Food and Bioprocess Technology
https://doi.org/10.1007/s11947-023-03039-8

RESEARCH

Pattern Recognition of GC-FID Profiles of Volatile Compounds in *Brandy de Jerez* Using a Chemometric Approach Based on Their Instrumental Fingerprints

María Guerrero-Chanivet[1,2] · **Fidel Ortega-Gavilán**[3] · **María Gracia Bagur-González**[3] · **Manuel J. Valcárcel-Muñoz**[2] · **María de Valme García-Moreno**[1] · **Dominico A. Guillén-Sánchez**[1]

Received: 20 October 2022 / Accepted: 16 February 2023

Abstract

Brandy de Jerez is a unique spirit produced in Southern Spain under Protected Geographical Indication “Brandy de Jerez” (PGI). Two key factors for the production of quality brandies are the original wine spirit and its aging process. They are significantly conditioned by specific variables related to the base wine and the distillation method employed to produce the wine spirit used to obtain a finally aged brandy. This final beverage is therefore strongly influenced by its production process. The chromatographic instrumental fingerprints (obtained by GC FID) of the major volatile fraction of a series of brandies have been examined by applying a chemometric approach based on unsupervised (hierarchical cluster analysis and principal component analysis) and supervised pattern recognition tools (partial least squares–discriminant analysis and support vector machine). This approach was able to identify the fermentation conditions of the original wine, the distillation method used to produce the wine spirit, and the aging process as the most influential factors on the volatile profile.

Keywords Distillation · GC-FID · Sulfur dioxide · Volatile compounds · Authentication · Wine spirits

Introduction

Brandy de Jerez is a unique spirit produced in Southern Spain under a Protected Geographical Indication “Brandy de Jerez” (PGI). As described in its Technical File (Consejería de Agricultura Pesca y Desarrollo Rural, 2018; Parlamento Europeo & Consejo de la Unión Europea, 2019), Brandy de Jerez can be elaborated from different types of wine spirits, as long as these wine spirits of medium and low alcoholic strength represent more than 50% of the alcoholic strength of the final brandy being recognized in this Technical File: (i) low alcohol content wine spirits, traditionally called “holandas,” with no more than 70% ABV (alcohol by volume); (ii) medium alcohol content wine spirits, with between 70 and 86% ABV; and (iii) high alcohol content wine spirits, traditionally known as wine distillates, with an alcohol content between 86 and 94.8% ABV.

The distillation method used to produce the wine spirit is decisive regarding its organoleptic characteristics (Balcerek et al., 2017; Spaho et al., 2013; Tsakiris et al., 2014). Two of the most commonly used distillation techniques to produce wine spirits are continuous column distillation (Spaho, 2017; Tsakiris et al., 2014; Xiang et al., 2020) and pot still distillation, which can be performed in one or two steps (Balcerek et al., 2017). This is one of the most relevant factors associated to the production of wine spirits. Thus, when pot still distillation is employed, fruity aromas (primary aromas) can be perceived and the “memory” of the raw material in the distilled product is more accentuated. On the other hand, the distillates obtained by column distillation are usually richer in higher alcohols, since the very nature of the distillation process separates the rest of the compounds to a greater extent (Spaho, 2017).

✉ María de Valme García-Moreno
valme.garcia@uca.es

[1] Departamento de Química Analítica, Facultad de Ciencias, Instituto de Investigación Vitivinícola y Agroalimentaria (IVAGRO), Universidad de Cádiz, Campus Universitario de Puerto Real, 11510 Puerto Real, Cádiz, Spain

[2] Bodegas Fundador S.L.U., C/ San Ildefonso, n° 3, 11403 Jerez de la Frontera, Cádiz, Spain

[3] Departamento de Química Analítica, Facultad de Ciencias, Universidad de Granada, Av. Fuentenueva S/N, 18071 Granada, Spain

Grape variety (Cacho et al., 2013; Xiang et al., 2020), fermentation conditions, and the oenological practices applied to obtain the wine to be distilled (Tsakiris et al., 2014; Xiang et al., 2020; Zierer et al., 2016) also have an influence on the character of the resulting wine spirit. Likewise, certain traditional oenological practices, such as the use of sulfur dioxide (International Organization of Vine & Wine, 2021), also affect the organoleptic properties of the wines (Korenika et al., 2020) and, as a consequence, that of the wine spirits produced from them (Tsakiris et al., 2014).

The character of brandies will also be shaped by another fundamental stage in its production process: aging. According to its Technical File, *Brandy de Jerez* must be aged in "properly seasoned cask wood." So casks are crucial elements in the aging of brandies, since the botanical origin of the wood and the thermal treatment it is subjected to during its manufacturing process, as well as the particular seasoning of the wood (Sánchez-Guillén et al., 2019), have a definite saying on the specific compounds involved in the process and that might be transferred to the brandy during its aging.

Among the compounds that constitute the volatile fraction of brandies, aldehydes, higher alcohols, and major esters are worth mentioning. These major volatile compounds that are found in brandies have their origin in the fermentation of the grape must (Berry & Slaughter, 2003; Swiegers et al., 2005; Valero et al., 2002) and, subsequently, their greater or lesser presence in the distillate will be dictated by the distillation method employed (Silva & Malcata, 1999). Thus, the distillation method will also determine the volatile fraction of the final brandies (García-Llobodanin et al., 2007; Hernandez-Gomez et al., 2003), given that the presence and concentration levels of these compounds in the aged product will depend on such method (Spaho et al., 2013).

Fingerprinting is a very powerful methodology that is increasingly used by the food industry (Bagur-González et al., 2015; Bikrani et al., 2019; Ortega-Gavilán et al., 2020; Pérez-Castaño et al., 2019). It presents multiple advantages, as it does not require the calibration or quantification of the compounds in a product to characterize it, since it is based on its whole signal. Instrumental fingerprints are highly specific indicators, as when obtained under certain analytical conditions they are unique to each sample, which allows generating a robust model for the rapid classification and/or quality control of the samples. The establishment of the robust model needs the use of data analysis approaches in order to (i) extract the maximum useful information, (ii) reduce the number of the variables of the system, and (iii) group and/or classify unknown samples with similar characteristics (Pastor et al., 2016, 2020; Psodorov et al., 2015; Szymańska, 2018).

Consequently, with the above said, this work intends to evaluate the potential of chromatographic fingerprinting applied to the volatile compounds present in 14- and 28-months-aged brandies obtained from different types of wine spirit under a chemometric approach based on the use of different patterns recognition techniques. The wine spirits were obtained from suitable for distillation wines produced under different fermentation conditions and distilled using different techniques in order to evaluate the impact of these raw materials on the aged product. For this reason, unseasoned casks were used for the aging process, so that the impact of particular the wine spirit used to elaborate the brandies could be determined, but without having cask seasoning as an additional variable.

Material and Methods

Samples

The wine spirits, the oak casks, and the premises where this study was carried out were provided by Bodegas Fundador S.L.U., a winery that belongs to the Protected Geographical Indication "Brandy de Jerez."

All the wines selected for the production of the wine spirits were suitable for distillation (volatile acidity of 0.28–0.51 g acetic acid/L and without organoleptic defects) from the Airén variety (Castilla La Mancha, Spain). Table 1 shows the seven types of wine distillates studied.

The total sulfur dioxide (SO_2) content of the wines selected for the production of the AG1, AG2, and AG3 wine spirits was less than 10 mg/L, while the total sulfur dioxide content of those selected for the production of the AG4 and AG6 wine spirits was 73 mg/L and that for the AG5 and AG7 wine spirits 36 mg/L.

The wines were distilled using four different distillation methods: double distillation in pot still, simple distillation in pot still, distillation with two pot stills in series, and continuous column distillation, resulting in seven types of wine spirits (AG1 to AG7).

