

## Modelling the ageing process: A novel strategy to analyze the wine evolution towards the expected features<sup>☆</sup>



Ana C. Pereira<sup>a,b</sup>, Maria J. Carvalho<sup>b,c</sup>, Andreia Miranda<sup>b,c</sup>, João M. Leça<sup>b,c</sup>, Vanda Pereira<sup>b,c</sup>, Francisco Albuquerque<sup>d</sup>, José C. Marques<sup>b,c</sup>, Marco S. Reis<sup>a,\*</sup>

<sup>a</sup> Chemical Process Engineering and Forest Products Research Centre (CIEQPF), Department of Chemical Engineering, University of Coimbra, Rua Sílvio Lima, 3030-790 Coimbra, Portugal

<sup>b</sup> Centre of Exact Sciences and Engineering, University of Madeira, Campus da Penteada, 9000-390 Funchal, Portugal

<sup>c</sup> Institute of Nanostructures, Nanomodelling and Nanofabrication (I3N), University of Aveiro, 3810-193 Aveiro, Portugal

<sup>d</sup> Madeira Wine Company, Rua Dos Ferreiros 82, 9000 Funchal, Portugal

### ARTICLE INFO

#### Article history:

Received 30 November 2015

Received in revised form 24 March 2016

Accepted 27 March 2016

Available online 2 April 2016

#### Keywords:

Wine ageing

Chemical characterization

Feature extraction

Classification models

Process monitoring and evaluation

### ABSTRACT

In this work we present a new strategy to monitor the wine evolution during the ageing process. More specifically, we validate a procedure for analyzing how wine evolves during the ageing process in relation to the desired and expected quality features and we apply the proposed methodology to the case of a Portuguese fortified wine, the Madeira wine, where we compare the wine evolution under two different ageing processes.

The approach developed consists on modeling samples labeled as aged reference wines (5 year old Madeira wines), produced from four different grape varieties, and then analyze how and in which extent young wines (up to 3 years old) come closer to the reference data set. The analysis is based on a comprehensive set of chemical data, including: polyphenolic composition, organic acids, reducing sugars, color and oenological parameters, commonly used as routine quality control information. The study considers several feature extraction methods, such as: Principal Components of Analysis (PCA), Independent Component of Analysis (ICA) and Partial Least Squares (PLS). The classification methodologies tested were: Linear Discriminant Analysis (LDA), nearest neighbor (*k*-NN) and Soft Independent Modelling by Class Analogy (SIMCA). The different options of preprocessing/feature extraction/classification were evaluated and compared using a Monte Carlo approach.

From our analysis, the best combination of feature extraction/classification methodologies was PLS/LDA, which presented a classification performance of approximately 90% for three out of the four classes modeled, and of about 78% for the remaining one. Regarding the wines monitored during the first 3 years, our analysis revealed that they indeed mature in relation to the five year old reference wines. Furthermore, for some wines, it is possible to detect differences between the two ageing processes analyzed.

This study is of particular importance for this type of wines, where the ageing process plays a central role for attaining the expected quality levels, implying significant risks and costs for local and industrial producers. Notwithstanding the specific case study presented, the strategy outlined can be extrapolated to other products with similar characteristics in terms of their monitoring and process control.

© 2016 Elsevier B.V. All rights reserved.

### 1. Introduction

In recent decades, chemometrics has yielded important tools for wine's safeguard and protection, in order to deal with possible adulteration practices and to verify product authenticity [1–5]. Pattern recognition techniques have been successfully employed in the definition of strategies to protect the wine designation of origin [6–10], to confirm the grape variety authenticity and the wine style typicity [11–14].

These studies are aligned with the endeavor of regulatory authorities to prevent and detect adulteration practices and also with the increasing demand of information by consumers, who are increasingly interested in knowing product details, such as where, when and how wines were produced [15]. Of particular interest from the producers' standpoint is the development of new winemaking process control tools. However, in this scope, patterns recognition techniques were explored to a lesser extent. In fact, one of the main concerns of wine producers is to guarantee that their wines will get the expected and desired features of their target market, ensuring the success of their investment. In this study, we developed a monitoring framework that assists them in this complex but critical task. Supervised pattern recognition techniques were explored with the purpose of creating a classification

<sup>☆</sup> Selected paper from the 9th Edition of the In Vino Analytica Scientia Symposium in Trentino Alto Adige (Italy), 14–17 July 2015.

\* Corresponding author.

E-mail address: [marco@eq.uc.pt](mailto:marco@eq.uc.pt) (M.S. Reis).

model for a specific set of aged wines, selected as reference wines, and then used to implement a strategy to verify if younger wines are evolving in accordance to the expected and desired characteristics.

Supervised pattern recognition techniques are commonly distinguished into two groups: one focused on discriminating among classes, like linear discriminant analysis (LDA), *k*-nearest neighbor (*k* – NN), decision trees, support vector machines (SVM) or artificial neural networks (ANN), and another group oriented towards modelling classes in an isolated way, the so-called one-class classification methods, for which soft independent modelling by class analogy SIMCA is a well-known example. The first group of discriminant techniques set up classification rules for a number of pre-specified classes that cover the entire space of analysis. These rules are then used for assigning class labels to new samples [25]. On the other hand, one-class modelling methods generate a separate model for each category, and define conditions under which such category should be assigned to future samples. Each new sample is then analysed in order to verify whether or not it is compatible with the characteristics of each class-specific model [26].

In contrast to supervised pattern recognition techniques, multivariate prediction methodologies have found broader applications in winemaking process control [16]. Usually, in this context, spectrometers are the analytical devices chosen to collect the data, which are then processed and used to follow target parameters during the winemaking process. These multivariate approaches are able to provide fast, in-line and reliable measurements in terms of a wide range of wine quality control specifications, such as alcohol content, total and volatile acidity, relative density, reducing sugars, tartaric, malic and lactic acid concentration [17–19]. The untargeted information from spectrometer devices has also been used to follow the wine ageing process [20]. In this particular case, the ageing time prediction estimates were compared with the real wine ageing time in casks and used to evaluate if wines showed the expected characteristics. Similar strategies for wine ageing control have also been proposed, but using other type of chemical information, namely the volatile and phenolic wine composition [21,22]. In spite of the lack of consensus regarding the type of chemical information to be used in studies of this nature, several works have demonstrated that the volatile profile is a good choice when it comes to distinguish between different grape varieties and the profile of trace elements is effective when the problem is to discriminate amongst possible geographical origin [23]. Concerning other classification purposes, namely wine classification by quality, brands or winemaking technology involved, the chemical data set used is quite diverse, including also the untargeted fingerprints [24]. In the present study, the chemical data set used comprises information from phenolic and organic acid composition, reducing sugar concentrations, two of the main furanic compounds, color data and other routine quality control parameters.

In this work we focus on the Madeira wine (MW) as a case study, which is an important Portuguese fortified wine, produced exclusively in the Madeira island (Designation of Origin, DO). The alcohol content of this fortified wine varies from 17 to 22% and its main characteristics are a result of the grape varieties from which they are produced and from the particular ageing process implemented. Four different styles of Madeira wines are considered in this study, namely sweet, medium sweet, medium dry, and dry wines, which are produced from *Malvasia*, *Boal*, *Verdelho* and *Sercial* grape varieties, respectively. These Madeira wines represent about 53% of all the wines sold with indication of age. In spite of this group of wines represent a small percentage in Madeira wines sales (7% of the 4.0 million of liters produced on average per year) they are recognized as the finest and highest quality Madeira wines.

