Microchemical Journal 204 (2024) 111039

Contents lists available at ScienceDirect

Microchemical Journal

journal homepage: www.elsevier.com/locate/microc

Applying multivariate curve resolution modelling combined with discriminant tools on near-infrared spectra for distinguishing between cheese varieties and stages of ripening

Elísabet Martín-Tornero [a,*], Isabel Durán-Merás [a], Mirta R. Alcaraz [a,b,c,*], Arsenio Muñoz de la Peña [a], Teresa Galeano-Díaz [a], Héctor C. Goicoechea [a,b,c]

[a] Departamento de Química Analítica, Facultad de Ciencias, Universidad de Extremadura, Avenida de Elvas, s/n, 06006 Badajoz, Spain
[b] Laboratorio de Desarrollo Analítico y Quimiometría (LADAQ), Cátedra de Química Analítica I, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, Ciudad Universitaria, 3000, Santa Fe, Argentina
[c] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Godoy Cruz 2290 CP C1425FQB, Buenos Aires, Argentina

ARTICLE INFO

Keywords:
Multivariate curve resolution
Torta del Casar
Queso de la Serena
Linear discriminant analysis
Quadratic discriminant analysis
Artificial neural networks

ABSTRACT

In this study, near-infrared (NIR) spectra were employed to monitor the ripening process of two kinds of soft cheese produced in the Extremadura region of Spain, manufactured by two different producers, "Torta del Casar" and "Queso de la Serena". Spectra were collected from the interior of the cheeses and the rind and analysed using appropriate chemometric techniques to distinguish between the two varieties and among different weeks of the maturation process. Different chemometric tools, including multivariate curve resolution with alternating least-squares (MCR-ALS), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), and feed-forward artificial neural networks (FF-ANN), were utilised, resulting in outstanding discrimination outcomes with sensitivity, precision, specificity, and accuracy reaching values c.a. 1.00 in optimal scenarios. More comprehensive information was acquired from the rind spectra analysis, indicating that the sampling process can be performed without disturbing the cheese in a non-destructive way. Remarkably, the capability to distinguish between various weeks of ripening for both cheeses could enable manufacturers to produce market-ready products earlier than the typically established timeline.

1. Introduction

Two renowned cheeses, crafted in the Extremadura region, Spain, using raw ewe milk, "Torta del Casar" (TC, from Cáceres) and "Queso de la Serena" (QS, from Badajoz), are distinguished by the prestigious protected designation of origin (PDO). These cheeses represent the essence of culinary excellence and cultural heritage and are a part of the appreciated soft to semi-soft Spanish cheese tradition. Controlled under specific regulations (EC, 1491/2003 for "Torta del Casar" [1] and 1107/96 for queso de la Serena [2]) these cheeses adhere to rigorous standards, ensuring authenticity and preserving the rich tradition of cheese-making in Extremadura.

During the ripening process of this kind of cheese, only a vegetable coagulant extracted exclusively from Cynara cardunculus is used, avoiding any additional starter culture. With the exclusive use of Merino ewe raw milk and vegetable rennet, this process bestows their distinctive flavour profiles and textures. Throughout this process, the cheese fats and proteins undergo a series of physical and chemical transformations, ultimately shaping the sensory characteristics of the finished product.

These cheeses are usually consumed after 60 days of ripening. However, an extended ageing period, generally up to 120 days, allows their flavours to deepen and complexities to unfold. In this regard, implementing online monitoring techniques to track the evolving composition of cheese constituents is a valuable tool for comprehending the ripening process. This approach not only aids in characterising cheeses but also enables effective discrimination between various varieties based on their unique development trajectories [3].

Near-infrared (NIR) spectroscopy could be considered one of the most broadly used process analytical technologies (PAT) in the dairy industry since it can detect molecular overtone and combination vibrations associated with C-H, N-H, O-H and S-H that are the most common groups in food constituent molecules [4]. Particularly in cheese

* Corresponding authors.
E-mail addresses: elisabetmt@unex.es (E. Martín-Tornero), malcaraz@fbcb.unl.edu.ar (M.R. Alcaraz).

https://doi.org/10.1016/j.microc.2024.111039
Received 21 May 2024; Received in revised form 18 June 2024; Accepted 19 June 2024
Available online 20 June 2024

manufacturing processes, NIR spectroscopy can be implemented to monitor parameters like moisture, protein, fat, and salt contents at various production stages. The strengths of this technique include its non-destructive nature, rapid analysis capabilities, cost-effectiveness, and ability to provide comprehensive compositional analysis.

In food analysis, pattern recognition methods emerge as a powerful tool that provides analysts with a qualitative or semi-quantitative binary response that aids the characterisation of samples. Such an approach does not entail quantifying individual chemical species but rather modelling measured data, compressing a segment of the composition of samples.

Pattern recognition techniques are categorised into two main subcategories: unsupervised and supervised methods. Unsupervised pattern recognition methods discover hidden patterns or structures in data without being told what to look for. On the other hand, supervised approaches teach to recognize patterns by providing labeled examples. They can be classified into discriminating or one-class modelling methods [5]. In particular, discriminant models are distinguished from class modelling methods by establishing linear or nonlinear boundaries between the modelled classes. Consequently, these techniques excel in addressing issues where all pertinent classes are clearly defined and adequately sampled, with balanced representation among them.

When developing a pattern recognition method, preliminarily assessing whether the data can effectively differentiate samples from different target classes is crucial. In addition, establishing the appropriate preprocessing data protocol and identifying outlier samples is an essential step in the method development. Principal component analysis (PCA) emerges as a simple, albeit reliable, exploratory method to evaluate the classificatory potential of the data.

One of the first stages of pattern recognition methods in multidimensional data analysis is data compression, PCA being the most utilised approach for this goal [6]. However, while increasing the complexity and the order of the data, other strategies can be implemented to compress or select variables, thereby reducing the data volume. For instance, the successive projection algorithm (SPA) is utilised when the number of variables must be reduced due to mathematical requirements [7]. Moreover, when second-order data are utilised, multivariate curve resolution with alternating least-squares (MCR-ALS) [8] or parallel factor analysis (PARAFAC) emerges as strategic approaches for data compression, with the advantage of getting chemically interpretable outcomes [9]. Pertinent examples of using MCR-ALS to model first-order data can be found in Ref. [10–12]. When using these algorithms, two kinds of outcomes are obtained: on the one hand, the scores are directly related to the contribution of the individual constituents, and on the other, the loadings are directly related to the chemical behaviour of the constituents, e.g., the spectral profiles. The latter aids in gaining deeper insights into the system and comprehending better its physicochemical characteristics. Moreover, leveraging information derived from concentration variations throughout a process has proven to be a potent tool for effectively modelling discrimination among various samples [13].

In supervised methods, the subsequent phase involves deciding whether to adopt a discriminating or modelling approach. Notably, considerable advancements have been achieved in distinguishing discriminant methods from modelling techniques within the domain of classification models. Discriminant methods, which employ multiclass (two or more) during the training phase, tend to exhibit greater effectiveness than stringent one-class models in generating models with higher specificity and sensitivity. Some of the most common discriminant methods are linear discriminant analysis (LDA) [14], Partial least-squares discriminant analysis (PLS-DA) [15] and quadratic discriminant analysis (QDA) [16] which follow a specific model, i.e., linear and quadratic, respectively. Nevertheless, in cases where no specific model is followed, a nonparametric approach, such as artificial neural networks (ANN), emerges as an alternative [17,18].

