, where a single procedure using the accuracy line to evaluate these figures of merit by generalizing ISO 11843 when using NIR-PLS in real scenarios in agri-food industries is shown. Nevertheless, it is a completely general procedure and can be used in any analytical context in which a PLS calibration is applied. As an example of its versatility, several analytical determinations were performed using different common food matrices in the agri-food industry (butter, flour, milk, yogurt, oil, and olives) for the quantification of protein, fat, salt, and two agrochemicals. Some results were a detection capability of 5.2% of fat in milk, 1.20 mg kg^{-1} for diflufenican, and 2.34 mg kg^{-1} for piretrin in olives when maximum limits were established at 5%, 0.6 mg kg^{-1}, and 0.5 mg kg^{-1} respectively. Also, 1.02% for salt in butter and 11.45%, 3.78%, and 2.65% for protein in flour, milk, and yogurt, respectively, were obtained when minimum limits were established at 1.2%, 12%, 4%, and 3% respectively. In all cases $\alpha = \beta = 0.05$.

Keywords: PLS calibration; NIR spectroscopy; contaminants; food adulterants; capability of detection; minimum detectable concentration; false positive; false negative; false compliance; false non-compliance

Academic Editor: Kaiqiang Wang

Received: 14 March 2025
Revised: 23 April 2025
Accepted: 24 April 2025
Published: 26 April 2025

Citation: Castro-Reigía, D.; García, I.; Sanllorente, S.; Ortiz, M.C.; Sarabia, L.A. Detection of Nutrients and Contaminants in the Agri-Food Industry Evaluating the Probabilities of False Compliance and False Non-Compliance Through PLS Models and NIR Spectroscopy. *Appl. Sci.* **2025**, *15*, 4808. https://doi.org/10.3390/app15094808

1. Introduction

Monitoring and preserving food quality is an increasingly critical aspect of food control that has direct repercussions on human health, also having impact on our economy. The problems of food adulterations make the food items used in daily life unsafe. Adulteration in food items can have a tremendous effect on human health without our knowledge. This effect has been extensively reviewed in reference [1]. On the other hand, the FDA estimates that food fraud affects 1% of the global food industry at a cost of approximately $40 billion per year [2].

The flexibility and adaptability of near infrared (NIR) spectroscopy in the agri-food industry is well known to assess the quality of raw materials and/or the final products. During food quality control using this technology, it is possible to perform the rapid and non-destructive determination of many quality characteristics of different foods and/or the detection of potential counterfeits or adulterations.

Technical advances in NIR instruments and the scalation of multivariate chemometric techniques have made this technology one of the most widely used in the analytical agri-food field, specifically in industry. To confirm this, bibliographic research was carried out over the past fifteen years, focusing, firstly, on NIR spectroscopy, secondly, on foods, and finally, on the food matrices of this paper: bibliographic research within "Article title, Abstract & Keywords" was carried out in Scopus using, as the keyword search, "NIR spectroscopy", having found a total of 31,376 publications. When also adding the word "foods", the number decreases to 2269. If focusing on the food matrices of this paper, the number of publications varies according to Figure 1. In those cases, the search was made by separately adding the words "butter, flour, milk, yogurt, olive oil, and olives" (which are the food matrices used in this work) to the previous search, having found a total of 357 publications.

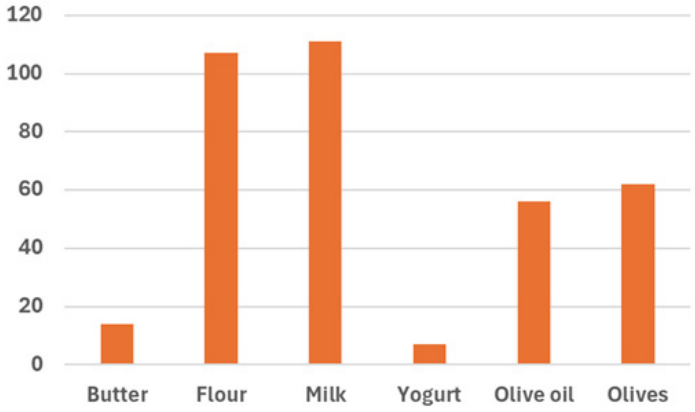

Figure 1. Number of publications found in Scopus by searching "NIR spectroscopy" and "foods" in each food matrix used in this work.

Among all the publications associated with food and NIR spectroscopy since the year 2010, 16% of them are related to the food matrices of this work; for instance, there are many publications referring to analyses on flour, the majority of them focused on detecting adulterations [3–5]. Studies on butter are less common but still relevant when quantifying adulterations [6,7]. Also, in references [8,9], different procedures based on NIR are presented to study the presence of different agrochemicals in olives.

The use of multivariate soft calibration models, such as the ones based on Partial Least Squares (PLS) coupled with NIR spectroscopy signals [10], are quite common in the agri-food industry; in fact, many of the previously cited studies used it. For many years, there have been different International Organization for Standardization (ISO) documents that protocolize how to calibrate using NIR and PLS depending on the analyte and the food matrix [11,12]. Although the importance of NIR spectroscopy and its advantages have been proven, this technique has some downsides, like its low sensitivity. For that matter, it is imperative to know, with certainty, the minimum discriminable concentration that can be detected with this calibration, that is, the capability of detection of the method or the capability of discrimination (given a target value of interest, whether it is a minimum or a maximum limit). All of this is precisely the main aim of this work, encouraged after having checked that within the latter bibliographic research mentioned, practically none of the publications that were found had used these figures of merit, only 20 (6 for flour [4,13–17],

12 for milk [17–29], and 1 for olives and olive oil [30]) out of 357 (the ones schematized in Figure 1) when adding the terms "capability of detection *or* detection limit" to the search. However, in no cases were the probability of false non-compliance and the probability of false compliance indicated. The capability of the detection of an analytical method evaluating the probabilities of false positives and false negatives according to ISO 11843-2 has already been implemented in reference [31]. Nevertheless, it has to be generalized when the signals are not of the order zero, as it was made in references [32,33] using different instrumental signals and PLS. In both cases, the generalization uses the accuracy line (predicted concentration using the PLS model versus the true concentration).

The present work shows the versatility of this unified procedure. It is applied to some analytical determinations in foods using NIR spectroscopy and PLS laying out all possible cases:

- Determine the quantity that can be ensured when maximum permitted limits are established by official regulations (as for agrochemicals or prohibited substances) or when minimum or maximum limits are established for a certain parameter in a food matrix by the industry itself to guarantee the quality of their products
- Ascertain with statistical guarantee the minimum amount that it is possible to discriminate in a certain analytical method.

Thus, all cases are a matter of determining the parameter of interest with a pre-established probability of committing false non-compliances or false compliances in the measurements (which, in the case of the capability of detection, will be false positives and false negatives, respectively).

The inclusion of the described methodology in the NIR-PLS calibrations would contribute to improving compliance with current food quality regulations, obtaining some benefits both for the industries and/or the consumers. Some practical aspects about this are described in Section 4.

2. Materials and Methods

2.1. Instrumentation and Experimental Methods

All the experiments were carried out in different Spanish food industries. As previously explained, the analyses were carried out using six different food matrices (butter, flour, milk, yogurt, olive oil, and olives), with the analyte studied in most of the examples being different (see Table 1). The measurement procedure was made in every case with the AONIR (AOTECH S.L. [34]) integrated solution for real-time NIR measurements, including an NIR sensor, a measurement platform, and the precise software to integrate the hardware with the model outcome for posterior real-time monitoring of the parameters that were going to be measured in this work. The spectrometer was configured equally in all cases so that NIR reflectance was registered in each food matrix in a wavelength range from 900 to 1670 nm (125 wavelengths, accounting for a pixel resolution of 6 nm and a spectral resolution of 12 nm) and 50 readings per spectrum with an integration time of 10.8 ms. Nevertheless, the number of samples, the sample replicates, and the spectral ones were not the same in all cases given the determining factor that the measurements were made in different industries where it is necessary to adapt to the times of reception of samples, production times, and availability, among others. The state of the samples was also considered; for example, as can be seen in Table 1, it was decided to average the eight spectra recorded for each olive sample, considering that the distribution of the agrochemicals could not be homogeneous. The same applies to the powdery-nature flour, for which replicates are higher than in the other matrices. In some cases, the samples were even measured on different days. The details of samples and sample replicates, along with the final data matrix for each case, can be found in Table 1.

After the NIR measurements, the samples were sent to an accredited laboratory in order to be measured using a certified reference method (the one indicated in Table 1 for each analyte in each food matrix).

Table 1. Reference method for each food matrix and each analyte, number of samples, data collection summary, and data matrix size.

Matrix	Analyte	Reference Method	N *	Sample Replicates	Spectral Replicates	Final Data Matrix	Data Matrix of the Prediction Set
Butter	Fat (%) (*w*/*w*) Salt (%) (*w*/*w*)	NMR Atomic absorption	11	2	3	66 × 125	24 × 125
Flour	Protein (%) (*w*/*w*)	Kjeldahl method	36	3	3 or 6	504 × 125	-
Milk	Fat (%) (*w*/*w*) Protein (%) (*w*/*w*)	FTIR FTIR	38	1 or 2	3	195 × 125	52 × 125
Yogurt	Fat (%) (*w*/*w*) Protein (%) (*w*/*w*)	Gravimetry Kjeldahl method	19	2 or 4	3	144 × 125	24 × 125
Olive oil	Refined olive oil (%) (*v*/*v*)	**	14	1 or 2	3	81 × 125	18 × 125
Olives	Both agrochemicals (mg kg^{-1})	GC-MS-MS QqQ	40	1	***	40 × 125	-

* N: number of samples; ** the mixtures were prepared, knowing the true concentration; *** 8 averaged spectra.

