Journal of Food Engineering 361 (2024) 111738

Contents lists available at ScienceDirect

Journal of Food Engineering

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

Bread fermentation monitoring through NIR spectroscopy and PLS-DA. Determining the optimal fermentation point in bread doughs

D. Castro-Reigía [a,b], I. García [b], S. Sanllorente [a], L.A. Sarabia [c], J.M. Amigo [d,e], M.C. Ortiz [a,*]

[a] Departamento de Química, Facultad de Ciencias, Universidad de Burgos; Plaza Misael Bañuelos s/n, Burgos 09001, Spain
[b] AOTECH (Advanced Optical Technologies S.L.), Escuela Ing. De Bilbao; Plaza Ingeniero Torres Quevedo, 1, 2°, 48013, Bilbao, Spain
[c] Departamento de Matemáticas y Computación, Facultad de Ciencias, Universidad de Burgos; Plaza Misael Bañuelos s/n, Burgos 09001, Spain
[d] IKERBASQUE, Basque Foundation of Science, Plaza Euskadi 5, 48009 Bilbao, Spain
[e] Department Analytical Chemistry, Faculty of Science and Technology, University of the Basque Country (UPV/EHU). Barrio Sarriena s/n. 48940 Leioa, Spain

ARTICLE INFO

Keywords:
Bread fermentation
NIR spectroscopy
Monitorization
PLS-DA
Sensitivity
Specificity

ABSTRACT

Dough fermentation is a fundamental step in bread manufacturing, commonly supervised by bakery experts. This work aims to develop an efficient technology based on near infrared spectroscopy (NIR) and partial least squares-discriminant analysis (PLS-DA) to determine the fermentation state of bread loaves in the bakery industry. Knowing the fermentation state of the loaves during the manufacturing will allow to act in the production line, avoiding introducing an inadequate loaf in the oven and reducing costs. For that, a new methodology is proposed that consists in reproducing the knowledge of a Master Baker through a NIR spectrometer. Regarding the sequentiality of the objects and the real applicability of the method, three different cases were proposed using PLS-DA, getting that, in the best case, the sensitivity of the prediction set for the unfermented doughs was 88% and for the fermented and over-fermented was 86%; whereas the specificities were all greater than 86%.

1. Introduction

Bread consumption in Spain is elevated. The Spanish government estimates that in 2020, bread consumption was an average of 33 kilos per inhabitant, which meant an increase of 5.5% compared to the previous year (Ministerio de Agricultura Pesca y Alimentación, 2020). These data demonstrate the necessity of high quality in this product and exhaustive control in its manufacturing since it is a crucial economic factor affecting the Spanish diet.

Loaves of fresh bread are highly consumed products. One of the key steps in their manufacture is the fermentation process of the dough since it is one of the stages that most influences the quality of the final product (Cauvain, 2015; Kulp, K., & Lorenz, 2003). As a critical control point of bread fermentation process, a bakery expert is always demanded to judge the ideality of doughs before baking by traditional methods. Although artificial intelligence and computational methods are being investigated in fermentation processes, human knowledge and experience continue to be much more used: bibliographic research was made in Scopus using as keywords search “monitoring” and “fermentation or leavening”, and 5107 results were found. If the keyword “spectroscopy” is also used (monitoring, fermentation or leavening, and spectroscopy), the results decrease to 584. And if the word “NIR” (near infrared) is included in the research (monitoring, fermentation or leavening, spectroscopy, and NIR) the publications are 158. The application of NIR spectroscopy as an effective technology has been proven to be useful in food process monitoring and quality assessment, (Jiang et al., 2018; Muncan et al., 2021; Jin et al., 2023), or also, to provide chemical insight of the different phenomena occurring in bread during storage (Amigo et al., 2019, 2021). Besides, if the word “bread” is used as well (monitoring, fermentation or leavening, spectroscopy, NIR, and bread), just two papers are found (Chang et al., 2021; Ulrici et al., 2008). Chang et al. (2021) monitored the fermentation of Chinese steamed bread by using Partial Least Squares (PLS) to eliminate the unrepresentative variables from preprocessed NIR spectra, but Partial Least Squares-Discriminant Analysis (PLS-DA) methodology was not applied. Moreover, the authors proposed the success rate classification to evaluate the models, whereas sensitivity and specificity will be used in our work. That is, not only the success rate capacity of the models will be evaluated, but also their ability to reject samples that do not belong to the true class. In the second paper, Ulrici et al. (2008) get information on the modifications that occur during the process in terms of kinetics and chemical properties that change along with the fermentation. If we

* Corresponding author. .
E-mail address: mcortiz@ubu.es (M.C. Ortiz).

https://doi.org/10.1016/j.jfoodeng.2023.111738
Received 15 June 2023; Received in revised form 26 August 2023; Accepted 7 September 2023
Available online 12 September 2023

move ahead with the searching adding the term PLS-DA (monitoring, fermentation or leavening, spectroscopy, NIR, bread, and PLS-DA), non-results are found, but using NIR spectroscopy combined with PLS-DA is widely common in the food industry to classify distinct types of one product or to detect fraud (Cozzolino, 2014; Guelpa et al., 2017; Liu et al., 2008; Morsy and Sun, 2013; Williams et al., 2009).

As demonstrated by the previous search, there is no evidence that this combination (NIR and PLS-DA) has ever been used to monitor fermentation processes in bread, although it has been used in the case of the cheese maturation process (Kang et al., 2020).

On the one hand, NIR spectroscopy provides the information necessary to interpret the characteristics of the sample being measured. NIR spectroscopy can provide rich and suitable information to describe the changes during food processing (Durek et al., 2014; Marques et al., 2016; Tian et al., 2021; Ulrici et al., 2008), as, for instance, dough bread fermentation since in that process, the carbohydrates, the moisture, the protein structure, and the gas production of the dough change in a significant way (Kulp, K., & Lorenz, 2003). Furthermore, NIR also has the advantages of non-destructive and real-time measurements, with no sample preparation and in most cases, low-cost devices. For all that, in this work, NIR spectroscopy was considered a viable choice for monitoring the bread fermentation process. On the other hand, PLS-DA will allow classifying between NIR spectra that correspond to unfermented doughs, spectra of correct-fermented doughs and over-fermented ones, evaluating, in each case, the sensitivity and specificity of the models.

In that sense, the main purpose of the present study was to establish a convenient and monitored method to determine the dough fermentation state by using an at-line NIR spectrometer and a software platform (to obtain the data) combined with PLS-DA to reproduce the experience of a Master Baker.

To analyze the feasibility of classifying the state of a dough as unfermented, fermented, or over-fermented by NIR signals, three cases have been considered: case 1, a classifier is made to separate the three classes using a PLS-DA model; case 2, as the classes have a temporal sequence, first a PLS-DA model is performed to discriminate the unfermented doughs from the other two states and then, a second PLS-DA is performed to discriminate between fermented and over-fermented doughs; finally, in case 3, the same sequential procedure is followed as in case 2 but with a distribution of the data in the training set and in the validation set more in accordance with the possible industrial application of the procedure.

This final methodology evidences the effectiveness of monitoring the fermentation process using as the model objects the individual spectra selected from all the bread loaves of each class, since later, the decision will be made using just one spectrum of each loaf. This investigation was made as a proof of concept to implement, shortly, a NIR spectrometer in the production line (fermentation chamber) of a bakery industry. The conception is to estimate the state of doughs during fermentation by taking just one single spectrum of every dough. In that way, it will be possible to detect any deficiency in the doughs before baking and, consequently, modify the conditions of the fermentation chamber (temperature and humidity) accordingly. For instance, if there is a continuing situation of unfermented loaves, the fermentation chamber temperature could be increased to speed the fermentation process and avoid those loaves to enter into the oven until they are ready. Or in the case of the existence of some over-fermented loaves, reject them and decrease the humidity and temperature levels to avoid more over-fermented samples.