In the present work, first-order NIR data were subjected to MCR-ALS analysis. The resulting outcomes served as inputs for various discriminant methods, including LDA, QDA, and feed-forward artificial neural networks (FF-ANN). These methods were employed to discriminate between the ripening stage and the origin of two PDO cheeses.

2. Materials and methods

2.1. Samples

Cheese samples were sourced from two distinct dairies representing the PDO regions in Extremadura, Spain: “Torta del Casar” PDO cheese, from Cáceres, and “Queso de la Serena” PDO cheese, from Badajoz.

For each type of soft cheese, a batch was randomly selected, and three samples were collected weekly throughout the ripening process until reaching day 60, considered suitable for commercialisation. In total, 30 samples of each cheese type were analysed.

2.2. Apparatus

A portable MicroNIR Pro 1700 (VIAVI, Santa Rosa, California, USA) equipped with two small tungsten light bulbs as a radiation source and a linear-variable filter (LVF) directly connected to a linear indium gallium arsenide (InGaAs) array detector was used for NIR spectra acquisitions.

Before spectra acquisition, an automatic dual-point calibration was accomplished using a Spectralon ceramic material as the white reference (100 % reflectance) and air as the dark current (0 % reflectance), placing the device, in this case, no closer than 0.5 m from any object.

The spectra acquisition was carried out in reflectance mode between 1100 and 1700 nm, with a 6 nm spectra resolution. A total of 50 scans were averaged for each spectrum. The measurements were made using the windowed collar as a protective accessory with a sapphire window placed at ~ 3 mm distance to the MicroNIR input to provide a consistent sample presentation in non-rigid materials.

Spectra were recorded using MicroNIR^{TM} Pro v.2.3 software (VIAVI, Santa Rosa, California, USA).

2.3. Sample spectra acquisition

The cheeses were analysed immediately after arrival at the laboratory. For each cheese, spectra were recorded at 10 different points of the rind and the centre of the cheese, as depicted in Fig. 1, performing several replicates. For the rind analysis, the spectra were directly acquired from the top of the cheese without the intervention of the piece, whereas to analyse the centre of the cheese, an approximately 1 cm-thick slice was extracted from the middle of the piece.

Spectra were collected over 7 weeks for TC (from day 1 to day 57 of the ripening process) and 10 weeks for QS (from day 1 to day 64 of the ripening process). This resulted in 517 NIR spectra obtained from the interior of TC samples, 402 from the rind of TC samples, 555 from the interior of QS samples, and 569 from the rind of QS samples. The fact of taking a number of different weeks and a variable number of spectra during the ripening process leads to the irregularity of the dimensions of the matrices generated to process the data.

2.4. Chemometric analysis

Data analysis was accomplished in MATLAB R2016B (The MathWorks Inc., USA). LDA, QDA, and FF-ANN were applied using the Classification toolbox for MATLAB freely available at https://michem.unimib.it/download/matlab-toolboxes/classification-toolbox-for-matlab/ [15]. MCR-ALS 2.0 toolbox was utilised for MCR-ALS analysis [8], which was freely downloaded from https://mcrals.wordpress.com/download/mcr-als-2–0-toolbox/. The SPA algorithm was implemented with the MATLAB codes acquired upon request from the authors [7]. Data preprocessing was implemented using codes written in MATLAB in our lab.

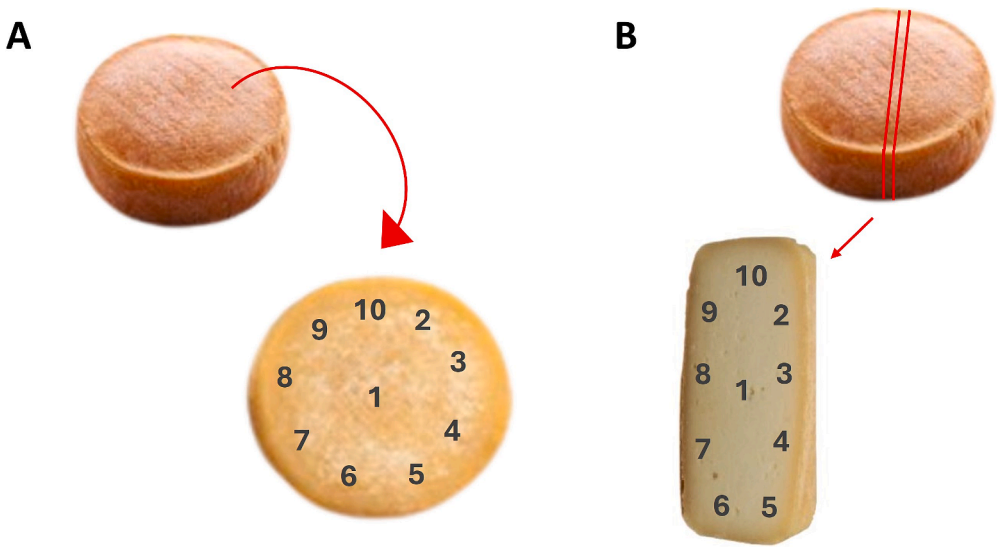

Fig. 1. Spatial zones of the cheese where spectra were recorded: (A) on the rind and (B) on a central slice of the product.

3. Theory

3.1. Multivariate curve resolution-alternating least squares

MCR-ALS is a family of algorithms that perform a bilinear decomposition of a two-dimensional matrix **D** ($I \times J$) containing the chemical responses of the system into the product of the two sub-matrices, **C** and $\mathbf{S}^{\mathrm{T}}$, as follows:

$$\boldsymbol{D} = \boldsymbol{C}\boldsymbol{S}^{T} + \boldsymbol{E} \tag{1}$$

In spectroscopy, this expression is tied to the premise of Lambert-Beer's law; thus, **C** and $\mathbf{S}^{\mathrm{T}}$ would contain information about the system constituent's concentration and spectral behaviour, respectively.

In the particular case of the NIR monitoring of the ripening process of a cheese, the matrix **D** comprises the NIR spectra (row) acquired for each ripening stage (columns). Hence, the columns of **C** ($I \times A$) would correspond to the abundance of the A individual components along the ripening process and the rows of $\mathbf{S}^{\mathrm{T}}$ ($A \times J$), the NIR spectral profiles of the A individual components. **E** ($I \times J$) is a matrix containing the models' residuals in all cases.

3.2. Linear and quadratic discriminant analysis

LDA and QDA are boundary discriminant methods used in machine learning and statistics, which aim to find boundaries that separate groups or classes of samples.

The LDA algorithm computes a separating surface between sample groups by establishing a linear discriminant function that maximises the ratio of between-class and within-class variances [14]. Classes are assumed to adhere to a multivariate normal distribution and be linearly separable.