As mentioned above, the AONIR platform developed by AOTECH S.L. [34] was used to obtain the recorded spectra in a .csv file, while PLS-Toolbox [35] working under MATLAB version 9.9.0 (R2020b) [36] was employed for fitting the PLS models. The decision limits and the capabilities of detection were calculated using the DETARCHI [31] program, whereas CDα and CDβ were estimated using a handmade program also working under MATLAB [36], which generalizes DETARCHI [31] to all the cases described in this work. These programs avoid using a table of the non-central t-distribution and automate obtaining graphs and CCα, CCβ, CDα, and CDβ values for any number of replicates and the values of the probabilities of false compliance and false non-compliance.

2.2. Statistical Method

When working with univariate linear calibrations, ISO 11843-2 [37] and the International Union of Pure and Applied Chemistry (IUPAC) [38] defined a procedure to determine the decision limit and the capability of detection of an analytical method, evaluating the probabilities of false positives and false negatives. Nevertheless, when a multivariate calibration is used, as mentioned in Section 1, this methodology has to be generalized because these documents only refer to univariate calibrations. The decision limits and the detection capabilities are always calculated using a hypothesis test obtained from the calibration model whether it is univariate or multivariate. In the following sections, the main aspects of this generalization for multivariate signals are explained (given that we are working with multivariate signals), both for the capability of detection and the capability of discrimination, and then, the global procedure followed in this work is summarized for PLS models with NIR signals.

2.2.1. Decision Limit and Capability of Detection at $x_0 = 0$ or for a Permitted Limit, $x_0 = PL$ with Multivariate Signals

NIR spectra combined with PLS regression can be used for the control of maximum permitted limits based on official regulations or the maximum and/or minimum limits in-

tentionally established to guarantee food quality. Regarding that, new figures of merit, CDα and CDβ had to be defined and evaluated, which are analogous to CCα and CCβ [37,38].

In order to formalize the notation, the definitions given by (EU) 2021/808 regulation [39] referring to a permitted limit (PL) were used. PL can be defined as the maximum residue limit, maximum level, or other maximum tolerance for substances established in any normative or regulation. On the one hand, the decision limit when PL = 0 (CCα) is "*the concentration limit at and above which it can be concluded with an error probability of α that a sample is non-compliant when it truly is*". On the other hand, the detection capability (CCβ) definition is different depending on if a PL is established or not. For prohibited or unauthorized substances, CCβ is "*the lowest concentration at which a method is able to truly detect or quantify, with a statistical certainty of* $1 - \beta$*, samples containing residues of prohibited or unauthorized substances*" whereas for the case of authorized substances with an established PL, it is defined as "*the concentration at which the analytical method is able to detect concentrations* below the permitted limit with a statistical certainty of $1 - \beta$".

CCα and CCβ were already established for PL = 0 by ISO 11843-2 [37] and IUPAC [38] and can be generalized for any other PL value, where they will be named as CDα and CDβ, respectively, to distinguish them from the terms already established in the regulatory framework mentioned above (when the calibration is univariate).

On the contrary, when the calibration is multivariate, and in particular, if for a given parameter, x, it is considered a minimum permitted (or established) limit, PL = x_0, the corresponding one-tailed hypothesis test should be outlined as in Equation (1). The word one-tailed specifies that the alternative hypothesis of the test is not an inequality.

$$\begin{gathered} H_0\text{: } x = x_0 \text{ (the parameter is greater than or equal to } x_0\text{, compliant sample)} \\ H_a\text{: } x < x_0 \text{ (the parameter is less than } x_0\text{, non-compliant sample)} \end{gathered} \quad (1)$$

For example, if it is wanted to guarantee that the fat content of a milk sample is at least 12%, then the minimum permitted limit is PL = x_0 = 12%. The hypothesis test allows one to decide, with a probability of false non-compliance (α) and false compliance (β), whether a new sample, based on its NIR spectrum, fulfils the minimum PL or not.

To build the hypothesis test of Equation (1), the procedure begins with a PLS calibration, $y = f_{PLS}(\mathbf{s})$, which assigns a concentration (y) to the NIR spectrum (**s**) of a sample of which the concentration measured using a reference accredited method is x. From the calibration data (x_i,y_i) i = 1, … I, the accuracy line $y = a + bx$, previously defined in Section 1, is built.

On another note, CDα is related to the probability of false non-compliance, α. That is, α is the probability of deciding that the tested sample is non-compliant when it truly was. In other words, α is the significance level in the hypothesis test of Equation (1). For a given α, the probability of a false compliant decision (β) is the probability of wrongly affirming that the tested sample has a value of the parameter greater than or equal to x_0, i.e., to conclude that it is compliant, when it was not. The capability of detection (CDβ) is the value of the parameter related to this decision and is computed as in Equation (2):

$$CD\beta = x_0 - \frac{\Delta(\alpha, \beta) w_{x_0} \hat{\sigma}}{\hat{b}} \quad (2)$$

where $\hat{\sigma}$ is the residual standard deviation of the accuracy line, $\hat{b}$ its slope, and w_{x_0} is defined according to Equation (3), where it can be observed that it depends on the position

of the standards in the accuracy line (x_i, i = 1, ... I), the number of sample replicates (r), and the number of standards (I).

$$w_{x_0} = \sqrt{\frac{1}{r} + \frac{1}{I} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{I}(x_i - \bar{x})^2}} \quad (3)$$

In Equation (2), $\Delta(\alpha, \beta)$ is the value of the non-centrality parameter of a non-central t-distribution that depends exclusively on the probabilities α and β and on the number of standards (I). This parameter models the changes in the probability distribution of the response when considering values lower than x_0. The distribution changes from being a Student's t with $I - 1$ freedom degrees to being a non-central Student's t distribution with the same degrees of freedom and with a non-centrality parameter equal to Δ.

The limit of decision, CDα, will be then calculated as in Equation (4),

$$CD\alpha = \frac{y_c - \hat{a}}{\hat{b}} \quad (4)$$

where $\hat{a}$ and $\hat{b}$ are the estimated intercept and slope of the accuracy line, respectively, and y_c, is the critical value that satisfies Equation (5) for a minimum permitted limit.

$$\alpha = \text{probability}\{y < y_c | x = x_0\} \quad (5)$$

Figure 2 shows a scheme of the hypothesis test that assesses, with an α probability of false non-compliance and a probability β of false compliance, a minimum permitted or established limit. The accuracy line is shown in the upper left corner of the figure. The concentration, x , of the calibration samples is represented on the abscissa axis, and the values obtained using the PLS model using the NIR spectra of those same samples are represented on the ordinate axis. The two arrows indicate the pair (x_0, y_0). The density function of the y values around y_0 is shown, and the probability α is the area marked in red and described by Equation (5), which defines the critical value y_c. With the recorded spectrum of a new sample, the corresponding y_p is obtained using the PLS calibration model. If this value is less than y_c, then it will be concluded that the sample is non-compliant with a probability α. Nevertheless, the probability of false compliance, β, is not being evaluated. As one moves toward lower values of concentration x, the mean of the density function decreases. At the value x_1 of the graph, there is a decrease in the mean of the density function and a larger one at x_2, and so on. In each one of these density functions, the probability of obtaining a value greater than the critical value, that is, β, has been marked in blue. Clearly, β decreases as x moves away from x_0. On the other hand, the power curve (Operating Characteristic Curve) of the test is shown at the bottom right of Figure 3. Again observing the accuracy line graph, it is obvious (and there is mathematical proof of this) that indicating β is equivalent to indicating the shift in concentration, because the mean value of the distribution is related to the concentration based on the accuracy line. Additionally, CDα is the value that corresponds to the accuracy line with the critical value y_c (Equation (4)), and it can be observed that for CDα, the probability of false compliance, β, is 0.5. That means that 50% of the times that the test is applied, a false compliant result will occur. This situation always occurs, even if α is small. In other words, if the test is not applied considering the requirement to guarantee the probability of false compliance, there will always be a probability equal to 0.5 of accepting that the sample is compliant when it is not. However, CDβ (Equation (2)) guarantees that probability.

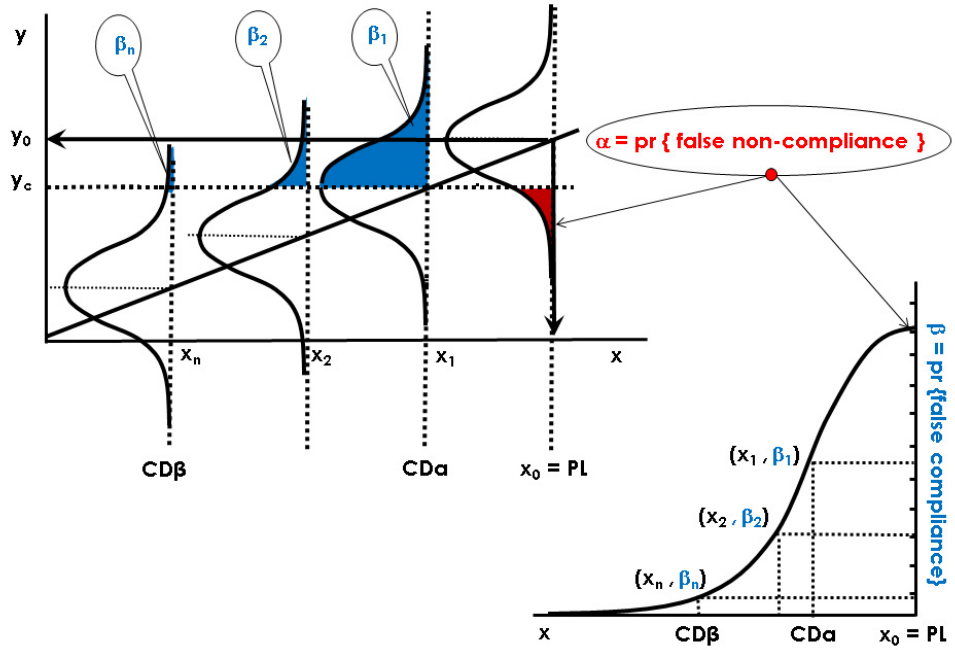

Figure 2. Schematic representation of CDα and CDβ linked to the accuracy line (upper left) and Operative Characteristic Curve (bottom right) for a minimum limit (x_0 = PL) according to the hypothesis test in Equation (1), where x is the true concentration of a sample, y is the obtained concentration based on the PLS model, and where y_c depends on the significance level of the test (α).