The wine spirits were hydrated to the aging alcoholic strength by using demineralized water in those cases where it was necessary. The wine distillates, used for the different experiments, met the technical specifications set out in the regulations governing Brandy de Jerez (Consejería de Agricultura Pesca y Desarrollo Rural, 2018; Parlamento Europeo & Consejo de la Unión Europea, 2019). Previous experiences internally carried out at the Bodegas Fundador, S.L.U. distillery, have demonstrated that when sulfur dioxide is incorporated to the winemaking process during fermentation (AG4 to AG7), the distillation of the wines can be carried out over a period of 1 to 6 months without any significant differences being observed in the brandies produced.

The brandies were aged in light- and medium-toasted 350-L oak casks (*Quercus alba*, *Quercus robur*, and *Quercus petraea*) (filled up to 335 L). Likewise, in order to evaluate

Table 1 Description of the types of spirits used and the experiences studied

Test[a]	SO_2 addition[b]	Time of wine distillation[c]	Distillation type	Alcohol content of spirits	Ageing graduation
AG1	No	1 month	Double distillation in pot still	70% ABV	55% ABV
AG2	No	1 month	Simple distillation in pot still	65% ABV	55% ABV
AG3	No	1 month	Distillation with two pot stills in series	65% ABV	55% ABV
AG4	Yes	6 months	Continuous column distillation	77% ABV	55% ABV
AG5	Yes	6 months	Distillation with two pot stills in series	65% ABV	55% ABV
AG6	Yes	6 months	Continuous column distillation	77% ABV	65% ABV
AG7	Yes	6 months	Distillation with two pot stills in series	65% ABV	65% ABV

[a] In these samples, the unaged starting spirits have also been analyzed

[b] In the wine fermentation

[c] Time after fermentation

the evolution of the brandies, three aging times were used: young brandies (unaged) T0, 14-month-old brandies T1, and 28-month-old brandies T2. All the samples were analyzed in duplicate.

Chemicals and Reagents

The compounds used for both the identification of aldehydes and higher alcohols peaks analyzed by GC-FID were supplied by Sigma-Aldrich (Saint Louis, MO, USA).

The hydroalcoholic mixtures of the compounds used for the identification of the peaks were prepared using ethanol, 99.8%, supplied by Sigma-Aldrich (Saint Louis, MO, USA), and ultrapure water (EMD-Millipore, Bedford, MA, USA).

GC-FID Analysis

An Agilent 7890B Gas Chromatograph (Agilent Technologies, Santa Clara, CA, USA) coupled to a flame ionization detector (GC-FID) was used to acquire the chromatograms. A DB-624 column (30 m × 250 μm × 1.4 μm, Agilent Technologies, Santa Clara, CA, USA) was employed to obtain the chromatograms corresponding to the aldehydes and higher alcohols. A CP-WAX 57 CB column (25 m × 250 μm × 0.2 μm, Agilent Technologies, Santa Clara, CA, USA) was used to obtain the chromatograms corresponding to the major esters. The samples were directly injected.

The GC-FID methodology used was the one previously described by Valcárcel-Muñoz et al. (2021).

Data Processing

The data were acquired using the software application OpenLAB CDS Chemstation (Agilent Technologies, Santa Clara, CA, USA). To obtain the chromatographic profiles, i.e., the instrumental fingerprints, all the chromatograms were exported into CSV format. For the construction of the two fingerprint matrices corresponding to the brandies under study, the first one related with the higher alcohols and aldehydes chromatograms and the second one to the major esters chromatograms, the procedure described by Bagur-González et al. (2015) was followed.

Two 232 × 8520 and 232 × 20,100 fingerprint matrices were obtained respectively for aldehydes and higher alcohols and for major esters. A representative fingerprint of each family of compounds has been included in the Online Resource 1 (Fig. OR1).

The data were preprocessed by means of MATLAB R2013b (Mathworks Inc., Natick, MA, USA), by applying the specific script known as "Medina" (version 14) (Pérez-Castaño et al., 2015) in accordance to the procedure described in previous works (Ortega-Gavilán et al., 2020; Pérez-Castaño et al., 2019). This script takes advantage of different functions in Matlab Bioinformatics Toolbox to filter, smooth, and correct the signal baseline and also to perform the normalization of the intensity values with respect to the intensity of the internal standard. As a last step, this script uses the "icoshift" algorithm to align the peaks in the chromatograms (Tomasi et al., 2011).

Previously to applying pattern recognition techniques, each matrix was mean centered using the PLS_Toolbox software, as a final pre-processing stage.

Results and Discussion

In order to evaluate the usability of the encrypted information in the instrumental fingerprints (that corresponded not only to clear markers but also to unknown compounds) to obtain information about the natural grouping trends of the heterogeneous samples set (i.e., seven wine spirits from different wines fermented either with or without the addition of sulfur dioxide, subjected to five different distillation methods, and aged at two alcoholic strengths), hierarchical cluster analysis (HCA) and principal component analysis (PCA) were applied. In addition, partial least squares–discriminant analysis (PLS-DA) and support vector machine (SVM) were employed as the tools to evaluate the discriminating/classifying suitability of the fingerprints used.

Unsupervised Pattern Recognition Methods

The analysis of the natural grouping trends could allow establishing a correlation between the data in the instrumental fingerprints and their impact on some of the experimental variables in the production process. This would lead to discerning which ones are of relevance with regard to the production of a quality Brandy de Jerez.

Hierarchical Cluster Analysis of the Chromatographic Fingerprints

First of all, an HCA was performed using the data matrices that had been previously defined (Fig. 1a, b). In this analysis, Ward's method and Euclidean distance were used as the linkage criterion and the measure of distance between pairs of observations respectively. To select the number of clusters, a $D_{\text{linkage}} = 2/3$ of the D_{max} was used as internal criterion. It could be observed that the brandies clustered naturally according to the addition or not of sulfur dioxide to the fermenting base wine.

Regardless of the instrumental fingerprint used (either higher alcohols and aldehydes or esters), the brandies grouped mainly into two large clusters: (i) the first one included those brandies whose wine spirit had been obtained from a wine that SO_2 had not been added to, and also whose SO_2 content was below 10 mg/L (cluster I); and (ii) the second one included those brandies produced from SO_2-treated wines and with a sulfur dioxide content ranging between 36 and 73 mg/L (cluster II).

In the case of higher alcohols and aldehydes (Fig. 1a), the degree of variability of the instrumental fingerprints of brandies from wines without SO_2 addition was much lower than the variability of those brandies whose base wines had SO_2 added. This fact is justified by the distance considered until nesting occurs, which is substantially shorter in this cluster, probably because of the lesser differences in SO_2 content.

Regarding the clustering of the brandies attending to their major esters content (Fig. 1b), it can be observed that those brandies from wines with added SO_2, a previous partial clustering takes place according to both their distillation method and their SO_2 content level as follows: serial pot

