Microchemical Journal 166 (2021) 106215

Contents lists available at ScienceDirect

Microchemical Journal

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

Review Article

ATR-MIR spectroscopy as a process analytical technology in wine alcoholic fermentation – A tutorial

Daniel Schorn-García [a], Julieta Cavaglia [a], Barbara Giussani [b], Olga Busto [a], Laura Aceña [a], Montserrat Mestres [a], Ricard Boqué [c,*]

[a] *Universitat Rovira i Virgili. Instrumental Sensometry (iSens), Department of Analytical Chemistry and Organic Chemistry, Campus Sescelades, Edifici N4, C/Marcel·lí Domingo s/n, Tarragona 43007, Spain*
[b] *Dipartimento di Scienza e Alta Tecnologia, Università Degli Studi Dell'Insubria, Via Valleggio 9, Como 22100, Italy*
[c] *Universitat Rovira i Virgili. Chemometrics, Qualimetrics and Nanosensors Group, Department of Analytical Chemistry and Organic Chemistry, Campus Sescelades, Edifici N4, C/Marcel·lí Domingo s/n, Tarragona 43007, Spain*

ARTICLE INFO

Keywords:
Infrared spectroscopy
PAT
Microfermentation
Wine
Multivariate process control

ABSTRACT

The goal of this article is to guide the reader through the critical points to be faced when monitoring a fermentation following a Process Analytical Technology (PAT) approach. To achieve this purpose Attenuated Total Reflectance – Mid-Infrared (ATR-MIR) spectroscopy coupled to chemometric techniques are proposed.

Each of the crucial steps (set up of microvinifications, sampling, spectroscopic analysis and chemometric data treatment) is deeply investigated, revealing how the sampling is decisive for the subsequent modeling phase, suggesting how to set parameters to obtain good quality signals, and explaining how to prepare the data for the chemometric modeling and to perform the calculations. The modeling strategies here presented, based mainly on basic chemometric tools such as principal component analysis and partial least square regression, proved to be effective to the purposes and affordable even for non-expert chemometric users.

The article shows, using real examples, how to obtain or predict several parameters from a fermentation data set – control of the fermentation evolution, prediction of oenological parameters during the alcoholic fermentation and detection of deviations from the normal operation condition.

1. Introduction

Wine production is based on the biochemical process called alcoholic fermentation, which implies the transformation of sugars (glucose and fructose) into ethanol and carbon dioxide carried out by yeast, usually of the genus *Saccharomyces*. It is a complex process due the release of many yeasts' metabolites at very different concentrations and that are related to physicochemical and organoleptic properties of the final product [1]. Thus, the control of these by-products has a substantial impact on the final quality of the wine. However, apart from the daily visual observation of the tanks, the parameters that are routinely and traditionally measured in the cellar to follow the alcoholic fermentation are simply the temperature, the density and the pH. Although there are already some simple devices on the market that can measure and predict these oenological parameters, they are not widely used. This is because, as in determining these parameters using other analytical methods, there is a time lag between obtaining the results and applying the corrective actions. Thus, when problems such as deviations, stuck or sluggish fermentations are detected, the delay in carrying out the corrective measures negatively affects the quality, which in the end means economic losses [2]. This is why there is a growing interest in an effective monitoring strategy based on a real-time approach [3].

The United States Food and Drug Administration defined Process Analytical Technologies (PAT) as "a mechanism to design, analyze, and control manufacturing pharmaceutical processes through the measurement of Critical Process Parameters which affect Critical Quality Attributes" [4]. This philosophy is well suited to any other industry where quality control is essential, such as the food industry. It should be noted that food samples are often complex chemical mixtures and, many times, easy to be altered. Therefore, the quality of a food product cannot be guaranteed by analyzing only the final product, but a control of the product is required throughout the production process by in-line or on-line measurements. Due to its feasibility and practicality, the PAT approach has already been applied to different sectors of the food

* Corresponding author.
E-mail address: ricard.boque@urv.cat (R. Boqué).

https://doi.org/10.1016/j.microc.2021.106215
Received 19 November 2020; Received in revised form 25 March 2021; Accepted 25 March 2021
Available online 29 March 2021
0026-265X/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

industry with satisfactory results (see e.g. [5–7] and references therein). As can be seen, in all cases, in the Food PAT approach it is imperative to fully understand the process and understand all the sources of variability to obtain reliable results. Concerning wine alcoholic fermentation, sources of variability come from both raw materials (grape variety, sanitary state of the grape, ripening and vintage) and the fermentation process (species and strain of microorganism, sudden temperature changes, excess or absence of oxygenation, amount of assimilable nitrogen or spoilages, among others).

As the sources of variation in biochemical processes can be many and very different, several parameters should be determined to detect the potential problems and, usually, these parameters should be determined in the different batches at several production points. Fortunately, the use of modern multi-analyzers greatly simplifies these determinations, although the large amount of data generated is usually difficult to manage. So, to extract valuable information these data should be processed through proper statistical/chemometric tools [8]. This is why the PAT approach often relies on the use of multivariate analysis in order to deal with the maximum amount of data to better control the process. Multivariate analysis allows both qualitative (classification or pattern recognition) and quantitative results (parameter prediction) through the analysis of the generated data [3].

Among the different instrumental techniques used for food process monitoring, the spectroscopy-based analyzers are taking on a special relevance. The rapidness of the data acquisition, the absence or minimal sample pretreatment and the portability of the instruments make spectroscopy an ideal tool in PAT approaches. This last characteristic is especially interesting because it allows moving the equipment to the sample instead of having to take the sample to the laboratory. In other words, with this type of equipment we can get the laboratory to the measurement point, thus saving time, improving the efficiency of the testing process and allowing a decision to be made on the spot.

Although NIR spectrometers have been widely used in food process control [9,10], ATR-MIR (Attenuated Total Reflectance – Mid Infrared) spectroscopy has gained popularity in food analysis as it is fast, non-destructive and environmentally friendly [11] and allows to surpass the water absorption limitations that other types of IR measurements present.

The aim of this tutorial is to help the reader understand the alcoholic fermentation of wine from the PAT point of view, and to provide a guidance on how this process can be controlled using ATR-MIR spectroscopy and what information can be obtained from such control. Following the PAT guidelines and in order to acquire knowledge about the process, this tutorial describes the fermentation process from the beginning and through microvinifications [4]. Thus, the sampling strategy, the analytical procedure, and the multivariate data pretreatment and modeling will be considered, explained and discussed and pros and drawbacks will be highlighted for each step. Real examples will be provided to illustrate the main ideas.

2. Experimental setting and sampling

On-line or in-line [8] measurements are the final goal in process monitoring, but often the difficult direct implementation of instrumentation in the cellar when dealing with high volumes, the need for control experiments and, sometimes, the lack of knowledge about key points in the process discourage the attempt. In this context, micro-scale analysis could be a suitable option for this purpose [4]. In addition, it should be borne in mind that the wine production in a cellar involves a limited number of fermentations, with a subsequent reduced number of batches that will be used to optimize the chemometric strategy. Instead, micro-scale fermentations generate a greater number of available samples to study the process variability, to choose the best strategy to analyze and control the process evolution and to finally facilitate the scale-up of the monitoring process in the cellar. This is even more important when studying complex bio-transformations such as alcoholic fermentation, as biological replicates in different containers have slightly different behaviors.