In another completely different situation when the established limit, PL = x_0 is a maximum, a different one-tailed hypothesis test is used (Equation (6)).

$$\begin{array}{c} H_0\text{: } x = x_0 \text{ (the parameter is smaller than or equal to } x_0\text{, compliant sample)} \\ H_a\text{: } x > x_0 \text{ (the parameter is greater than } x_0\text{, non-compliant sample)} \end{array} \quad (6)$$

Thus, the capability of detection, CDβ (analogous to CCβ when $x_0 = 0$ and named x_d by ISO 11843-2 [37]), is defined as in Equation (7).

$$CD\beta = x_0 + \frac{\Delta(\alpha, \beta) w_{x_0} \hat{\sigma}}{\hat{b}} \quad (7)$$

In this case (when having a maximum permitted limit, x_0), w_{x_0} is estimated with Equation (3), and CDα is obtained equally as for a minimum permitted limit, with Equation (4), but in this case, y_c is the value that satisfies Equation (8).

$$\alpha = \text{probability}\{y > y_c | x = x_0\} \quad (8)$$

Figure 3 shows a scheme of the hypothesis test (Equation (6)) that assesses, with an α probability of false non-compliance and a probability β of false compliance, a maximum permitted or established limit. The interpretation of this figure is completely analogous, point by point, to the one of Figure 2, except for the reason that since the PL is a maximum, the inequality of the alternative hypothesis is the opposite of the one in Equation (1). This implies that the displacement of the mean value of the density function goes towards greater values than in PL = x_0. This is reflected by the change in the sign in Equation (7) regarding Equation (2). Equation (7) allows one to obtain the CDβ for a predetermined probability of false non-compliance, α.

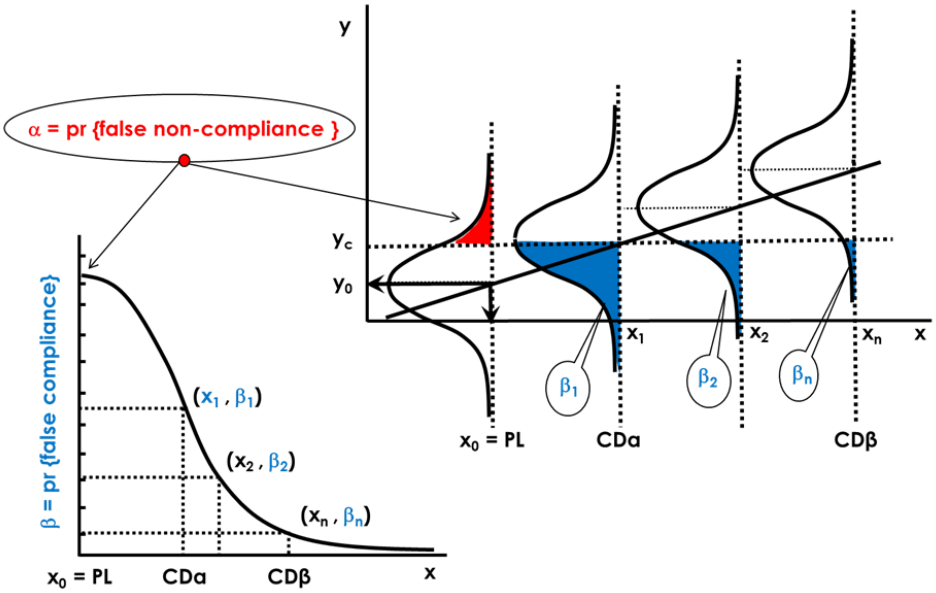

Figure 3. Schematic representation of CDα and CDβ linked to the accuracy line (upper right) and Operative Characteristic Curve (bottom left) for a maximum limit (x_0 = PL) according to the hypothesis test in Equation (6), where x is the true concentration of a sample, y is the obtained concentration based on the PLS model, and where y_c depends on the significance level of the test (α).

2.2.2. Capability of Discrimination or Multivariate Sensitivity

The sensitivity of an analytical method can be defined as the smallest amount of substance or analyte that can be accurately measured. So, an increased sensitivity means a higher increment of the response for the same amount or concentration of an analyte. To fix the notation, consider the interval of concentrations $(x_0 - \varepsilon, x_0 + \varepsilon)$ around a value x_0. The sensitivity (the increase in the response of the method in that interval) should be defined in terms of what probability that a value outside the interval is significantly different from x_0 or not is obtained. If only the quotient of the increment in the response regarding the increment of the concentration is considered (e.g., the slope of an univariate calibration), the inherent uncertainty in any determination in the decision is omitted, and it also depends on the units in which the response is expressed. Given that the hypothesis tests have already been formalized so that, given a predetermined value (x_0), it is possible to decide whether it is significantly greater or less than x_0, it is reasonable to generalize the situation to the case of deciding whether the concentration is significantly different from x_0. Consequently, given a nominal concentration, PL = x_0, in order to know the performance of an analytical procedure in samples with a similar concentration, the minimum discriminable concentration (or multivariate sensitivity) is defined as the smallest concentration of the analyte in a sample that can be distinguished with a probability $1 - \beta$ from x_0. Some extended details about the multivariate capability of discrimination can be consulted in reference [33]. In process quality control, univariate charts are used to control a single parameter or several parameters at the same time with multivariate charts. Both cases are formalized as a hypothesis test with the evaluation of probabilities α and β. Unlike control charts, the multivariate sensitivity approach proposed in this work does not imply taking the decision based on the parameter measurement (e.g., the fat % of a milk sample). Instead, an NIR spectrum of the sample is recorded, PLS calibration is then applied, and the result is used to determine whether the sample fulfills the quality standards using multivariate analytical sensitivity. That is, the decision is not made with a direct measurement of the controlled parameter but rather indirectly through its NIR spectrum.

Formally, the compliance of an analytical procedure corresponds to $x = x_0$, while, if $x \neq x_0$, the procedure is not compliant. Therefore, a two-tailed hypothesis test can be outlined as in Equation (9).

$$\begin{gathered} H_0\text{: } x = x_0 \text{ (the parameter is equal to } x_0\text{, compliant sample)} \\ H_a\text{: } x \neq x_0 \text{ (the parameter is smaller or greater than } x_0\text{, non-compliant sample)} \end{gathered} \tag{9}$$

This text formalizes the definition of the minimum discriminable concentration considering that a given concentration will be different from x_0 if it is greater than $x_0 + \varepsilon$ or if it is less than $x_0 - \varepsilon$. In terms of the previous section, $x_0 + \varepsilon$ is a maximum PL and $x_0 - \varepsilon$ is a minimum PL. Nevertheless, both α and β must be distributed within both limits. Therefore, each of the CDβ and the corresponding critical values with $\alpha/2$ and $\beta/2$ are determined as indicated in Equations (10)–(12), for which the graphical interpretation can be made observing Figure 4. In summary, the capability of detection, CDβ, has, in this case, a minimum value, $CD\beta_{lower}$, and a maximum value, $CD\beta_{upper}$, evaluated using Equation (10).

$$\begin{aligned} CD\beta_{lower} &= x_0 - \frac{\Delta(\alpha/2,\beta/2)w_{x_0}\hat{\sigma}}{\hat{b}} \\ CD\beta_{upper} &= x_0 + \frac{\Delta(\alpha/2,\beta/2)w_{x_0}\hat{\sigma}}{\hat{b}} \end{aligned} \tag{10}$$

In the same way, the decision limit, CDα, is divided into two values depending on whether it is above or below the PL value. These values are calculated using Equation (11).

$$\begin{aligned} CD\alpha_{lower} &= \frac{y_{cl} - \hat{a}}{\hat{b}} \\ CD\alpha_{upper} &= \frac{y_{cu} - \hat{a}}{\hat{b}} \end{aligned} \tag{11}$$

where y_{cl} and y_{cu} are the values that satisfy Equation (12) for a discrimination problem.

$$\begin{aligned} \alpha/2 &= \text{probability}\{y < y_{cl} | x = x_0\} \\ \alpha/2 &= \text{probability}\{y > y_{cu} | x = x_0\} \end{aligned} \tag{12}$$

Figure 4 is a schematic representation of the elements that make up the capability of discrimination. Methodologically, it is analogous to the cases in which there is a minimum or a maximum permitted/established limit (Figures 2 and 3) but distributing the probabilities of false compliance and false non-compliance bilaterally regarding y_0. As in the previous cases, x is the true concentration of a sample, y, the concentration obtained based on the PLS calibration, and y_{cl} and y_{cu} depend on the significance level (α) of the test.

Figure 4 shows that the two critical values from Equation (12) are established with the density function of the response in x_0, assigning the probability $\alpha/2$ in each one of their tails (marked in red in the figure). As the concentration x moves away from the value x_0 based on higher or by lower values, the probability of false compliance decreases (the sum of the two areas marked in blue in the figure). In a similar way as in Figure 2 (but in this case, bilaterally) if the interval [$CD\alpha_l$, $CD\alpha_u$] is considered as the capability of discrimination regarding x_0, the probability of incorrectly stating that the concentration is x_0 is β, which is practically 1. On the contrary, by expanding the interval around x_0, or what is the same, decreasing β, that is the probability of incorrectly stating that the concentration is different from x_0.