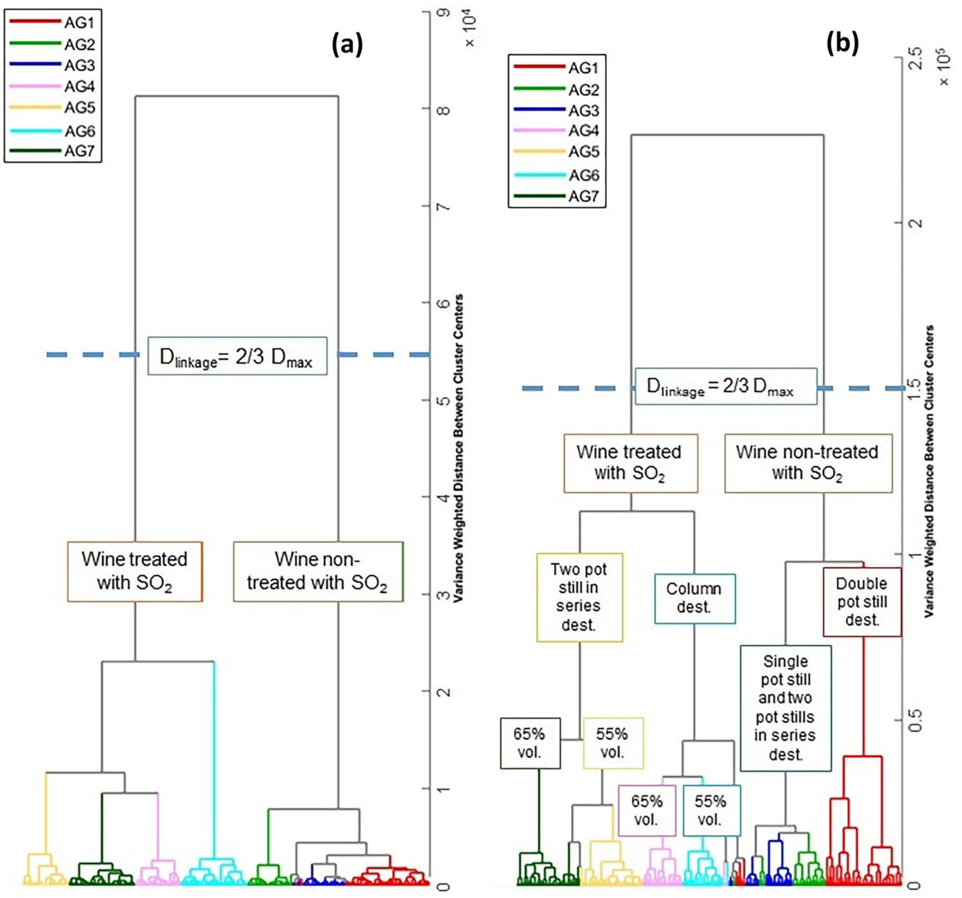


Fig. 1 Dendrograms from the HCA of the analyzed brandies using aldehydes and higher alcohols fingerprints matrix (**a**) and the major ester fingerprints matrix (**b**)

stills (brandies AG5 and AG7) with 36 mg/L of SO_2 and column distillation (brandies AG4 and AG6) with 73 mg/L of SO_2. These sulfur dioxide contents have an influence on the initial clustering of the brandies with different aging alcoholic strengths (55% ABV or 65% ABV).

Therefore, it should be noted that, regardless of the aging time in the cask, the use of SO_2 during the wine fermentation and the distillation method used to obtain the wine spirit are the factors with the most significant influence on the volatile fraction of the brandies. In other words, the starting raw material (wine and wine spirit) has a greater impact on the brandies' major volatile compound content of the aged brandies than the aging process itself.

Principal Component Analysis

PCA of Aldehydes and Higher Alcohols Chromatographic Fingerprints When PCA was applied to the matrix of the instrumental fingerprints of the higher alcohols and aldehydes, 3 principal components (PCs) were obtained which explained 98.12% of the variance of the model for the brandies. PC1 explained 94.76% of the total variance of the system, while the other two principal components explained respectively 2.61% and 0.75% of the remaining variance.

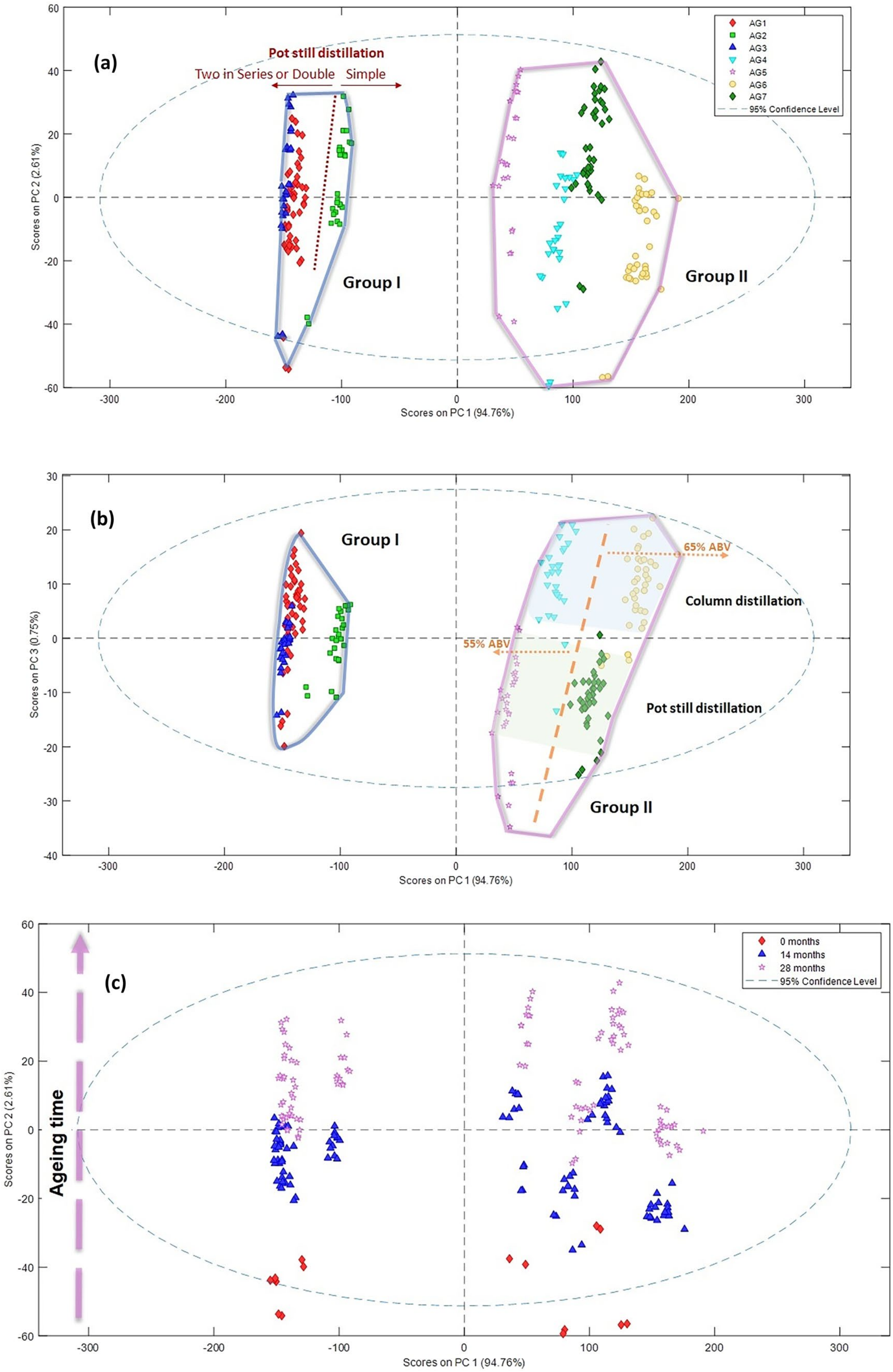
Figure 2a illustrates the scores received by the brandies in the space of the first two components (PC2 vs. PC1). In this figure, it can be observed that, similarly to what occurs when HCA is applied, the brandies are once again grouped according to the addition of SO_2 during the fermentation of the base wine. It can also be seen that the brandies from the wines without SO_2 addition received negative scores for PC1 (group I), while the brandies from the wines which had SO_2 added scored positively for this component (group II). A common trend to separate some brandies as potential "outliers" can be observed in both groups. These scores correspond to the unaged brandies, which could explain this behavior. In addition, with respect to the brandies in group I, a second effect attributable to the number of times the distillation is carried out in the pot stills can be observed. Thus, certain differences among brandies from wine spirit obtained by double or serial distillation (AG1 and AG3 respectively) and the brandies from wine spirit obtained by single distillation (AG2) can be observed.