2. Materials and methods

2.1. NIR spectroscopy. Spectrophotometer configuration and measurements

The experiments were carried out in an industry of the bakery sector in Spain. IPASA-Sanbrandan (Ipasa-Sanbrandán) was born in A Coruña (Spain) in 1970 as the union of the most active industrial bakers of the city, that decided to unify their small companies into a new one of larger dimensions and with higher technological development. This work is a result of that continued development. Nowadays, the company manufactures and distributes the product all over the country, being one of the most prominent in the sector.

The measurements were carried out at-line on four different days to explore better the possible variations in manufacturing the loaves of bread. Each day, the spectra of the evolution of the three loaves were taken during the three stages of the fermentation process, having measures of unfermented doughs (Fig. 1a), of fermented doughs (Fig. 1b) and over-fermented doughs (Fig. 1c). These figures make clear the differences in the state of the doughs during the fermentation process, at least in terms of their volume.

The experimental procedure was made with the AONIR integrated solution for real-time NIR measurements (AOTECH S.L. (AOTECH)), including a NIR sensor , a measurement platform, and the precise software to integrate the hardware with the model outcome for real-time monitoring and controlling of the fermentation. The spectra were recorded throughout the entire fermentation process of the bread dough loaves in a fermentation chamber with temperature and humidity control (the temperature oscillated between 30.00 °C and 34.84 °C, while the humidity was 75%). Once the optimum fermentation time was achieved, spectra were still being taken to obtain data on the spectral behaviour of the over-fermented loaves of bread. Measurements were taken in triplicate and approximately every 3 min. That is, every 3 min, 3 spectra were recorded for each of the samples measured. As aforementioned, while the spectral measurements were being made, a Master Baker was aware of the evolution of the fermentation in the loaves of bread to ensure the reference status of the fermentation (unfermented, fermented and over-fermented).

The spectrometer was configured manually so that NIR reflectance was measured in a wavelength range from 900 to 1670 nm (125 wavelengths, accounting for a spectral resolution of 6 nm), with a spectrum reading interval of 1 s, and 50 readings per spectrum with an integration time of 10.8 ms.

As mentioned in the previous section, this work was divided into three cases. Since the measurements were made on four different days, approximately two-thirds of the samples from each day were used to build the models of cases 1 and 2 (8 loaves), and the remaining third part of the samples (4 loaves) was used as a prediction external set for validation. However, in the third case, the calibration and validation were made using individual spectra of each one of the loaves instead of using all the recorded spectra of the samples. That is, the same spectra registered and used in the two previous cases have been organised in a different way (around 70% of the data were used to build the calibration set, and a 30% for the prediction), having in both sets spectra of the three stages of the fermentation process (preventing calibration and prediction sets from sharing replicates). As a result, in case 1, the model was built using a data matrix in calibration of dimensions 677 × 125 (where for each of the 677 analysed samples, the measured reflectance at 125 different wavelengths was recorded) and was validated with a data matrix of dimensions 343 × 125. The training set is formed by the data of two complete doughs for each day and the validation by the other dough, as shown in Fig. 2a. This figure shows in solid color the samples of each class used in calibration and in the same dashed color those used in prediction. The color indicates the fermentation state of the dough, i.e. the class to which the spectra belongs to, red, green, and blue corresponding to unfermented, fermented, and over-fermented doughs, respectively. The squares represent the replicates carried out at each measurement time.

In case 2, a sequential decision was made, so two different models were built using two different matrices. In the first one, the data matrix has the same dimensions as in case 1, but grouped into two classes: unfermented versus fermented and over-fermented for the first PLS-DA model. In the second one, the model was built using a data matrix of

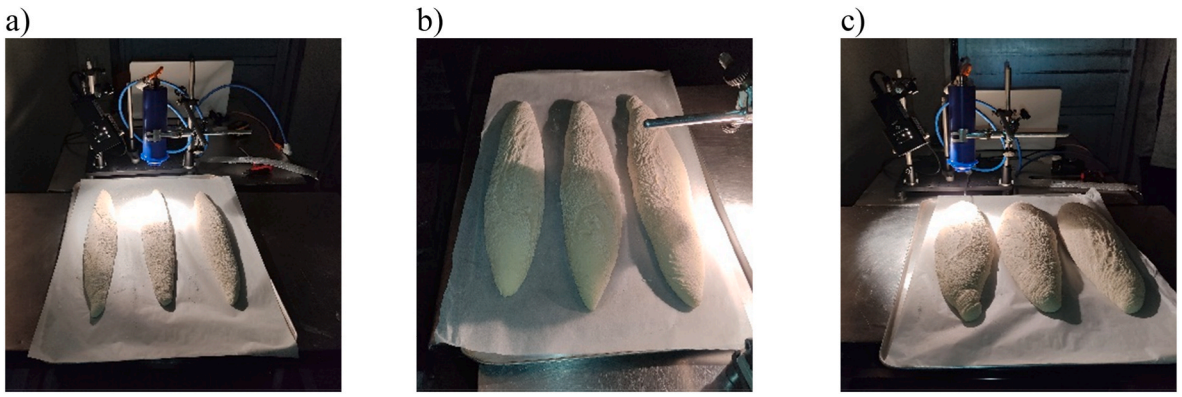

Fig. 1. Set up for the measurements and samples a) Unfermented doughs, b) Fermented doughs and c) Over-fermented doughs.

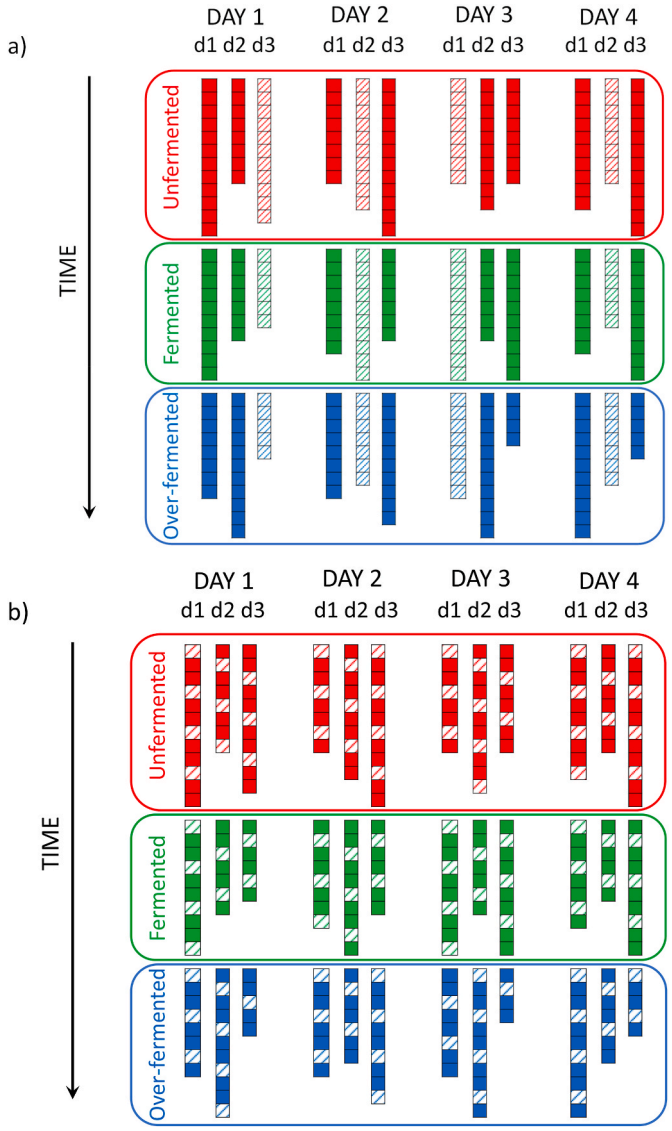

Fig. 2. Scheme of the assignment of the recorded spectra in the calibration and validation sets a) for cases 1 and 2, and b) for case 3. On each day, the doughs have been coded as d1, d2 and d3. The squares symbolize the replicate spectra that are always considered together in the assignment either to the calibration set or to the prediction set. The calibration set is built with the spectra of the solid squares, while the daseh ones form the prediction set.

dimensions 356 × 125 for calibration and dimensions 186 × 125 for prediction to discriminate the fermented loaves from the over-fermented. The same assignment of spectra in calibration and prediction shown in Fig. 2a is maintained.