Madeira wines can follow two typical ageing processes: the typical ageing process in casks, called the *Canteiro* ageing process, in which wines are stored in casks placed in warmed lofts heated by the sun (the temperature can range from 20 to 30 °C), and the *Estufagem* ageing process, in which the wine is heated and kept in stainless steel vats at a

temperature of about 45 °C at least during three months. The aim of this second process is to accelerate wine ageing. The *Estufagem* process is not a common practice for the high quality Madeira wines that are under study in the present work. However, some attempts are being made in order to evaluate the impact of this ageing process for this selected group of wines. Specifically, the goal is to verify if the *Estufagem* process can indeed bring any advantage to these wines and evaluate the best way to conduct it for each style of wines. This study addresses this issue, by first developing a procedure to follow the ageing process and to assist in decision-making about wine compliance on the basis of a large variety of intrinsic wine characteristics.

This paper is organized as follows. In Sections 2.1 and 2.2 the data collection process is presented and the data sets are described, as well as wine chemical characterization carried out. In Section 2.3, we present the data analysis methodologies used, including those employed with the purpose for preprocessing, feature extraction and classification. The evaluation of classification models is presented in Section 3, together with the proposed approach to follow the evolution of the ageing process. This approach is applied to the analysis of younger wines that have been monitored during their ageing process. Finally, a brief discussion of the issue addressed in this study is presented in Section 4 and the main conclusions summarized in Section 5.

## 2. Materials and methods

### 2.1. Sample collection and data sets

Data collection for this study contemplated two different sets of samples. The first one comprise wines from the 2007 and 2008 harvests, collected and analysed after 5 years of ageing in casks. These wines were selected by an oenologist as reference wines of a five year old Madeira wine. The selection was made from wines of two harvests in order to incorporate in our analysis the inter-harvest variability. It comprises wines from the four white grape varieties recommended for Madeira wine production: *Malvasia*, *Boal*, *Verdelho* and *Sercial*. *Verdelho* and *Sercial* are produced in smaller quantities (together they represent about 35% of sales), and therefore they are also represented by a smaller number of samples than the remaining wines. More specifically, the following numbers of samples were collected and analyzed: 13 *Malvasia* wines (7 from the 2007 harvest and 6 from 2008), 9 *Boal* wines (5 from the 2007 harvest and 4 from 2008), 4 *Verdelho* and 4 *Sercial* wines (2 from 2007 and the other two from 2008, in both cases). The typical features of these wines can be summarized as follows. *Malvasia* grape variety is commonly used to produce sweet wines (residual sugars can range between 99 and 140 mg/L). These wines are frequently described as having rich honey notes, raisin and caramel aromas and present a dark brown color. *Boal* wines are medium sweet wines (residual sugars can vary between 80–100 mg/L). These wines typically have nutty and vanilla characteristics, flavors of smoke and dried fruit aromas, and are quite dark (mahogany) in color. *Verdelho* is a medium sweet wine, with sugar level ranging from 60 up to 81 mg/L. These are the more aromatic Madeira wines produced from white grape varieties, showing hints of caramel flavors and moderate acidity. *Sercial* is known to produce excellent dry wines (residual sugar varies from 25 to 61 mg/L), standing out for its mineral tones, subtle fruit and nut flavors, and high acidity [27].

The second group of samples is formed by wines from the 2011 harvest, each one produced exclusively from one of the four grape varieties (*Malvasia*, *Boal*, *Verdelho* and *Sercial*). These wines follow the common vinification practices [27] and they were separated in two sets. For the first set, wines were stored in casks and aged according to the *Canteiro* process, while for the second set wines were submitted to the *Estufagem* ageing process. Both sets of wines were monitored during the first three years of ageing. Samples were collect at the end of the *Estufagem* process, in the fourth month and then at the end of the first, second and third years of ageing.

## 2.2. Analytical methods

### 2.2.1. Polyphenols and furanic compounds

The identification and quantification of polyphenols and furanic compounds was performed by direct injection (20  $\mu\text{L}$ ) in a Waters Alliance Liquid Chromatograph (Milford, MA, USA) equipped with an auto-injector (Waters 2695) and a photodiode array detector (Waters 2996) system. Briefly, the chromatographic separation employs three mobile phases: A—10 mM of phosphate buffered at pH 2.70 with phosphoric acid; B—acetonitrile and C—methanol. The gradient program varied from 100% aqueous mobile phase to 60% organic phase in 58 min followed by 12 min of re-equilibration. The mobile phase was set to a flow rate of 1.0 mL/min and the column Atlantis® T3 (4.6  $\times$  250 mm id; 5  $\mu\text{m}$ ; Milford, MA, USA) thermostated at 30  $^{\circ}\text{C}$  [28].

The identification of compounds was done according to their retention times, analysing the UV–VIS spectra (scanned from 200 nm to 750 nm) and also confirmed by spiking samples with pure analytes. Quantification was carried out on 210 nm, for hydroxybenzoic acid, hydroxybenzaldehyde and flavan-3-ol, and on 280, 315 and 360 nm, 280 nm for furans, hydroxycinnamic acid and flavonol, respectively, and according to an external standard calibration curve previously validated. A total of 27 phenolic compounds (12 hydroxycinnamic acids, 8 hydroxybenzoic acids, 4 flavan-3-ols, 3 flavonols, 1 stilbene) and 2 furanic compounds were identified and quantified.

### 2.2.2. Organic acid and reducing sugars

The quantification of organic acids and reducing sugar was carried out using the same system described in Section 2.2.1. In this case, the column employed for the separation of organic acids was a Hi-Plex H (300  $\times$  7.7 mm; 8  $\mu\text{m}$ ; Agilent Technologies, U.S.A.). The analysis was also performed by direct injection (10  $\mu\text{L}$ ) using an isocratic elution with the following mobile phase: 0.0025 M of sulphuric acid. The flow rate was set to 0.6 mL/min and the column thermostated at 65  $^{\circ}\text{C}$ . Organic acids were detected at 210 nm while reducing sugars, glycerol and ethanol were detected by refractive index. The individual compounds were identified following the same strategy described in the previous section.

### 2.2.3. CIELab and routine quality control parameters

Samples were analyzed on the dual beam spectrophotometer Shimadzu UV–Vis 2600, using quartz cuvettes with optical thickness of 10 mm and ultrapure water as blank (OIV 2006). The parameters  $L^*$ ,  $a^*$ , and  $b^*$  were determined from the transmittance from 380 to 770 nm at 5 nm intervals, considering the illuminant D65 (daylight source) and 10° standard observer. Chromaticity ( $C^*$ ) and Hue ( $H$ ) were calculated according to the following equations, respectively [29]:

$$C^* = (a^{*2} + b^{*2})^{1/2} \quad (1)$$

$$H = \tan^{-1}\left(\frac{b^*}{a^*}\right) \quad (2)$$

The quality control parameters used in this study are: density (mg/L), total and volatile acidity (g/L), pH and total of reducing sugars (g/L). These measurements were obtained from Infrared Analyser FTIR and UVVis (Bacchus 3 Multispec).

## 2.3. Data analysis: supervised pattern recognition

According to the sequence of steps presented in the introduction of this article, the core of data analysis activities is centered on the following tasks: preprocessing; feature extraction; classification and performance evaluation. In this work, several options were considered for each task, in order to find out the best overall combination for the conditions found in this case study. The different alternatives are briefly

presented and described in the following paragraphs of this section. More detailed information can be found elsewhere in the literature [25,30–33]. Models validation was done using a Monte-Carlo approach.