For the **X** matrix of dimensions $I \times J$, which contains the J variables measured on I samples, and the dummy matrix **Y** of dimensions I (samples) $\times$ K (number of categories), the optimal representation is achieved by maximising the ratio of the between-class variance (**Bc**) matrix to the within-class variance (**Wc**) matrix. The latter matrices can be expressed as:

$$\boldsymbol{Bc} = (k-1)^{-1}\boldsymbol{A}^{T}\boldsymbol{Y}\left(\boldsymbol{Y}^{T}\boldsymbol{Y}\right)^{-1}\boldsymbol{Y}^{T}\boldsymbol{X} \tag{2}$$

$$\boldsymbol{Wc} = (I-k)^{-1}\left[\boldsymbol{X}^{T}\boldsymbol{X} - (k-1)\boldsymbol{Bc}\right] \tag{3}$$

It can be demonstrated that the canonical variate (CV) scores encapsulate the progressively maximised ratio of between-groups variance to within-groups variance, which are obtained by principal component analysis of the matrix ($\mathbf{Wc}^{-1}\mathbf{Bc}$) and projection of the data matrix **X** onto the first loadings.

The samples are subsequently visualised in a two or three-dimensional space delineated by the first canonical variate (CV) scores for each sample. It is imperative to acknowledge that the dimensionality of matrix **X** is $I \times J$. Consequently, before employing LDA, variable selection (with SPA for example) or data compression (by PCA or MCR-ALS), as will be shown below, and utilisation of scores should be executed to ensure that the condition $I > J$ is met.

Unlike LDA, QDA does not assume equal covariance matrices across classes, allowing for more flexibility in modelling further complex relationships between variables. Thus, each class is modelled by its multivariate normal distribution, estimating parameters from a set of training data. QDA captures nonlinear relationships between predictors, being more flexible than LDA.

The QDA classification score (Qij) is estimated using the variance–covariance matrix for each class k and an additional natural logarithm term, as follows:

$$Q_{ik} = (\mathbf{x}_i - \overline{\mathbf{x}}_k)^{\mathrm{T}}\Sigma_k^{-1}(\mathbf{x}_i - \overline{\mathbf{x}}_k) + log_e|\Sigma_k| - 2log_e\pi_k \tag{4}$$

where $\mathbf{x}_i$ is an unknown measurement vector for a sample, $\overline{\mathbf{x}}_k$ is the average measurement vector of class k, $\boldsymbol{\Sigma}_k$ is the variance–covariance matrix of class k, and $log_e|\boldsymbol{\Sigma}_k|$ is the natural logarithm of the determinant of variance–covariance matrix $\boldsymbol{\Sigma}_k$. The prior probability (π_k), pooled covariance matrix ($\boldsymbol{\Sigma}_{pooled}$) and variance–covariance matrix ($\boldsymbol{\Sigma}_k$) are calculated as follows [16].

$$\pi_k = \frac{N_k}{N} \tag{5}$$

$$\Sigma_{pooled} = \tfrac{1}{N}\textstyle\sum_{k=1}^{K} N_k\Sigma_k \tag{6}$$

$$\Sigma_k = \frac{1}{N_k}\sum_{i=1}^{N_k}(\mathbf{x}_i - \overline{\mathbf{x}}_k)(\mathbf{x}_i - \overline{\mathbf{x}}_k)^{\mathrm{T}} \tag{7}$$

where N_k is the number of objects of class k, N is the total number of objects in the training set, and K is the total number of categories, as was commented above.

3.3. Feed-forward artificial neural networks

FF-ANN emulate the functioning of neurons in the human brain, making them a powerful tool in machine learning. They are structured

into layers, each comprising interconnected nodes with activation functions. Input vectors are fed into the network through the input layer, which then passes them through one or more hidden layers. Here, the actual computation occurs through a network of weighted connections. Various learning rules are employed in neural networks, with back-propagation (BP) being the focus of this study, in which the learning is supervised and unfolds with each epoch, where the network processes new input patterns through forward activation flow and adjusts weights via backward error propagation [17].

The fundamental approach of BP learning involves updating network weights and biases along the direction where the performance function decreases most rapidly, namely the negative gradient. The following equation succinctly captures one iteration of this algorithm:

$$\boldsymbol{w}_{k+1} = \boldsymbol{w}_k - \alpha_k \boldsymbol{g}_k \tag{8}$$

where $\mathbf{w}_k$ is a vector containing the current weights and biases, $\boldsymbol{g}_k$ is the current gradient, and α_k is the learning rate. In this study, gradient descent with momentum was utilised, applying the mean square error (MSE) as the performance function. The latter figure quantifies the average squared discrepancy between the network outputs and the actual output, serving as a crucial metric for optimisation.

In the basic gradient descent algorithm, adjustments to weights and biases align with the negative gradient of the performance function. Incorporating momentum into gradient descent often accelerates convergence. Momentum enables the network to not only respond to local gradients but also capture recent trends in the error surface, facilitating navigation through shallow local minima towards the global minimum. To integrate momentum into back-propagation learning, weight updates are computed as the sum of a fraction of the previous weight change and the new change suggested by the back-propagation rule. In this work, the inputs to feed the FF-ANN were the scores obtained by the MCR-ALS procedure described above, unlike what is usually done using PCA scores. However, the networks were also trained using PCA scores, but the results (not shown) were of lower quality than those obtained using the MCR-ALS scores, a fact that was repeated when using LDA and QDA.

3.4. Successive projection algorithm

As commented above, given the **X** matrix of dimensions $I \times J$, which contains the J variables measured on I samples, variable selection should be executed before employing LDA or QDA to ensure that the condition $I > J$ is met. This variable selection process starts by considering each of the J variables in the dataset **X**. Then an ordered chain of variables Kc is built, ensuring that each selected variable minimises collinearity with the previously chosen ones. Collinearity is assessed by calculating the correlation between the column vectors of matrix **X**. Importantly, a constraint is set: no more than Kc variables can be included in the chain. From each of the J chains, it is possible to extract Kc subsets of variables. These subsets consist of one up to Kc elements, arranged in the order they were selected during the variable selection process. Hence, a total of $J \times Kc$ subsets of variables can be generated. To determine the optimal subset, a cost function represented by the average risk G of misclassification through LDA when utilising the evaluated subset of variables, is employed. For further elaboration, please refer to Ref. [7].

3.5. Figures of merit

Classification outcomes can be depicted using a confusion matrix, a square matrix with dimensions ($K \times K$), where K represents the number of classes. This matrix provides insights into the correspondence between actual and predicted classifications, assuming a classifier that assigns each sample exclusively to one of the available classes. The element nKk in the matrix denotes the count of samples belonging to class K and assigned to class k. Diagonal elements (e.g., nKk) indicate correct predictions, while off-diagonal elements signify misclassifications. Binary classification, prevalent in numerous scenarios, simplifies complex problems into yes/no outcomes. In this context, samples are usually categorised as positive or negative, condensing the confusion matrix into a succinct 2×2 numerical table, as will be observed when discriminating between both types of cheeses. On the other hand, when discriminating between sampling weeks, confusion matrices of 7×7 for TC and 10×10 for QS were generated (see below).