In case 3, the sequential decision was also applied (as in case 2), therefore, two PLS-DA models were built. The main difference is the distribution of the spectra between the calibration and the prediction sets, but keeping in mind the real industrial application of the procedure. However, individual spectra from the three different fermentation states of the doughs were used. That is, the objects of the model are selected spectra of all the doughs that form the class. The distribution between calibration and prediction sets has been made as shown in Fig. 2b. In each class, one out of three consecutive groups of the replicated spectra (dashed square in Fig. 2b) was selected to build the prediction set, including always carefully the replicates (represented by each square in Fig. 2b) in the same set, either in calibration, or in prediction. In this way, it is guaranteed that both, in calibration and in prediction, the spectra are being recorded uniformly over time, collecting the instrumental variations by means of replicated spectra, temporal variation in the fermentation process, and the variations between different loaves.

The model for the first decision was built using a data matrix of 678 × 125 for calibration set, and 342 × 125 for the prediction one. For the second decision of case 3, the dimensions of the calibration data matrix were 359 × 125 and 186 × 125 for prediction.

The first three columns of Table 1 summarise the dimensions and the data arrangement of the three cases.

2.2. PLS-DA and how to evaluate the models

Depending on the field of application of the hypothesis test, the terminology has a considerable discrepancy. Therefore, the elements that will be handled in it are summarised below to ease the reading of this paper.

A classification rule to assign objects to one class can be expressed as a hypothesis test (Ortiz et al., 2006; Valencia et al., 2022). In the present

Table 1
Data matrix dimensions, number of latent variables, and permutation test results of the models.

Case	Model	Data matrix dimensions	LV	Test 1*	Test 2**	Test 3***
1	A	677 × 125	7	<0.001	<0.001	0.005
2	B	677 × 125	6	<0.001	<0.001	0.005
	C	356 × 125	10	<0.001	<0.001	0.005
3	D	678 × 125	4	<0.001	<0.001	0.005
	E	359 × 125	6	<0.001	<0.001	0.005

*Pairwise Wilcoxon signed rank test.
*Pairwise signed rank test.
*Randomisation *t*-test.

investigation, there are three classes of objects: C_1, unfermented loaf; C_2, fermented loaf and C_3, over-fermented loaf. It is aimed to build a rule that allows classifying a bread loaf into one of the three classes, C_j, based on the spectra registered with them. I_j spectra registered with loaves of class C_j, $j = 1, 2, 3$ are available. These spectra arrange one matrix $\mathbf{X}_{I\times N}$, being $I = I_1+I_2+I_3$ the total number of spectra and N the number of wavelengths of each spectrum. The matrix $\mathbf{Y}_{I\times 3} = (y_{ij})$ is used to identify each class. That matrix is defined by

$$y_{ij} = \begin{cases} 0, \text{if the spectrum i} - \text{th is not from a loaf of } C_j \\ 1, \text{if the spectrum i} - \text{th is from a loaf of } C_j \end{cases} \quad i = 1, \ldots, I, j = 1, 2, 3 \tag{1}$$

This codification of classes is known as "one versus all" because each column of **Y** identifies one of the classes with a "one" versus the other two, identified with a "zero".

For the purpose of building the decision rule, or what is the same, the hypothesis test

$$\rightarrow\rightarrow \quad \begin{array}{l} H_0 : \text{the loaf does not belong to the class } C_j \\ H_1 : \text{the loaf belongs to the class } C_j \end{array} \tag{2}$$

It is necessary to assign a probability distribution to the spectra of the C_j class, and to its complement $C(C_j) = \bigcup_{i\neq j} C_i$.

With the $\{\mathbf{X}, \mathbf{Y}\}$ data a statistical index $g_j(\mathbf{x})$ is built. This $g_j(\mathbf{x})$ allows building a distribution of probability F_{1j} for each C_j class and other F_{0j} for its complement $C(C_j)$, $j = 1, 2, 3$. From them, the critical value of the test, CV_j, is calculated. So, if for a given spectrum $\mathbf{x}_0$ the condition $g_j(\mathbf{x}_0) > CV_j$ is fulfilled, then H_0 will be rejected, and the loaf will be assigned to C_j. The critical values are determined regarding the probabilities α y β that are defined as follows:

- α is the probability of assigning a loaf to C_j when it does not belong to C_j (Type I error). Formally $\alpha = p\{\textit{reject } H_0 | H_0 \textit{ is true}\} = 1 - F_{0j}(CV_j)$.
- β is the probability of not assigning a loaf to the class C_j when it belongs to C_j (Type II error). Formally $\beta = p\{\textit{accept } H_0 | H_0 \textit{ is false}\} = F_{1j}(CV_j)$.
- The sensitivity of the C_j class is defined as the probability of correctly assigning the objects of the C_j class. Using the hypothesis test notation, this probability is equal to $p\{\textit{reject } H_0 | H_0 \textit{ is false}\} = 1 - \beta = 1 - F_{1j}(CV_j)$.
- The specificity of the C_j class is defined as the probability of correctly rejecting the objects that do not belong to the C_j class. According to the previous definition of α, this probability is equal to $p\{\textit{accept } H_0 | H_0 \textit{ is true}\} = 1 - \alpha = F_{0j}(CV_j)$.

Both probabilities evaluate the quality of the decision rule, which will be as better as the sensitivities and specificities are closer to 1. However, once the $\{\mathbf{X}, \mathbf{Y}\}$ data set and the distributions of probability F_{0j} and F_{1j} for each C_j class are fixed, the probabilities α and β depend on the critical value CV_j and when changing it, one increases and the other decreases. Occasionally, particularly if there is no cost associated with each error, $\alpha = \beta$ is considered.

The relation between the collected spectra in **X** and the correspondent responses in **Y**, according to Eq. (1) is made by means of a PLS regression. This PLS regression builds sequentially K pars of lineal combinations of the variables of **X** and **Y**. PLS exploits the covariation between predictor variables and response variables and tries to find a new set of latent variables (LVs) in **X**-block and **Y**-block that maximally relate to them. In other words, the covariance between the extracted LVs should be maximised as

$$max_{\mathbf{w}_1\mathbf{w}_2} \mathbf{w}_1^T \mathbf{X}^T \mathbf{Y} \mathbf{w}_2 \tag{3}$$

where T stand for transpose and $\mathbf{w}_i$, $i = 1, 2$ are the weight vectors of the first latent variable in **X**-block and **Y**-block, respectively. After that, **X** and **Y** are deflated by subtracting the contribution of the first LV and the procedure is iterated to build the K LV. The value of K is determined by cross-validation (CV). In this case, PLS is the lineal application $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), g_3(\mathbf{x}))$ of **X** in **Y**, so each one of the component functions, g_j, will take close values to one and zero for the spectra of the C_j and the $C(C_j)$ classes, respectively. These values are the predicted ones for each column of **Y**, that is to say, $\widehat{y}_j(\mathbf{x}) = g_j(\mathbf{x})$ for any spectrum of the space defined by **X**.