Figure 2b displays the graphical representation of the scores received by the brandies in the PC3 vs. PC1 space. Considering the distribution of the brandies' score along the PC1 space, once again, the same main grouping that attends to the SO_2 treatment of the base wine (groups I and II) can be observed. In the case of group II, a new subgrouping is observed, which can be explained attending to the distillation process. Thus, the brandies from the wine spirit obtained through column distillation (AG4 and AG6) had positive PC3 scores, while those brandies from distillates obtained by means of serial pot stills (AG5 and AG7) were in the most negative area of this component. Furthermore, in this same group, it could be also observed that as the alcohol content of the aging wine spirit increases, the higher PC1 positive scores are given to the brandies. The new described trends could be explained attending to those fingerprint regions that have minimum variations on the total variance, i.e., minimum differences among fingerprints.

When the PCA corresponding to the aging times was completed, it could be observed that the brandies represented in the new PC2 vs. PC1 plane could be differentiated according to their aging time in the casks (Fig. 2c). Thus, each of the different types of brandies considered received increasingly positive PC2 scores as aging time was longer. This new PCA corroborates not only the scores received by those brandies that were further away from the rest of the clusters observed in the previous model, but also allows to identify each type of starting wine spirit used to produce the brandy (Fig. 2a, b).

If the loading plots for each component (Fig. OR2 in Online Resource 1) are analyzed based on Fig. OR2a, it can be observed that the PC1 loadings would allow identifying those variables that explain the grouping of the brandies according to the use of SO_2 during the fermentation of the base wine. The variables associated to the areas of the fingerprint in which the isoamyl alcohols (3-methyl-1-butanol and 2 methyl-1-butanol) appear have a positive influence on this component. Isobutanol, n-propanol, methanol, acetaldehyde, and its corresponding diethyl-acetal also contribute to these groups. Furthermore, the sub-groupings related to the alcoholic degree of the distillates to be aged can also be caused by the aforementioned compounds. On the other hand, the greater positive contribution of the acetaldehyde and diethyl acetal areas and the negative contribution of the isoamyl alcohols to the PC1 loadings indicate that the scores received by the brandies are affected by the alcoholic strength of the distillates.

An analysis of the loadings plot that corresponds to PC2 (Fig. OR2b) reveals that aging is reflected by the positive trend in the area of the fingerprint where ethyl acetate and n-butanol appear. In addition, with the passing of time, variations take place in the areas corresponding to acetaldehyde, n-propanol, and diethyl acetal. The initial ethyl acetate content in wine spirits depends, on the one hand, on the addition of SO_2 to the wine used to obtain the distillate, and on the other, on the distillation method used, where a greater or lesser separation of the head compounds determines its content (Balcerek et al., 2017; Louw & Lambrechts, 2012; Xiang et al., 2020). Likewise, ethyl acetate is involved in numerous esterification reactions between acetic acid (generated during aging) and ethanol. Guerrero-Chanivet et al. (2020) proved that wood is also capable of transferring acetic acid into the wine spirit being aged, thus accounting for


(a)
Pot still distillation
Two in Series or Double
Simple
Group I
Group II
AG1
AG2
AG3
AG4
AG5
AG6
AG7
95% Confidence Level
Scores on PC 1 (94.76%)
Scores on PC 2 (2.61%)
(b)
Group I
Group II
65% ABV
55% ABV
Column distillation
Pot still distillation
Scores on PC 1 (94.76%)
Scores on PC 3 (0.75%)
(c)
Ageing time
0 months
14 months
28 months
95% Confidence Level
Scores on PC 1 (94.76%)
Scores on PC 2 (2.61%)

◂**Fig. 2** Score plots of the brandies in the plane PC2 vs. PC1 (**a**), the plane PC3 vs. PC1 (label information according to the wine spirits employed) (**b**), and the plane PC2 vs. PC1 changing the label information according to their aging time (**c**)

the fact that brandies have a higher acetic acid content than young wine spirit. This implies that a greater amount of ethyl acetate is generated by esterification as time goes by, which makes this compound a marker of brandies' age.

Finally, Fig. OR2c (loading plot related to PC3) shows the positive influence on the areas of the fingerprints corresponding to acetaldehyde, and its corresponding diethyl acetal, n-propanol, ethyl acetate, and isobutanol, in addition to the negative influence on the area of the fingerprint corresponding to isoamyl alcohols and methanol, would explain the impact attributable to the distillation method on the production of brandies. The positive influence of the instrumental fingerprint region where acetaldehyde and its corresponding diethyl acetal appear can be explained by the fact that, when using SO_2-added wine for the production of Brandy de Jerez, the resulting distillates have more head, i.e., they are richer in these compounds. If this is taken into account, it would also explain how in Fig. 2a, brandies AG4 and AG6, obtained from column distillates, are further right than brandies AG5 and AG7, which had been obtained by pot still distillation (where the wines with low SO_2 and consequently a low acetaldehyde content could only be subjected to a limited removal of their heads if the aromatic quality of the distillate was to be preserved). Contrarily, when column distillation is used, some compounds, such as aldehydes or methanol, are more abundant in the head. Although distillation columns are more versatile, they are also used to distill wines with higher sulfur dioxide content. In turn, the wine spirits obtained from distillation columns are richer in higher alcohols in comparison to those obtained in a more traditional manner such as pot stills (where once the "heart" fraction has been obtained, the following fraction, called "tail," continues presenting some higher alcohol content).

PCA of Major Esters Chromatographic Fingerprints By applying PCA to the matrix of the instrumental fingerprints relative to the major esters, 10 principal components (PCs) were obtained that explained 95.44% of the model variance attributed to the brandies. Since the PCA model is strongly influenced by the use of SO_2 in the base wine (as is the case for aldehydes and higher alcohols), in order to find natural groups that explained the influence of other variables, it was necessary to turn to the representations of other PC scores.

Figure 3 shows the scores received by the brandies in the PC4 (8.14% of the variance) vs. PC1 (40.93% of the variance) space. This figure again exhibits two groups according to whether or not the base wine was treated or not with SO_2 during its fermentation. As in the previous section, group I presented negative PC1 values, while group II showed positive or negative values close to 0 for this component. In this case, the brandies in group I exhibit a greater dispersion of the scores, with sample AG1 presenting a noteworthy dispersion.

Furthermore, as in Fig. 2c, it can be observed that those brandies with a final alcohol content of around 65% ABV presented positive or slightly negative values for this component regardless of the distillation method used, while the brandies with an alcohol volume of around 55% ABV presented negative values for this component. In this case, no groupings were observed based on their aging process.

In order to evaluate which areas of the chromatographic instrumental fingerprints of the major esters exerted the greatest influence on the clusters that had been observed, the loading plots corresponding to the PC1 (Fig. OR3a) and to the PC4 (Fig. OR3b) components were examined. Thus, the groupings of the brandies displayed in Fig. 3 can be explained by the positive influence that the areas of the fingerprint where ethyl lactate, hexanol, diethyl succinate, and 2-phenylethanol appear have on PC1. The areas where ethyl hexanoate, ethyl octanoate, ethyl decanoate, ethyl dodecanoate, and ethyl tetradecanoate are present also contribute to this grouping through their negative influence on this PC; these compounds being esters derived from fatty acids are responsible for the fruity aromas of brandy. This fact corroborates the actual impact from the use of SO_2 during the fermentation of the base wines as a key factor for this clustering, since the brandies that come from the wine spirits that had not been treated with SO_2 present a greater dispersion, as they contain yeast residues that would result in ester-rich distillates. This fact, together with the use of pot stills, means that very little head is removed, which means that the distillates are very rich in fruity compounds with a high aromatic complexity.