Various established class indices, such as sensitivity, specificity, and precision, can be derived from this matrix. These metrics offer insights into the classification performance for each individual class. However, it is crucial not to rely solely on any single measure to assess the predictive capability of a model. Since each metric captures different aspects of the overall classification performance, simultaneously evaluating them is essential for a comprehensive assessment of classification quality [19]. Figures of merit are computed using the following equations:

a) Sensitivity (Sn): is also identified as true positive rate (TPR), and defined as the ratio between true positive (TP) and the total number of positive samples ($TP + FN$), where FN is the number of false negative samples:

$$Sn = \frac{TP}{TP + FN} \tag{9}$$

b) Specificity (Sp): is also known as true negative rate (TNR), and termed as the ratio between true negative (TN) and the total number of negative samples ($TN + FP$), where FP is the number of false positive samples:

$$Sp = \frac{TN}{TN + FP} \tag{10}$$

c) Precision (Pr): is also known as positive predictive rate (PPR), and calculated as the ratio of TP and the total number of samples predicted as positive:

$$Pr = \frac{TP}{TP + FP} \tag{11}$$

d) Accuracy (Acc): is defined as the ratio of the sum of TP and TN over the total number of samples:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{12}$$

Three widely recognised class-specific metrics (sensitivity, specificity, and precision) serve to gauge the classification performance within each class. Computed individually for each class, these metrics encapsulate distinct facets of classification accuracy. Sensitivity measures a classifier's capability to accurately identify.

The precision of a class reflects its purity, indicating the classifier's proficiency in minimising erroneous predictions within that specific class. Conversely, the specificity of a class signifies the classifier's aptitude in discriminating against samples from other classes.

4. Results

4.1. NIR spectra and preprocessing

One of the most common artefacts on NIR spectra is the scattering effects caused by sample heterogeneity or particle size variations. These effects introduce undesired variations in the data unrelated to the system's chemical properties. Therefore, data needs to be subjected to some corrections before data modelling to remove or minimise the effects of these artefacts. The most common preprocessing procedure utilised in NIR spectra to achieve this goal is the multiplicative scattering correction (MSC). Fig. 2 shows the spectra obtained from the centre of three samples (see Fig. 1B) of QS at the beginning of the ripening process before and after applying MSC. As can be seen, the preprocessing corrects the dispersions observed in Fig. 2A, giving more homogeneity to

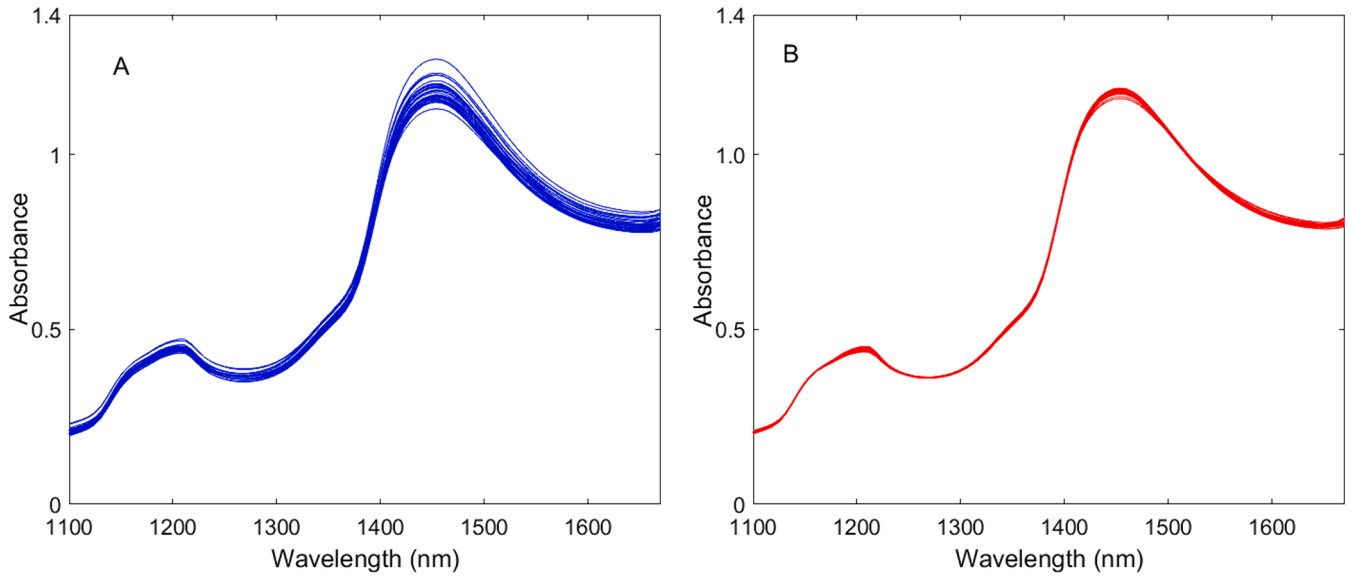

Fig. 2. . Several spectra gathered in different parts of the centre of three samples of Queso de la Serena at the beginning of the ripening process. A) Raw spectra. B) After application of MSC.

the set of spectra taken on the same day of the maturation process.

An analysis of the pretreated spectra allowed us to match with reports that shed light on the composition of the cheeses. According to da Silva Medeiros et al. [3], the different absorption peaks present in the NIR spectra obtained for cheeses can be attributed to a) O–H stretching in the second and first overtone of the water band (1450 nm), b) C – H stretching in the first and second overtone of aliphatic chains of fat (CH, CH2, CH3; 1130 – 1240 nm) and unsaturated fatty acids (HC = CH; 1400 and 1660 nm), and c) N – H elongation in the first overtone of proteins (1510 nm).

4.2. MCR-ALS modelling

The initial step of the process consisted of building up four augmented column-wise matrices **D** from the individual spectra by setting one on top of the other and keeping the column vector space in common. This procedure was carried out for the four sets built with signals registered in the centre and on the surface of two cheeses. The dimensions of these data matrices were a) TC, inside (402 × 96) and surface (517 × 96); and b) QS, inside (555 × 96) and surface (569 × 96). From now on, these matrices will be called $\mathbf{D}_{TCc}$, $\mathbf{D}_{TCr}$, $\mathbf{D}_{QSr}$ and $\mathbf{D}_{QSc}$, respectively.

The first step of MCR-ALS analysis involves determining the number of components that explain an acceptable level of total variance within the four individual **D** matrices described above. In this study, a compromise was made by selecting 99 % of the total explained variance, balancing the number of components against computational time considerations. The number of components in each matrix was determined using the singular value decomposition (SVD) algorithm. This process suggested 4 principal components (PCs) for $\mathbf{D}_{TCc}$ and $\mathbf{D}_{QSc}$ and 5 PCs for $\mathbf{D}_{TCr}$ and $\mathbf{D}_{QSr}$. This fact primarily suggests some variations in the chemical composition of the different parts of the cheeses.

Spectral initial estimates obtained through the 'purest variables' algorithm, with a noise level set at 0.1 [20], were used to initiate the ALS procedure. Then, non-negativity constraints were applied to concentration and spectra during the iterative optimisation to obtain chemically meaningful solutions.

Figure SM1 displays a screen generated by the graphical interface, following the application of MCR-ALS on the $\mathbf{D}_{TCc}$ matrix containing 402 spectra recorded from the rind of a TC sample throughout the ripening process. As can be seen, five components were identified, showcasing varying trends.