The next step is to fit the distributions of probability F_{0j} and F_{1j} to the $\widehat{y}_j(\mathbf{x})$ values for the spectra $\mathbf{x} \in C(C_j)$ and C_j respectively. These two distributions define the probabilities associated with the null and the alternative hypothesis from the test of Eq. (2). For each class, a normal distribution will be fitted for F_{0j} and F_{1j} and with them, the CV_j value corresponding to $\alpha = \beta$ is obtained. Each spectrum is assigned to one of the three classes through the obtained three critical values. If they belong to the acceptance region of two (for example, $g_j(\mathbf{x}_0) < CV_j$, $j = 1, 2$), it will be assigned to the one with the highest probability of belonging, that is, to the C_j such that F_{1j} is greater. In this way, all spectra are assigned to a single class, and none remain unassigned.

Once the assignment of all the objects has been made, a matrix of sensitivities and specificities $\mathbf{S} = (s_{ij})$ is obtained. **S** is a square matrix of dimensions equal to the number of classes that are being classified. The fraction of objects of the current class *i* correctly assigned to it by PLS-DA is noted on the diagonal of the matrix, and it is the estimation of the sensitivity of each class. Each s_{ij} element off the diagonal is the estimation of the specificity of the *i* class in relation to the *j* class. A classification model is as better as the closer all the values of the **S** matrix are to 1.

It is usual to calculate the sensitivities and specificities in prediction, either through CV, or by applying the classifier to an external set $\{\mathbf{X}_p, \mathbf{Y}_p\}$, independent from the training set. The estimated values in prediction will usually be worse than in fitting. Nevertheless, if both estimations are similar, the classifier will be considered stable in prediction.

2.3. Software

The AONIR platform developed by AOTECH (*AOTECH. Advanced Optical Technologies*) was used to record the spectra, while PLS-Toolbox (ver. 8.8.1, Eigenvector Research Inc. 196 Hyacinth Road, Manson, WA 98831) (Wise, B.M., Gallagher, N.B., Bro, R., Shaver, J.M., Winding, W., Koch, 2022) working under MATLAB version 9.9.0 (R2020b) update 8 (The MathWorks Inc, 2022) was employed for fitting PLS-DA models.

3. Results and discussion

3.1. NIR spectra and preprocessing

The main ingredients of the bread loaves are flour and water, among other components like yeast. Flour contains considerable amounts of carbohydrates (Cauvain, 2015; Hoseney, 1994), being starch the major constituent. During fermentation, yeasts take up nutrients in doughs, especially carbohydrates, for growth. Its metabolism leads to the formation of glucose and provides the necessary carbon dioxide to leaven the dough along with ethanol, which plays a role in conditioning flour proteins (Kulp, K., & Lorenz, 2003). So, since the composition of the samples was changing during the fermentation process, so should did the spectra. Considering all this, Fig. 3 can be explained as follows: Fig. 3a shows the NIR spectra of one of the samples during the whole fermentation process, differentiating between unfermented in red, fermented in green, and over-fermented in blue. As it can be seen, variations in intensity are observed, but apparently, the shape of the spectrum is almost the same during the three phases of the process. The most remarkable band is the one between 1400 and 1480 nm in the raw spectra, possibly related to the first overtone of the symmetric and asymmetric vibration stretch of the water molecule, the second overtone

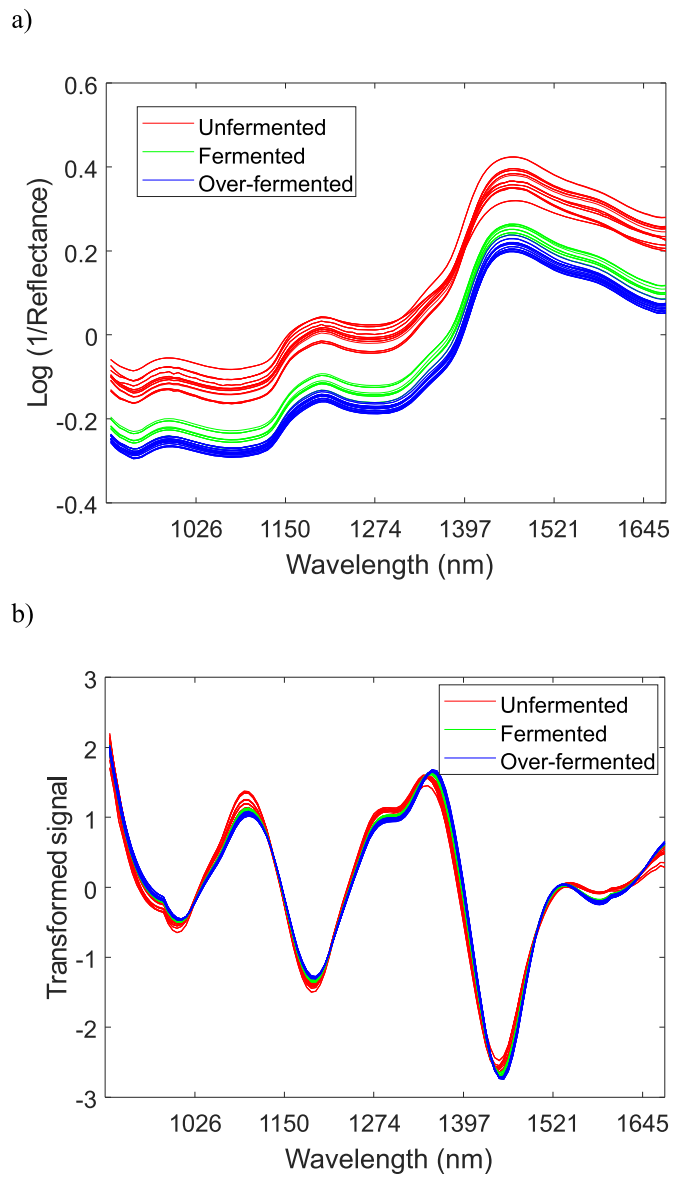

Fig. 3. a) NIR spectra of one of the loaves during the whole fermentation process b) Same spectra but preprocessed (Savitzky-Golay with a window width of 25 points using a second-degree polynomial and a second derivative, followed by SNV). In red, the spectra and preprocessed spectra of unfermented doughs, in green the fermented ones, and in blue, the over-fermented.

of the *O*–H bond of sugars, and the stretching vibration of the first overtone of the *N*–H bond (related with proteins). Nevertheless, there are two important bands at around 990 nm and 1200 nm. The first one, corresponding to the first overtone of the *O*–H bond of sugars and the second overtone of the *C*–H stretching of carbohydrates, and possibly overlapped with the second overtone of the *N*–H bond (1020 nm). The second one, is related to the first overtone of the symmetric stretch, the bending mode, and the asymmetric stretch of the water molecule). In general, the intermediate NIR (1050–1250 nm) is a complex region, primarily containing the second overtone of the *C*–H stretching vibrations. Some gentle bands at around 1350–1380 nm can also be observed, related to the stretching and deformation vibration modes of the *C*–H bonds corresponding to the $–CH_2$ and the methyl group of the carbohydrates, and the first overtone of the *O*–H stretching of sugars (Workman and Weyer, 2007).

On the other hand, Fig. 3b shows the same spectra but preprocessed. Combining different mathematical preprocessing methods is necessary when working with IR signals (Mas et al., 2020; Oca et al., 2012; Schoot et al., 2020). In this work, the results were obtained using the second derivative and SNV in order to baseline corrections and resolution enhancement. Specifically, the raw spectra were preprocessed by Savitzky-Golay with a window width of 25 points using a second-degree polynomial and a second derivative, and afterwards by standard normal variate (SNV), since after testing some other different combinations of preprocessing methods, particularly, the inverse (SNV + 2nd derivative), it was found to be the best one.