### 2.3.1. Preprocessing

Preprocessing methods aim at removing unwanted variation that might impair the interpretation or predictive ability of the model, without compromising chemically relevant information [34]. Therefore, prior to implement the subsequent stages of feature extraction and modelling, data should be pretreated. Mean centering followed by unit scaling (usually known as autoscaling) is the most common approach for preprocessing, where all variables are put on a comparable footing regarding its variation. This is an important aspect when subsequent methodologies are scale-dependent, as happens with PCA and PLS. However, when variables arise in groups and each group potentially brings different information to the analysis, this simple preprocessing scheme is no longer adequate. The reason is the following: as all variables have equal weight in the analysis, the groups having more variables tend to dominate the analysis [35]. This distortion should therefore be corrected, namely through block scaling. There are two types of block scaling: soft and hard block scaling. Soft block scaling down-weight variables of largest blocks, by scaling them in such a way that the sum of the variable's variances (after completed the scaling) equals the square root of the number of variables in a particular block (Eq. (3)):

$$\sum(\text{variable's variance}) = \sqrt{\text{number of variables in a block}} \quad (3)$$

On the other hand, hard block scaling totally eliminates the differences between block sizes, as the variables in each block are scaled so that the sum of their variances is unity:

$$\sum(\text{variable's variance}) = 1 \quad (4)$$

The scaling weight used in soft and hard block scaling is  $1/(\text{number of variables in a block})^{1/4}$  and  $1/(\text{number of variables in a block})^{1/2}$ , respectively [36].

Both soft and hard block scaling approaches were considered in this study. The blocks formed were the following ones: (i) polyphenols together with furanic compounds; (ii) the organic acids; (iii) the reducing sugars; and (iv) all the remaining measurements were grouped in a separate group.

### 2.3.2. Feature extraction

Chemometric applications easily involve many variables or features, not all of them equally valuable for the purposes of the analysis. A preliminary stage of feature selection is therefore highly recommendable, or even fundamental, for conducting the modelling stages (in this case, supervised pattern recognition), as some methods cannot even cope with the situations where there are more variables than training samples or to deal with rank deficient systems [37–39]. Well-known examples include classifiers like LDA and QDA (Linear and Quadratic Discriminant Analysis) [40]. One common way to circumvent this scenario is to compute some features from all the variables collected (the raw data) and use this reduced set of transformed variables, with much lower dimensionality, to perform the analysis as the new predictors. Examples of methods used for feature extraction include Principal Components of Analysis (PCA), Independent Components of Analysis (ICA) and Partial Least Squares (PLS). The retained components or latent variables in these methods are then feed to classifiers. The basic and common idea to all of these methodologies is to apply some kind of a transformation to the data in order to reveal the essential data structure through a simpler and more suitable representation. PCA, ICA and PLS, have in common the fact that they only use linear transformations of

data, determined by means of the algebraic concepts of ED (eigenvalue decomposition) and SVD (singular value decomposition). One difference between these methods relies on the fact that PCA and PLS work better under Gaussian scenarios, and therefore only the second order statistics are considered, whereas ICA imposes constraints of statistical independence on the covariance of the sources, which involves extending the analysis to higher-order moments [41–43].

PCA, ICA and PLS have been successfully used to simplify and extract knowledge from large data sets, by removing the redundancy between variables and mitigating noise sources, while retaining the important information to be further used as inputs for classification models [44].

The general model structures and the notational conventions to be adopted in this work for these three methods, are referred below. Considering a data matrix  $\mathbf{X}(n \times m)$ , the PCA, ICA and PLS decompositions take the following forms, respectively:

$$\text{PCA} : \mathbf{X} = \mathbf{T} \cdot \mathbf{P}^T + \mathbf{E} \quad (5)$$

$$\text{ICA} : \mathbf{X} = \mathbf{A} \cdot \mathbf{S} + \mathbf{E} \quad (6)$$

$$\begin{aligned} \text{PLS} : \mathbf{X} &= \mathbf{T} \cdot \mathbf{P}^T + \mathbf{E}_x \\ \mathbf{Y} &= \mathbf{T} \cdot \mathbf{Q}^T + \mathbf{E}_y \end{aligned} \quad (7)$$

By analogy, the matrix  $\mathbf{S}(k \times m)$  in ICA corresponds to the loading matrix  $\mathbf{P}(m \times k)$  in PCA and PLS, while the values in matrix  $\mathbf{A}(n \times k)$  (called the mixing matrix) can be related to score vectors  $\mathbf{T}(n \times k)$  [45].  $\mathbf{E}(n \times m)$  stands for the residual matrix. As can be seen, PCA and ICA only make use of the information on predictor variables while PLS relates the predictor variables  $\mathbf{X}$  with a dependent variable,  $\mathbf{Y}$ , by maximizing the covariance between these two groups of variables.

### 2.3.3. Classification

In this work, we have considered the following classification methodologies: LDA,  $k$ -NN and SIMCA. The selection was based on the characteristics of collected data, where the number of samples is small when compared to the number of variables collected. This imposes limitations on the type of classifiers to adopt, which should be structurally simple and robust. Therefore, only linear and algorithmically simple approaches were selected, such as these ones referred above.

LDA is one variant of discriminant analysis, the simplest of all possible classification methods based on the Bayes formula. As the name suggests, the discrimination boundaries obtained from this technique are linear. In LDA, each class is assumed to follow a multivariate normal distribution with equal class-specific covariance matrices [46]. The latter assumption means that the variability within each group has the same structure and the only difference between groups is their mean. Class labels are assigned after computing the *a posteriori* probabilities for the observations to belong to the different classes (or some monotone transformation of it, such as the logarithmic transformation), by selecting the one maximizing this quantity. For the simpler case where no *prior* probabilities are considered, the measured vector  $\mathbf{x}$  is allocated to a class  $j$  with the smallest value of the squared Mahalanobis distance,  $D_{ij}^{\text{LDA}}$ :

$$D_{ij}^{\text{LDA}} = (\mathbf{x} - \bar{\mathbf{x}}_j)^t \hat{\Sigma}^{-1} (\mathbf{x} - \bar{\mathbf{x}}_j) \quad (8)$$

being  $\hat{\Sigma}$  the estimated covariance matrix obtained by pooling the covariance matrices across the  $G$  groups (each one formed by  $N_j$  samples,  $N = \sum_{j=1}^G N_j$ ), computed as follows:

$$\hat{\Sigma} = (N - G)^{-1} \sum_{j=1}^G (N_j - 1) \hat{\Sigma}_j \quad (9)$$

As mentioned above, the LDA criterion can also be expressed in terms of the underlying probability density function, that is, the probability of a sample  $\mathbf{x}$  to belong to class  $j$ . This is expressed by the following equation, obtained under the assumption that all prior probabilities are identical:

$$P(\mathbf{x}|j) = \left( \frac{1}{(2\pi|\hat{\Sigma}|)^{1/2}} \right) \exp\left(-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}}_j)^t \hat{\Sigma}^{-1} (\mathbf{x} - \bar{\mathbf{x}}_j)\right) \quad (10)$$

LDA is generally considered to be a robust and versatile approach in the area of classification, often leading to good and reliable results when the separation boundaries are of a simple nature [25].

The  $k$ -NN method is based on the determination of the distances between an unknown object and objects of the training set. Usually the Euclidean distance  $D_{ij}^{\text{kNN, Eucl}}$  is used, computed as follow:

$$D_{ij}^{\text{kNN, Eucl}} = \left[ \sum_r (x_i(r) - \bar{x}_j(r))^2 \right]^{0.5} \quad (11)$$

The class assigned to the unknown object is established through a majority voting approach among the  $k$  training objects that are closer to it. To avoid the possibility of ties,  $k$  is usually chose to be an odd number. In this study, it was considered to be  $k = 3$ . This non-parametric methodology does not assume any underlying probability distribution for the observations and is not restricted to linearly separable data sets, as happens with LDA [46].