2.1. Grape must characteristics

The importance of the must quality is obvious and, therefore, the legislation allows wineries to adjust some basic parameters of this raw material such as acidity, sugar content or assimilable nitrogen content [12]. When this process is transferred to micro-scale fermentations, they can be produced with natural grape juice or concentrated grape must (CGM). The latter, with a sugar concentration of about 800 g·L^{-1}, allows better preservation of the samples before dilution, as it reduces the activity of microorganisms due to osmotic pressure [13]. Moreover, whereas natural grape juice is a seasonal product, CGM can be purchased at any time of the year. The CGM must be properly diluted to reach a sugar concentration that provides the desired alcoholic strength (in a winery the usual concentration is between 170 and 200 g·L^{-1}) [1]. Apart from sugar, the pH value should also be adjusted with some acid allowed by legislation, such as tartaric acid, as the pH range of the wine is set between 2,8 and 4,2 [13,14]. Finally, it is also necessary to control the Yeast Assimilable Nitrogen (YAN), as it is a limiting factor for yeast growth. Its concentration may range from 140 to 500 mg·L^{-1}, with an optimal interval from 200 to 350 mg·L^{-1}. A deficit in YAN concentration leads to unfinished fermentations or slower fermentation kinetics; and a higher concentration would lead to the generation of unpleasant aromas [15]. The addition can be made using ammonium salts or commercial YAN supplements, as it is done in wineries, which apart from inorganic and organic nitrogen content, contain other compounds such as vitamins that enhance yeast fermentation [15].

2.2. Inoculation

Once the diluted CGM or natural must have the right composition, the next step is to inoculate the yeast. This can be performed by adding liquid culture media or dry active yeasts to each fermenting container, the latter option being more complicated because it involves an additional step for rehydration. In any case, the temperature difference between the must and the liquid inoculum must be carefully controlled because the suppliers recommend a difference of about 5 °C and never exceed 10 °C difference. As for the inoculum concentration, this value depends on the desired fermentation process. Thus, to reproduce the natural population in the grapes to give rise to a spontaneous fermentation it should range between 10^{3} and 10^{5} CFU·mL^{-1} and to reproduce the inoculum population used in wineries to ensure the imposition of the yeast of interest it should be between 10^{6} and 10^{7} CFU·mL^{-1}. It should be noted that, although the main yeast species used to inoculate in wineries is *Saccharomyces* due to its high fermentation capacity, it is known that non-*Saccharomyces* yeasts can also influence the course of fermentation and the character of the resulting wine. Recent studies have evaluated the use of controlled mixtures of both species [1,16].

2.3. Fermentation temperature

Among the physical parameters, temperature is possibly the most important because the alcoholic fermentation is an exothermic process and yeasts only have proper metabolic activity within optimal temperature ranges. Thus, this parameter deeply influences the kinetics of fermentation, so that when the temperature rises 10 °C, the speed of the kinetics doubles [13]. Therefore, the fermentations must take place at a controlled temperature, being the recommended values between 10 °C and 32 °C. These values differ when it comes to white or red must: in white wine, the temperature should be kept below 18 °C for aroma retention, but for red must the temperature should be above 25 °C to improve the color extraction from pomace [13]. A slight increase in kinetics can also be obtained by agitation; however, to reproduce real conditions in a winery, it is not necessary.

2.4. Sampling times

Sampling time is crucial in a process control approach and there are several points to consider:

- alcoholic fermentation follows a sigmoid function, due to the yeast adaptation time at the beginning and the stressful conditions at the end [13,17]. Therefore, to properly define the fermentation curve, a representative sampling at all different phases of the fermentation process is necessary;
- sampling times must be frequent enough to allow corrective actions to be taken in case of deviation from the normality of the process and, from a practical oenological point of view, this usually means less than 12 h between samplings;
- sampling must reproduce actual conditions, so it should be noted that when using at-line analysis, sampling requires the physical removal of a fraction of the sample to analyze it. This issue is not really crucial for spectroscopic techniques, as the analysis is very fast and requires a small fraction of the sample;
- when applying chemometrics, the sampling design will condition future samplings, as the same number of samplings or even the same sampling frequency will be required to take decisions using the statistical model. If possible, samples should be taken at a fixed frequency throughout the fermentation, but this is not obvious. An effective way to overcome this problem, especially when historical data are used, will be shown in section 5.4.

2.5. Sample pretreatments

One of the main advantages of using vibrational spectroscopy is that little or no sample pre-treatment is required. However, as sampling is carried out from the fermenting must with many microorganisms, to avoid biochemical changes in the samples, it is necessary to remove the microorganisms. This removal can be achieved using physical or chemical pretreatments. In the case of physical pretreatments, it is possible to distinguish between partial removal by centrifuging the sample and separating the supernatant, or total removal by filtering the sample through a 0,45 µm (yeast sterilization) or 0,22 µm (yeast and bacteria sterilization) diameter filter. The chemical pretreatment usually applied is the addition of sodium fluoride (NaF) as an oxidative metabolism suppressor for yeast [17]. For right-after infrared and standard physicochemical analysis, partial removal is a good option (see e.g. [18] and references therein).

3. Measurements and data collection

The infrared spectral region is characterized by the vibration of the molecular bonds. In the case of the mid-infrared region, approximately between 4000 and 400 cm^{-1}, the spectrum obtained results from fundamental vibrations, which facilitates its interpretation compared to the NIR spectrum. However, it has to be pointed out that MIR spectroscopy shows its full usefulness when the ATR sampling technique is applied. This is because ATR allows to overcome the limitations that imply the large absorptions of water bands [19]. In this tutorial we suggest the use of an ATR-FT-MIR portable spectrometer to monitor the wine fermentation process. This equipment offers the possibility of measuring in real-time and on-site, which also means a reduction in the costs of materials, transport and storage of samples. In addition, portable spectroscopic instruments are cheaper, have lower operating costs and require less power consumption than their benchtop-equivalents [20]. However, there are also some drawbacks to consider when using a portable ATR-FT-MIR spectrometer. In fact, they tend to record noisier spectra than the benchtop ones and are sometimes considered less accurate and less reliable. Nevertheless, depending on the samples and on the problem to be addressed, this higher noise does not necessarily imply poorer performances when data are properly treated with chemometric tools [see e.g. [21]].

From all these considerations, it follows that there are different parameters that condition the spectra obtained. This is why in this tutorial we propose the use of an ATR-FT-MIR portable spectrometer (4100 ExoScan FTIR, Agilent, California, USA) focusing the attention on the parameters to be controlled and optimized to register good quality spectra.

Regarding the usual values of spectral resolution and number of acquisitions, these range between 2 and 32 cm^{-1} and from 32 to 512 scans, respectively. In general terms, the lower the first parameter and the higher the second, the better the signal quality in terms of signal-to-noise ratio. Nevertheless, the time of analysis must also be taken into account, as it increases when increasing the number of scans to get the best signal, so a compromise is needed between the two [22]. In addition, when a 2 cm^{-1} resolution is used, the spectra obtained often show a larger baseline noise that requires a strong smoothing treatment before data analysis with the subsequent loss of information.

The MIR region between 1800 and 900 cm^{-1} has been reported to be the fingerprint of the alcoholic fermentation signals, as the main biochemical bonds absorb in this area [23], so many studies have focused on this region to build the models [18,24,25]. However, there are other important bands in the spectrum related to water and ethanol that can also be used to build the models, as shown by Cavaglia et al. when, to predict the pH, the whole spectral range was needed to obtain good results [18].

Other important points to be considered are the cleaning of the ATR crystal and the background collection. In addition to good equipment maintenance, cleaning must be done after each measurement to avoid carry-over effects between samples [6]. It is crucial especially at the beginning of the alcoholic fermentation, when the sugar concentration values are at their maximum, making the sample more viscous. However, the great solubility of sugars, acids and alcohols in water makes cleaning with deionized water quite successful. It is important to evaluate cleanliness based on the results obtained with the background spectra recorded after each cleaning step.