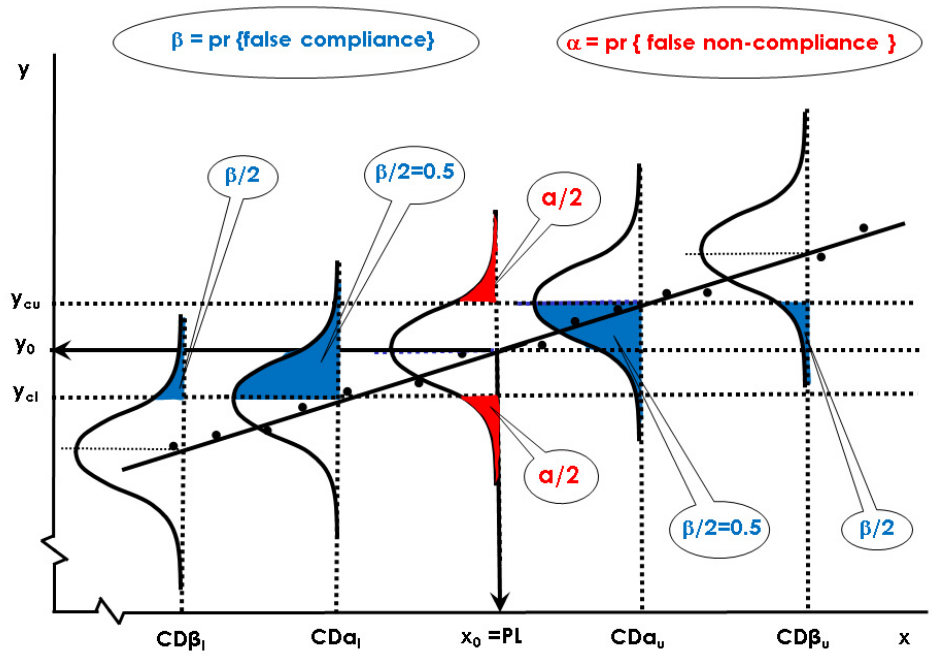

Figure 4. Schematic representation of $CD\alpha_1$, $CD\alpha_u$, $CD\beta_1$, and $CD\beta_u$ linked to the accuracy line for a discriminant amount in (x_0 = PL) according to the hypothesis test in Equation (9).

2.2.3. Global Procedure for Multivariate Calibration to Guarantee CCα and CCβ (or CDα and CDβ) with NIR Spectroscopy and a PLS Model

To determine CCα and CCβ (or CDα and CDβ), a multivariate PLS calibration model is built using the NIR spectra as predictor variables and the concentration of the analyte or substance as a response. Particularly, the matrix size of the predictor variables of this work could be seen in Table 1.

In order to carry out the complete procedure, the specific following steps were followed in this work:

1. To build a PLS model for each parameter, y = f(**X**), (i) first, the predictors were preprocessed by applying the standard normal variate (SNV) followed by a first (D1) or a second (D2) derivative and a second-degree polynomial (also varying the window size from 9 and 15) depending on the case of study. Then, both the predictors and the responses were mean-centered (all the details can be consulted for each data set and each case in Table 2). (ii) The number of latent variables was selected through cross validation. (iii) The samples with a standardized residue greater than 3 (in the absolute value) or with both Q residuals and T^2 Hotelling values larger than their corresponding threshold values at a 95% confidence level were removed (outliers). (iv) Steps (ii) and (iii) were repeated until no outliers were detected;
2. The accuracy line was then built by means of a least squares regression, representing the predicted values obtained with the PLS models (y) versus the true concentration (x) obtained using the reference method specified in Table 1 for each case of study. In this way, the predicted and true concentrations are linked by means of a linear model. The characteristics of every constructed accuracy line can be found in Table 3, whereas their graphical representations can be seen in Figure S1 in the Supplementary Material;
3. Using the data resulting from the accuracy lines, CCα and CCβ (or CDα and CDβ) were calculated for probabilities of both a false positive and false negative (or false non-compliance and false compliance) of 0.05, regarding the definitions in Sections 2.2.1 and 2.2.2.
4. The final results, after applying this global procedure, can be consulted in Section 3.

3. Results

3.1. PLS Calibration

According to the experimental data described in Table 1, a PLS calibration was built for each analyte individually. The data were pretreated as indicated in Table 2. In every case, the standard normal variate (SNV) was applied to reduce the (physical) variability between samples due to scattering. It was also applied to the adjustment for baseline shifts between samples. Except for the protein determination in yogurt, a derivative was also always applied after the SNV. The first derivative was used for corrections of the baseline, and the second derivative to eliminate both the baseline and linear trends. The former has been used for the butter model, whereas in the others, the second derivative was used. The derivatives have been applied using the Savitsky–Golay method with a second-degree polynomial and a window width varying from 7 to 15 points, as shown in Table 2. The derivative order, the polynomial degree, and the window width have been chosen after testing several different combinations to improve the PLS model, which is why in the case of yogurt, the preprocessing method is different for protein (being just the SNV the one that provided the best results). A review of preprocessing methods for NIR spectra can be consulted in reference [40]. Also, in Figure S2 of the Supplementary Material, these explained effects can be observed by comparing the raw spectra with the pre-processed ones for each food matrix. Centering the NIR spectra after preprocessing them has a huge impact on the PLS calibration models. In a qualitative way, PLS can be described as a regression technique that searches for directions in the space of predictor variables such that they explain the greatest variance and, at the same time, that they are as highly correlated as possible with the response. As a consequence, if the predictor variables (preprocessed spectra) are not centered, the direction of greatest variability will be the one that connects the vector of means of the data with the origin of the coordinates. From the perspective of the PLS regression, it is highly recommended to center the preprocessed spectra, which is what has been performed in every case studied in this work (see Table 2). The number of the LV of the models, also indicated in Table 2, was selected according to the minimum value of Root Mean Squares Error of Cross-validation (RMSEC), obtained by applying the venetian blinds method in all cases, since there were many objects in random order. Three samples per blind were selected to avoid the replicate sample trap with a variable number of splits, depending on the size of the data matrix. In the case of olives, the number of samples per blind was just one, given that each spectrum was obtained as an average of eight spectra originating from one olive. The explained variance of the models ranged between 90.32% and 98.60%, both for the predictors and for the response, having found few outliers. After the models were built, the Root Mean Square Error of Prediction (RMSEP) was calculated using an external prediction set to reflect the predictive ability of the model. As can be seen in Table 2 that the RMSECV and RMSEP values are quite similar, what indicates the stability of the models and their predictive ability.

Table 2. PLS model results for each food matrix and each modeled analyte.

Matrix	Analyte	Preprocess [1]	LV	Variance of x-Block (%)	Variance of y-Block (%)	Out. [2]	RMSEC	RMSECV	RMSEP
Butter	Fat (%)	SNV + 1D (2, 11) + MC	4	94.98	95.10	-	0.295	0.437	0.317
	Salt (%)	SNV + 1D (2, 11) + MC	6	98.68	96.46	-	0.084	0.175	0.175
Flour	Protein (%)	SNV + 2D (2, 13) + MC	7	97.00	95.15	20	0.278	0.357	-
Milk	Fat (%)	SNV + 2D (2, 13) + MC	6	98.32	96.11	3	0.112	0.135	0.172
	Protein (%)	SNV + 2D (2, 13) + MC	4	96.77	90.32	6	0.105	0.117	0.085

Table 2. *Cont.*

Matrix	Analyte	Preprocess [1]	LV	Variance of x-Block (%)	Variance of y-Block (%)	Out. [2]	RMSEC	RMSECV	RMSEP
Yogurt	Fat (%)	SNV + 2D (2, 9) + MC	6	99.31	98.60	7	0.292	0.360	0.315
	Protein (%)	SNV + MC	7	99.95	96.40	2	0.177	0.207	0.215
Olive oil	Refined olive oil (%)	SNV + 2D (2, 15) + MC	4	94.99	94.13	5	2.896	3.620	2.872
Olives	Diflufenican (mg kg^{-1})	SNV + 2D (2, 7) + MC	7	97.09	94.90	2	0.317	0.484	-
	Piretrin (mg kg^{-1})	SNV + 2D (2, 7) + MC	7	92.09	95.93	-	0.971	1.644	-

[1] SNV: standard normal variate; iD (y, z): where iD refers to the "first" or "second" "derivative", y is the polynomial degree, and z is the window size; MC: mean centering. [2] Out: outliers. The number indicated in each row of this column refers to the total number of outlier spectra.

The decision limit and the capability of detection are invariant for linear transformation of the response [33]. Therefore, the concentration calculated using an univariate calibration model (accuracy line) can be employed instead of using the univariate analytical response. Even if the calibration model is multivariate, the accuracy line can always be calculated. In fact, not only with a multivariate PLS calibration model, but also with any other function of calibration, can the accuracy line be used (i.e., principal component regression, regression on a parallel factor analysis decomposition -PARAFAC-, neural networks...). It is suitable for univariate and multivariate signals (excitation emission matrices or hyperspectral images among others). Therefore, this approach facilitates probabilistic quantification in the application of false compliance and false non-compliance criteria for any type of sample and analytical signal. Considering all of this, once the PLS calibration modes were obtained (of which characteristics were previously explained and indicated in Table 2), the accuracy lines (predicted concentration based on the PLS models versus the real concentration of the analyte) were constructed. The accuracy lines were extensively used to evaluate the calibration models, but for the purpose of this work, only the necessary data for building them and to apply the methodology developed in Section 2 were collected in Table 3. So, for each of the PLS models in Table 2, the estimates of the intercept, slope, and residual standard deviation of the corresponding accuracy line are noted in Table 3. The concentration ranges of the analytes in each food matrix are also indicated. The *p*-values that also figure into that table indicate that there is no evidence to reject the null hypothesis of the test; that is, there is a significant linear relation between the parameters at a confidence level of 95%. The graphical representations of these accuracy lines can be observed in Figure S1 in the Supplementary Material.

Table 3. Number of samples, N, intercept, and slope of the accuracy lines, and *p*-value for the significance test.