In SIMCA, each category is independently modelled using PCA, with the possibility to use different numbers of principal components for the various classes. In general, data belonging to a given class can be well approximated through a low dimensional PCA model. In this study, to obtain the appropriate dimensionality of each PCA model, we have used the rule that the retained principal components should account for, at least, 80% of the total class data variability [47]. In SIMCA, an unknown sample  $i$  is assigned to the class  $j$  if the distance measure  $D_{ij}^{\text{SIMCA}}$  is less than  $\sqrt{2}$  [48]:

$$D_{ij}^{\text{SIMCA}} = \left( Q_r^2 + (T_r^2)^2 \right)^{0.5} \quad (12)$$

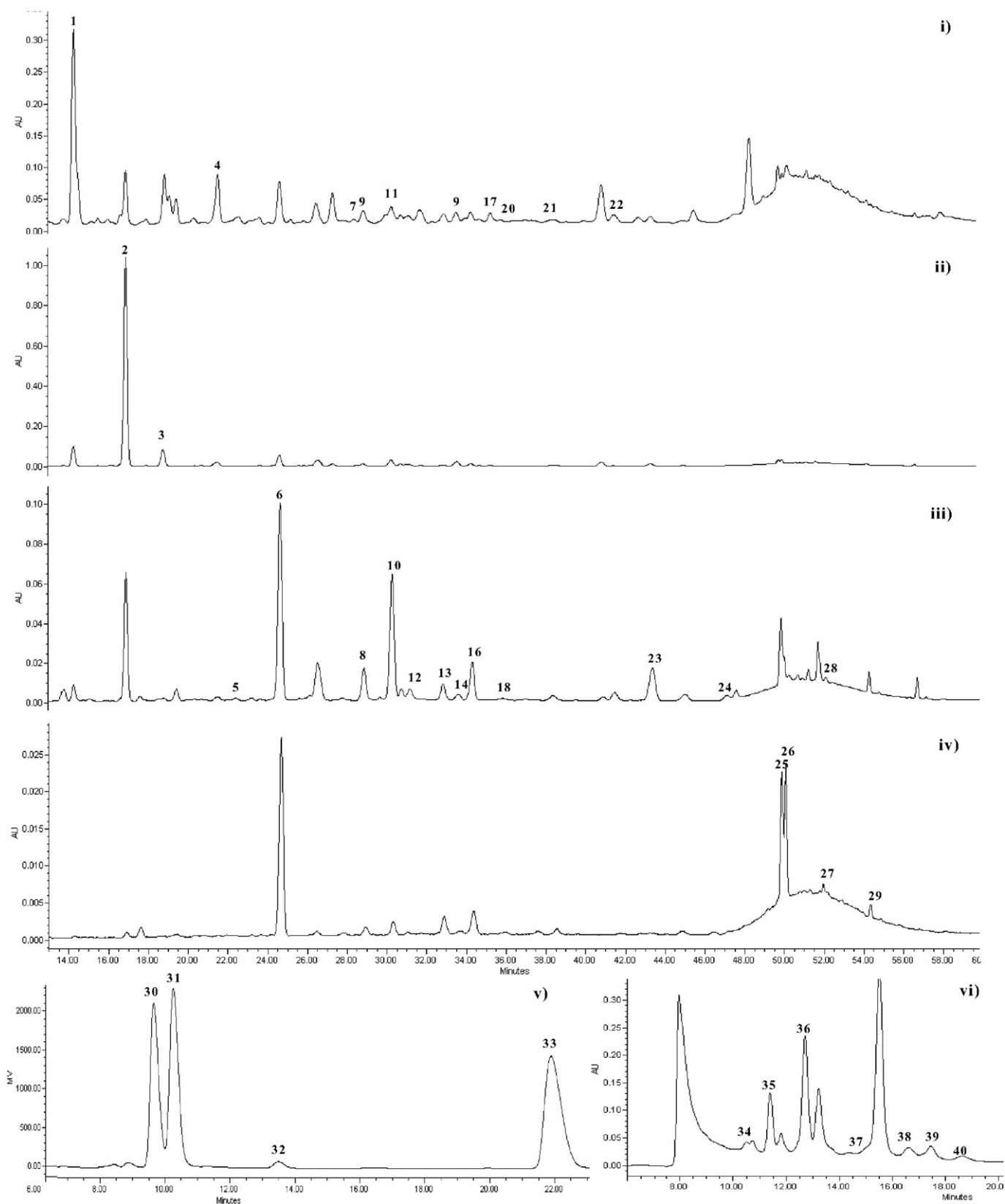
$$\begin{aligned} \text{being } Q_r &= \frac{Q}{Q_{0.95}} \\ T_r^2 &= \frac{T^2}{T_{0.95}^2} \end{aligned}$$

This distance measure gives equal weighting to the distance in the model space and in residual space,  $Q$ . This version, used in the PLS toolbox (Eigenvector Research, Inc.), presents some minor differences from the original SIMCA proposal introduced by Wold [49]. Details and discussion of its potential advantages and drawbacks can be found in reference [48]. If the distance computed according to Eq. (12) is not lower than  $\sqrt{2}$ , then the sample may be an outlier or belonging to a new group other than the original classes in the training set [46,50].

### 2.3.4. Evaluation

A critical step in the development of a classification approach regards its evaluation and final validation. Model evaluation and validation is carried out to verify whether the assessment made with the classifier is consistently good and robust, namely when applied to unknown samples not used during its development [41]. Monte Carlo Cross-Validation procedures are often used in this regard [21,51]. In brief terms, this type of procedure consists of randomly selecting a subset of samples, which temporarily performs the role of the training set and are used to build a classification model. This random division was performed in order ensuring that the test set was represented at least by  $(N_j/2 - 1)$  samples of each class (stratified random sampling).





**Fig. 1.** HPLC-DAD/IR chromatograms representative of a 5 year old *Malvasia* Madeira wine, where one can analyze polyphenols, quantified at i) 210 nm, ii) 280 nm, iii) 315 nm, iv) 360 nm, v) individual reducing sugars, ethanol and glycerol (IR) and vi) organic acids (quantified at 210 nm). (1) gallic acid; (2) HMF; (3) furfural; (4) protocatechuic acid; (5) Unk 1; (6) caftaric; (7) (–)-epigallocatechin; (8) cis-coutaric; (9) p-hydroxybenzoic acid; (10) trans-coutaric; (11) (+)-catechin; (12) trans-fertaric; (13) unk 2; (14) unk 3; (15) vanillic acid (not presented in this sample); (16) caffeic acid; (17) syringic acid; (18) unk 4; (19) (–)-Epigallocatechin gallate; (20) (–)-epicatechin; (21) vanillin; (22) syringaldehyde; (23) p-coumaric acid; (24) ferulic acid; (25) unk 5; (26) ellagic acid; (27) myricetin; (28) trans-resveratrol; (29) quercetin; (30) glucose; (31) fructose; (32) glicerol; (33) ethanol; (34) citric acid; (35) tartaric acid; (36) malic acid; (37) succinic acid; (38) lactic acid; (39) formic acid; (40) acetic acid.

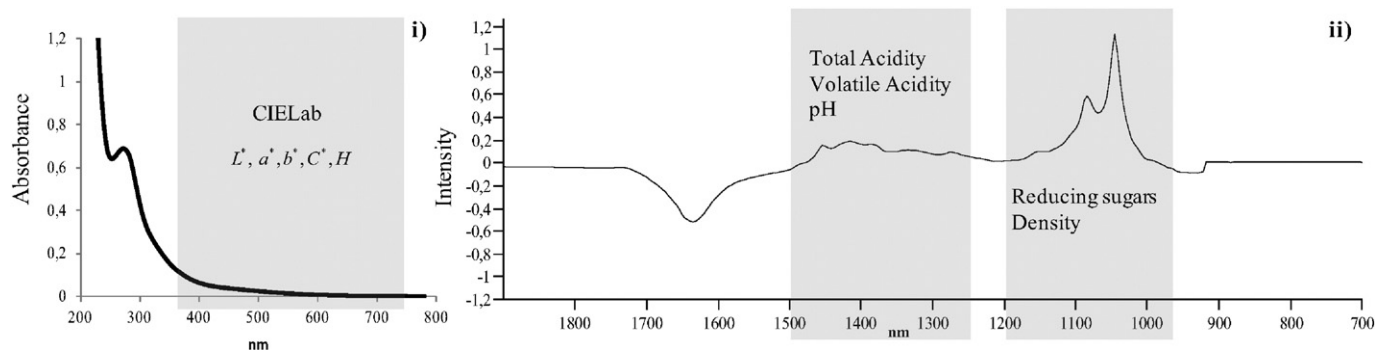


Fig. 2. Typical UV-Vis i) and Infrared spectrum ii) of a 5 year old *Malvasia* Madeira wine, from which the CIELab and quality control parameters were estimated, respectively.