Concerning background collection, two strategies have been described in the literature. In the first one, the background is collected with the crystal empty, providing what is known as an air background [26]. In the second option the background is collected by using deionized water so that the signal obtained allows to subtract the water contribution from the sample spectrum [27]. To our experience an air background has proven to be effective [22].

Finally, it is important to remark that, as IR lamps heats up over time, an equilibrium time is necessary. To be more precise, when switching on the equipment, the spectral intensity changes slowly in such a way that, in order to obtain a stable signal, in the case of our portable ExoScan FTIR instrument we have to wait approx. 100 min, as can be seen in Fig. S1 [12]. Stabilization time is highly dependent on the FTIR equipment and can be lower for equipment with temperature control.

4. Data analysis

4.1. Spectra pre-processing

The raw signal must be preprocessed to make it more suitable for multivariate analysis and to obtain good quality chemometrics models. It has been shown that the application of optimal pre-processing is critical for PAT, as the resulting models can increase their performance by up to 25% [28]. In the case of FTIR spectroscopic data, the main objective of pre-processing is to ensure that the signal follows the Beer's law as much as possible [11]. This means removing or minimizing all sources of variation not related to the process studied and this includes removing noise, offsets and baseline drifts and light scattering effects. Typical pre-processing methods used for MIR spectra are divided into two groups: 1) scatter correction methods, which include Standard Normal Variate (SNV), Multiplicative Scatter Correction (MSC) and

baseline correction; and 2) derivation methods, which include Savitzky-Golay smoothing/derivative as the most used [26,28]. SNV is often used in ATR-MIR spectra to remove the variability due to physical aspects or equipment characteristics, which may generate scatter. Savitzky-Golay smoothing coupled with derivative is useful to reduce baseline noise and simultaneously emphasize small peaks when needed [26]. In our case it is advantageous because the spectra show high absorbance peaks due to water, sugar and ethanol depending on the fermentation time, while peaks related to compounds indicating an abnormal fermentation (e.g. lactic, acetic or malic acids) are smaller and quite overlapped. The smoothing window must be carefully evaluated because a too severe smoothing can remove useful information contained in the spectra.

4.2. Model building

Once the data are preprocessed, multivariate data analysis is applied to extract the information from the data, in this case from the spectra. Data analysis aims at finding relationships between samples and variables, detecting trends in the data and also detecting anomalous samples [29]. The different multivariate analysis methods available can be classified, depending on their purpose, into exploratory, classification or prediction methods, among others.

Exploratory analysis is usually performed as a first step prior to more complex data analysis to provide some preliminary information, such as the presence of sample groupings or the existence of outliers, or to evaluate the precision (repeatability) of the measurements. The gold-standard exploratory technique is principal component analysis (PCA). PCA builds a set of reduced principal components (PCs), which are linear combinations of the original spectra (**X** matrix) and keep the maximum information contained in them. This way, simple two-dimensional plots (scores and loadings) can be plotted to visualize the data.

Classification techniques aim at finding a criterion to assign a sample to a predefined class (or category), each class representing a group of samples sharing specific characteristics [30]. Some of the most used classification techniques are Linear Discrimination Analysis (LDA), Partial Least Squares – discriminant analysis (PLS-DA) and *k* nearest neighbours (*k*NN).

Finally, quantitative predictive modelling aims at correlating a set of spectra (**X** matrix) to one or some properties (**Y** matrix). A training set of samples is used to build and optimise the model (where **X** and **Y** are known) and the final objective is to predict the **Y** properties of a new and unknown set of samples [6]. The multivariate predictive technique most commonly used is Partial Least Squares Regression (PLSR). Real examples using the three model categories are shown from Section 5.2 to 5.4.

4.3. Model validation

In all methods, a critical step is the selection of the optimal number of latent variables (LVs). In the case of PCA, the simplest way is to choose those that explain a given percentage of the variance in the spectra (i.e. 95%), although in some cases some LVs accounting for the remaining variance may be also informative. Cross-validation strategies can also be applied (see below). In the case of classification models, and in particular of PLS-DA models, the choice of the optimal number of LVs is performed based on the maximum number of correctly classified samples. i.e. accuracy, in the validation set. However, depending on the problem at hand, sometimes a balance between sensitivity (proportion of out-of-control samples correctly classified) and specificity (proportion of under control samples correctly classified) is sought [31]. Finally, for prediction models, the optimal number of latent variables is selected based on the minimum value of prediction error, estimated for the validation set.

The selection of the optimal number of latent variables is performed during method validation. There are basically two approaches for method validation, depending on the number of samples available: cross-validation and test set validation. Cross-validation is usually applied when the number of samples is low. In cross-validation the original dataset is split in different blocks. Then each block is left-out once at a time, the model is built with the rest of the samples and the left-out block is predicted. This process is repeated for each block and for different LVs. Test-set validation consists on leaving out a subset (usually 30–40%) of the original data, build the model with training set and decide the optimal number of LVs based on the prediction of the test set. The test samples must fall within the calibration range and have similar physicochemical characteristics. Finally, the number of misclassified samples (for PLS-DA) or the residuals, that is the difference between predicted and measured values (for PLSR) are calculated and a global estimator of the validation error is obtained. For classification models, these estimators are accuracy, sensitivity and specificity, and for prediction models we have the Root Mean Square Error of Cross-Validation (RMSECV) or the Root Mean Square Error of Prediction (RMSEP), the latter for test set validation.

Cross-validation is an important part of model building to avoid over-optimistic models [32]. There are different types of cross-validation, which are distinguished from each other depending on the pattern used when selecting the samples to validate the model. As examples, leave-one-out cross-validation is mainly used for small data matrices, as just one sample per time is used to test. For time-organized data, venetian blinds are useful to asses non-temporarily errors while contiguous block asses temporal stability. Finally, typical errors in CV setup include making only one split of the data, remove small groups of samples or splitting unnatural replicates into calibration and test groups.

4.4. Variable selection

Nowadays analyzers provide hundreds of data for every sample in a very short time, but sometimes there may be spectral regions not suitable for modelling, that is, not related to the problem at hand. Variable selection (VS) is then an important part in method validation, as it allows detecting the specific spectral regions of interest and produce better models. VS can be performed using algorithms or using chemical knowledge. In the latter case, the operator knows that a certain region of the spectra could be attributed to a given family of molecules. In many cases, a combination of both strategies is applied to obtain better results. There are many algorithms available for VS, such as genetic algorithms, interval PLS, recursive PLS, selectivity ratio or variable importance in projection (VIP), to name just a few (see [33] and references therein). All algorithms seek to obtain the best combination of variables, and at this point validation is fundamental (as explained in the Model Validation section) to avoid overfitting (the model just explains the actual samples and no future samples).

5. Alcoholic fermentation monitoring. Real examples

5.1. Data arrangement

The data collected from a fermentation process consist of a series of spectra measured at certain time points and can be structured in a two-dimensional matrix, **X** ($K \times J$), where K are the sampling times and J the wavenumbers, which are usually expressed as hours from the inoculation time (time 0).

If data are available for several samples, then data are structured in a three-dimensional matrix, **X** ($I \times J \times K$), where I are the samples. This three-way matrix includes the process variability and from it, a NOC (Normal Operation Conditions) space can be defined. It is worthwhile to note that the three-way matrix can be generated only if the sampling times are the same, that is, when spectra are collected at the same time after the start of fermentation, or at the same time intervals.