On the one hand, when Fig. OR3b is more deeply examined, the sub-groupings that can be observed in the SO_2 brandies group and that are based on the final alcoholic strength of the brandies would be explained by the positive influence from the areas of the fingerprints corresponding to ethyl hexanoate, ethyl lactate, ethyl octanoate, ethyl decanoate, ethyl dodecanoate, 2-phenylethanol, and ethyl hexadecanoate, in addition to the negative influence from the fingerprint areas corresponding to hexanol, furfural, and diethyl succinate. On the other hand, the positive influence of the former on PC4 may explain the distribution of the subgroups observed according to the aging alcoholic strength within the group of brandies produced from SO_2-added wines. Furthermore, low alcoholic brandies have a lower pH, which results in a higher hydrolysis of the esters at 55% ABV than at 65% ABV. This fact explains why the brandies in group I present positive PC4 values equal to those of the brandies in group II, which have an alcohol content of around 65% ABV.

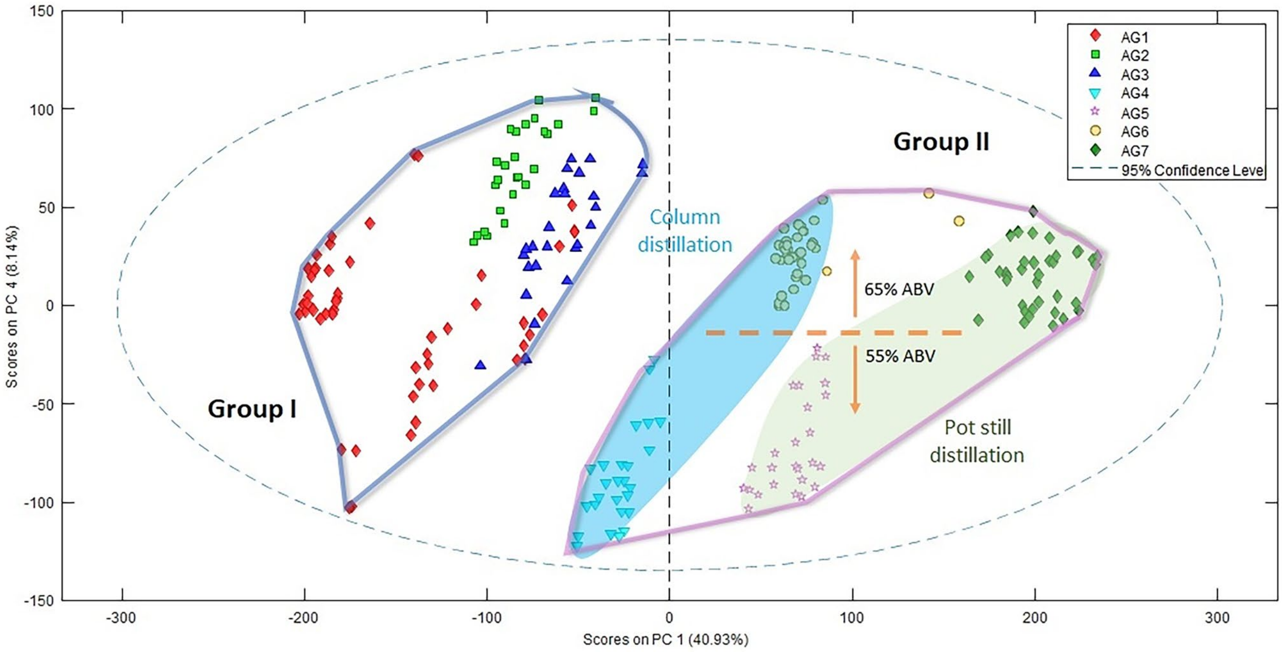


Fig. 3 PC4 vs PC1 score plot of the brandies analyzed labeled according to the wine spirits employed in the elaboration process

Supervised Pattern Recognition Techniques

In order to verify that the groupings obtained by unsupervised pattern recognition allow the experimental variables studied to be regarded as classificatory variables, different binary (one input class) discrimination/classification models were developed using both PLS-DA and SVM.

In all the cases, the original instrumental fingerprints matrices were divided into two subsets: (i) the first one, constituted by a matrix of 163 instrumental fingerprints used to establish the model and for internal cross-validation (for PLS DA Venetian blinds, data Split 10, and leave one out for SVM); and (ii) the second subset constituted by a matrix of 69 instrumental fingerprints used for the external validation of the models at the prediction stage. This subset was obtained by applying Kennard-Stone's algorithm.

Partial Least Squares–Discriminant Analysis

PLS-DA Model According to the Use of Sulfur Dioxide During the Base Wine Fermentation Stage This model was built using the matrix of the instrumental fingerprints corresponding to the major esters, while considering the use of SO_2 to produce the base wines of the young wine spirits as the input class. Three latent variables were selected, which explains 66.06% of the total variance in the matrix of the instrumental fingerprints used in the training stage, and 93.90% of the total variance of the class.

When the binary classification plot (Fig. OR4a) was examined, it could be seen that the model allows the use of the instrumental fingerprints of the major esters to discriminate/classify and correctly predict brandies according to the use of SO_2 for the production of the base wine that gives rise to the initial distillates. The quality metrics of the proposed model are shown in Table OR1 in Online Resource 2.

PLS-DA Model According to the Distillation Method This model was built using the matrix of the instrumental fingerprints corresponding to the aldehydes and major alcohols, while considering the use of distillation columns as input class. For this purpose, 3 latent variables that explained 97.90% of the instrumental fingerprint matrix total variance and 93.85% of the total variance of the class were selected.

When examining the classification graph (Fig. OR4b), it can again be observed that the established model allows the use of the instrumental fingerprint of aldehydes and major alcohols to discriminate/classify and correctly predict the brandies according to the type of distillation method employed to obtain the initial wine spirits. The quality metrics of the proposed model are shown in Table OR2.

It can be seen that for these two models all the parameters related to the error measurement at all the stages (calibration, cross-validation, and prediction/external validation) used in model development are close to 0; meanwhile, sensibility and specificity are equal to 1. These facts show the goodness of the model, which is confirmed by data contained in confusion tables.

PLS-DA Model According to the Aging Time of the Brandies Considering the aging time of the brandies as an input class, new discrimination/classification models were developed. For this purpose, the instrumental fingerprint of higher alcohols and aldehydes, which was affected by the aging time variable, was used. Binary models (one input class) were developed according to the following values: T0 (0 months); T1 (14 months), and T2 (28 months). The binary classification plots obtained for the three models are shown in Fig. 4a–c.

The model for the discrimination between young and aged wine spirits (T0–Not T0) was constructed based on 11 latent variables that explained 99.70% of the variance of the instrumental fingerprints of the samples and 78.04% of the variance of the modeled class. By looking into the binary classification plot of this model (Fig. 4a), we can see that both the samples used as the training set as well as the prediction set appear to be correctly assigned to this modeled class. This fact demonstrates once again that the variation experienced in the chromatographic fingerprint of aldehydes and higher alcohols allows a clear discrimination between young brandies and aged brandies.

The model to discriminate between 14-months-aged wine spirits (T1–Not T1) and the rest of the samples was constructed from 15 latent variables that explained 99.90% of the variance of the samples' instrumental fingerprints and 75.40% of the variance of the modeled class. A higher variability among the different samples can be observed in the binary classification plot of this model (Fig. 4b). In fact, in this figure, the misclassification of some of the samples used for both calibration and prediction (indicated in the figure by a blue arrow) can be observed. It is also worth noting that both, the target class as well as the rest of the samples, present certain proximity to the threshold. The samples that were misclassified correspond to both young wine spirit (T0) and brandies aged for a period of 28 months. This fact suggests that after 14 months of aging, the chromatographic instrumental fingerprints of some of the samples either did not evolve with respect to the initial young brandies or underwent a series of aging processes that made them more similar to brandies aged for a longer period of time.