Then, this model is used to assign classes for the samples which have been set aside from the training data set and the percentage of correctly classified samples is calculated. This procedure was repeated  $i$  times for each classification technique, considering the three scenarios of scaling and an increased number of  $l$  components or latent variables, for PCA, ICA and PLS, respectively. Parameter  $l$  range from 1 to 10, and the number of runs,  $i$ , was set to 1000 in order to ensure that all samples were set aside approximately the same number of times. Finally, the  $i$  stored values of percentage of correctly classified samples are averaged in order to determine the average success rate of each classification technique, leading to the performance criteria for the classification models: the global error (regardless of the class) and the classification rate by class. The 95% bootstrap confidence intervals for these quantities were also estimated, in order to provide information regarding the prediction accuracy of the reported performances.

### 3. Results

#### 3.1. Development and preliminary assessment of the classifiers

This work aims at presenting a new and simple strategy to monitor the wine evolution during the ageing process. The strategy consists on first developing a classification model that discriminates among different styles of wines, with the same ageing time—the 5 year old Madeira wine—and then, use these reference models to monitor how new wines evolve towards their expected maturity state. The wine chemical information analyzed and used in the development of the classifiers were: the concentration of 27 polyphenols, 2 furanic compounds, glucose, fructose and the total of reducing sugars, 7 organic acids, the glycerol and ethanol % (v/v), CIELab parameters and the routine quality control parameters, as presented in Figs. 1 and 2. The data set to be analyzed consist of 49 variables for 29 different samples analyzed in duplicate

( $\mathbf{X}(58 \times 49)$ ). The discussion of this data set in terms of wine chemical composition, namely their differences both concerning wine style (or grape varieties) and ageing times lies outside the scope of this paper.

All data analysis tasks were performed in the computational platform MatLab (The Mathworks, Inc., version 7.9.0.529, R2009b). Data pre-processing, PCA and PLS algorithms were implemented by home-made scripts, while the ICA approach where carried out using the fast-ICA algorithm, freely accessible on the Internet (<http://www.cis.hut.fi/projects/ica/fastica/>). Details about the later algorithm can be found elsewhere [52].

Data analysis firstly contemplated the analysis of the relative performance of each classifier tested using three different feature extraction methods and three scenarios of block scaling. The analysis of all these different combinations of preprocessing/feature extraction/classifiers is needed, since there is no rule of thumb in terms of which the best data pretreatment as well as the optimum way to proceed with the dimensional reduction of data (feature extraction) can be a priori established for a given classifier. In practice, the best combination should be selected as the one optimizing a suitable performance metric, such as the classification error found under cross-validation conditions. This was the approach followed in this work, given its robustness, stability and widespread use.

In Fig. 3, it is represented the correct classification rate for the LDA and  $k$ -NN classifiers, as a function of the number of components used and regarding all three scenarios of block-scaling considered. SIMCA performance was evaluated separately, since it requires fewer components by class (less or equal to three). The correct classification rate is the mean of  $i$  simulations computed according to Monte Carlo approach described in Section 2.3.4.

It was found that when no block scaling is used, the differences between PCA, ICA and PLS are small for any of the classifiers tested: either using LDA (the continuous line) or  $k$ -NN (the dotted line). Indeed,

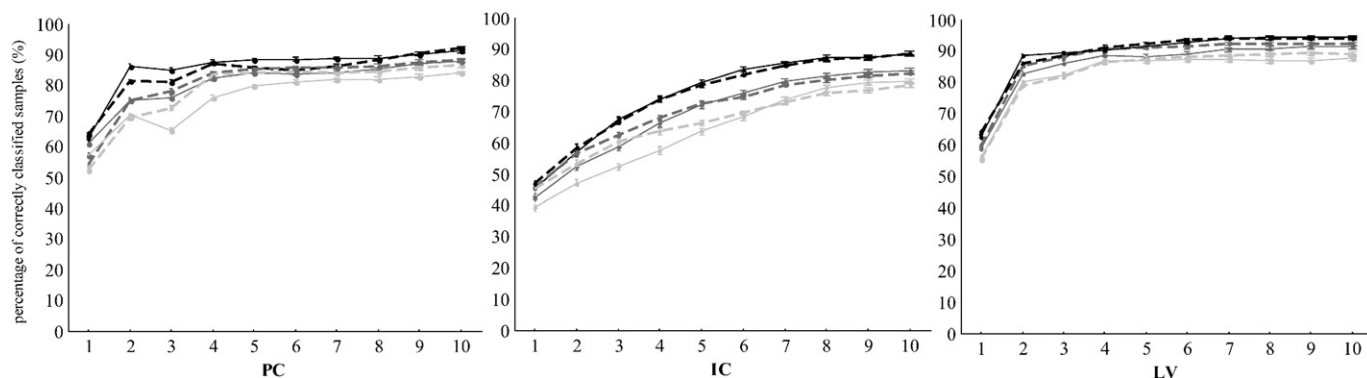
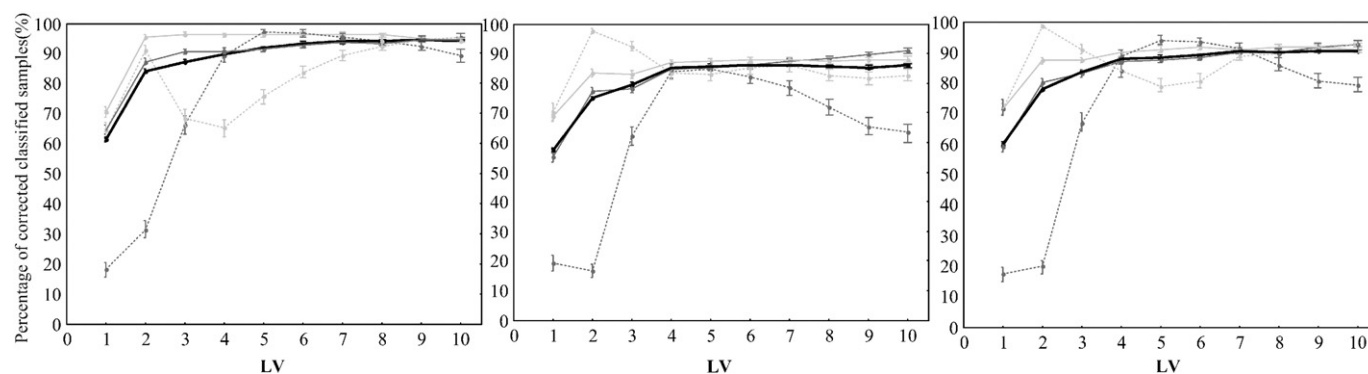


Fig. 3. Overall correct classification of LDA (solid line) and  $k$ -NN (dotted line) in the three scenarios of scaling: no scaling (black line), soft block scaling (dark grey line) and hard block scaling (light grey line), using three dimension-reduction methodologies: (a) PCA, (b) ICA and (c) PLS.