An analysis of the MCR-ALS spectral profiles obtained for the rind of a TC cheese (Fig. 3) enables the identification of some specific bands related to previously mentioned compounds, whose variation during the cheese ripening process allows us to formulate the hypothesis that it is possible to discriminate between both types of cheese and, even more, among weeks of cheese involved in the process. For example, the band corresponding to water, which can be assumed to be the blue band (Fig. 3), has its centre at 1450 nm, corresponds to the blue concentration profile of Fig. SM1. As can be seen in the later figure, the concentration decreases with time, indicating the effect of the maturation process in making the product firmer.

After decomposition, the areas under the concentration profiles for each component at every sampling date were obtained from the MCR-ALS scores matrices. Each value represents the area of the *a*th component on the *i*th sampling day. Fig. 4 illustrates the progression of the five components during the ripening process obtained from the MCR-ALS

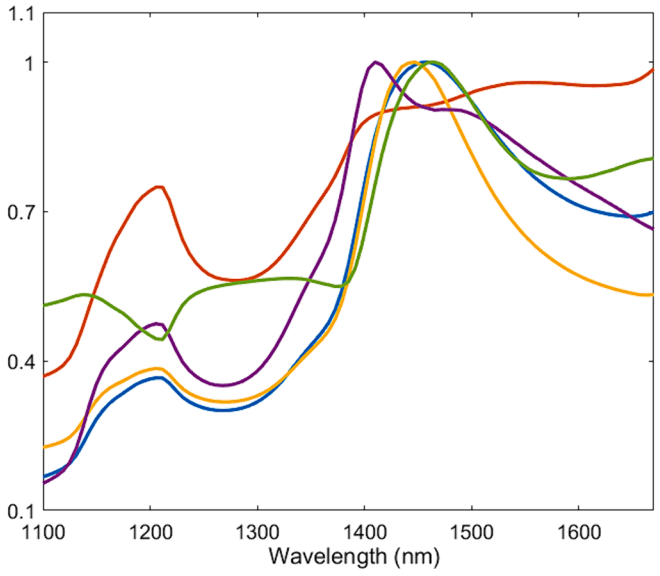

Fig. 3. Spectral profiles retrieved with MCR-ALS modelling, corresponding to NIR spectra taken on the surface of a Torta del Casar cheese sample.

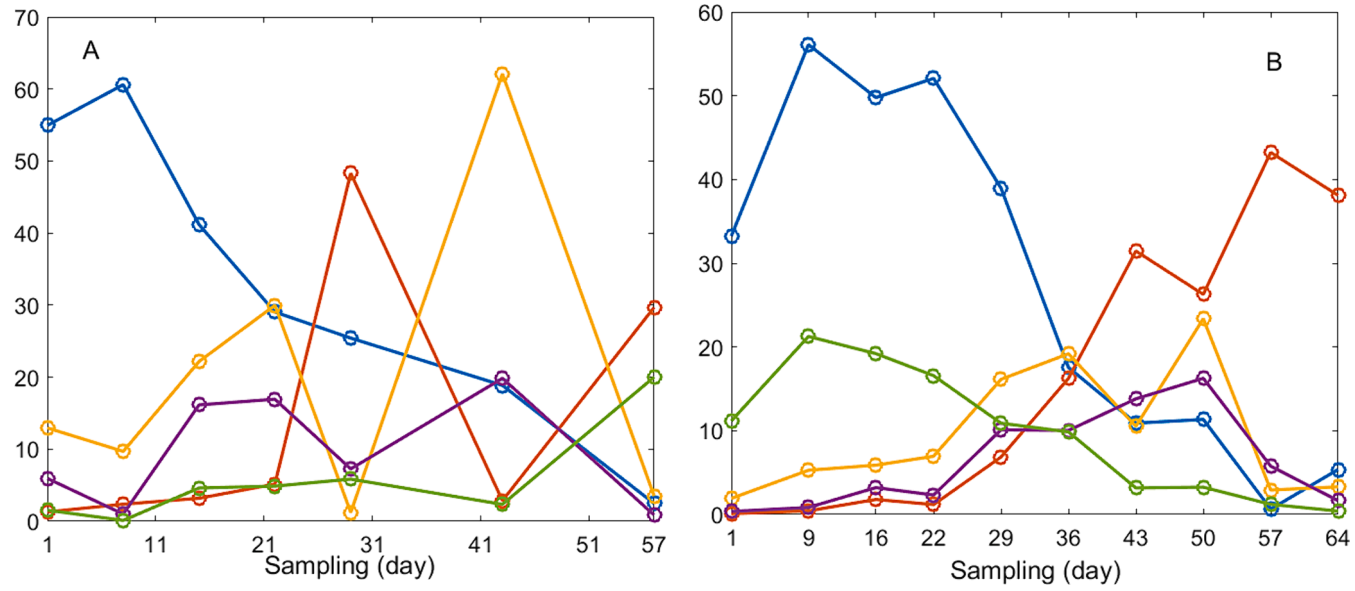

Fig. 4. Evolution of the five components during the ripening process extracted from the spectra recorded on the surface of both kinds of cheese (A: Torta de Casar and B: Queso de la Serena).

analysis of the data obtained from the rind of both cheeses. One evident fact is the rising of some components while others decrease with the ripening process of about 60 days, as commented above for the water.

A visual examination of the profiles depicted in Fig. 4A-B suggests that the discernible variability, evident in the evolution of the main compounds, could serve as crucial background knowledge. This variability may enable the differentiation between the two types of cheeses and among the different weeks of ripening when analysed using appropriate discrimination algorithms.

4.3. Application of an unsupervised pattern recognition tool: PCA

An unsupervised algorithm, PCA, was applied to perform an exploratory study of the data sets composed of corrected spectral values (after MSC application) and scores gathered by MCR-ALS modelling of the later spectra. Four matrices corresponding to TC were analysed: a) TC_spectra_inside (460 × 96 wavelengths), b) TC_spectra_surface (402 × 96 wavelengths), c) TC_MCR_4components_inside (460 × 4), and d) TC_MCR_5components_surface (402 × 5). Similarly, other four matrices corresponding to QS samples were subjected to PCA analysis: a) QS_spectra_inside (555 × 96 wavelengths), b) QS_spectra_surface (569 × 96 wavelengths), c) QS_MCR_4components_inside (555 × 4), and d) QS_MCR_5components_surface (569 × 5). The first three principal components in the PCA models captured c.a. 95 % of the total variance in all the cases. Figure SM2 shows the representation of the samples in the space defined by the PC1 vs. PC2 for the MCR-ALS scores of both cheeses corresponding to measurements on the surface. (TC = 402 samples, and QS = 569 samples). As evident, while there are slight differences between the two types of cheese, variations are noticeable across different weeks for both types. However, it should be kept in mind that an unsupervised tool is being applied, which only provides preliminary information, allowing us to think about the good performance of supervised models, especially to differentiate weeks of maturation. On the other hand, it is also evident that powerful tools should be applied when performing a supervised analysis with the diverse types of data used to differentiate between cheese varieties.