Depending on the goal of the investigation, the three-way matrix **X** ($I \times J \times K$) is usually re-arranged (unfolded) either time-wise ($IK \times J$) or batch-wise ($I \times JK$) [9]. In the following sections we will discuss what can be obtained when monitoring a fermentation process by ATR-FT-

MIR spectroscopy and the different unfolding and some of the most used modeling strategies necessary to achieve the goals.

5.2. Alcoholic fermentation monitoring

The evolution of the fermentation can be visualized in an easy-to-see way using unfolded PCA (Principal Component Analysis) modeling. PCA compresses the information contained in the original variables (i.e. spectra) into principal components (PCs) to better visualize, usually through bidimensional plots, trends in the samples during the process, or detect abnormal data (outliers).

In this case, the data matrix is organized with samples in the rows and wavelengths in the columns, and in case of several batch samples the time-wise unfolding ($IK \times J$) is employed. Prior to unfolded PCA, the data matrix is column mean centered, so the average considers every sample and time.

The main changes during alcoholic fermentation are usually reflected in the score values of the first PC, as it contains the maximum variability of the samples. It has been shown by Buratti et al. [17] that the first PC plotted versus time fits the fermentation kinetic evolution (Eq. (1)), as the general form of the Gompertz equation.

$$C = C_{\infty} \cdot exp\left\{ - exp\left[\frac{\mu_{max} \cdot e}{C_{\infty}}(\lambda - t) + 1\right]\right\}$$

μ_{max} representation of maximum specific growth rate, λ length of lag phase, C_{∞} curve asymptote, depending of the compounds considered, ethanol or sugars, C is the concentration at time t or the concentration of sugars consumed, respectively.

A similar trend is observed by visual inspection of the sigmoid functions of the first PC versus time and density versus time [see e.g. [18,34] and Fig. 1]. This behavior is confirmed by a good mathematical correlation between the scores of the first PC and the density values.

To obtain the evolution of the scores over time shown in Fig. 1a, we followed 5 alcoholic fermentations inoculated with $3 \cdot 10^6$ CFU·mL^{-1} of *Saccharomyces cerevisiae*. The dimension of the data matrix was (5 × 845 × 48), which was time-wise unfolded into a matrix of (240 × 845). The unfolded matrix was preprocessed using a 15-points smoothing, then SNV and finally column mean centering. After this, the PCA model was built and the trend observed in the PC1 score values clearly recalls the evolution of density during the fermentation process, as can be seen on the graph of density measurements (Fig. 1b) throughout the process. Fig. 1c shows that a good linear correlation (>0.99%) is obtained between the mean values of the scores on first PC and the mean density for the times in which both measurements were carried out. After PCA modelling new samples can be qualitative evaluated with this methodology. For that purpose, the same data unfolding is performed, data are centered using the mean of the calibration data and new samples are finally projected onto the PCA model.

While PCA returns qualitative information, Multivariate Statistical Process Control (MSPC) evaluates, with statistical limits, whether the process is running under control. The scores and the residuals of the PCA model are used to build Hotelling T^2 and Q charts, respectively [35].

To reproduce the process conditions in a small-scale laboratory experiment, different fermentation batches must be monitored, and samples measured over time following the rules mentioned above. It is also possible to consider additional sources of variability, such as the initial sugar concentration or the ripening state of the grapes. A PCA model is then built, in which scores (from one or more PCs) and residuals are used as parameters to define the statistical limits (samples are considered to belong to a normal distribution) of the control charts (T^2 and Q) and to monitor them for future batches.

Once the model has been built, and the statistical limits have been calculated, new samples from an ongoing process are projected onto the model to determine whether the process is running under normal operation conditions (NOC). For that, the projected samples must have

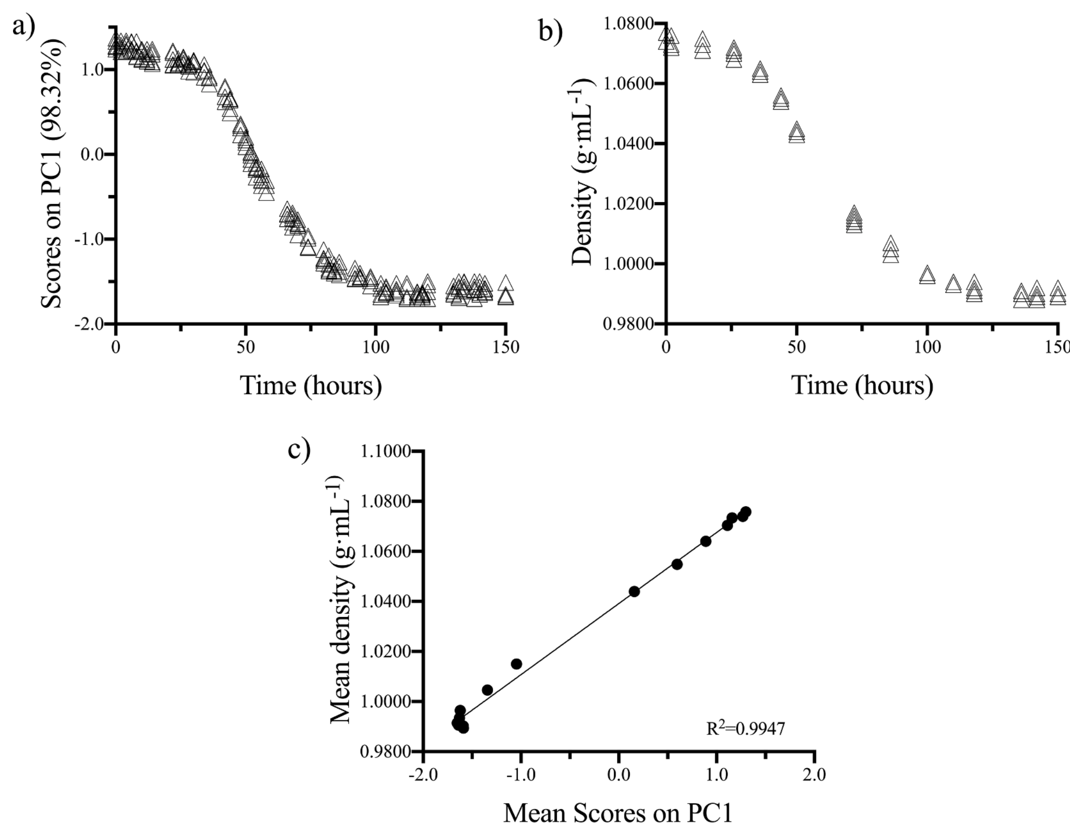

Fig. 1. PC1 score values plotted versus time (a), density plotted versus time (b) during alcoholic fermentation and mean PC1 score values plotted versus mean density (c).

the same spectral range and must have been pretreated in the same way as the modelling samples. New data are centered with the mean obtained from calibration data. Literature shows a wide range of applications and types of control charts [18,36,37] and references therein]; for example, Cavaglia et al. have shown that building control charts with various statistical parameters (PC1 and PC2 scores values [18], Q residuals and Hoteling T^2 [34]; or evolution over time) allows detecting unwanted subprocesses, such as malolactic fermentation. Malolactic fermentation is related to the presence of unwanted microorganisms, lactic acid bacteria, which exist in the ecology of the vineyard and also in the winery [38]. These microorganisms, like other contaminants (yeasts and acetic acid bacteria), produce their metabolites in low amounts, making it difficult to detect them using the whole spectral range. This is the reason why, when it comes to suspected contamination (mainly when the tracked parameters are in the minority), it is necessary to focus on the specific region of the substrate, the product or both [18,39]. An example of this application is shown in Fig. 3.