Matrix	Analyte	N	Analyte Range		Intercept	Slope	s_{yx}	*p*-Value *
			Min	Max				
Butter	Fat (%)	66	81.10	86.60	4.109	0.951	0.293	<0.0001
	Salt (%)	66	0.00	1.20	0.008	0.965	0.083	<0.0001
Flour	Protein (%)	484	9.41	14.58	0.575	0.952	0.272	<0.0001
Milk	Fat (%)	192	3.65	6.16	0.166	0.961	0.110	0.0058
	Protein (%)	190	3.09	4.27	0.339	0.904	0.100	<0.0001
Yogurt	Fat (%)	137	0.1	9.4	0.038	0.986	0.292	<0.0001
	Protein (%)	142	2.8	6.4	0.137	0.964	0.174	<0.0001
Olive oil	Refined olive oil (%)	76	61	100	4.794	0.941	2.847	<0.0001
Olives	Diflufenican (mg kg^{-1})	38	0.00	3.42	0.047	0.949	0.281	<0.0001
	Piretrin (mg kg^{-1})	40	0.00	11.40	0.126	0.9593	0.869	<0.0001

* H_0: the regression is not significant versus H_a: the regression is significant.

With these data, it is possible to determine the capability of detection and/or the capability of discrimination as explained in the following section.

3.2. Estimation of the Capability of Detection and the Capability of Discrimination

In addition to the stringent official regulations imposed within the agri-food sector (where the maximum permitted limits of some residues are established, as in the case of some agrochemicals in olives in this work), sometimes, some companies internally establish some standards (target values) that ensure the quality of their products, either to fulfill legal requirements or for internal reasons, for example, because the value of the characteristic determines the subsequent stages of processing. In this sense, it is crucial to know the level of certainty with which these limits are not exceeded, whether they are minimum or maximum values, that is, to know the minimum discriminable quantity with a probability of false compliance (β) and false non-compliance (α) in order to decide whether to accept a sample or not and continue with the process. In the following, some possible real scenarios that may occur in the agri-food industry are presented.

The results of applying the global procedure described in Section 2.2 can be seen in Tables 2–4, and graphically, for some of the Operative Characteristic Curves of the applied tests, in Figures 5–8. In these figures, the Operating Characteristics Curves are represented, β on ordinate axis versus the true concentration on the abscissa axis. In each one of them, the dashed line indicates the value of CDβ (or CCβ) on the x-axis for a probability β = 0.05 (in the y-axis), when the probability $\alpha = 0.05$ is fixed. In Section 2.2.2., the use of the operative curve of the test to discriminate a concentration has been described. This allows one to evaluate the performance of an NIR-PLS calibration in relation to another different one obtained with another different instrument or measuring in another matrix. Note that by setting the values of α and β and the number of replicates, r, the interval [$CD\beta_i$, $CD\beta_u$] allows one to immediately obtain the best option in terms of the discrimination capability. This approach is of great interest in scenarios such as food safety and food fraud control, as shown in this section. The curves were plotted for the case of different future replicates (r = 1, r = 3, r = 6 and r = ∞, the theoretical value in the case of infinite samples). It can be observed that as more replicates are performed, the values obtained for the capability of detection and/or the capability of discrimination become closer to the desired PL value. This happens as a consequence of the fact that the expressions CCα, CCβ, CDα, and CDβ contain the factor w_{x0}, which also contains the summand 1/r, which decreases as the number of replicates, r, increases. In some cases, increasing the number of replicates may allow a limit to be fulfilled without modifying any other aspect of the measurement system. However, it is not always possible to perform a high number of replicates in certain industrial processes; it will always depend on the conditions imposed by each industry. In this work, r = 3 replicates were considered. It can also be seen in every representation that for the same value of r, the greater the $|x - x_0|$ distance, the lower the value of β is. In other words, the lower the risk of committing false compliances (or false negatives) is in an analysis.

In particular, Figure 5 shows these Operative Curves when the discriminable amount wants to be detected; for instance, for fat in butter, for the 85% w/w established value, the two-tailed test applied is the one in Equation (9). As can be seen in Table 4, this test, as it is a two-tailed test, has a lower CDβ and an upper one that encompass the range [84.27, 85.73]. That is, the capability of discrimination of this method of analysis in the 85% value is bounded by these values, guaranteeing the measure of fat via NIR spectroscopy with a probability of false non-compliance (α) and false compliance (β) of 0.05.

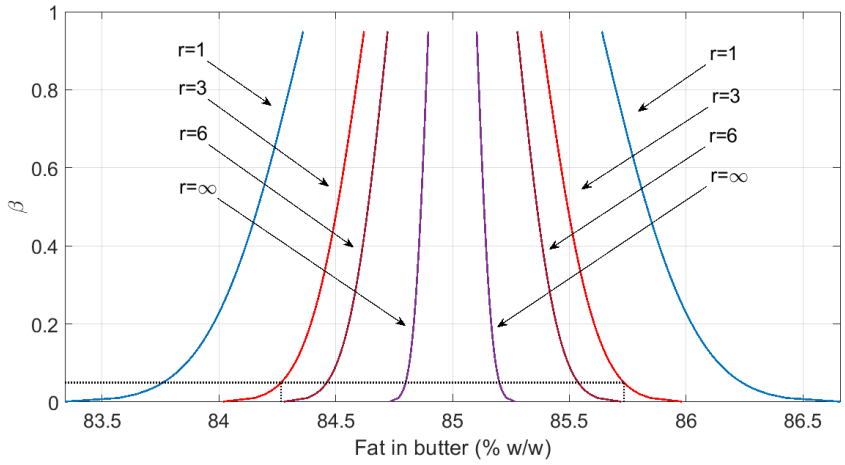

Figure 5. Capability of discrimination of operative curves at α = 0.05 for fat percentage27 in butter samples having 85% as the PL.

Table 4. Decision limit, capability of detection, and minimum discriminable concentration for each matrix and each analyte obtained for probabilities of a false positive (or false non-compliance) and false negative (or false compliance) α = β = 0.05 and for r = 3.

Matrix	Analyte	N	Range		PL = x_0	y_c	CCα	CCβ	CDα	CDβ
			Min	Max						
Butter	Fat (%)	66	81.10	86.60	85 *	84.94/85.64	-	-	[84.63, 85.36]	[84.27, 85.73]
	Salt (%)	66	0.00	1.20	0 1.2 **	0.091 1.077	0.086 -	0.171 -	- 1.11	- 1.02
Flour	Protein (%)	484	9.41	14.58	12 **	11.733	-	-	11.73	11.45
Milk	Fat (%)	192	3.65	6.16	5 ***	5.079	-	-	5.11	5.22
	Protein (%)	190	3.09	4.27	4 **	3.856	-	-	3.89	3.78
Yogurt	Fat (%)	137	0.1	9.4	0	0.322	0.290	0.580	-	-
	Protein (%)	142	2.8	6.4	3 **	2.867	-	-	2.83	2.65
Olive oil	Refined olive oil (%)	76	61	100	80 **	77.30	-	-	77.03	74.07
Olives	Diflufenican (mg kg^{-1})	38	0.00	3.42	0 0.6 ***	0.336 0.901	0.304 -	0.604 -	- 0.90	- 1.20
	Piretrin (mg kg^{-1})	40	0.00	11.40	0 0.5 ***	1.016 1.492	0.928 -	1.844 -	- 1.42	- 2.34

(*) discrimination limit; (**) minimum permitted or established limit; (***) maximum permitted or established limit.

On the other hand, Figure 6 shows the Operative Curves for a minimum limit. The corresponding test from Equation (1) is applied to determine the minimum amount of protein in flour, in this case, 12% (*w*/*w*). It can be observed that, despite having a minimum of a 12% (*w*/*w*) of protein as a target value, with the PLS calibration model built, only 11.45% (*w*/*w*) can only be assured with probabilities α and β equal to 0.05.

On the contrary, when working with a maximum limit, as in the case of an established maximum value of 5% (*w*/*w*) of fat in milk samples, the applied hypothesis test is the one in Equation (6). Figure 7 shows the Operative Curves for this specific case, concluding that the value that can be obtained to guarantee both probabilities of a false non-compliance and false compliance (α and β) of 0.05 is 5.22% (*w*/*w*).

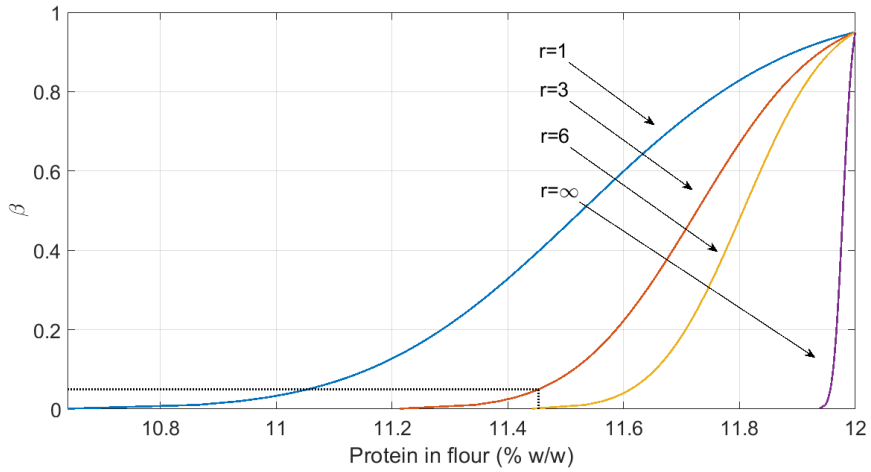

Figure 6. Capability of discrimination of curves for an established minimum value of a 12% (w/w) protein content in flour samples at $\alpha = 0.05$.