PCA analysis was also implemented to detect outliers, i.e. samples that are somewhat disturbing or unusual, which can be outright erroneous. In this regard, just one sample was detected to be an outlier. It was confirmed using the influence plot, which is obtained by plotting Q against Hotelling's T^2 residuals [6]. The first one is the sum squared residuals of each sample to look for samples that are not well-described by the PCA model. The second one is a diagnostic statistic that can be seen as an extension of the *t*-test and can also be applied to the scores of a PCA model. The mentioned plot allows us to detect samples that are far from the rest of the samples. In the present case, only two samples (QS # 337, and TC # 340) had to be removed. Fig. 3SM shows the influence plot in which TC samples are marked in red, while QS samples are marked in blue.

4.4. Application of discriminant tools to differentiate both cheese varieties

The Kennard-Stone (K-S) algorithm was applied to split the sets of spectra corresponding to samples taken both inside and on the surface of the cheeses. The splitting was performed to generate training and validation sets for testing the discrimination algorithms. To perform the partitioning, matrices of (1015 × 4) and (971 × 5) were built by pooling the MCR-ALS scores obtained by analysing the spectra of cheeses recorded inside and, on the surface, respectively. In the case of samples taken inside the cheeses, a training set of 750 samples was generated (350 for TC and 400 for QS). The validation set consisted of 265 samples (110 for TC and 155 for QS). On the other hand, in the case of cheese rind samples, a training set of 700 samples was created (300 for TC and 400 for QS). The validation set comprised 271 samples (102 for TC and 169 for QS). The procedure was repeated for matrices containing the complete spectra and the variables selected with the SPA algorithm (see below).

To implement the discrimination algorithms, four vectors were constructed containing the numbers 1 or 2 according to the class (1 for TC and 2 for QS). The dimensions of the two vectors for measurements inside of the cheeses were: training = 750 × 1, and validation = 265 × 1. The dimensions of the two vectors for measurements on the surface of the cheeses were: training = 700 × 1, and validation = 271 × 1.

To compare and evaluate the classification ability among the models built with LDA, QDA and FF-ANN using MCR-ALS scores, SPA variables and whole NIR spectra, the following figures of merit (see the description above) were analysed: sensitivity (*Sn*), specificity (*Sp*), precision (*Pr*) and accuracy (*Acc*). In addition, the evaluation of the confusion matrix obtained for training and prediction results in every case was considered.

For training evaluation, the Venetian blind cross-validation (VBCV) technique was implemented, which comprises splitting the dataset into

multiple partitions. The data is divided into a fixed number of folds (in this case, five), each of them containing an equal number of samples from each class present in the dataset. This guarantees that each partition is representative of the overall distribution of the data. After that, the model is trained on one partition and tested on the other, and this process is repeated for each pair of folds.

Table 1 displays the performance metrics achieved through the application of LDA, QDA, and FF-ANN. In the subsequent table, only the outcomes derived from the analysis of matrices constructed with MCR-ALS scores are provided. This is due to the unsatisfactory results obtained when employing modelling on matrices constructed with variables selected using SPA (six variables) or with the complete NIR spectra compressed with PCA, which are not shown here.

As depicted in the table, the outcomes achieved with LDA are unsatisfactory. However, a notable enhancement is observed with the application of the nonlinear QDA algorithm. This indicates the potential for further improvement through the utilisation of FF-ANN. Indeed, implementing FF-ANN yields excellent results. Notably, the most notable improvements are observed when modelling data generated from spectra collected from the surface of the cheeses. This enhancement is likely attributed to MCR-ALS modelling capturing an additional component on the cheese surface compared to the interior. This discovery holds significant interest as it implies that sample destruction for measurements may not be necessary, offering notable economic benefits, particularly for high-priced products such as these. The FF-ANN modelling was implemented using one hidden layer containing five neurons, a learning rate of 0.01 and a moment (alpha) equal to 0.5. The optimum model was reached after 300 epochs.

The confusion matrices for the three mentioned models are depicted in Table 2. It is evident that only the FF-ANN model enables a highly accurate differentiation between the two types of cheese.

NIR data were subjected to PCA, and PC scores were used to train the ANN. This approach served as a reference methodology aiming at proving the efficiency of the MCR-ALS-ANN strategy in the classification analysis. PCA-ANN and MCR-ALS-ANN classification results were compared, and they were highly similar in terms of classification efficiency. This comparison helped to assert that MCR-ALS provides useful information for classification purposes, with the addition of qualitative spectral details that enable a more comprehensive physicochemical evaluation of the system under study.

A previous study recently published reports a successful distinction between the two types of cheese according to the PDO by employing excitation-emission matrix (EEM) fluorescence spectroscopy with parallel factor analysis (PARAFAC) and LDA modelling techniques. Notably, the best precision in differentiation was reached in samples with 36 or more days of maturation [9]. Interestingly, in the work presented here, NIR spectra obtained inside and on the rind of two cheeses were used to differentiate between the two varieties. The outcomes of our study highlight that richer information is obtained when measurements are taken from the cheese rind, indicating a non-destructive sampling approach. Furthermore, it was possible to differentiate between weeks of ripening for both cheeses, which could allow the manufacturers to obtain products ready to be marketed before the date that is usually established.

It is worth mentioning that PLS-DA was used to model NIR data, as it is one of the most widely used algorithms for discrimination analysis with multivariate data. However, the results were similar to those obtained with LDA, shedding light on the fact that linear-based approaches were unsuitable for these systems.

4.5. Application of discriminant tools to differentiate among weeks of maturation

To discriminate between weeks of ripening, it was decided to work only with data obtained on the rind of the cheeses, since there is no need to destroy the sample and it adds more value to the process, especially from the point of view of cost and speed. In the case of TC, the 402 samples were divided into 280 and 122 for training and validation, respectively. In turn, as the aim was to discriminate between 7 different weeks, the training set consisted of 40 samples for each week, while the validation set was composed of 17, 18, 19, 17, 17, 17 and 17 samples, respectively.

In the case of QS, the 569 samples were split into 400 and 169, respectively. In turn, as in this case we wanted to discriminate between 10 different weeks, the training set was composed of 40 samples for each week, while the validation set comprised 7, 14, 19, 18, 20, 21, 18, 17, 17 and 18 samples for the consecutive weeks.

To implement the discrimination algorithms, four vectors were constructed: a) for TC, containing the numbers 1, 2, 3, 4, 5, 6 and 7 according to the seven weeks, with dimensions of (280 × 1) and (122 × 1) for training and validation sets, respectively; and b) for QS, containing the numbers 1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 according to the ten weeks, with dimensions of (400 × 1) and (169 × 1) for training and validation sets, respectively.

It is well documented that one of the most appropriate strategies for multi-class discrimination is to construct several models to discriminate one class against the rest, utilizing only 1 index for the target class and 0 indexes for the rest [21–23]. This strategy was implemented to discriminate between weeks by building 7 models for TC and 10 for QS. Here, it is worth mentioning that the Classification toolbox requires the information comprised in a vector, as mentioned above, to label the N classes, which then automatically constructs a dummy matrix containing 1 index for the target class and 0 index for the rest. For instance, in the 7-class system, the dummy matrix includes 7 columns (one for each class) using binary code to indicate the target class, e.g., the (0 0 1 0 0 0 0) codification refers to the class 3. Consequently, the model predicts output values between 0 and 1, which can be considered the likelihood of belonging to the target class. If the classification is successful, the outputs of samples belonging to a target class will be close to the class index. As seen in Figure SM4, all samples belonging to class 1 render outputs close to 1, samples of class 2 render outputs close to 2, and so on.