**Fig. 4.** Percentage of correct classification per class (*Malvasia*—solid dark grey line, *Boal*—solid light grey line, *Verdelho*—dotted dark grey line and *Sercial*—dotted light grey line) of LDA (solid line) preceded by PLS and without using block scaling. Black line is the global error presented in Figure 3b.

the differences rely more in the number of components to be used to achieve similar performances. PCA and ICA require more components to achieve similar performances of the PLS model with only 4 latent variables, requiring 5 or 9 latent variables, respectively.

Accordingly, it is the combination of PLS together with LDA which seems to lead to the most parsimonious model in these conditions. More specifically, the LDA classifier constructed over the 4 first PLS scores presents a classification rate of 90%, when no block scaling is used. The superior feature extraction ability of PLS when compared to the other two methodologies tested is likely to be related with the supervised nature of its algorithm. As mentioned in Section 2.3.2, ICA was implemented using the fastICA algorithm. In future studies, other ICA algorithms will also be tested, such as the Joint Approximation Diagonalization and Eigenmatrices (JADE) [45] or Mutual Information Least Dependence Components Analysis (MILCA) [46] in order to verify if further improvements in ICA performance can be achieved. Also, as ICA models are not nested (such PCA or PLS), the Monte Carlo approach may need some adaptation for dealing with this feature.

In terms of block scaling effect, no improvements in classifier performance were verified in this study, meaning that soft (grey line) and hard block (blue line) scaling should be avoided in data pre-treatment because it might partially remove interesting information. The classifier performance was also evaluated in terms of the correct classification rate by class. This analysis as only considered the best classifier in terms of global error found initially, namely the LDA/PLS combination (Fig. 4). The main difference between the three scenarios of scaling is the fact that block scaling seems to provide a more balanced classifier, that is, with a four latent variable model, all classes are recognized with similar capability. In this scenario, the correct classification rate ranges from 80 to 90%, contrarily to no block scaling, for which performances lower than 70% were observed for *Verdelho* and *Sercial* classes.

Regarding SIMCA, the classifier was validated in terms of only 3 classes, since the *Verdelho* and *Sercial* wine samples are represented in reduced number. It was decided to gather these samples and considered them in one group, since they have similar properties—both are dry wines. In this case, the best classification performance was obtained when soft block scaling was used (Table 1). Here, it is clear the importance of decreasing the weight of variables belonging to the larger blocks, namely the polyphenol and furanic compounds, and provide

to the smaller groups of variables a similar importance. For this data set, this assumes particular interest since reducing sugars are expected to have an important discriminant power concerning the different groups of wines (sweetness degree).

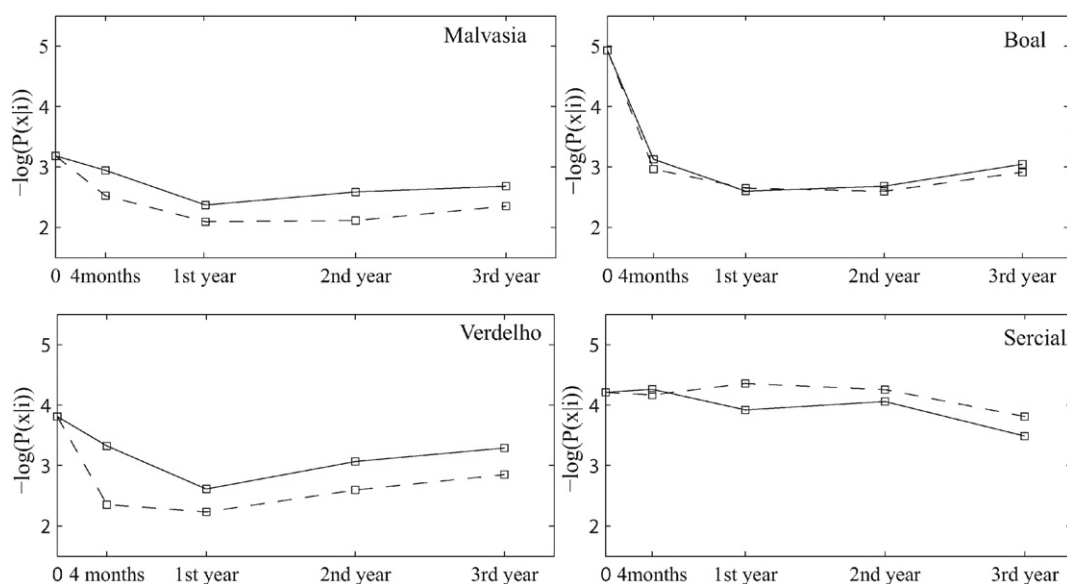
However, SIMCA does not improve the results previously obtained by LDA and therefore the linear discriminant analysis classifier is the one adopted to proceed with the current study, where it will be applied to a second independent data set, composed by the wines from the 2011 harvest.

### 3.2. Evaluation of the selected classifier using an independent data set with wines from a new harvest

Once established the classification model, the interest is to evaluate how it behaves when applied to new wines from another harvest. Our analysis focused in monitoring the evolution of wines produced in 2011 harvest in relation to 5 year old reference wines. For such, we have used a distance metric based on the probability to belong to this ageing class, as computed by the LDA classifier. More precisely, we have analyzed the minus logarithm of the probability,  $-\log(P(\mathbf{x}|j))$ : the lower this value is, the closer samples are to the reference wines. These values are represented in Fig. 5. The solid line correspond to wine samples stored in casks and the dot line correspond to wines submitted to the *Estufagem* ageing process during the first four months and then transferred and stored in casks. From these results, it can be concluded that the wines from 2011 harvest are evolving towards the desired direction, as they gradually come closer to the reference wines during the ageing process. In general, the most significant evolution is observed during the first year of ageing, being more evident for *Boal* and *Verdelho* wines. For *Malvasia* wines, the evolution during this period may not be so evident because the initial state of the wine is already closer to that of the reference wines. This proximity is, in part, due to the evolution that occurs during the period between the end of the winemaking process and the beginning of ageing, which takes about six months. During this period wine changes occur and probably are more significant for *Malvasia* wines. Comparing both ageing process, it was found that the differences between the two processes are not the same for the four styles of wines. In particular, the effect of *Estufagem* process seems to be more evident for *Malvasia* and *Verdelho* wines,

**Table 1**  
SIMCA results expressed in terms of percentage of correct classification, misclassification and no classification.

	No scaling			Soft scaling			Hard scaling		
	Correct classification	Misclassification	No classification	Correct classification	Misclassification	No classification	Correct classification	Misclassification	No classification
Malvasia	67.8	0.0	32.2	89.9	0.0	10.1	78.9	7.5	21.1
Boal	40.4	1.5	58.1	61.8	15.8	22.4	61.2	9.5	38.8
Medium and Dry wines	16.4	0.0	83.6	54.8	0.0	45.2	63.2	16.0	23.8



**Fig 5.** Evaluation of wines from 2011 harvest in relation to the reference wines (i—Malvasia, ii—Boal, iii—Verdelho and iv—Sercial) regarding the ageing processes considered in this work (dotted line—the *Estufagem* ageing process, solid line—the *Canteiro* ageing process).

since the values of  $-\log(P(\mathbf{x}|j))$  decrease faster than the corresponding values for the wines stored only in casks. For *Boal* wines, it is not observed a significant difference between both ageing processes. Notwithstanding, these differences can become more evident if other type of chemical information is considered. Regarding *Sercial* wines, differences between *Estufagem* and *Canteiro* were observed just after the fourth month, and in this particular case *Canteiro* wines come closer to reference wines faster than *Estufagem* wines. However, some caution is needed here, since this result can be also related to the lower number of available samples to model the *Sercial* class, and as a consequence of the lower capability of the classifier to recognize wines from this category.