Finally, mean centering (MC) of the data matrix was applied before PLS-DA modelling, while the response was, at the same time, autoscaled.

3.2. Classification models

As indicated before, the responses of PLS-DA where codified using zeros and ones, in such way that a 1 indicates a sample belongs to a class, and a 0 indicates that it does not (Eq. (1)). So, in case 1 a PLS for a multi-response **Y** can be performed, knowing that the class of the unfermented doughs, C_1, had the (1,0,0) code, the fermented ones, C_2, (0, 1, 0), and the over-fermented, C_3, the code (0, 0, 1). This was made to create a response with a variable for each class.

But in the cases 2 and 3, the response variable **Y** has two columns, because only two classes are being compared. In the first decision, the loaves of C_1 *vs* $C_2 \bigcup C_3$ are classified and, according to Eq. (1), the unfermented loaves are coded with (1,0) and the remaining ones with (0,1). In the second decision, the loaves of C_2 *vs* C_3, are classified codifying the fermented loaves with (1,0) and the over-fermented with (0,1).

The optimal number of LVs was chosen by Venetian blinds CV procedure, considering the classification error in calibration (C) and CV. This CV method was used since as the samples were measured in real time during the whole fermentation process, it could be useful to estimate errors in the method from non-temporal sources. In addition, it is a simple and fast method to use. Also, as explained in section 2.2., the models were evaluated using the sensitivity and specificity for each class.

3.2.1. Case 1. PLS-DA applied to three classes

A PLS-DA model with three different classes was built: unfermented bread, correctly fermented and over-fermented doughs. As it can be seen in Table 1 (case 1), the absence of overfitting of the model (which needed 7 LVs) has been evaluated by doing three permutation tests (50 iterations) using the residuals in CV, because they are more sensitive to detect overfitting. All the p-values reported in Table 1 are less than 0.005. The model fitted for each response is distinguishable from one created randomly, at least at a confidence level of 0.995, which is much higher than usual 0.95.

The sensitivity and specificity of the model were evaluated. Fig. 4a represents sensitivity and specificity in a graphical form, referred to as a Receiver Operating Characteristic (ROC) curve. The sensitivity versus 1-specificity is shown for each of the three class models as a function of the selected threshold. For the three classes, the ROC curves in fitting and CV are practically equal and no difference is perceived graphically. If the ROC curve reaches the upper left corner means that at some threshold, the specificity could be perfect without losing any sensitivity (100% classification success). The threshold selected by the PLS-DA algorithm is shown as a red circle on each estimated and CV ROC curves. The area under the ROC curve (AUC) is used as a measure of the validity of a classifier, which will be better the closer AUC is to one. The AUC value is 0.99, 0.86 and 0.96 for the unfermented, fermented, and over-fermented loaves, respectively. These values can be considered sufficiently high. Nevertheless, they do not describe the discriminant power or capacity of the built PLS model. The critical values, CV_j, j = 1, 2, 3, calculated as explained in section 2.2., are showed as a vertical line in Fig. 4b. These values are 0.46, 0.27 and 0.41 for the unfermented, fermented, and over-fermented loaves, respectively. With them, the sensitivity and 1-

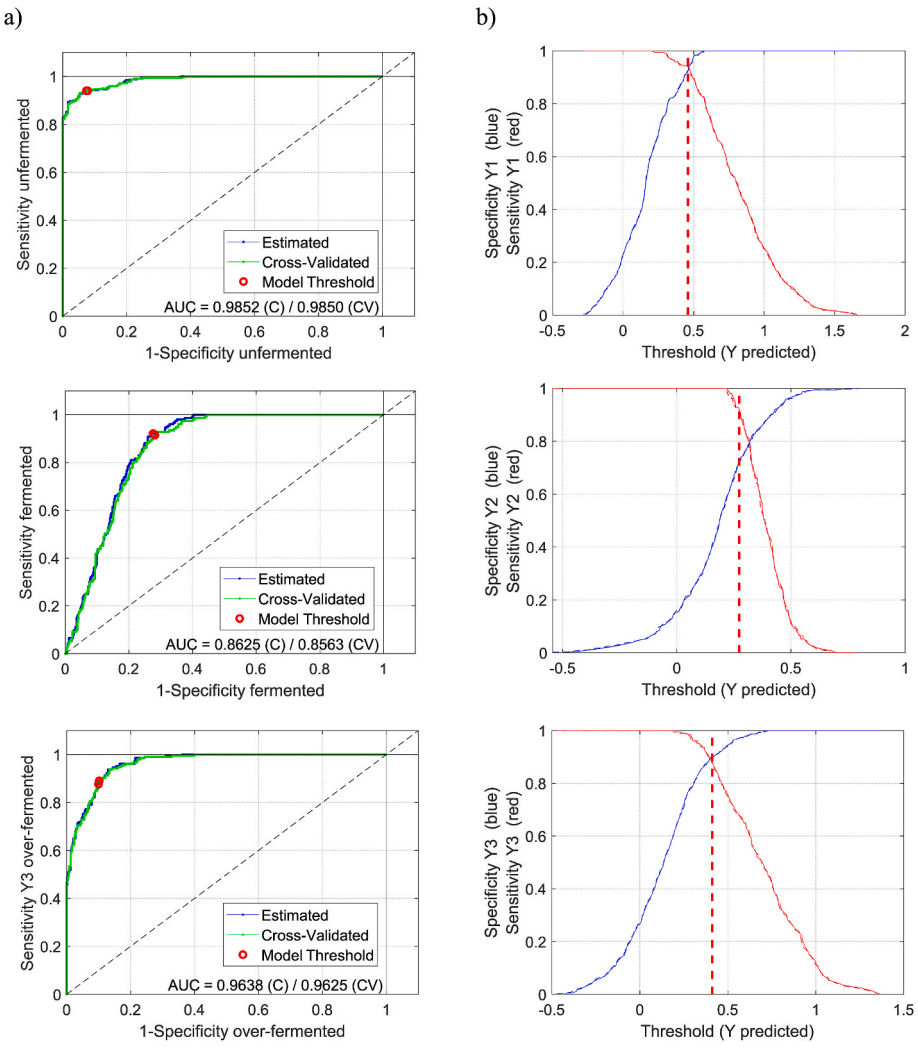

Fig. 4. a) ROC curves and b) sensitivity and specificity variations for case 1 (model A). The vertical red dashed line represents the model threshold. The specificity for each class is represented in each graphic in blue, the solid line in calibration and the dashed one in CV. The sensitivity for each class is represented in each graphic in red, the solid line in calibration and the dashed one in CV.

specificity values (also labelled with the red circle) on each estimated and cross-validated ROC curves in Fig. 4a are obtained. For instance, Fig. 4a for class 1 (unfermented bread loaves) has a 94% of classification success (94% of sensitivity) and almost 93% of specificity. That means that in training (and CV), the fitted PLS-DA model can not only classify mostly unfermented dough bread loaves as such, but also the ones that are not unfermented in a different class. The CV results are quite similar to the calibration ones, therefore, the models could be considered stable.

Fig. 4b also shows the sensitivity (as $1 - F_{1j}$) and specificity (as F_{0j}), being F_{0j} y F_{1j} the probability distributions of the $\hat{y}_j(\mathbf{x})$, j = 1, 2, 3 values. The vertical dashed red line indicates the CV_j selected by the PLS-DA algorithm. Each CV_j is calculated by fitting a normal distribution to the frequencies of the $\hat{y}_j(\mathbf{x})$ values. In the case of the fermented loaves this value differs from the point where the empirical distributions F_{02} and F_{12} intersect. This discrepancy may contribute to the decrease of the specificity of C_2 (fermented loaves) from 0.92, regarding C_1, until 0.72, regarding C_3. Remember (section 2.2.) that for each C_j class, the specificity has been calculated by means of CV_j j = 1, 2, 3, using its complement set $C(C_j)$.