Fig. 2 shows the evolution over time of Q residuals (from a PCA model built with 2 PCs) obtained from 10 normal alcoholic fermentations and 4 fermentations intentionally contaminated with lactic acid bacteria at a concentration of $2.5{\cdot}10^6$ CFU·mL^{-1}. The best PCA model was obtained when using specifically the region of organic acids from 1250 to 1089 cm^{-1}. It may be explained as the metabolism of lactic acid bacteria consists of converting malic acid into lactic acid which absorb in that specific region. The data matrix obtained was of dimension (14 × 62 × 34) and it was time-wise unfolded into an array of (506 × 62). After unfolding, the data matrix was preprocessed by 15-points smoothing, SNV and all fermentations were mean centered.

To detect process deviations, PCA-based control charts can be applied. However, to meet the process control requirements and assigning the cause of the deviation the contribution plots are required [35]. Contribution plots show the main regions of the spectra that are used to differentiate normal and abnormal process samples, as shown in Fig. 2b. Contribution plots are obtained as the multiplication of the loadings by the spectra, which increases the information of the spectra related to specific variables. Therefore, for a certain spectral region, deviations related to specific compounds can be detected. As it can be seen in the contribution plot (right), each type of fermentation showed a different trend.

5.3. Prediction of physicochemical parameters

As the IR spectrum contains information about the main species present in the sample, it is possible to correlate the changes in the IR spectrum with the evolution of the process, i.e. changes in the reacting

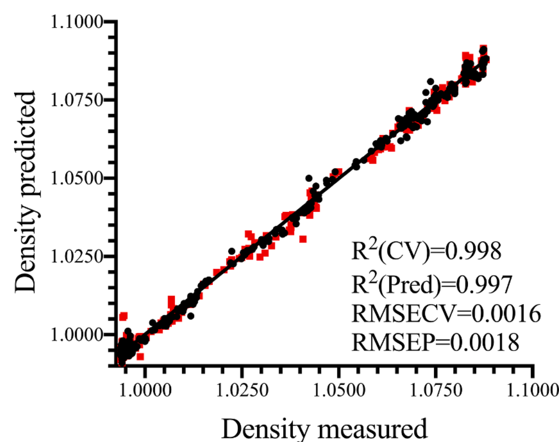

Fig. 3. PLS model to predict density (g·mL^{-1}) along alcoholic fermentation.

species. For this purpose, Partial Least Square (PLS) regression is usually applied. The PLS regression method works like PCA, compressing the spectral information into a few latent variables, but also correlating this spectral information to a **Y** matrix that contains the values of a physicochemical parameter to be predicted. Typical oenological parameters in wineries are density, pH, volatile acidity, total acidity, and initial YAN concentration, but it has been shown that is also possible to predict other properties, such as organoleptic attributes, polyphenols and other minor compounds [25].

To build a PLS model, it is necessary to use a reference or standard analytical method to estimate the *y* values, in the same way that the use of standards is required in many other analytical methods. Therefore, to monitor and control the conventional parameters, it is necessary to analyze at the early stages every sample with both methods, the standard and the infrared.

To build the PLS model a time-wise unfolding (*IK* × *J*) is necessary to statistically correlate the spectrum at a given time with the measured property. This also means that the samples must have been analyzed by the reference methods of analysis, which often are time consuming. This limiting factor implies a decrease in the number of points that can be taken to obtain correlation between a property and the spectrum, as it is illustrated in Fig. 1.

When a new sample is analyzed, its similarity to the calibration data is evaluated through the model, which predicts the parameter value. Projected data are spectrally preprocessed as the calibration samples and centered using the mean of the calibration data. The prediction ability of the calibration models is assessed by the Root Mean Square

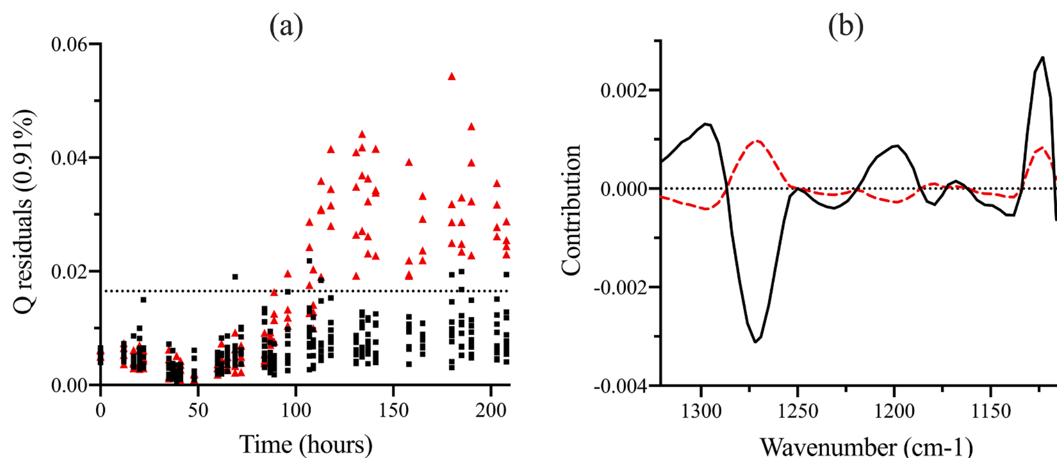

Fig. 2. PCA based control chart (left) and the mean contribution plot (right) for normal process samples (squares) and deviated process samples (triangles and dashed line).

Error of Prediction (RMSEP) or standard error of prediction (SEP). In addition, when the number of samples is not enough to split the samples into calibration and validation sets, it is possible to evaluate the prediction of the model through Root Mean Square Error of Cross-Validation (RMSECV). It is important to note that an averaged error (along the whole range) is provided by this methodology [25].

The prediction of wine parameters has been extensively studied by varying vintage and grape varieties in order to cover typical oenological parameter ranges [40]. This methodology should be considered in the context of process control, to obtain robust models, since the samples must cover the full range of variation in a given process and also between processes with different initial characteristics [41]. The prediction of sugars (glucose and fructose) allows building a control chart to plot this parameter versus time. This has been shown by several authors, obtaining good prediction errors: 5 $g \cdot L^{-1}$ and 2 $g \cdot L^{-1}$ at the end of the fermentation [41], or even 10.9 $g \cdot L^{-1}$ with a portable ATR-MIR [22]. Other parameters typically predicted when monitoring the alcoholic fermentation are density and pH. They have been predicted using a portable ATR-MIR with an error of prediction of 0.0014 $g \cdot mL^{-1}$ and 0.07, respectively [18]. Other important parameters, such as phenolic compounds, anthocyanins and flavonoids, which are important in red wine alcoholic process control, have also been predicted using unfolded PLS regression [42].

It is also possible to build a control chart through the prediction of a property using the unfolded PLS algorithm, and to establish a critical limit, from a legal or quality perspective, to decide what samples are under or out-of-control (Fig. 3).

The above mentioned unfolded PLS regression model (shown in Fig. 3) is an example of the density prediction ability of ATR-MIR. The example uses a **X** calibration data matrix of (580 × 845) that was correlated to a **Y** matrix of (580 × 1). As it can be seen both matrixes have the same rows as each sample used for prediction must be analyzed by standard physicochemical analysis, in this case by a portable densimeter. Also, in Fig. 3, **X** test data matrix of (528 × 845) that also has an **Y** matrix of (528 × 1) were used to stablish prediction error. The prediction ability of the models is deeply dependent on the number of samples used, as more samples means a better estimation of the sample variability. Moreover, having more samples allows to improve the validation of the models.