At this phase, it should also be highlighted that the goal of *Estufagem* ageing process is not to decrease the ageing time, but to promote, for the same ageing period (in this case, 5 years), additional quality and distinctive characteristics for Madeira wines. As stated in the introduction, *Estufagem* ageing process is not a commonly practice for this particular type of Madeira wines. However, some studies have been carried out to evaluate the possible benefits of this peculiar and unique ageing process for Madeira wines. This study provides a contribute in this regard, since it demonstrates how to compare both processes using different types of information as well as reveals that in terms of the wine chemical composition studied, it seems to be advantageous to submit *Malvasia* and *Verdelho* Madeira wines to *Estufagem* ageing process, during four months at 45 °C.

#### 4. Discussion

The results presented in the previous section reveal the potential of the strategy outlined to follow the wine ageing evolution. The case study addresses the particular case where the target was defined to be the 5 year old Madeira wine, but the strategy can be extrapolated to other times of ageing and other wines as well.

This study follows previous works [20,21,53] that have been made in order to provide advanced tools for wine producers to monitor the evolution of their wines. This issue is of particular importance for wines which undergo long ageing periods, during which they acquire their typical features and the expected quality patterns. Very significant costs are involved in this phase and therefore additional tools that can improve the effectiveness of monitoring and assessment process,

bringing more information and knowledge and improving decision-making, are relevant and very opportune.

An additional goal inherent to the present study was also to demonstrate how a relevant data analysis can be very informative for wine producers. The information available regarding wine composition can easily be overlooked unless used in an integrated way, with the proper tools that are able to address the underlying data structure and extract the relevant trends. This is not aligned with the progress in wine chemical characterization and advanced data analysis, and therefore additional efforts should be made to improve this situation and raise the quality standards of this important industrial sector. As previously stated in Section 3.1, other feature extraction methodologies can also be tested in the future. The same applies to additional classifiers, including those having non-linear structures. However, those are not recommended in this study, given the limited number of samples available. Nevertheless, LDA revealed to be very effective in discriminating the different styles of Madeira wine, with 5 year old, contemplated in the case study.

Another aspect worthwhile considering in the future is using several class-specific models for different time of ageing to monitor the ageing process. This may bring added value for cases where the target is set to a longer ageing period (10–20 years), as ageing trends become more complex in these situations requiring more structure to be properly monitored. However this is not the case for the current case study, for which a simpler approach, based on a single reference model, was sufficient to capture the essential of the ageing dynamics.

In summary, together with the results of particular interest for the Madeira wine sector, namely the comparison of *Estufagem* and *Canteiro* ageing processes taking into account different wine chemical data sets, the present work proposes and validates a new tool for monitoring the evolution of wine ageing, in a simple and informative way.

#### 5. Conclusions

In this paper we address the issue of monitoring the wine ageing process. Wine physicochemical data yielding information of a different and complementary nature was analyzed. More specifically, samples were analyzed in terms of polyphenol composition, furanic and reducing sugars, organic acids, color parameters and routine quality control parameters. Data analysis involved data pretreatment, feature extraction (dimensionality reduction) techniques,



classification methods and the evaluation/validation of all the combinations of options tested.

The results achieved show that it is indeed possible to monitor wine evolution during the ageing process, towards the desired characteristics. Our study also revealed the benefits of applying the *Estufagem* ageing process to the high-quality Madeira wines.

## Acknowledgments

Ana C. Pereira, Maria J. Carvalho acknowledge Agência Regional para o Desenvolvimento da Investigação Tecnologia e Inovação (ARDITI) for the financial support through their grants in the scope of the project with reference 002248/2014/131, and João M. Leça, in the scope of the project with reference 002376/2013/131.

The authors acknowledge the financial support given by FEDER (+ *Conhecimento* program, project IMPACT II, MADFDR-01-0190-FEDER-000010) and also to Madeira Wine Company, S.A. for supplying all wine samples analyzed.