Table 1
Quality performance features for the modelling to discriminate between "Torta del Casar" (Class 1) and "Queso de la Serena" (Class 2).

Modelling Quality performance features[a]								
	Validation during the training				Validation with test set			
	Sn	*Sp*	*Pr*	*Acc*	*Sn*	*Sp*	*Pr*	*Acc*
	Cheese rind samples							
MCR- LDA	0.91/0.90	0.90/0.91	0.87/0.93	0.90	0.72/0.90	0.90/0.72	0.81/0.84	0.83
MCR- QDA	1.00/0.97	0.97/1.00	0.96/1.00	0.98	1.00/0.97	0.97/1.00	0.95/1.00	0.98
MCR- ANN	1.00/1.00	1.00/1.00	1.00/1.00	1.00	1.00/1.00	1.00/1.00	1.00/1.00	1.00
	Cheese inside samples							
MCR- LDA	0.57/0.95	0.95/0.57	0.90/0.72	0.77	0.37/0.99	0.99/0.37	0.95/0.72	0.75
MCR- QDA	0.57/0.94	0.94/0.57	0.90/0.71	0.77	0.22/0.99	0.99/0.22	0.96/0.68	0.70
MCR-ANN	0.96/0.94	0.94/0.96	0.94/0.97	0.95	0.93/0.99	0.99/0.93	0.99/0.96	0.97

[a] *Sn*: sensitivity; *Sp*: specificity; *Pr*: precision; *Acc*: accuracy. *Sn*, *Sp* and *Pr* results presented in the following format: Class 1/Class 2.

Table 2
Confusion matrix for the modelling to discriminate between "Torta del Casar" (Class 1) and "Queso de la Serena" (Class 2).

Modelling	Validation during the training				Validation with the test set			
	Real/ Predicted	Class 1	Class 2	Not assig [a]	Real/ Pred	Class 1	Class 2	Not assig [a]
	Cheese rind samples							
LDA-MCR	Class 1	272	28	0	Class 1	73	29	0
	Class 2	40	360	0	Class 2	17	152	0
QDA-MCR	Class 1	299	1	0	Class 1	102	0	0
	Class 2	12	388	0	Class 2	5	164	0
ANN-MCR	Class 1	300	0	0	Class 1	102	0	0
	Class 2	0	400	0	Class 2	0	169	0
	Cheese inside samples							
LDA-MCR	Class 1	199	151	0	Class 1	85	9	6
	Class 2	21	379	0	Class 2	4	160	1
QDA-MCR	Class 1	199	151	0	Class 1	87	8	5
	Class 2	21	379	0	Class 2	3	161	1
ANN-MCR	Class 1	333	13	4	Class 1	88	7	5
	Class 2	23	377	0	Class 2	1	164	0

[a] Not assig: not assigned to any class.

In this scenario, it is interesting to note that the data generated from NIR measurements, and subsequently processed with MCR/ALS, enables precise differentiation among the various weeks of ripening when modelled using the LDA algorithm. Surprisingly, even more robust indicators are achieved compared to those obtained with nonlinear algorithms such as QDA and ANN.

For the reasons mentioned above, Table 3 shows the confusion matrices corresponding to the modelling of test data of TC and QS with the LDA algorithm. Interestingly, the values computed for all the figures of merit, i.e. *Sn, Sp, Pr* and *Acc* for Class 1 to Class 7 in the case of TC were equal to 1.00. On the other hand, the values computed for the ten weeks of QS were not as favourable, particularly during the initial four weeks. As can be seen in Table 3, the performance is exceptional for TC. On the contrary, while the results for QS are not as conclusive, the lower accuracy during the initial phase of the process suggests that the outcomes remain acceptable. It is especially significant considering that the richness of information could potentially streamline the overall duration of the process. In other words, it is crucial to differentiate the final weeks effectively. These findings could significantly impact production processes. With straightforward NIR spectral measurements, a staple tool in food quality laboratories, it becomes feasible to ascertain the completion of the process and the readiness of the product for market release.

5. Conclusions

The current investigation showcases the efficacy of NIR spectroscopy, in conjunction with proper chemometric modelling, as a potent means to monitor soft cheeses like "Torta del Casar" PDO and "Queso de la Serena" PDO. Leveraging NIR spectra as unique sample fingerprints offers distinct advantages, notably in their facile acquisition, bypassing the need for sample treatment and facilitating real-time monitoring for effective quality control measures.

Interestingly, the MCR-ALS decomposition of NIR spectra yields scores whose modelling using appropriate algorithms has demonstrated the ability to discriminate not only between both types of cheeses but among different maturation weeks with a high degree of accuracy. Consequently, it can be concluded that this methodology holds promise

Table 3
Confusion matrix corresponding to the test sets for the modelling to discriminate between weeks of ripening for "Torta de Casar" (Class 1 to Class 7) and "Queso de la Serena" (Class 1 to Class 10).

Real/ Predicted	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9	Class 10	Not assig[a]
"Torta de Casar"											
Class 1	17	0	0	0	0	0	0	–	–	–	0
Class 2	0	18	0	0	0	0	0	–	–	–	0
Class 3	0	0	19	0	0	0	0	–	–	–	0
Class 4	0	0	0	17	0	0	0	–	–	–	0
Class 5	0	0	0	0	17	0	0	–	–	–	0
Class 6	0	0	0	0	0	17	0	–	–	–	0
Class 7	0	0	0	0	0	0	17	–	–	–	0
"Queso de la Serena"											
Class 1	1	0	0	0	0	0	0	0	0	0	6
Class 2	0	14	0	0	0	0	0	0	0	0	0
Class 3	0	0	0	0	0	0	0	0	0	0	19
Class 4	0	0	0	1	0	0	0	0	0	0	17
Class 5	0	0	0	0	20	0	0	0	0	0	0
Class 6	0	0	0	0	0	21	0	0	0	0	0
Class 7	0	0	0	0	0	0	18	0	0	0	0
Class 8	0	0	0	0	0	0	0	17	0	0	0
Class 9	0	0	0	0	0	0	0	0	17	0	0
Class 10	0	0	0	0	0	0	0	0	0	18	0

[a] Not assig: not assigned to any class.

as a viable option for authentication purposes in the final product.

CRediT authorship contribution statement

Elísabet Martín-Tornero: Writing – review & editing, Investigation, Formal analysis, Conceptualization. **Isabel Durán-Merás:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Mirta R. Alcaraz:** Writing – review & editing, Formal analysis. **Arsenio Muñoz de la Peña:** Writing – review & editing, Methodology, Funding acquisition. **Teresa Galeano-Díaz:** Writing – review & editing, Funding acquisition. **Héctor C. Goicoechea:** Writing – original draft, Supervision, Formal analysis, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

Financial support was provided by the Ministerio de Ciencia e Innovación of Spain (Project PID2020-112996 GB-I00 financed by MCIN/AEI/10.13039/501100011033) and Junta de Extremadura (Project IB20016) co-financed by Fondos Europeos de Desarrollo Regional. H.C.G. and M.R.A thank to program María Zambrano (Ministerio de Universidades of Spain).