It is also possible to detect process deviations if some compounds, such as acetic acid, are predicted. Acetic acid is a compound generated by yeast in order to obtain energy in sub-optimal process conditions, such as Yeast Assimilable Nitrogen (YAN) shortfall or when there is an increase in process temperature. Urtubia et al. predicted sugars, ethanol, glycerol, succinic and acetic acids and showed that the small errors obtained made possible to detect the miss-behavior from a temperature-gradient fermentation and a YAN shortfall fermentation [43].

5.4. Prediction of fermentation evolution

The spectral information can also be used to predict the time points of the fermentation as alcoholic fermentation is a bioprocess evolving over time. Cozzolino et al. have showed that it is possible to predict the time course of a wild fermentation, with an error of 1.21 days [44]. This methodology would be applicable to detect stuck and sluggish fermentations, as the predicted value of time course of fermentation would be smaller than the actual one. From a practical oenological point of view, an error of more than one day is not a suitable solution. Therefore, another methodology, such as the biological time, is needed to decrease the error.

Biological time, firstly introduced by Jorgensen et al. [45], allows detecting slight differences in the fermentation process due to the metabolism of yeasts. First, the data is relativized in a scale from 0 to 1, afterwards PLS regression is applied, and the calibration equation is used to predict the time of the spectra used. This circular approach allows determining the biological time, due to small differences in sugar consumption along fermentations. Then, PLS is re-applied with the predicted biological time. The scores convey into an alcoholic fermentation control chart and a 95% confidence limits are calculated. The calibration model will be reused in future fermentations and the scores will be plotted in the control chart to determine how normal the time involved is.

Cavaglia et al. have shown that the biological time approach described above provides good results to detect fermentations with small yeast assimilable nitrogen concentration (YAN). The control chart showed that a YAN fermentation was in the 0.6 of the biological time at the time it should be 1 [22].

As shown in Fig. S2, through biological time it is possible to obtain the confidence intervals of PLS factor 1. The dimension of the spectral data matrix was (20 × 845 × 9) and it was time-wise unfolded into a matrix of (180 × 845). The unfolded spectral matrix was preprocessed using an 11-points smoothing, then SNV and finally column mean centered. The **Y1** matrix (180 × 1) that contained the times of the alcoholic fermentation process was relativized from 0 to 1. The predicted parameters were obtained in the **Y2** matrix (180 × 1). The prediction was re-built with the **Y2** matrix, obtaining the biological time in the **Y3** matrix. Besides, the confidence intervals were calculated with the values of PLS factor 1 using **Y3** [22,45]. A new spectral data matrix (27 × 845) of 3 nitrogen deficient alcoholic fermentations was projected onto the model to predict their biological time, and its representation (Fig. S2) showed slower fermentation kinetics.

5.5. Detection of deviations from NOC

To determine whether the samples are under or out-of-control, it is possible to apply PLS-DA. In PLS-DA models, a regression is performed between the spectra (**X** matrix) and a **y**-vector containing a dichotomous variable (typically 0 and 1) that expresses the type of process, under and out-of-control, respectively. This methodology can improve the detection of deviations, as low concentrated compounds, which are under the quantification threshold, could not be predicted but the overall spectral changes are sufficient to detect deviation. With this methodology, local and batch-wise unfolding approaches may be used.

Local unfolding is used to determine if the behavior of a sample in a specific time or time gap is under-control. If the sampling pattern is reproducible over the different measured processes, that is, if sampling is always performed at the same time during the fermentation with a given time interval (e.g. each k hours), a local k time unfolding is used summoning the same time for all the samples. In wineries, the sampling pattern may not be standard over years, so in order to use historical data it is necessary to use a gap time for the models. We have shown that an 8-hour approach improves process control and allows corrective actions to be taken [34].

To increase the classification performance it is even possible to apply a moving window approach, firstly introduce by Camacho et al. [46]. In this approach, a k number of times are gathered, and PLS-DA is applied (from time n to time $n+k$). The models are built by moving the times used, from time $\mathrm{n}+1$ to time $\mathrm{n}+\mathrm{k}+1$. This approach is partially batch-wise ($I \times JK'$) unfolded, where K' is a subset of K. Besides, using the batch-wise unfolding the result will express if the whole process is under NOC conditions.

The performance evaluation of the PLS-DA models is done using a discrimination threshold (between 0 and 1), which is calculated from the probability of classification error of the samples in the classes. The discrimination threshold is calculated taking into account a Gaussian distribution of the predicted classes, and the y value at which the two curves converge is the discrimination threshold [47] (see blue and red dashed lines in Fig. 4). The optimization of the discrimination threshold is based on the Receiver Operating Characteristic (ROC) curve, which is a graphical representation of the specificity and sensitivity variation as a function of the threshold. Specificity is the ratio of true positives to the total of true positives plus false negatives; and selectivity is the ratio of

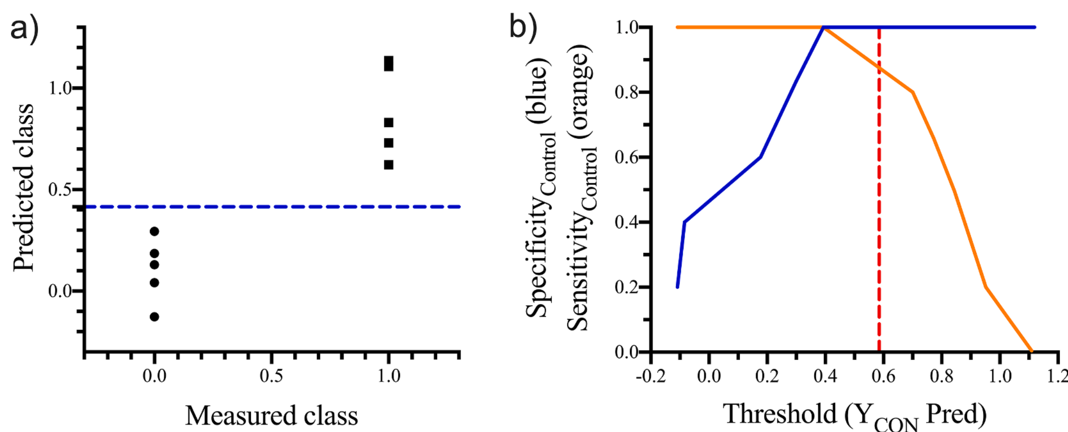

Fig. 4. a) PLS-DA model for under control (zero) and out-of-control (one) samples. b) Receiver Operating Characteristic curve for the under control class.

true negatives to the total of true negatives plus false positives Cavaglia et al. have shown that local approaches to distinguish YAN shortfall [22] and lactic acid bacteria contamination [18] provided 100% correct classification of the classes. However, when applying the moving window approach in a YAN shortfall process, the lack of sampling points did not show better results [22].

The PLS-DA model shown in Fig. 4a was built using 5 normal alcoholic fermentations and 5 fermentations that were intentionally contaminated by co-inoculating a population of $1{\cdot}10^6$ $CFU{\cdot}mL^{-1}$ lactic acid bacteria, to reproduce typical contamination at the start of the alcoholic fermentation. Data used in this *k* model is from 213 h, as in this point the 50% of malic acid was consumed. Data matrix was (10 × 62), as the model focused only in the organic acids' region from 1250 to 1089 cm^{-1}.

Fig. 4b shows the ROC curve for the under control class. The threshold for under control samples is 1-threshold for out-of-control samples. This is why the threshold is calculated to maximize both sensitivity and selectivity for the two categories. In the example, under control class threshold is 0.5849 and for out-of-control class is 0,4151 (which is the one showed in Fig. 4a). In the example, specificity for both classes in the optimized thresholds is 1; and selectivity is 0,925 for out-of-control class (data not shown) and 0.875 for control class.