## References

- [1] S. Preys, E. Vigneau, G. Mazerolles, V. Cheynier, D. Bertrand, Multivariate prototype approach for authentication of food products, *Chemom. Intell. Lab. 87* (2007) 200–207.
- [2] P. Oliveri, G. Downey, Multivariate class modeling for the verification of food-authenticity claims, *TrAC Trends Anal. Chem.* 35 (2012) 74–86.
- [3] J. Riedl, S. Esslinger, C. Faulstich-Hassek, Review of validation and reporting of non-targeted fingerprinting approaches for food authentication, *Anal. Chim. Acta* 885 (2015) 17–32.
- [4] F. Marini, R. Bucci, A.L. Magri, A.D. Magri, Authentication of Italian CDO wines by class-modeling techniques, *Chemom. Intell. Lab. 84* (2006) 164–171.
- [5] M. Forina, M. Casale, P. Oliveri, 4.04—Application of chemometrics to food chemistry, in: S.D.B.T. Walczak (Ed.), *Comprehensive Chemometrics*, Elsevier, Oxford 2009, pp. 75–128.
- [6] L. Liu, D. Cozzolino, W.U. Cynkar, M. Gishen, C.B. Colby, Geographic classification of Spanish and Australian tempranillo red wines by visible and near-infrared spectroscopy combined with multivariate analysis, *J. Agric. Food Chem.* 54 (2006) 6754–6759.
- [7] Y.B. Monakhova, R. Godelmann, T. Kuballa, S.P. Mushtakova, D.N. Rutledge, Independent components analysis to increase efficiency of discriminant analysis methods (FDA and LDA): application to NMR fingerprinting of wine, *Talanta* 141 (2015) 60–65.
- [8] Y.B. Monakhova, R. Godelmann, A. Hermann, T. Kuballa, C. Cannet, H. Schäfer, M. Spraul, D.N. Rutledge, Synergistic effect of the simultaneous chemometric analysis of <sup>1</sup>H NMR spectroscopic and stable isotope (SNIF-NMR, <sup>18</sup>O, <sup>13</sup>C) data: application to wine analysis, *Anal. Chim. Acta* 833 (2014) 29–39.
- [9] M. Forina, P. Oliveri, H. Jäger, U. Römisch, J. Smeyers-Verbeke, Class modeling techniques in the control of the geographical origin of wines, *Chemom. Intell. Lab. 99* (2009) 127–137.
- [10] E.I. Geana, R. Popescu, D. Costinel, O.R. Dinca, R.E. Ionete, I. Stefanescu, V. Artem, C. Bala, Classification of red wines using suitable markers coupled with multivariate statistic analysis, *Food Chem.* 192 (2016) 1015–1024.
- [11] A.C. Pereira, M.S. Reis, P.M. Saraiva, J.C. Marques, Aroma ageing trends in GC/MS profiles of liqueur wines, *Anal. Chim. Acta* 659 (2010) 93–101.
- [12] A. Jos, I. Moreno, A.G. González, G. Repetto, A.M. Cameán, Differentiation of sparkling wines (cava and champagne) according to their mineral content, *Talanta* 63 (2004) 377–382.
- [13] J.E. Welke, V. Manfroi, M. Zanus, M. Lazzarotto, C. Alcaraz Zini, Differentiation of wines according to grape variety using multivariate analysis of comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometric detection data, *Food Chem.* 141 (2013) 3897–3905.
- [14] J. Saurina, Characterization of wines using compositional profiles and chemometrics, *TrAC Trends Anal. Chem.* 29 (2010) 234–245.
- [15] M.E. Alañón, M.S. Pérez-Coello, M.L. Marina, Wine science in the metabolomics era, *TrAC Trends Anal. Chem.* 74 (2015) 1–20.
- [16] R.G. Brereton, Introduction to multivariate calibration in analytical chemistry, *Analyst* 125 (2000) 2125–2154.
- [17] 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 (2011) 181–186.
- [18] M. Friedel, C.-D. Patz, H. Dietrich, Comparison of different measurement techniques and variable selection methods for FT-MIR in wine analysis, *Food Chem.* 141 (2013) 4200–4207.
- [19] Q. Wang, Z. Li, Z. Ma, L. Liang, Real time monitoring of multiple components in wine fermentation using an on-line auto-calibration Raman spectroscopy, *Sensors Actuators B Chem.* 202 (2014) 426–432.
- [20] A.C. Pereira, M.S. Reis, P.M. Saraiva, J.C. Marques, Development of a fast and reliable method for long- and short-term wine age prediction, *Talanta* 86 (2011) 293–304.
- [21] A.C. Pereira, M.S. Reis, P.M. Saraiva, J.C. Marques, Madeira wine ageing prediction based on different analytical techniques: UV–vis, GC–MS, HPLC–DAD, *Chemom. Intell. Lab.* 105 (2011) 43–55.
- [22] A. Rudnitskaya, S.M. Rocha, A. Legin, V. Pereira, J.C. Marques, Evaluation of the feasibility of the electronic tongue as a rapid analytical tool for wine age prediction and quantification of the organic acids and phenolic compounds. The case-study of Madeira wine, *Anal. Chim. Acta* 662 (2010) 82–89.
- [23] A. Versari, V.F. Laurie, A. Ricci, L. Laghi, G.P. Parpinello, Progress in authentication, typification and traceability of grapes and wines by chemometric approaches, *Food Res. Int.* 60 (2014) 2–18.
- [24] D. Serrano-Lourido, J. Saurina, S. Hernández-Cassou, A. Checa, Classification and characterisation of Spanish red wines according to their appellation of origin based on chromatographic profiles and chemometric data analysis, *Food Chem.* 135 (2012) 1425–1431.
- [25] T. Næs, T. Isaksson, T. Fearn, T. Davies, A User-Friendly Guide to Multivariate Calibration and Classification, NIR Publications, Chichester, 2002.
- [26] P. Oliveri, G. Downey, Chapter 13—Discriminant and class-modelling chemometric techniques for food PDO verification, in: G. Miguel de la, G. Ana (Eds.), *Comprehensive Analytical Chemistry*, Elsevier 2013, pp. 317–338.
- [27] T. Elliott, *The Wines of Madeira*, Trevor Elliott Publishing, Gosport, 2010.
- [28] V. Pereira, F. Albuquerque, J. Cacho, J. Marques, Polyphenols, antioxidant potential and color of fortified wines during accelerated ageing: the Madeira wine case study, *Molecules* 18 (2013) 2997.
- [29] M. Carvalho, V. Pereira, A. Pereira, J. Pinto, J. Marques, Evaluation of wine colour under accelerated and oak-cask ageing using CIE Lab and chemometric approaches, *Food Bioprocess Technol.* 8 (2015) 2309–2318.
- [30] J.E. Jackson, *A User's Guide to Principal Components*, John Wiley & Sons, Inc., New York, 2004.
- [31] K.H. Esbensen, P. Geladi, *Principal Component Analysis: Concept, Geometrical Interpretation, Mathematical Background, Algorithms, History, Practice*, Comprehensive Chemometrics, Elsevier, 2009 211–226.
- [32] M. Andersson, A comparison of nine PLS1 algorithms, *J. Chemom.* 23 (2009) 518–529.
- [33] G. Wang, Q. Ding, Z. Hou, Independent component analysis and its applications in signal processing for analytical chemistry, *TrAC Trends Anal. Chem.* 27 (2008) 368–376.
- [34] J. Trygg, J. Gabrielsson, T. Lundstedt, 2.01—Background estimation, denoising, and preprocessing, in: S.D.B.T. Walczak (Ed.), *Comprehensive Chemometrics*, Elsevier, Oxford 2009, pp. 1–8.
- [35] M. Lapinsh, P. Prusis, S. Uhlén, J.E.S. Wikberg, Improved approach for proteo-chemometrics modeling: application to organic compound–amine G protein-coupled receptor interactions, *Bioinformatics* 21 (2005) 4289–4296.
- [36] L. Eriksson, E. Johansson, U. Academy, N. Kettaneh-Wold, S. Wold, *Multi- and Megavariate Data Analysis: Principles and Applications*, Umetrics Academy, 2001.
- [37] T. Næs, U. Indahl, A unified description of classical classification methods for multicollinear data, *J. Chemom.* 12 (1998) 205–220.
- [38] T. Næs, B.-H. Mevik, Understanding the collinearity problem in regression and discriminant analysis, *J. Chemom.* 15 (2001) 413–426.
- [39] C.F. Dormann, J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J.R.G. Marquéz, B. Gruber, B. Lafourcade, P.J. Leitão, T. Münkemüller, C. McClean, P.E. Osborne, B. Reineking, B. Schröder, A.K. Skidmore, D. Zurell, S. Lautenbach, Collinearity: a review of methods to deal with it and a simulation study evaluating their performance, *Ecography* 36 (2013) 27–46.
- [40] W.-S. Lin, C.-M. Yang, B.-J. Kuo, Classifying cultivars of rice (*Oryza sativa* L.) based on corrected canopy reflectance spectra data using the orthogonal projections to latent structures (O-PLS) method, *Chemom. Intell. Lab.* 115 (2012) 25–36.
- [41] E. Cubero-Leon, R. Peñalver, A. Maquet, Review on metabolomics for food authentication, *Food Res. Int.* 60 (2014) 95–107.
- [42] F. Westad, M. Kermit, 2.14—Independent component analysis, in: S.D.B.T. Walczak (Ed.), *Comprehensive Chemometrics*, Elsevier, Oxford 2009, pp. 227–248.
- [43] L. De Lathauwer, B. De Moor, J. Vandewalle, An introduction to independent component analysis, *J. Chemom.* 14 (2000) 123–149.
- [44] L.A. Berrueta, R.M. Alonso-Salces, K. Héberger, Supervised pattern recognition in food analysis, *J. Chromatogr. A* 1158 (2007) 196–214.
- [45] D. Jouan-Rimbaud Bouveresse, A. Moya-González, F. Ammari, D.N. Rutledge, Two novel methods for the determination of the number of components in independent components analysis models, *Chemom. Intell. Lab.* 112 (2012) 24–32.
- [46] B.K. Lavine, W.S. Rayens, 3.15—Classification: basic concepts, in: S.D.B.T. Walczak (Ed.), *Comprehensive Chemometrics*, Elsevier, Oxford 2009, pp. 507–515.
- [47] I.T. Jolliffe, *Principal Component Analysis*, Springer, 2002.
- [48] C. Durante, R. Bro, M. Cocchi, A classification tool for N-way array based on SIMCA methodology, *Chemom. Intell. Lab.* 106 (2011) 73–85.
- [49] W. Svante, S. Michael, SIMCA: a method for analyzing chemical data in terms of similarity and analogy, in: B.R. Kowalski (Ed.), *Chemometrics: Theory and Application*, American Chemical Society 1977, pp. 243–282.
- [50] B.K. Lavine, Pattern recognition, *CRC Crit. Rev. Anal. Chem.* 36 (2006) 153–161.
- [51] M. Forina, P. Oliveri, S. Lanteri, M. Casale, Class-modeling techniques, classic and new, for old and new problems, *Chemom. Intell. Lab.* 93 (2008) 132–148.
- [52] A. Hyvarinen, Fast and robust fixed-point algorithms for independent component analysis, *IEEE Trans. Neural Netw.* 10 (1999) 626–634.
- [53] A.C. Pereira, M.S. Reis, P.M. Saraiva, J.C. Marques, Analysis and assessment of Madeira wine ageing over an extended time period through GC–MS and chemometric analysis, *Anal. Chim. Acta* 660 (2010) 8–21.