The position of each calculated $\hat{y}_j(\mathbf{x})$ value concerning the CV_j value is showed in Fig. 5. In Fig. 5a it can be observed that most (91%) of the $\hat{y}_1(\mathbf{x})$ values (red diamonds) are greater than $CV_j = 0.46$ (dashed red line). That means they belong to the critical region of the test and therefore, H_0 is rejected, so, the corresponding loaves are classified correctly as unfermented. Moreover, a 7% of the fermented loaves (represented as green squares) also fulfil $\hat{y}_1(\mathbf{x}) > 0.46$ and consequently, they are classified mistakenly as unfermented. Regardless, none of the over-fermented loaves are classified as unfermented.

Analysing Fig. 5c it can be seen that just a 75% of the over-fermented loaves are correctly assigned, 14% of the fermented loaves and less than 1% of the unfermented are incorrectly assigned as over-fermented. That is, C_3 has high specificities (0.86 and 0.99) although its sensitivity decreases to 0.75.

Concerning the class of fermented loaves (C_2), a low sensitivity can be observed (Fig. 5b). Just 79% of the fermented loaves are precisely classified, while 18% of the over-fermented and 8% of the unfermented are also misclassified as fermented. All of them are in the critical region $\{\mathbf{x}|\hat{y}_2(\mathbf{x}) > 0.27\}$ of the correspondent hypothesis test.

The sensitivity and specificity values of this first PLS-DA model are in Table 2. As seen, quite good model results are obtained for sensitivity and specificity in the unfermented doughs, although it could be better for the case of over-fermented ones and mainly for the ones with an optimal fermentation point. The main problem is that many over-fermented doughs are classified as well-fermented, and many of the fermented ones are classified by the model as unfermented and over-fermented. In that same table (Table 2), the results obtained using an

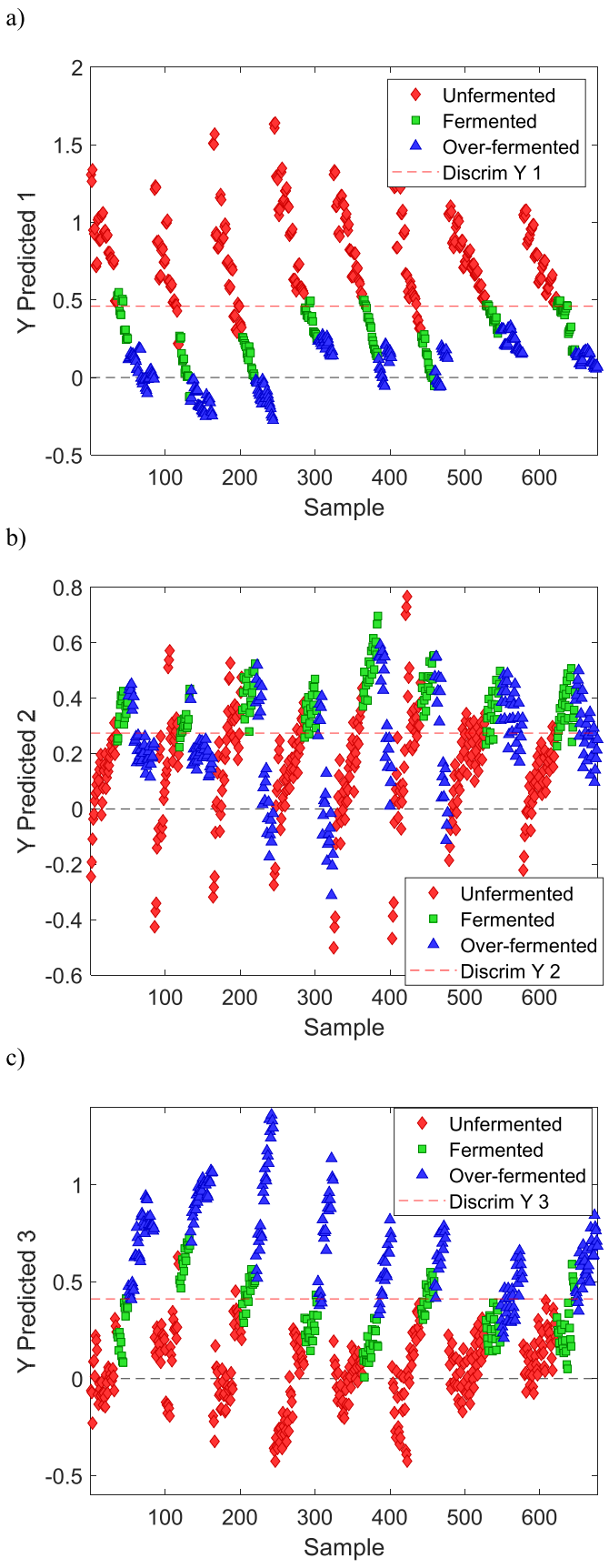

Fig. 5. Sample classification in calibration for Case 1 (model A).

Table 2
Sensitivity and specificity results for Case 1 (Model Ausing PLS-DA with three classes

		True class		
		Unfermented	Fermented	Over-fermented
Predicted class	Calibration			
	Unfermented	0.9117	0.9281	1.0000
	Fermented	0.9211	0.7843	0.7524
	Over-fermented	0.9905	0.8562	0.7524
	Prediction			
	Unfermented	0.8854	0.9643	1.0000
	Fermented	0.9236	0.3810	0.9608
	Over-fermented	0.9618	0.4167	0.9608

external prediction set of loaves are shown. The problem is notably exacerbated, the PLS-DA model only classifies correctly 38% of the fermented loaves, and 42% of the fermented loaves are misclassified as over-fermented. It can be seen that the PLS-DA model is not stable since the sensitivity and specificity results are quite different from the calibration ones.

Also, with this example, it is confirmed that a prediction external set for validation is adequate to assess the model (Esbensen and Geladi, 2010) since, as we can see in Fig. 4a, the CV results are similar to the calibration ones, but when we make real predictions, the results are not that good.

With these poor results in prediction, particularly related to the classes C_2 y C_3, is appropriate to consider that fermentation is a continuous process that evolves and sorts the classes ($C_1 < C_2 < C_3$). In that sense, it is reasonable for a sequential decision to discriminate first C_1 from $C_2 \bigcup C_3$ and then, C_2 from C_3.

3.2.2. Case 2. sequential decision

In this second case a two-step decision will be tested. Two models were built, one for each decision, and once again, the absence of overfitting of these models was evaluated (Table 1).

Firstly, a first decision to separate the spectra of the unfermented samples from the rest was made by following the same methodology as in section 3.2.1. A PLS-DA model was fitted (model B of case 2, see row 2 of Table 1), establishing this time only two classes, C_1 and $C_2 \bigcup C_3$. In this case, the PLS will predict a number for each sample that will be 1 if the sample is on the class, or on the contrary, will be 0 if that is not the case. The model will not predict exactly a 1 or a 0. Therefore, a limit will be established by the PLS-DA algorithm. Above that limit, the sample will be estimated as 1; below that limit, it will be predicted as 0. That is to say, it will be established, as explained section 2.2., a discriminant rule.

With the fitted PLS model and the discriminant rule established, the obtained results for the first decision of this case in terms of sensitivity and specificity are in Table 3. The results are shown both in calibration and in prediction, and they are quite better than the previous ones, since they guarantee a sensitivity of 95% for unfermented loaves, but also, only 5% of loaves that do not belong to the unfermented category have been classified as unfermented. Furthermore, it is verified that in difference with case 1, this model is stable in prediction, because after the validation set was predicted, the results remain similar.