6. Conclusions

This tutorial provides practical methodologies that can be used to study wine alcoholic fermentation at a laboratory microscale with ATR-FT-MIR spectroscopy and following PAT recommendations. The speed and portability of current FT-MIR equipment are characteristics that wineries demand, so this technique can have a wide application in this field. In that sense, many of the given recommendations can be extended to other vibrational techniques.

It should be noted that the combination of spectroscopy with chemometrics allows to obtain many features, almost at the same time, from the same dataset such as the monitoring of the fermentation, the prediction of relevant parameters and even the detection of deviations.

Finally, this tutorial can help the oenological researcher, who is not usually familiar with chemometrics, not to get lost among all the chemometric approaches available in the literature. Thus, this report explains the chemometric techniques, without deepening into chemometric algorithms but keeping the scientific rigor, which can be applied in each case followed by a discussion of the results that can be obtained.

Funding

This project has been possible with the support of the Ministerio de Ciencia e Innovación, the Agencia Estatal de Investigación (AEI) (project PID2019-104269RR-C33); and with the support of the Ministerio de Ciencia, Innovación y Universidades, the AEI and the European Social Fund (FEDER) (project AGL2015-70106-R). This publication has been possible with the support of the Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement de la Generalitat de Catalunya (2020 FISDU 00221; Schorn-Garcia, D.). Cavaglia, J. would also like to thank the Catalan Agency for Management of University and Research Grants (AGAUR) for the FI grant (FI_B100154).

CRediT authorship contribution statement

Daniel Schorn-García: Investigation, Methodology, Visualization, Writing - original draft. **Julieta Cavaglia:** Investigation, Methodology, Writing - review & editing. **Barbara Giussani:** Conceptualization, Methodology, Supervision, Writing - review & editing. **Olga Busto:** Conceptualization, Funding acquisition. **Laura Aceña:** Resources, Project administration. **Montserrat Mestres:** Conceptualization, Methodology, Supervision, Writing - review & editing. **Ricard Boqué:** Conceptualization, Funding acquisition, Formal analysis, Software, Supervision, Writing - review & editing.

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.microc.2021.106215.

References

[1] P. Ribéreau-Gayon, D. Dubourdieu, B. Donèche, A. Lonvaud, Handbook of Enology Volume 1. The Microbiology of Wine and Vinifications, 2nd ed., 2006. 10.1002/0470010363.fmatter.

[2] A. Urtubia, G. Hernández, J.M. Roger, Detection of abnormal fermentations in wine process by multivariate statistics and pattern recognition techniques, J. Biotechnol. 159 (4) (2012) 336–341, https://doi.org/10.1016/j.jbiotec.2011.09.031.

[3] D. Cozzolino, State-of-the-art advantages and drawbacks on the application of vibrational spectroscopy to monitor alcoholic fermentation (beer and wine), Appl. Spectrosc. Rev. 51 (4) (2016) 302–317, https://doi.org/10.1080/05704928.2015.1132721.

[4] Guidance for Industry PAT — A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance, FDA Off. Doc. (2004) 16.

[5] R.E. Jerome, S.K. Singh, M. Dwivedi, Process analytical technology for bakery industry: a review, J. Food Process Eng. 42 (2019) 1–21, https://doi.org/10.1111/jfpe.13143.

[6] F. van den Berg, C.B. Lyndgaard, K.M. Sørensen, S.B. Engelsen, Process analytical technology in the food industry, Trends Food Sci. Technol. 31 (1) (2013) 27–35, https://doi.org/10.1016/j.tifs.2012.04.007.

[7] Y.Y. Pu, C. O'Donnell, J.T. Tobin, N. O'Shea, Review of near-infrared spectroscopy as a process analytical technology for real-time product monitoring in dairy processing, Int. Dairy J. 103 (2020) 104623, https://doi.org/10.1016/j.idairyj.2019.104623.

[8] N.D. Lourenço, J.A. Lopes, C.F. Almeida, M.C. Sarraguça, H.M. Pinheiro, Bioreactor monitoring with spectroscopy and chemometrics: a review, Anal. Bioanal. Chem. 404 (4) (2012) 1211–1237, https://doi.org/10.1007/s00216-012-6073-9.

[9] J.U. Porep, D.R. Kammerer, R. Carle, On-line application of near infrared (NIR) spectroscopy in food production, Trends Food Sci. Technol. 46 (2) (2015) 211–230, https://doi.org/10.1016/j.tifs.2015.10.002.

[10] S. Grassi, C. Alamprese, Advances in NIR spectroscopy applied to process analytical technology in food industries, Curr. Opin. Food Sci. 22 (2018) 17–21, https://doi.org/10.1016/j.cofs.2017.12.008.

[11] D.W. Sun, Infrared Spectroscopy for Food Quality Analysis and Control, 1st ed.,, Elsevier, 2009.

[12] International Organization of vine and wine (OIV), International Oenological Codex, 2019.

[13] A. Aranda, E. Matallana, M. del Olmo, Saccharomyces yeasts I: primary fermentation, in: Mol. Wine Microbiol., 2011, pp. 1–31. 10.1016/B978-0-12-375021-1.10001-3.

[14] International Organization of vine and wine (OIV), Chemical Acidifcation of Musts, in: Int. Code Oenological Pract., 2015, pp. II.2.1-8.

[15] Sally.-Jean. Bell, Paul.A. Henschke, Implications of nitrogen nutrition for grapes, fermentation and wine, Aust. J. Grape Wine Res. 11 (3) (2005) 242–295, https://doi.org/10.1111/j.1755-0238.2005.tb00028.x.

[16] B. Padilla, J.V. Gil, P. Manzanares, Past and future of non-Saccharomyces yeasts: from spoilage microorganisms to biotechnological tools for improving wine aroma complexity, Front. Microbiol. 7 (2016) 1–20, https://doi.org/10.3389/fmicb.2016.00411.

[17] S. Buratti, D. Ballabio, G. Giovanelli, C.M.Z. Dominguez, A. Moles, S. Benedetti, N. Sinelli, Monitoring of alcoholic fermentation using near infrared and mid infrared spectroscopies combined with electronic nose and electronic tongue, Anal. Chim. Acta. 697 (1-2) (2011) 67–74, https://doi.org/10.1016/j.aca.2011.04.020.

[18] J. Cavaglia, D. Schorn-García, B. Giussani, J. Ferré, O. Busto, L. Aceña, M. Mestres, R. Boqué, ATR-MIR spectroscopy and multivariate analysis in alcoholic fermentation monitoring and lactic acid bacteria spoilage detection, Food Control. 109 (2020) 106947, https://doi.org/10.1016/j.foodcont.2019.106947.

[19] A. Subramanian, L. Rodriguez-Saona, Fourier Transform Infrared (FTIR) Spectroscopy, 1st ed., Elsevier Inc., 2009. 10.1016/B978-0-12-374136-3.00007-9.

[20] R.A. Crocombe, Portable spectroscopy, Appl. Spectrosc. 72 (12) (2018) 1701–1751, https://doi.org/10.1177/0003702818809719.

[21] G. Gorla, M. Mestres, R. Boqué, J. Riu, D. Spanu, B. Giussani, ATR-MIR spectroscopy to predict commercial milk major components: a comparison between a handheld and a benchtop instrument, Chemom. Intell. Lab. Syst. 200 (2020) 103995, https://doi.org/10.1016/j.chemolab.2020.103995.