Finally, Fig. 4c shows the binary classification plot of the model that had been constructed taking T2 as input class. This model was developed by selecting 13 latent variables which explained 99.85% of the total variance of the samples and 82.80% of the total variance of the class. As with the model built for the T0 class, we can consider that this model discriminates/classifies correctly, even though a certain confusion between some samples could be detected during the cross-validation and prediction stages. This confusion among samples was again attributable to the brandies that had been aged for 14 months which, in this case, exhibited a greater similarity with the target class.

The quality metrics of the different models can be seen in Tables OR3 to OR5. In general terms, it should be pointed out that the best metrics are obtained from the model established to distinguish between young and aged spirits (Table OR3), whereas the worst are those arising from the model where the mid-aging time is considered.

Support Vector Machine

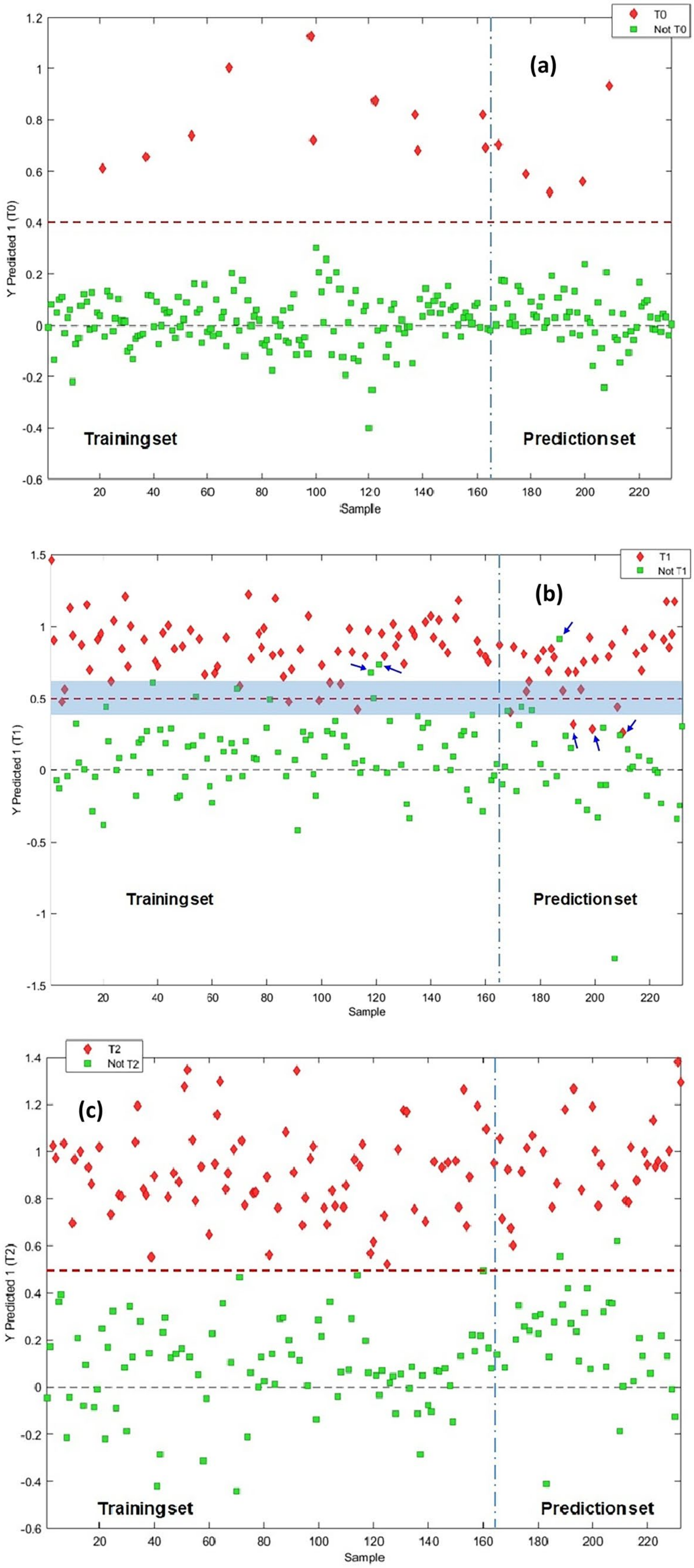
The SVM models generated were uncompressed, i.e., without any reduction of the data dimensions, applying the radial basis function (RBF) as Kernel algorithm and using the PLS_Toolbox default values for the gamma and cost parameters. As in the PLS-DA models, two strategies were followed to evaluate the classifying ability of the experimental variables that give rise to the different natural groupings that were identified from both HCA and PCA: (i) using the instrumental fingerprints of the major esters to determine the effect from the use of distillates derived from wines with and without added SO_2 on the brandies, and (ii) using the instrumental fingerprints of aldehydes and higher alcohols to evaluate both distillation method and aging time. All the models developed were one input class.

SVM Model Attending to the Use of Sulfur Dioxide During the Base Wine Fermentation Stage Using the class prediction probability graph (Fig. OR5a), it can be appreciated that all the samples were successfully classified as in or out of the target class, similar to the corresponding PLS-DA model. However, the model obtained by SVM exhibited a lower dispersion when predicting class membership (target and non-target classes). In all the cases, the samples present values close to 1 for the target class and close to 0 for the non-target class. The quality metrics of the proposed model are presented in Table OR6.

SVM Model According to the Distillation Method When analyzing the class prediction probability graph (Fig. OR5b), it can again be observed that the model established allows the use of the instrumental fingerprint of aldehydes and major alcohols to correctly classify and predict brandies produced according to the distillation method. As with the previous model, SVM again shows less dispersion of the classification results with respect to the equivalent PLS-DA model. The quality metrics of the proposed model are included in Table OR7, with similar results with respect to the previous model (section A).

In the same fashion as PLS-DA, it can be observed that all the parameters related to the error measurement at all the stages used in model development are close to 0; meanwhile, sensibility and specificity are equal to 1. Once again, these facts show the goodness of the model, which is confirmed by data contained in confusion tables.

Fig. 4 Binary classification plots obtained from the PLS-DA models T0 – Not T0 (**a**), T1 – Not T1 (**b**), and T2 – Not T2 (**c**)