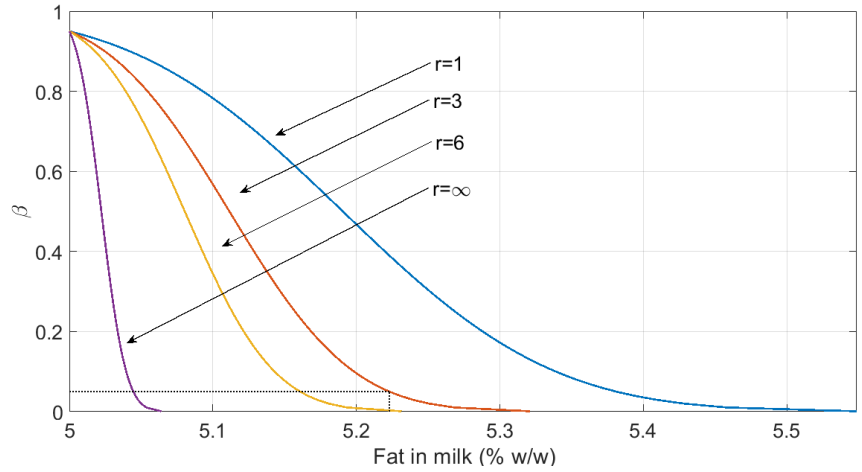

Figure 7. Capability of discrimination of operative curves for a maximum established value of a 5% (w/w) fat content in milk samples at $\alpha = 0.05$.

Lastly, Figure 8 shows the Operative Curves for PL = 0 mg kg^{-1} when determining the agrochemical diflufenican in olives. Once again, the hypothesis test applied was also the one in Equation (6), but in this case, the evaluated probabilities are the false positive and the false negative (instead of false non-compliance and false compliance), α and β, respectively, because PL = 0. It should be remembered that in this case, a false positive (α) corresponds to mistakenly stating that an olive sample has the agrochemical (a concentration of diflufenican greater than zero), while a false negative (β) is an erroneous statement that the olive does not contain diflufenican. Reducing the probability of a false negative, in this case, it is important to ensure the absence of toxic or prohibited substances in food and articles intended to be in contact with foods and consumers. It is also necessary to consider that false positives could cause significative economic losses for manufacturers. Thus, it is critical to consider the risk of committing them, that is, the probability of committing both errors in the decision to accept or reject a sample as compliant. This is reflected in current legislation [39], in international regulations [37], and in the chemical analysis [38], but always for determinations with univariate signals. The proposal of the present work for multivariate signals is consistent with the aforementioned standards, and its application is not limited to the agrifood sector. There are many products that must not contain certain substances, for example, lead in gasoline, migrant products, or substances from plastic packaging, such as bisphenol.

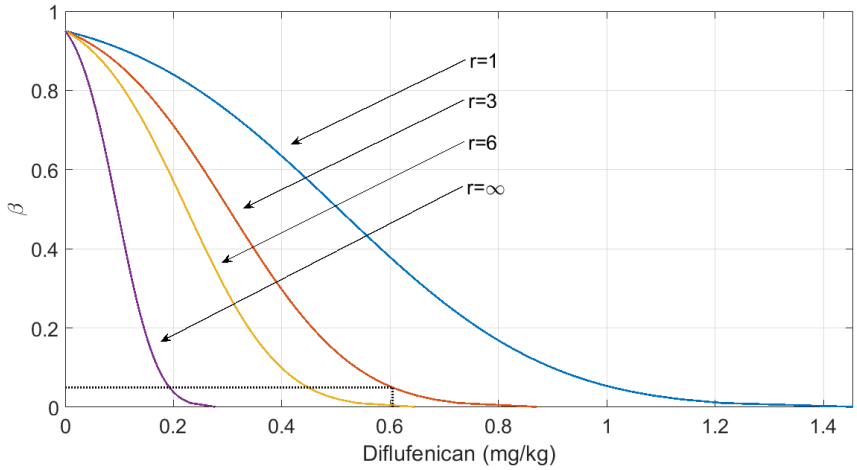

Figure 8. Capability of detection operative curves for diflufenican for a PL = 0 mg kg^{-1} at $\alpha = 0.05$.

Table 4 summarizes the final global results for each case estimated for r = 3 for every parameter measured. In that same table, the calibration ranges of the accuracy lines and the PL = x_0 employed in each case are also shown. The y_c signal is the value of the ordinate axis indicated in Figures 2–4, which allows one to determine CCα (or CDα). Then, the intersection of β = 0.05 with the Operative Curve (for r = 3) in Figures 5–8 is the one that allow one to determine the detection capability CCβ (or CDβ) with the previously fixed probabilities of α and β at 0.05 and the corresponding equation in each case.

From the table above, it is important to highlight that the CCβ for diflufenican is above the permitted limit established by the EU; nevertheless, it must be considered that the values are calculated for false positive and false negative probabilities of 5% and r = 3. It can be seen, for example, in Figure 8, how if the number of spectral replicates is increased, for example, up to r = 6, it is possible to reach a CCβ of 0.450 mg kg^{-1} for diflufenican.

4. Discussion

4.1. Contributions and Practical Implications

The increasing use of NIR spectroscopy and PLS calibration models in the agri-food industry is undeniable. The necessity of establishing analytical sensitivity and some permitted limits with specific assurance probabilities is also evident. However, little effort has been devoted to this issue, as concluded based on the bibliographic review made. In most of the few publications that deal with this matter, a limit of detection (LOD) or a limit of quantification (LOQ) is calculated. Usually, a confidence interval is calculated for the signal at a zero concentration (PL = 0) by calculating k times the standard deviation of the signal in the blank (k = 3 for LOD and k = 10 for LOQ). The interval is then used to decide that a sample does not contain the analyte when its signal is within the interval and that it contains the analyte when it is outside of it. Regarding this, the mistake is to use a confidence interval instead of a hypothesis test, because with the confidence interval, only the probability of a false positive (α) can be estimated, but the probability of a false negative (β) is not evaluated, which, as it has already been shown, is 0.5 for the LOD or LOQ (as can be seen in Figure 2 with a PL = 0).

What has not been found in any publication is the use of the terms of false non-compliance and false compliance for permitted limits. Using the concept of a hypothesis test and the accuracy line, the unified methodology presented in this work is a versatile and easy-to-apply tool.

The hypothesis tests proposed in Section 2.2 are one-tailed (Equations (1) and (6)) or two-tailed (Equation (9)) for the mean y_0, which is the accuracy line for PL = x_0. Precisely,

the structure of the data allows for a calculation of the density function of any value y that is outside the acceptance region of the test, that is, outside the intervals [y_c, y_0], [y_o, y_c] or [y_{cl}, y_{cu}] for a minimum permitted or established limit, for a maximum permitted or established limit, or for the capability of discrimination, respectively. This density function includes the changes in the standard deviation $\hat{\sigma}$ when increasing the distance $|y - y_0|$, which is calculated using the factor w_{x_0} (Equation (3)). The corresponding CDβ (Equations (2), (7), (9), and (10)) is obtained using the accuracy line. In this way, all the key elements of each problem are being considered. Knowing the parameters of this distribution makes it possible to compute the probability of false compliance, β, for each x different from PL = x_0. This information is important and is the differential contribution when considering the decision of accepting or rejecting the PL as a hypothesis test (instead of using an interval of confidence for PL). The probability of false non-compliance, α, manages the length of the intervals [y_c, y_0], [y_0, y_c], or [y_{cl}, y_{cu}]. The smaller the value of α, the greater their length, and therefore, the greater the distance between the PL and CDα. If a sample is considered compliant when y belongs to the corresponding interval, or what is the same, when the quantity is between PL and CDα (or between CDα_l y CDα_u), there would be a risk of false compliance, β, which can reach the value of "1" in the case of discrimination, as observed theoretically in the diagram of Figure 4 and for the case of the fat % determination in butter samples, for which the Operating Curves are seen Figure 5.

Considering all the above and applying it to the studied cases of this work, it can be observed that, for example, Figure 8 shows the capability of detection (false negative, because it has been calculated for PL = 0 mg kg^{-1} of diflufenican), which, regarding the number of future replicate samples that could be made, could have been 1 mg kg^{-1} if only one replicate was made or it could be improved to 0.42 mg kg^{-1} by performing six replicates, instead of the 0.6 mg kg^{-1} obtained for r = 3. The three cases have probabilities of a false positive and false negative of 0.05.

A similar effect can be seen for the case of the determination of fat in butter (Figure 5), for which the Operating Curves show the discrimination capability having 85% (w/w) as the PL value. Its interval for r = 3 is [84.27, 85.73], while if only one replicate was made, the interval would be larger [83.75, 86.25], once again maintaining both probabilities of false non-compliance and false compliance at 0.05.

On another note, the change in the food matrix implies modifications in the NIR signals and consequently in the CDα and CDβ values. The PLS calibration and the accuracy of the analytical method to establish the reference values will also influence these values of CDα and CDβ. For example, consider the value of CDβ for the minimum established limit of the percentage of protein in flour, in milk, and in yogurt (values in the last column of rows 4, 6, and 8 of Table 4). All the CDβ values are calculated for r = 3 and with α = β = 0.05. As the PL is different for each case (12%, 4%, and 3%), to compare them, CDβ will be noted as a percentage regarding the PL (95.42%, 94.59%, and 88.33%, respectively).

In the case of the protein percentage in yogurt, we will have a distance of 88.33% from the PL = 3%, so that the probabilities of false compliance and non-compliance are equal to 0.05.

All of these results depend on the characteristics of the NIR measurement in each case; specifically, it depends on the quotient $\hat{\sigma}/\hat{b}$ from Equation (2), which for protein in flour, milk, and yogurt are 0.09, 0.11, and 0.18, respectively (data calculated from rows 4, 6, and 8 of Table 3). This quotient $\left(\hat{\sigma}/\hat{b}\right)$ is the definition of analytical sensitivity and is invariant for linear transformations of the response in any linear regression based on least squares and characterizes the accuracy line.

The possibility of knowing the probability of false compliance for a predetermined probability of false non-compliance opens a new scenario in the agri-food processes because,

once the cost C_1 and C_2 of considering a sample as a false compliant or false non-compliant have been assessed, respectively, the expected costs of the erroneous decision will be $\beta \times C_1 + \alpha \times C_2$. Regarding all of this, the usefulness of being able to calculate both probabilities based on all the NIR system characteristics is quite evident (obtained from Equations (2), (7), and (10) for the different possible cases).