[22] J. Cavaglia, B. Giussani, M. Mestres, M. Puxeu, O. Busto, J. Ferré, R. Boqué, Early detection of undesirable deviations in must fermentation using a portable FTIR-ATR instrument and multivariate analysis, J. Chemom. 33 (2019) e3162, https://doi.org/10.1002/cem.3162.

[23] D. Cozzolino, Sample presentation, sources of error and future perspectives on the application of vibrational spectroscopy in the wine industry, J. Sci. Food Agric. 95 (5) (2015) 861–868, https://doi.org/10.1002/jsfa.2015.95.issue-510.1002/jsfa.6733.

[24] S. Bureau, D. Cozzolino, C.J. Clark, Contributions of Fourier-transform mid infrared (FT-MIR) spectroscopy to the study of fruit and vegetables: a review, Postharvest Biol. Technol. 148 (2019) 1–14, https://doi.org/10.1016/j.postharvbio.2018.10.003.

[25] C.A.T. dos Santos, R.N.M.J. Páscoa, J.A. Lopes, A review on the application of vibrational spectroscopy in the wine industry: from soil to bottle, Trends Anal. Chem. 88 (2017) 100–118, https://doi.org/10.1016/j.trac.2016.12.012.

[26] A. Rinnan, L. Norgaard, F. van den Berg, J. Thygesen, R. Bro, S.B. Engelsen, Data pre-processing, in: D.-.-W. Sun (Ed.), Infrared Spectrosc. Food Qual. Anal. Control 1st ed., Elsevier Ltd, 2009, pp. 29–50.

[27] Z. Wu, E. Xu, J. Long, Y. Zhang, F. Wang, X. Xu, Z. Jin, A. Jiao, Monitoring of fermentation process parameters of Chinese rice wine using attenuated total reflectance mid-infrared spectroscopy, Food Control. 50 (2015) 405–412, https://doi.org/10.1016/j.foodcont.2014.09.028.

[28] Å. Rinnan, F.V.D. Berg, S.B. Engelsen, Review of the most common pre-processing techniques for near-infrared spectra, TrAC - Trends Anal. Chem. 28 (10) (2009) 1201–1222, https://doi.org/10.1016/j.trac.2009.07.007.

[29] N. Kumar, A. Bansal, G.S. Sarma, R.K. Rawal, Chemometrics tools used in analytical chemistry: an overview, Talanta 123 (2014) 186–199, https://doi.org/10.1016/j.talanta.2014.02.003.

[30] L.A. Berrueta, R.M. Alonso-Salces, K. Héberger, Supervised pattern recognition in food analysis, J. Chromatogr. A. 1158 (1-2) (2007) 196–214, https://doi.org/10.1016/j.chroma.2007.05.024.

[31] D. Ballabio, F. Grisoni, R. Todeschini, Multivariate comparison of classification performance measures, Chemom. Intell. Lab. Syst. 174 (2018) 33–44, https://doi.org/10.1016/j.chemolab.2017.12.004.

[32] F. Westad, F. Marini, Validation of chemometric models - a tutorial, Anal. Chim. Acta 893 (2015), https://doi.org/10.1016/j.aca.2015.06.056.

[33] T. Mehmood, K.H. Liland, L. Snipen, S. Sæbø, A review of variable selection methods in Partial Least Squares Regression, Chemom. Intell. Lab. Syst. 118 (2012) 62–69, https://doi.org/10.1016/j.chemolab.2012.07.010.

[34] J. Cavaglia, D. Schorn-García, B. Giussani, J. Ferré, O. Busto, L. Aceña, M. Mestres, R. Boqué, Monitoring wine fermentation deviations using an ATR-MIR spectrometer and MSPC charts, Chemom. Intell. Lab. Syst. 201 (2020) 104011, https://doi.org/10.1016/j.chemolab.2020.104011.

[35] S. Bersimis, J. Panaretos, S. Psarakis, Multivariate statistical process control charts and the problem of interpretation: a short overview and some applications in industry, Qual. Reliab. Eng. Int. 23 (2009) 517–543. http://arxiv.org/abs/0901.2880.

[36] T.A. Catelani, J.R. Santos, R.N.M.J. Páscoa, L. Pezza, H.R. Pezza, J.A. Lopes, Real-time monitoring of a coffee roasting process with near infrared spectroscopy using multivariate statistical analysis: a feasibility study, Talanta 179 (2018) 292–299, https://doi.org/10.1016/j.talanta.2017.11.010.

[37] T.C. Ávila, R.J. Poppi, I. Lunardi, P.A.G. Tizei, G.A.G. Pereira, Raman spectroscopy and chemometrics for on-line control of glucose fermentation by Saccharomyces cerevisiae, Biotechnol. Prog. 28 (6) (2012) 1598–1604, https://doi.org/10.1002/btpr.v28.610.1002/btpr.1615.

[38] E. Vaudano, A. Costantini, E. Garcia-Moruno, Microbial ecology of wine, in: Model. Microb. Ecol. Foods Quant. Microbiol. Food Process., 2017, pp. 547–559. 10.1002/9781118823071.ch28.

[39] S. Grassi, Microbial food fermentations: innovative approach using infrared spectroscopy, Università degli Studi di Milano. Milan (Italy) (2013), https://doi.org/10.1146/annurev.genet.37.110801.142807.

[40] D. Cozzolino, W. Cynkar, N. Shah, P. Smith, Feasibility study on the use of attenuated total reflectance mid-infrared for analysis of compositional parameters in wine, Food Res. Int. 44 (1) (2011) 181–186, https://doi.org/10.1016/j.foodres.2010.10.043.

[41] V. Bellon, Fermentation control using ATR and an FT-IR spectrometer, Sensors Actuators B. Chem. 12 (1) (1993) 57–64, https://doi.org/10.1016/0925-4005(93)85015-3.

[42] V. Di Egidio, N. Sinelli, G. Giovanelli, A. Moles, E. Casiraghi, NIR and MIR spectroscopy as rapid methods to monitor red wine fermentation, Eur. Food Res. Technol. 230 (6) (2010) 947–955, https://doi.org/10.1007/s00217-010-1227-5.

[43] A. Urtubia, J.R. Pérez-correa, F. Pizarro, E. Agosin, Exploring the applicability of MIR spectroscopy to detect early indications of wine fermentation problems, Food Control. 19 (4) (2008) 382–388, https://doi.org/10.1016/j.foodcont.2007.04.017.

[44] D. Cozzolino, C. Curtin, The use of attenuated total reflectance as tool to monitor the time course of fermentation in wild ferments, Food Control. 26 (2) (2012) 241–246, https://doi.org/10.1016/j.foodcont.2012.02.006.

[45] P. Jørgensen, J.G. Pedersen, E.P. Jensen, K.H. Esbensen, On-line batch fermentation process monitoring (NIR) - introducing "biological process time", J. Chemom. 18 (2) (2004) 81–91, https://doi.org/10.1002/(ISSN)1099-128X10.1002/cem.v18:210.1002/cem.850.

[46] J. Camacho, J. Picó, A. Ferrer, The best approaches in the on-line monitoring of batch processes based on PCA: does the modelling structure matter? Anal. Chim. Acta 642 (1-2) (2009) 59–68, https://doi.org/10.1016/j.aca.2009.02.001.

[47] N.F. Pérez, J. Ferré, R. Boqué, Calculation of the reliability of classification in discriminant partial least-squares binary classification, Chemom. Intell. Lab. Syst. 95 (2) (2009) 122–128, https://doi.org/10.1016/j.chemolab.2008.09.005.