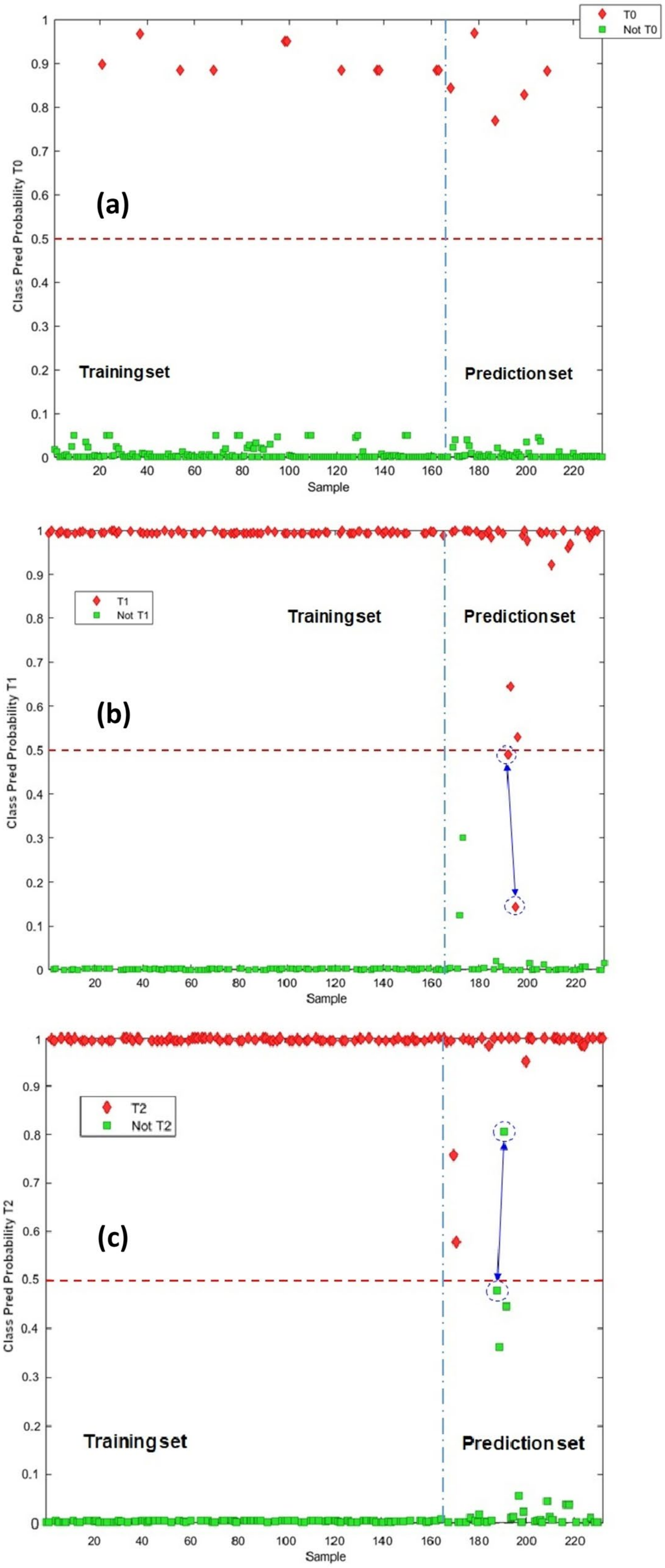


Fig. 5 Classification prediction probability plots obtained from the SVM models: T0 (0 months) – Not T0 (14 months and 28 months) model (**a**), T1 (14 months) – Not T1 (0 months and 28 months) model (**b**), and T2 (28 months) – Not T2 (0 months and 14 months) (**c**)

SVM Models According to the Aging Time of the Brandies Similar to the PLS-DA model, these classification models were developed based on the aging time of the brandies by means of SVM. In the same way, the models were established as binary models (one input class) based on the following classes: T0 (0 months), T1 (14 months), and T2 (28 months).

Figure 5a–c shows the respective class prediction probability graphs obtained for each case.

Based on the classification model for the unaged wine spirit (Fig. 5a), it could be verified that both the samples used as the training set and those used for the prediction set appeared to be correctly classified. Once again, the dispersion of the results was considerably lower than those shown in the corresponding PLS-DA model. This fact remained invariable for the rest of the models developed.

Regarding the brandies with 14 months of aging (Fig. 5b), it can be clearly seen that some misclassified samples appeared. The dispersion of the results is much greater compared to the corresponding PLS-DA model, where the misclassified samples are not only dispersed but also close to the correctly classified samples.

Finally, we can see from Fig. 5c how class predictions improved with respect to the PLS-DA model, given that only one sample was misclassified and another one was scored very close to its classification threshold. These two samples coincided with those that had been misclassified using the previous model (T1). The quality parameters of these three models are shown in Tables OR8 to OR10. In general terms, it was observed that the models created using SVM were more reliable than those ones developed by PLS-DA.

Conclusion

The chemometric study of the instrumental fingerprints obtained for the volatile fractions corresponding to aldehydes and higher alcohols as well as the major esters showed that for the experimental variables studied, i.e., the use of SO_2 during the fermentation process of the wines that give rise to the wine spirit, the distillation method applied to the base wines and the aging time of the wine spirit allowed us to discriminate and classify the brandies studied.

In addition, the different areas of the instrumental fingerprints allowed us to rank them in order of importance. Thus, by means of the unsupervised techniques used (HCA and PCA), it was demonstrated that the most influential variable identified was the use of SO_2 during the fermentation stage of the base wine, followed by the distillation method employed to obtain the wine spirit to be aged and, finally, the aging time in the cask. Nevertheless, although to a lesser extent, some differences can also be observed according to the alcoholic strength of the wine spirit to be aged.

By applying supervised techniques such as PLS-DA and SVM, we have obtained models to discriminate (PLS-DA) and classify (SVM) the brandies according to the aforementioned most influential variables. Furthermore, the models generated by means of SVM were more reliable in terms of the quality of their metrics.

This research confirms that the use of SO_2 during the fermentation process of the base wine to be used for the production of brandies has an impact on the major volatile compound profiles of the final products. Even after 28 months of aging, SO_2 is still an important variable to be taken into account when selecting wines to be distilled for the production of Brandy de Jerez.

Finally, it should be emphasized that the fingerprinting methodology has been proven to be really suitable for the analysis of this type of matrices since it allows to take into account not only the compounds that have been identified and considered to be the most influential markers, but also those compounds that, without having to be identified, allow the fingerprints to be associated to the brandies studied. This should be considered as a very useful feature, in terms of internal control, for brandy producers.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11947-023-03039-8.

Author Contribution María Guerrero-Chanivet, Manuel J. Valcárcel-Muñoz, M. Valme García-Moreno, and Dominico A. Guillén-Sánchez carried out the conceptualization, methodology, and formal analysis of the samples. María Guerrero-Chanivet, Fidel Ortega-Gavilán, and M. Gracia Bagur-González wrote the main manuscript text and prepared all the figures. All the authors reviewed the manuscript.

Funding Funding for open access publishing: Universidad de Cádiz/CBUA. This research has been supported by the University of Cadiz and Bodegas Fundador, S.L.U. (ref.: OT2019/108, OT2020/128 and OT2021/076). The authors would like to thank the University of Cadiz and Bodegas Fundador, S.L.U., for the industrial pre-doctoral contract TDI-8-18, granted to the author Guerrero Chanivet, M.

Data Availability Not applicable.

Declarations

Conflict of Interest The authors declare no competing interests.

References

Bagur-González, M. G., Pérez-Castaño, E., Sánchez-Viñas, M., & Gázquez-Evangelista, D. (2015). Using the liquid-chromatographic-fingerprint of sterols fraction to discriminate virgin olive from other edible oils. *Journal of Chromatography A, 1380*, 64–70. https://doi.org/10.1016/j.chroma.2014.12.052

Balcerek, M., Pielech-Przybylska, K., Patelski, P., Dziekońska-Kubczak, U., & Strąk, E. (2017). The effect of distillation conditions and alcohol content in 'heart' fractions on the concentration of aroma volatiles and undesirable compounds in plum brandies. *Journal of the Institute of Brewing, 123*(3), 452–463. https://doi.org/10.1002/jib.441

Berry, D. R., & Slaughter, J. C. (2003). Alcoholic beverage fermentations. In A. G. Lea & J. R. Piggott (Eds.), *Fermented beverage production* (2nd ed., pp. 25–38). Kluwer Academic/Plenum Publishers.