4.2. Directions of Future Research

Once the general framework for obtaining a discrimination or a permitted/established limit with specified assurance probabilities has been established, the effect of the number of calibration samples and their distribution in relation to the presented problem (limit or discrimination) will be analyzed. The effect on CDα and CDβ based on the two $w_{x_0}\hat{\sigma}$ factors in Equation (3) is complex because one can only act on the number of recorded spectra and their characteristics. This issue is important given that with PLS calibrations, it is necessary to have the response variable measured using a reference method, or the samples need to be prepared knowing their concentration in a laboratory, which can be expensive and time-consuming. In both cases, it is relevant to optimize the calibration design, reducing the experimental effort of off-line (or even at-line) samples but maintaining the quality of the parameters CDα and CDβ. For this task, it will be important to explore the feasibility of using synthetic data generation to design and analyze calibration sets. Experimentally generated data are never sufficient to study the complexities generated based on the complex relationships between NIR spectra and the figures of merit evaluated in this work. Several neural network architectures can create new synthetic data, with the problem at hand including, but not limited to, recurrent neural networks (RNNs), variational autoencoders (VAEs), and generative adversarial networks (GANs). These neural networks have been extensively used with images; nevertheless, their use with NIR data is relatively recent, for example to detect skin cancer [41] to improve palm vein authentication performance [42] or to generate high-resolution (HR) spectra from low-resolution (LR) spectra [43].

Another aspect of this same issue is the predictive maintenance of the calibration over time, since it is assumed that the NIR spectroscopy control of agri-food processes must remain with the lowest possible cost and could also be performed by controlling these figures of merit (CDα and CDβ when PL $\neq$ 0, or CCα and CCβ when PL = 0). It could be useful to adapt strategies such as periodically reducing a (optimized) percentage of calibration samples and replacing them with new ones while maintaining the characteristics of the PLS model. From the perspective of periodically recorded NIR spectra, tools such ASCA (ANOVA Simultaneous Component Analysis) could also help to identify timing points at which a calibration should be updated. Another line of approach is to identify these changes in the space of latent variables of the PLS model. Finally, a highly interesting option when having a mathematical model between predictor variables (NIR spectra) and the figures of merit mentioned (as is in the case of this work) is the use of sequential adaptive methods, which, although very computationally demanding, are effective in practice. These methods, based on Bayesian methods, would propose how the new calibration samples should be used to increase the information already provided by all the samples measured.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/app15094808/s1, Figure S1: Accuracy lines of the models in the manuscript: (**a**) for fat in butter; (**b**) for salt in butter; (**c**) for protein in flour; (**d**) for refined oil; (**e**) for fat in milk; (**f**) for protein in milk; (**g**) for fat in yogurt; (**h**) for protein in yogurt; (**i**) for diflufenican in olives; (**j**) for piretrin in olives; Figure S2: Raw and preprocessed spectra of every food matrix in this work: (**a**) raw spectra of butter samples; (**b**) preprocessed spectra of butter samples; (**c**) raw spectra of flour samples; (**d**) preprocessed spectra of flour samples; (**e**) raw spectra of milk samples; (**f**) preprocessed spectra of

milk samples; (**g**) raw spectra of yogurt samples; (**h**) preprocessed spectra of yogurt samples; (**i**) raw spectra of oil samples; (**j**) preprocessed spectra of oil samples; (**k**) raw spectra of olive samples with diflufenican; (**l**) preprocessed spectra of olive samples with diflufenican; (**m**) raw spectra of olive samples with piretrin; (**n**) preprocessed spectra of olive samples with diflufenican.

Author Contributions: Conceptualization, D.C.-R., S.S., M.C.O. and L.A.S.; methodology, D.C.-R. and I.G.; software, L.A.S. and M.C.O.; validation, D.C.-R., I.G., S.S., M.C.O. and L.A.S.; formal analysis, D.C.-R., M.C.O. and L.A.S.; investigation, D.C.-R.; resources, I.G.; data curation, D.C.-R., S.S. and M.C.O.; writing—original draft preparation, D.C.-R. and M.C.O.; writing—review and editing, D.C.-R., S.S., M.C.O. and L.A.S.; supervision, I.G. and L.A.S.; project administration, I.G.; funding acquisition, I.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by MINISTERIO DE INDUSTRIA Y TURISMO, under project ESPECTROLIVE, grant number AEI-010500-2023-232, and by EUROPEAN COMISSION, under project GreenSME, grant number 101058613.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data sets presented in this article are not readily available due to technical limitations.

Acknowledgments: The authors thank Lugar daVeiga S.L.L., Lácteos De Moeche S.L., and Fundación CITOLIVA/INOLEO for allowing the measurements in their facilities and the entities for the financial support.

Conflicts of Interest: The authors S. Sanllorente, M.C. Ortiz, and L.A. Sarabia declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results. The authors D. Castro-Reigía and I. García declare the following conflicts of interest: the work on the manuscript arises as a part of an industrial doctoral thesis, where the student (Castro-Reigía) pursued his studies at the University of Burgos in collaboration with the company AOTECH, in accordance with the Spanish Royal Decree 576/2023 and the correspondent agreement signed between the University of Burgos and AOTECH, where García (CEO) acts as the doctoral supervisor of the student in the company.

Abbreviations

The following abbreviations are used in this manuscript:

NIR	Near infrared
PLS	Partial Least Squares
ISO	International Organization for Standardization
MIR	Medium infrared
IUPAC	International Union of Pure and Applied Chemistry
NMR	Nuclear magnetic resonance
FTIR	Fourier transform infrared
GC-MS-MS	Gas chromatography coupled to mass spectrometry
QqQ	Triple quadrupole
PL	Permitted limit/established limit
SNV	Standard normal variate
LV	Latent variable
CCα	Decision limit
CCβ	Capability of detection
CDα	Decision limit for a permitted limit different from 0
CDβ	Capability of discrimination for a permitted limit different from 0

References

1. FDA Releases New Food Fraud Webpage. Available online: https://www.food-safety.com/articles/7424-fda-releases-new-food-fraud-webpage (accessed on 16 April 2025).
2. Choudhary, A.; Gupta, N.; Hameed, F.; Choton, S. An Overview of Food Adulteration: Concept, Sources, Impact, Challenges and Detection. *Int. J. Chem. Stud.* **2020**, *8*, 2564–2573. [CrossRef]
3. Casarin, P.; Giopato Viell, F.L.; Good Kitzberger, C.S.; dos Santos, L.D.; Melquiades, F.; Bona, E. Determination of the Proximate Composition and Detection of Adulterations in Teff Flours Using Near-Infrared Spectroscopy. *Spectrochim. Acta. Part A Mol. Biomol. Spectrosc.* **2025**, *334*, 125955. [CrossRef] [PubMed]
4. Zaukuu, J.L.Z.; Nkansah, A.A.; Mensah, E.T.; Agbolegbe, R.K.; Kovacs, Z. Non-Destructive Authentication of Melon Seed (Cucumeropsis Mannii) Powder Using a Pocket-Sized near-Infrared (NIR) Spectrophotometer with Multiple Spectral Preprocessing. *J. Food Compos. Anal.* **2024**, *134*, 106425. [CrossRef]
5. Caballero-Agosto, E.R.; Sierra-Vega, N.O.; Rolon-Ocasio, Y.; Hernandez-Rivera, S.P.; Infante-Degró, R.A.; Fontalvo-Gomez, M.; Pacheco-Londoño, L.C.; Infante-Castillo, R. Detection and Quantification of Corn Starch and Wheat Flour as Adulterants in Milk Powder by Near- and Mid-Infrared Spectroscopy Coupled with Chemometric Routines. *Food Chem. Adv.* **2024**, *4*, 100582. [CrossRef]
6. Medeiros, M.L.d.S.; Freitas Lima, A.; Correia Gonçalves, M.; Teixeira Godoy, H.; Fernandes Barbin, D. Portable Near-Infrared (NIR) Spectrometer and Chemometrics for Rapid Identification of Butter Cheese Adulteration. *Food Chem.* **2023**, *425*, 136461. [CrossRef]
7. Kazazić, S.; Gajdoš-Kljusurić, J.; Radeljević, B.; Plavljanić, D.; Špoljarić, J.; Ljubić, T.; Bilić, B.; Mikulec, N. Comparison of GC and NIR Spectra as a Rapid Tool for Food Fraud Detection: Case of Butter Adulteration with Different Fat Types. *J. Food Process. Preserv.* **2021**, *45*, e15732. [CrossRef]
8. Salguero-Chaparro, L.; Gaitán-Jurado, A.J.; Ortiz-Somovilla, V.; Peña-Rodríguez, F. Feasibility of Using NIR Spectroscopy to Detect Herbicide Residues in Intact Olives. *Food Control* **2013**, *30*, 504–509. [CrossRef]
9. Castro-Reigía, D.; García, I.; Sanllorente, S.; Sarabia, L.A.; Ortiz, M.C. Differentiating Five Agrochemicals Used in the Treatment of Intact Olives by Means of NIR Spectroscopy, Discriminant Analysis and Compliant Class Models. *Microchem. J.* **2024**, *206*, 111550. [CrossRef]
10. Næs, T.; Isaksson, T.; Fearn, T.; Davies, T. *A User-Friendly Guide to Multivariate Calibration and Classification*; NIR Publications: Chichester, UK, 2002.
11. *ISO 21543:2020 (IDF 201:2020)*; Milk and Milk Products-Guidelines for the Application of Near Infrared Spectrometry, 2nd ed. International Organization for Standardization: Geneva, Switzerland, 2020.
12. *ISO 12099:2017*; Animal Feeding Stuffs, Cereals and Milled Cereal Products-Guidelines for the Application of Near Infrared Spectrometry, 2nd ed. International Organization for Standardization: Geneva, Switzerland, 2017.
13. Unuvar, A.; Boyaci, I.H.; Yazar, S.; Koksel, H. Rapid Detection of Common Wheat Flour Addition to Durum Wheat Flour and Pasta Using Spectroscopic Methods and Chemometrics. *J. Cereal Sci.* **2023**, *109*, 103604. [CrossRef]
14. Liang, K.; Huang, J.; He, R.; Wang, Q.; Chai, Y.; Shen, M. Comparison of Vis-NIR and SWIR Hyperspectral Imaging for the Non-Destructive Detection of DON Levels in Fusarium Head Blight Wheat Kernels and Wheat Flour. *Infrared Phys. Technol.* **2020**, *106*, 103281. [CrossRef]
15. Tyska, D.; Mallmann, A.; Gressler, L.T.; Mallmann, C.A. Near-Infrared Spectroscopy as a Tool for Rapid Screening of Deoxynivalenol in Wheat Flour and Its Applicability in the Industry. *Food Addit. Contam. Part A Chem. Anal. Control. Expo. Risk Assess.* **2021**, *38*, 1958–1968. [CrossRef]
16. Chepkoech, B.; Wafula, E.N.; Sila, D.N.; Orina, I.N. Prediction of Retinol in Fortified Maize Flour Using Fourier Transform-Near Infrared Spectroscopy. *Curr. Res. Nutr. Food Sci.* **2024**, *12*, 384–396. [CrossRef]
17. Sitorus, A.; Lapcharoensuk, R. Exploring Deep Learning to Predict Coconut Milk Adulteration Using FT-NIR and Micro-NIR Spectroscopy. *Sensors* **2024**, *24*, 2362. [CrossRef] [PubMed]
18. Cattaneo, T.M.P.; Holroyd, S.E. The Use of near Infrared Spectroscopy for Determination of Adulteration and Contamination in Milk and Milk Powder: Updating Knowledge. *J. Near Infrared Spectrosc.* **2013**, *21*, 341–349. [CrossRef]
19. Balabin, R.M.; Smirnov, S.V. Melamine Detection by Mid- and near-Infrared (MIR/NIR) Spectroscopy: A Quick and Sensitive Method for Dairy Products Analysis Including Liquid Milk, Infant Formula, and Milk Powder. *Talanta* **2011**, *85*, 562–568. [CrossRef]
20. Smirnov, S.V. Neural Network Based Method for Melamine Analysis in Liquid Milk. In Proceedings of the 2011 Fourth International Conference on Intelligent Computation Technology and Automation (ICICTA 2011), Shenzhen, China, 28–29 March 2011; Volume 2, pp. 999–1002. [CrossRef]
21. Qian, S.; Qiao, L.; Xu, W.; Jiang, K.; Wang, Y.; Lin, H. An Inner Filter Effect-Based near-Infrared Probe for the Ultrasensitive Detection of Tetracyclines and Quinolones. *Talanta* **2019**, *194*, 598–603. [CrossRef] [PubMed]
22. Basangar, S.; Kumar, A.; Pal, D.; Jha, K.; Kumar, V. Enhanced Detection of Milk Fat Adulteration Using SMS Fiber Sensor. *Fiber Integr. Opt.* **2024**, *44*, 124–142. [CrossRef]