After the unfermented loaves of bread were differentiated with success from the rest of the loaves, a second decision was made through a new PLS-DA model (model C of case 2, see row 3 of Table 1). This decision consists of discriminating the fermented loaves from the over-fermented ones. To do that, the same PLS-DA procedure applied until now will be used again, obtaining a new discriminant rule that allows the results of the second part of Table 3.

As it can be seen, the obtained results are worse than for the case of

Table 3
Sensitivity and specificity for both decision of case 2 (Model B and Model C).

Decision 1 (Model B)			
		True class	
		Unfermented	Fermented + Over-fermented
Predicted class	Calibration		
	Unfermented	0.9528	0.9304
	Fermented + Over-fermented	0.9528	0.9304
	Prediction		
	Unfermented	0.9363	0.9731
	Fermented + Over-fermented	0.9363	0.9731

Decision 2 (Model C)			
		True class Fermented	Over-fermented
Predicted class	Calibration		
	Fermented	0.9185	0.8738
	Over-fermented	0.9185	0.8738
	Prediction		
	Fermented	0.7857	0.7451
	Over-fermented	0.7857	0.7451

Table 4
Sensitivity and specificity for both decision of case 3 (Model D and Model E).

Decision 1 (Model D)			
		True class	
		Unfermented	Fermented + Over-fermented
Predicted class	Calibration		
	Unfermented	0.9313	0.9162
	Fermented + Over-fermented	0.9313	0.9162
	Prediction		
	Unfermented	0.8788	0.9091
	Fermented + Over-fermented	0.8788	0.9091

Decision 2 (Model E)			
		True class Fermented	Over-fermented
Predicted class	Calibration		
	Fermented	0.9167	0.8621
	Over-fermented	0.9167	0.8621
	Prediction		
	Fermented	0.8642	0.8571
	Over-fermented	0.8642	0.8571

unfermented doughs (C_1) *vs* fermented and over-fermented ($C_2 \bigcup C_3$) (model B). High sensitivity and specificity values are achieved in calibration, but not so good in prediction, which denotes that the model is not stable. The most remarkable thing is the low values for specificity in prediction for the over-fermented class. This fact raises the importance of considering the specificity of the models mentioned in the introduction because a 100% classification success rate can be achieved (the perfect sensitivity), but at the same time, it is possible to have bad specificity values. That could mean that, although all the samples are being classified in their class, so are the samples that do not belong to that class and therefore, a bad classification is being made.

Yet, if the overall results are observed, it is confirmed that the sequential decision is the most reasonable option and is the one that allows to achieve better results if compared with the three classes model. Also, if it is considered that in a near future, the measurements in the industry will be taken independently (in individual bread loaves in a punctual moment during the fermentation process), instead of monitoring the whole fermentation process, a more adequate approach should be proposed according to the applicability of the procedure. Regarding that, as explained in the following section, a similar methodology with the exact same data was applied but changing the organisation of the data.

3.2.3. *Case 3. sequential decision using individual spectra*

In this last case, alternatively, to build the model and the external prediction set with all the spectra of each one of the loaves, it was decided to select independent spectra of different loaves as explained in section 2.1. So, two new models were built, the first one (Model D) to separate the unfermented loaves from the others, and a second one (Model E), to differentiate between fermented and over-fermented loaves. By doing so, the results in Table 4, that are considerably better than the previous ones (mostly because they are stable in prediction), were obtained.

One of the main objectives of this study was to make a proof of concept to check if a NIR sensor could monitor the fermentation process of bread loaves doughs in the production line of the bakery industry. With this study and with the results obtained in case 3, it has been proved the effectiveness of applying a sequential decision and a spectra selection regarding the sequentiality of the objects and the real application of the procedure, respectively.

With the methodology applied in case 3, the sensitivity (considering the prediction results) for the class 1 (unfermented doughs) is 88%, for the class 2 (fermented) is 86%, and for C_3 (over-fermented) is also 86%, while the specificities are all greater than 86%. With these results, it is possible to detect high percentages of defects in the loaves in such a way that the correctly fermented loaves enter the oven when they should. In the case of over-fermented loaves, they could be prevented from entering the oven, saving time and energy in baking. Furthermore, in case that the sensor detects several over-fermented loaves continuously, it would indicate that the conditions of the fermentation chamber are inadequate, but the temperature and humidity could be decreased to prevent this from occurring. The same would happen in the case of loaves that are not yet fermented, but in this case, the temperature and humidity would be increased, and it would be prevented that those loaves that the sensor classifies as not fermented could reach the correct fermentation state before entering the oven. In this way, the knowledge of an expert baker is reproduced to a high percentage (minimum of 86% both for sensitivity and specificity in prediction) and consequently, its work, and the production costs, are reduced. With this system, even the cost of yeast and flour could be reduced, since quantities could be readjusted depending on how the fermentation process progresses, adjusting the quantities or the conditions for the following loaves.

4. Conclusions

The work developed in this study has demonstrated the possibility of monitoring the industrial fermentation process of bread loaves through NIR spectroscopy combined with chemometrics. It has been proven that implementing this technology in-line is conceivable since it seems that it considerably reproduces the perception of one expert of the bakery industry. However, the technology allows to take action in the fermentation process by modifying the temperature or the humidity in the fermentation chamber, avoiding the entry of defective products in the oven. That, will allow the reduction of energetic consumption and timing or wasting yeasts and flour. Once these results were presented to the industry where the measurements were made, they accept to implement the sensor in their production line.

Funding

This work has been funded by the Ministerio de Industria, Comercio y Turismo under Project PHOTONICS4BAKERY (AEI-010500-2021b-111) and Consejería de Educación of JCyL under Project BU052P20 cofinanced with Regional European Funds.

CRediT authorship contribution statement

D. Castro-Reigía: Investigation, Conceptualization, Data curation, Writing – original draft, Reviewing and Editing. **I. García:** Investigation, Supervision, Funding acquisition. **S. Sanllorente:** Visualization, Supervision, Validation. **L.A. Sarabia:** Data curation, Conceptualization, Visualization, Supervision, Validation, Writing – original draft. **J.M. Amigo:** Conceptualization, Supervision, Reviewing and Editing. **M.C. Ortiz:** Data curation, Supervision, Reviewing and Editing, Funding acquisition.

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

The authors do not have permission to share data.

Acknowledgements

The authors thank IPASA-Sanbrandan (INDUSTRIALES PANADEROS AGRUPADOS S.A.) for allowing the measurements in their facilities and the entities for financial support.

References

Amigo, J.M., del Olmo, A., Engelsen, M.M., Lundkvist, H., Engelsen, S.B., 2019. Staling of white wheat bread crumb and effect of maltogenic α-amylases. Part 2: monitoring the staling process by using near infrared spectroscopy and chemometrics. Food Chem. 297, 124946 https://doi.org/10.1016/j.foodchem.2019.06.013.

Amigo, J.M., Olmo, A. del, Engelsen, M.M., Lundkvist, H., Engelsen, S.B., 2021. Staling of white wheat bread crumb and effect of maltogenic α-amylases. Part 3: spatial evolution of bread staling with time by near infrared hyperspectral imaging. Food Chem. 353 https://doi.org/10.1016/j.foodchem.2021.129478.

AOTECH. Advanced Optical Technologies. (n.d.). Retrieved April 30, 2023, from https://www.aotech.es/.

Cauvain, S., 2015. Technology of breadmaking. In: Technology of Breadmaking. https://doi.org/10.1007/978-3-319-14687-4.