References

[1] Regulation (CE) 1491/2003 of European Commission of 25 of August 2003: Registration of the protected designation of origin "Torta del Casar" in the register of protected designations of origin and the protected geographic indications., (n. d.).
[2] Regulation (CE) 1107/96 of European Commission of 12 of June 1996: Registration of geographical indications and designations of origin under the procedure laid down in Article 17 of Council Regulation (EEC) No 2081/92, (n.d.).
[3] M.L. da S. Medeiros, A. Freitas Lima, M. Correia Gonçalves, H. Teixeira Godoy, D. Fernandes Barbin, Portable near-infrared (NIR) spectrometer and chemometrics for rapid identification of butter cheese adulteration, Food Chem 425 (2023) 136461. https://doi.org/10.1016/j.foodchem.2023.136461.
[4] L. Čurda, O. Kukačková, NIR spectroscopy: A useful tool for rapid monitoring of processed cheeses manufacture, in, J Food Eng (2004) 557–560, https://doi.org/10.1016/S0260-8774(03)00215-2.
[5] A. de Araújo Gomes, S.M. Azcarate, I. Špánik, L. Khvalbota, H.C. Goicoechea, Pattern recognition techniques in food quality and authenticity: A guide on how to process multivariate data in food analysis, TrAC - Trends in Analytical Chemistry 164 (2023) 117105, https://doi.org/10.1016/j.trac.2023.117105.
[6] R. Bro, A.K. Smilde, Principal component analysis, Anal. Methods 6 (2014) 2812–2831, https://doi.org/10.1039/c3ay41907j.
[7] M.J.C. Pontes, R.K.H. Galvão, M.C.U. Araújo, P.N.T. Moreira, O.D.P. Neto, G. E. José, T.C.B. Saldanha, The successive projections algorithm for spectral variable selection in classification problems, Chemom. Intel. Lab. Syst. 78 (2005) 11–18, https://doi.org/10.1016/j.chemolab.2004.12.001.
[8] J. Jaumot, R. Gargallo, A. De Juan, R. Tauler, A graphical user-friendly interface for MCR-ALS: A new tool for multivariate curve resolution in MATLAB, Chemom. Intel. Lab. Syst. 76 (2005) 101–110, https://doi.org/10.1016/j.chemolab.2004.12.007.
[9] E. Martín-Tornero, I. Durán-Merás, A. Muñoz de la Peña, T. Galeano-Díaz, Fiber optic fluorescence as non-invasive tool to monitor the ripening process of cheeses: Torta del casar and Queso de la Serena, LWT 199 (2024) 116141, https://doi.org/10.1016/j.lwt.2024.116141.
[10] R.C. Castro, D.S.M. Ribeiro, J.L.M. Santos, R.N.M.J. Páscoa, Comparison of near infrared spectroscopy and Raman spectroscopy for the identification and quantification through MCR-ALS and PLS of peanut oil adulterants, Talanta 230 (2021) 122373, https://doi.org/10.1016/j.talanta.2021.122373.
[11] S.K. Karimvand, A. Pahlevan, S.V. Zade, J.M. Jafari, H. Abdollahi, Multivariate curve resolution-soft independent modelling of class analogy (MCR-SIMCA), Anal Chim Acta 1291 (2024) 342205, https://doi.org/10.1016/j.aca.2024.342205.
[12] S.V. Zade, H. Abdollahi, The classification performance of multivariate curve resolution-discriminant analysis: A comparative study, Microchem. J. 191 (2023) 108867, https://doi.org/10.1016/j.microc.2023.108867.
[13] G.G. Siano, I.S. Pérez, M.D.G. García, M.M. Galera, H.C. Goicoechea, Multivariate curve resolution modeling of liquid chromatography-mass spectrometry data in a comparative study of the different endogenous metabolites behavior in two tomato cultivars treated with carbofuran pesticide, Talanta 85 (2011) 264–275, https://doi.org/10.1016/j.talanta.2011.03.064.
[14] E.K. Kemsley, A genetic algorithm (GA) approach to the calculation of canonical variates (CVs), Trends Anal. Chem. 17 (1998) 24–34.
[15] D. Ballabio, V. Consonni, Classification tools in chemistry. Part 1: linear models, PLS-DA, Analytical Methods 5 (2013) 3790–3798, https://doi.org/10.1039/C3AY40582F.
[16] S.J. Dixon, R.G. Brereton, Comparison of performance of five common classifiers represented as boundary methods: Euclidean Distance to Centroids, Linear Discriminant Analysis, Quadratic Discriminant Analysis, Learning Vector Quantization and Support Vector Machines, as dependent on data structure, Chemom. Intel. Lab. Syst. 95 (2009) 1–17, https://doi.org/10.1016/j.chemolab.2008.07.010.
[17] Z. Ramadan, P.K. Hopke, M.J. Johnson, K.M. Scow, Application of PLS and Back-Propagation Neural Networks for the estimation of soil properties, Chemom. Intel. Lab. Syst. 75 (2005) 23–30, https://doi.org/10.1016/j.chemolab.2004.04.009.
[18] W. Wu, B. Walcmk, D.L. Massart, S. Heuerding, F. Erni, I.R. Last, K.A. Prebble, Artificial neural networks in classification of NIR spectral data: Design of the training set, Chemometrics and Intelligent Laboratory Systems 33 (1996) 35–46.
[19] A. Ríos, D. Barceló, L. Buydens, S. Cárdenas, K. Heydorn, B. Karlberg, K. Klemm, B. Lendl, B. Milman, B. Neidhart, R.W. Stephany, A. Townshend, A. Zschunke, M. Valcárcel, Quality assurance of qualitative analysis in the framework of the European project "MEQUALAN", Accred. Qual. Assur. 8 (2003) 68–77, https://doi.org/10.1007/s00769-002-0556-x.
[20] W. Windig, J. Guilment, Interactive Self-Modeling Mixture Analysis, Anal Chem 63 (1991) 120–1232, https://doi.org/10.1021/ac00014a016.
[21] T. Ke, X. Ge, F. Yin, L. Zhang, Y. Zheng, C. Zhang, J. Li, B. Wang, W. Wang, A general maximal margin hyper-sphere SVM for multi-class classification, Expert Syst Appl 237 (2024) 121647, https://doi.org/10.1016/j.eswa.2023.121647.
[22] G. Psaltakis, K. Rogdakis, M. Loizos, E. Kymakis, One-vs-One One-vs-Rest, and a novel Outcome-Driven One-vs-One binary classifiers enabled by optoelectronic memristors towards overcoming hardware limitations in multiclass classification, Discov Mater 4 (2024) 7, https://doi.org/10.1007/s43939-024-00077-7.
[23] R. Voccio, C. Malegori, P. Oliveri, F. Branduani, M. Arimondi, A. Bernardi, G. Luciano, M. Cettolin, Combining PLS-DA and SIMCA on NIR data for classifying raw materials for tyre industry: A hierarchical classification model, Chemom. Intel. Lab. Syst. 250 (2024) 105150, https://doi.org/10.1016/j.chemolab.2024.105150.