23. Yin, M.; Liu, C.; Ge, R.; Fang, Y.; Wei, J.; Chen, X.; Chen, Q.; Chen, X. Paper-Supported near-Infrared-Light-Triggered Photoelectrochemical Platform for Monitoring Escherichia Coli O157:H7 Based on Silver Nanoparticles-Sensitized-Upconversion Nanophosphors. *Biosens. Bioelectron.* **2022**, *203*, 114022. [CrossRef]
24. Scholl, P.F.; Bergana, M.M.; Yakes, B.J.; Xie, Z.; Zbylut, S.; Downey, G.; Mossoba, M.; Jablonski, J.; Magaletta, R.; Holroyd, S.E.; et al. Effects of the Adulteration Technique on the Near-Infrared Detection of Melamine in Milk Powder. *J. Agric. Food Chem.* **2017**, *65*, 5799–5809. [CrossRef]
25. Henn, R.; Kirchler, C.G.; Grossgut, M.E.; Huck, C.W. Comparison of Sensitivity to Artificial Spectral Errors and Multivariate LOD in NIR Spectroscopy—Determining the Performance of Miniaturizations on Melamine in Milk Powder. *Talanta* **2017**, *166*, 109–118. [CrossRef]
26. Mabood, F.; Jabeen, F.; Ahmed, M.; Hussain, J.; Al Mashaykhi, S.A.A.; Al Rubaiey, Z.M.A.; Farooq, S.; Boqué, R.; Ali, L.; Hussain, Z.; et al. Development of New NIR-Spectroscopy Method Combined with Multivariate Analysis for Detection of Adulteration in Camel Milk with Goat Milk. *Food Chem.* **2017**, *221*, 746–750. [CrossRef] [PubMed]
27. Fernández Pierna, J.A.; Abbas, O.; Lecler, B.; Hogrel, P.; Dardenne, P.; Baeten, V. NIR Fingerprint Screening for Early Control of Non-Conformity at Feed Mills. *Food Chem.* **2015**, *189*, 2–12. [CrossRef] [PubMed]
28. Huan, J.; Liu, Q.; Fei, A.; Qian, J.; Dong, X.; Qiu, B.; Mao, H.; Wang, K. Amplified Solid-State Electrochemiluminescence Detection of Cholesterol in near-Infrared Range Based on CdTe Quantum Dots Decorated Multiwalled Carbon Nanotubes@reduced Graphene Oxide Nanoribbons. *Biosens. Bioelectron.* **2015**, *73*, 221–227. [CrossRef] [PubMed]
29. Huang, Y.; Min, S.; Duan, J.; Wu, L.; Li, Q. Identification of Additive Components in Powdered Milk by NIR Imaging Methods. *Food Chem.* **2014**, *145*, 278–283. [CrossRef]
30. Vanstone, N.; Moore, A.; Martos, P.; Neethirajan, S. Detection of the Adulteration of Extra Virgin Olive Oil by Near-Infrared Spectroscopy and Chemometric Techniques. *Food Qual. Saf.* **2018**, *2*, 189–198. [CrossRef]
31. Sarabia, L.; Ortiz, M.C. DETARCHI: A Program for Detection Limits with Specified Assurance Probabilities and Characteristic Curves of Detection. *TrAC Trends Anal. Chem.* **1994**, *13*, 1–6. [CrossRef]
32. Ortiz, M.C.; Sarabia, L.A.; Herrero, A.; Sánchez, M.S.; Sanz, M.B.; Rueda, M.E.; Giménez, D.; Meléndez, M.E. Capability of Detection of an Analytical Method Evaluating False Positive and False Negative (ISO 11843) with Partial Least Squares. *Chemom. Intell. Lab. Syst.* **2003**, *69*, 21–33. [CrossRef]
33. Ortiz, M.C.; Sarabia, L.A.; Sánchez, M.S. Tutorial on Evaluation of Type I and Type II Errors in Chemical Analyses: From the Analytical Detection to Authentication of Products and Process Control. *Anal. Chim. Acta* **2010**, *674*, 123–142. [CrossRef]
34. AOTECH. Advanced Optical Technologies S.L. Available online: https://www.aotech.es/ (accessed on 16 April 2025).
35. Wise, B.M.; Gallagher, N.B.; Bro, R.; Shaver, J.M.; Winding, W.; Koch, R.S. PLS_TOOLBOX 9.1.1. 2022. Available online: https://eigenvector.com/software/pls-toolbox/ (accessed on 25 April 2025).
36. The mathworks I. MATLAB, MATLAB Version: 9.9.0 (R2020b). 2020. Available online: https://mathworks.com (accessed on 25 April 2025).
37. *International Standard ISO 11843-2*; Capability of Detection. Methodology in the Linear Calibration Case. International Organization for Standardization: Geneva, Switzerland, 2000.
38. Inczedy, J.; Lengyel, T.; Ure, A.M.; Gelencser, A.; Hulanicki, A. *IUPAC, Compendium of Analytical Nomenclature*; Blackwell: Hoboken, NJ, USA; Oxford, UK, 1998; pp. 31–40.
39. EU Commission. Implementing Regulation (EU) 2021/808 of 22 March 2021 on the Performance of Analytical Methods for Residues of Pharmacologically Active Substances Used in Food-Producing Animals and on the Interpretation of Results as Well as on the Methods To. *Off. J. Eur. Union* **2021**, *180*, L180/84–L180/109.
40. Rinnan, Å.; van den Berg, F.; Engelsen, S.B. Review of the Most Common Pre-Processing Techniques for near-Infrared Spectra. *TrAC Trends Anal. Chem.* **2009**, *28*, 1201–1222. [CrossRef]
41. Loss, F.P.; da Cunha, P.H.; Rocha, M.B.; Zanoni, M.P.; de Lima, L.M.; Nascimento, I.T.; Rezende, I.; Canuto, T.R.P.; de Paula Vieira, L.; Rossoni, R.; et al. Skin Cancer Diagnosis Using NIR Spectroscopy Data of Skin Lesions in Vivo Using Machine Learning Algorithms. *Biocybern. Biomed. Eng.* **2024**, *44*, 824–835. [CrossRef]
42. Huang, J.; Huang, J.; Lin, H.; Luo, D.; Kang, W. A GAN-Based Data Augmentation Method for Palm Vein Authentication. In Proceedings of the Biometric Recognition 18th Chinese Conference, CCBR 2024, Nanjing, China, 22–24 November 2024; Yu, S., Jia, W., Shu, X., Yuan, X., Gui, J., Tang, J., Eds.; Part I, Lecture Notes in Computer Science. Springer: Singapore, 2025; Volume 15352, pp. 142–152. [CrossRef]
43. Tai, M.H.; Hsu, C.C. Reconstructing Spectral Shapes with GAN Models: A Data-Driven Approach for High-Resolution Spectra from Low-Resolution Spectrometers. *Chemom. Intell. Lab. Syst.* **2025**, *258*, 105333. [CrossRef]