Bikrani, S., Jiménez-Carvelo, A. M., Nechar, M., Bagur-González, M. G., Souhail, B., & Cuadros-Rodríguez, L. (2019). Authentication of the geographical origin of margarines and fat-spread products from liquid chromatographic UV-absorption fingerprints and chemometrics. *Foods, 8*(11), 588. https://doi.org/10.3390/foods8110588

Cacho, J., Moncayo, L., Palma, J. C., Ferreira, V., & Culleré, L. (2013). The impact of grape variety on the aromatic chemical composition of non-aromatic Peruvian pisco. *Food Research International, 54*(1), 373–381. https://doi.org/10.1016/j.foodres.2013.07.019

Consejería de Agricultura Pesca y Desarrollo Rural. (2018). Orden *de 28 de junio de 2018, por la que se aprueba el expediente técnico de Indicación Geográfica "Brandy de Jerez", 127 Boletin Oficial de la Junta de Andalucía 19 a 20*. Retrieved October 20, 2022, from https://www.juntadeandalucia.es/boja/2018/127/3

García-Llobodanin, L., Achaerandio, I., Ferrando, M., Güell, C., & López, F. (2007). Pear distillates from pear juice concentrate: effect of lees in the aromatic composition. *Journal of Agricultural and Food Chemistry, 55*(9), 3462–3468. https://doi.org/10.1021/jf0633589

Guerrero-Chanivet, M., Valcárcel-Muñoz, M. J., García-Moreno, M. V., & Guillén-Sánchez, D. A. (2020). Characterization of the aromatic and phenolic profile of five different wood chips used for ageing spirits and wines. *Foods, 9*(11), 1613. https://doi.org/10.3390/foods9111613

Hernandez-Gomez, L. F., Ubeda, J., & Brions, A. (2003). Melon fruit distillates: comparison of different distillation methods. *Food Chemistry, 82*(4), 539–543. https://doi.org/10.1016/S0308-8146(03)00008-6

OIV, International Organization of Vine and Wine. (2021). Sulfiting (Oeno 3/04), in: International Code of Oenological Practices. OIV, Paris, France, p. II.1.1:18–19.

Korenika, A. M. J., Biloš, J., Kozina, B., Tomaz, I., Preiner, D., & Jeromel, A. (2020). Effect of different reducing agents on aromatic compounds, antioxidant and chromatic properties of sauvignon blanc wine. *Foods, 9*(8), 996. https://doi.org/10.3390/foods9080996

Louw, L., & Lambrechts, M. G. (2012). *Grape-based brandies: Production, sensory properties and sensory evaluation*. Woodhead Publishing Limited. https://doi.org/10.1533/9780857095176.3.281

Ortega-Gavilán, F., Valverde-Som, L., Rodríguez-García, F. P., Cuadros-Rodríguez, L., & Bagur-González, M. G. (2020). Homogeneity assessment of reference materials for sensory analysis of liquid foodstuffs. The virgin olive oil as case study. *Food Chemistry, 322*, 126743. https://doi.org/10.1016/j.foodchem.2020.126743

Parlamento Europeo, & Consejo de la Unión Europea. (2019). Reglamento (UE) 2019/787 del Parlamento Europeo y del Consejo de 17 de abril de 2019. *Diario Oficial de la Unión Europea, L130*(III), 17–27.

Pastor, K., Ačanski, M., Vujić, Đ, & Kondić-Špika, A. (2016). Binary simple sugar profiling in corn and small grain flour authentication using GC/EI-qMS approach. *Chromatographia, 79*(21–22), 1553–1559. https://doi.org/10.1007/s10337-016-3159-0

Pastor, K., Ilić, M., Vujić, D., Jovanović, D., & Ačanski, M. (2020). Characterization of fatty acids in cereals and oilseeds from the Republic of Serbia by gas chromatography–mass spectrometry (GC/MS) with chemometrics. *Analytical Letters, 53*(8), 1177–1189. https://doi.org/10.1080/00032719.2019.1700270

Pérez-Castaño, E., Medina-Rodríguez, S., & Bagur-González, M. G. (2019). Discrimination and classification of extra virgin olive oil using a chemometric approach based on TMS-4,4′-desmetylsterols GC(FID) fingerprints of edible vegetable oils. *Food Chemistry, 274*, 518–525. https://doi.org/10.1016/j.foodchem.2018.08.128

Pérez-Castaño, E., Ruiz-Samblás, C., Medina-Rodríguez, S., Quirós-Rodríguez, V., Jiménez-Carvelo, A. M., Valverde-Som, L., et al. (2015). Comparison of different analytical classification scenarios: application for the geographical origin of edible palm oil by sterolic (NP) HPLC fingerprinting. *Analytical Methods, 7*(10), 4192–4201. https://doi.org/10.1039/C5AY00168D

Psodorov, Đ, Ačanski, M., Psodorov, D., Vujić, Đ, & Pastor, K. (2015). Determination of the content of buckwheat and wheat flours in bread using GC-MS and multivariate analysis. *Journal of Food and Nutrition Research, 54*(2), 179–183.

Sánchez-Guillén, M. M. M. M., Schwarz-Rodríguez, M., Rodríguez-Dodero, M. C. C., García-Moreno, M. V. V., Guillén-Sánchez, D. A. D. A., & García-Barroso, C. (2019). Discriminant ability of phenolic compounds and short chain organic acids profiles in the determination of quality parameters of Brandy de Jerez. *Food Chemistry, 286*, 275–281. https://doi.org/10.1016/j.foodchem.2019.02.006

Silva, M. L., & Malcata, F. X. (1999). Effects of time of grape pomace fermentation and distillation cuts on the chemical composition of grape marcs. *European Food Research and Technology, 208*(2), 134–143. https://doi.org/10.1007/s002170050390

Spaho, N. (2017). Distillation techniques in the fruit spirits production. *Distillation - Innovative Applications and Modeling*. https://doi.org/10.5772/66774

Spaho, N., Dürr, P., Grba, S., Velagić-Habul, E., & Blesić, M. (2013). Effects of distillation cut on the distribution of higher alcohols and esters in brandy produced from three plum varieties. *Journal of the Institute of Brewing, 119*(1–2), 48–56. https://doi.org/10.1002/jib.62

Swiegers, J. H., Bartowsky, E. J., Henschke, P. A., & Pretorious, I. S. (2005). Yeast and bacterial modulation of wine aroma and flavour. *Australian Journal of Grape and Wine Research, 11*(2), 139–173. https://doi.org/10.1111/j.1755-0238.2005.tb00285.x

Szymańska, E. (2018). Modern data science for analytical chemical data – a comprehensive review. *Analytica Chimica Acta, 1028*, 1–10. https://doi.org/10.1016/j.aca.2018.05.038

Tomasi, G., Savorani, F., & Engelsen, S. B. (2011). icoshift: an effective tool for the alignment of chromatographic data. *Journal of Chromatography A, 1218*(43), 7832–7840. https://doi.org/10.1016/j.chroma.2011.08.086

Tsakiris, A., Kallithraka, S., & Kourkoutas, Y. (2014). Grape brandy production, composition and sensory evaluation. *Journal of the Science of Food and Agriculture, 94*(3), 404–414. https://doi.org/10.1002/jsfa.6377

Valcárcel-Muñoz, M. J., Guerrero-Chanivet, M., García-Moreno, M. V., Rodríguez-Dodero, M. C., & Guillén-Sánchez, D. A. (2021). Comparative evaluation of Brandy de Jerez aged in American oak barrels with different times of use. *Foods, 10*, 288. https://doi.org/10.3390/foods10020288

Valero, E., Moyano, L., Millan, M., Medina, M., & Ortega, J. (2002). Higher alcohols and esters production by Saccharomyces cerevisiae. Influence of the initial oxygenation of the grape must. *Food Chemistry, 78*(1), 57–61. https://doi.org/10.1016/S0308-8146(01)00361-2

Xiang, X. F., Lan, Y. Bin., Gao, X. T., Xie, H., An, Z. Y., Lv, Z. H., et al. (2020). Characterization of odor-active compounds in the head, heart, and tail fractions of freshly distilled spirit from Spine grape (Vitis davidii Foex) wine by gas chromatography-olfactometry and gas chromatography-mass spectrometry. *Food Research International, 137*, 109388. https://doi.org/10.1016/j.foodres.2020.109388

Zierer, B., Schieberle, P., & Granvogl, M. (2016). Aroma-active compounds in bartlett pears and their changes during the manufacturing process of Bartlett pear brandy. *Journal of Agricultural and Food Chemistry, 64*(50), 9515–9522. https://doi.org/10.1021/acs.jafc.6b04612