Chang, X., Huang, X., Xu, W., Tian, X., Wang, C., Wang, L., Yu, S., 2021. Monitoring of dough fermentation during Chinese steamed bread processing by near-infrared spectroscopy combined with spectra selection and supervised learning algorithm. J. Food Process. Eng. 44 (9) https://doi.org/10.1111/jfpe.13783.

Cozzolino, D., 2014. An overview of the use of infrared spectroscopy and chemometrics in authenticity and traceability of cereals. Food Res. Int. 60, 262–265. https://doi.org/10.1016/j.foodres.2013.08.034.

Durek, J., Ghadiri Khozroughi, A., Fröhling, A., Schlüter, O., Knorr, F., Mader, A., Goodarzi Boroojeni, F., Zentek, J., Knorr, D., Bolling, J.S., 2014. Effects of thermally treated broiler feed with different organic acid levels on resulting meat composition and parameters related to meat quality. Innovative Food Sci. Emerging Technol. 26, 397–405. https://doi.org/10.1016/j.ifset.2014.05.001.

Esbensen, K.H., Geladi, P., 2010. Principles of proper validation: use and abuse of re-sampling for validation. J. Chemometr. 24 (3–4), 168–187. https://doi.org/10.1002/cem.1310.

Guelpa, A., Marini, F., du Plessis, A., Slabbert, R., Manley, M., 2017. Verification of authenticity and fraud detection in South African honey using NIR spectroscopy. Food Control 73, 1388–1396. https://doi.org/10.1016/j.foodcont.2016.11.002.

Hoseney, R.C., 1994. In: Principles of Cereal Science and Technology, vol. 2. American Association of Cereal Chemists (AACC).

Ipasa-sanbrandán. Retrieved April 30, 2023, from. http://www.sanbrandan.com/.

Jiang, H., Mei, C., Li, K., Huang, Y., Chen, Q., 2018. Monitoring alcohol concentration and residual glucose in solid state fermentation of ethanol using FT-NIR spectroscopy and L1-PLS regression. Spectrochim. Acta Mol. Biomol. Spectrosc. 204, 73–80. https://doi.org/10.1016/j.saa.2018.06.017.

Jin, S., Sun, F., Hu, Z., Li, Y., Zhao, Z., Du, G., Shi, G., Chen, J., 2023. Online quantitative substrate, product, and cell concentration in citric acid fermentation using near-infrared spectroscopy combined with chemometrics. Spectrochim. Acta Mol. Biomol. Spectrosc. 285, 121842 https://doi.org/10.1016/j.saa.2022.121842.

Kang, R., Zhao, M., Fagan, C.C., Methven, L., Oruna-concha, M.J., Colm, P.. Monitoring of cheese maturation using near infrared- hyperspectral imaging (NIR-HIS). Written for presentation at the 2020 ASABE Annual International Meeting Sponsored by ASABE 3–10. https://10.13031/aim.202001555.

Kulp, K., Lorenz, K. (Eds.), 2003. Handbook of Dough Fermentations, vol. 127. Crc Press.

Liu, L., Cozzolino, D., Cynkar, W.U., Dambergs, R.G., Janik, L., O'Neill, B.K., Colby, C.B., Gishen, M., 2008. Preliminary study on the application of visible-near infrared spectroscopy and chemometrics to classify Riesling wines from different countries. Food Chem. 106 (2), 781–786. https://doi.org/10.1016/j.foodchem.2007.06.015.

Marques, E.J.N., De Freitas, S.T., Pimentel, M.F., Pasquini, C., 2016. Rapid and non-destructive determination of quality parameters in the "Tommy Atkins" mango using a novel handheld near infrared spectrometer. Food Chem. 197, 1207–1214. https://doi.org/10.1016/j.foodchem.2015.11.080.

Mas, C., Rubio, L., Valverde-Som, L., Sarabia, L.A., Ortiz, M.C., 2020. Impact of the pretreatment of ATR-FTIR signals on the figures of merit when PLS is used. Chemometr. Intell. Lab. Syst. 201, 104006 https://doi.org/10.1016/j.chemolab.2020.104006.

Informe del Consumo de Alimentación en España, 2020. https://www.mapa.gob.es/es/alimentacion/temas/consumo-tendencias/informe-anual-consumo-2020-v2-nov2021-baja-res_tcm30-562704.pdf.

Morsy, N., Sun, D.W., 2013. Robust linear and non-linear models of NIR spectroscopy for detection and quantification of adulterants in fresh and frozen-thawed minced beef. Meat Sci. 93 (2), 292–302. https://doi.org/10.1016/j.meatsci.2012.09.005.

Muncan, J., Tei, K., Tsenkova, R., 2021. Real-time monitoring of yogurt fermentation process by aquaphotomics near-infrared spectroscopy. Sensors 21 (1), 1–18. https://doi.org/10.3390/s21010177.

Oca, M.L., Ortiz, M.C., Sarabia, L.A., Gredilla, A.E., Delgado, D., 2012. Prediction of Zamorano cheese quality by near-infrared spectroscopy assessing false non-compliance and false compliance at minimum permitted limits stated by designation of origin regulations. Talanta 99, 558–565. https://doi.org/10.1016/j.talanta.2012.06.035.

Ortiz, M.C., Sarabia, L., García-Rey, R., De Castro, M.D.L., 2006. Sensitivity and specificity of PLS-class modelling for five sensory characteristics of dry-cured ham using visible and near infrared spectroscopy. Anal. Chim. Acta 558 (1–2), 125–131. https://doi.org/10.1016/j.aca.2005.11.038.

Schoot, M., Kapper, C., van Kollenburg, G.H., Postma, G.J., van Kessel, G., Buydens, L.M. C., Jansen, J.J., 2020. Investigating the need for preprocessing of near-infrared spectroscopic data as a function of sample size. Chemometr. Intell. Lab. Syst. 204, 104105 https://doi.org/10.1016/j.chemolab.2020.104105.

The MathWorks Inc, 2022. MATLAB version: 9.9.0 (R2020b). The MathWorks Inc, Natick, Massachusetts. https://www.mathworks.com.

Tian, X.Y., Aheto, J.H., Bai, J.W., Dai, C., Ren, Y., Chang, X., 2021. Quantitative analysis and visualization of moisture and anthocyanins content in purple sweet potato by Vis–NIR hyperspectral imaging. J. Food Process. Preserv. 45 (2), 1–16. https://doi.org/10.1111/jfpp.15128.

Ulrici, A., Vigni, M.L., Durante, C., Foca, G., Belloni, P., Brettagna, B., De Marco, T., Cocchi, M., 2008. At-line monitoring of the leavening process in industrial bread making by near infrared spectroscopy. J. Near Infrared Spectrosc. 16 (3), 223–231. https://doi.org/10.1255/jnirs.781.

Valencia, O., Ortiz, M.C., Ruiz, S., Sánchez, M.S., Sarabia, L.A., 2022. Simultaneous class-modelling in chemometrics: a generalization of Partial Least Squares class modelling for more than two classes by using error correcting output code matrices. Chemometr. Intell. Lab. Syst. 227 https://doi.org/10.1016/j.chemolab.2022.104614.

Williams, P., Geladi, P., Fox, G., Manley, M., 2009. Maize kernel hardness classification by near infrared (NIR) hyperspectral imaging and multivariate data analysis. Anal. Chim. Acta 653 (2), 121–130. https://doi.org/10.1016/j.aca.2009.09.005.

Wise, B.M., Gallagher, N.B., Bro, R., Shaver, J.M., Winding, W., Koch, R., 2022. PLS Toolbox 8.8.1. Eigenvector Research Inc.

Workman Jr., J., Weyer, L., 2007. Practical Guide to Interpretive Near-Infrared Spectroscopy. CRC press. https://doi.org/10.1201/9781420018318.