Postprint of Food Chemistry Volume 230, 1 September 2017, Pages 108–116

DOI 10.1016/j.foodchem.2017.02.118

CHARACTERIZATION AND AUTHENTICATION OF SPANISH PDO WINE VINEGARS USING MULTIDIMENSIONAL FLUORESCENCE AND CHEMOMETRICS

Rocío Ríos-Reina[a], Saioa Elcoroaristizabal[b], Juan A Ocaña-González[c], Diego L. García-González[d], José M. Amigo[e], Raquel M. Callejón[a,*]

[a]Área de Nutrición y Bromatología, Facultad de Farmacia, Universidad de Sevilla, C/P. García González n°2, E- 41012, Sevilla, Spain

[b] Chemical and Environmental Engineering Department, Faculty of Engineering, University of the Basque Country, Alameda de Urquijo s/n, E-48013 Bilbao, Spain

[c] Departamento Química Analítica, Facultad de Química, Universidad de Sevilla, C/P. García González s/n, E-41012 Sevilla, Spain

[d] Instituto de la Grasa (CSIC), Campus University Pablo de Olavide - Building 46, Ctra. de Utrera, km. 1 E– 41013, Sevilla, Spain

[e]Chemometric Analytical Technologies, Department of Food Sciences, Faculty of Science, University of Copenhagen, Rolighedsvej 30, Frederiksberg CDK-1958, Denmark

Running Title: Characterization and authentication of Spanish PDO vinegars by fluorescence and chemometrics.

*Author to whom correspondence should be sent.

E-mail: rcallejon@us.es

Abstract

This work assesses the potential of multidimensional fluorescence spectroscopy combined with chemometrics for characterization and authentication of Spanish Protected Designation of Origin wine vinegars. Seventy-nine vinegars of different categories (aged and sweet) belonging to the Spanish PDOs *"Vinagre de Jerez", "Vinagre de Montilla-Moriles"* and *"Vinagre de Condado de Huelva"*, were analyzed by excitation-emission fluorescence spectroscopy. A visual assessment of fluorescence landscapes pointed out different trends with vinegar categories. PARAllel FACtor analysis (PARAFAC) extracted the potential fluorophores and their values in the PDO vinegars. This information, coupled with different classification methods (Partial Least Square Discrimination Analysis "PLS-DA" and Support Vectors Machines "SVM"), was able to discriminate the wine vinegar category within each PDO, for which SVM models obtained better results (>92% of classification). In each category, SVM also allows the differentiation between PDOs. The proposed methodology could be used as an analysis method for the authentication of Spanish PDO wine vinegars.

1. Introduction

Vinegar is a product used worldwide as a condiment and food preserving agent, obtained by a double fermentation process (alcoholic and acetic fermentation) of sugary and starchy substrates (FAO, 1998). Vinegar can be produced by different methods and raw materials (such as malt, apple, rice, etc.), among which, wine vinegar is the most commonly produced and consumed vinegar in Mediterranean countries and Central Europe (Polo & Sanchez-Luengo, 1991).

For many years, wine vinegar has been considered as a low-cost secondary product spontaneously derived from wine production. However, in recent years wine vinegar has become a valued food product much appreciated in gastronomy. As a result, the demand for high-quality wine vinegars has significantly increased over the last years. In this framework, Spain is one of the major producers of high-quality wine vinegars, including three of the five types of vinegar registered in Europe (Council Regulation (EC) No 510/2006) with a "Protected Designation of Origin" (PDO): *"Vinagre de Jerez", "Vinagre de Montilla-Moriles"* and *"Vinagre de Condado de Huelva"* (Table I Supplemental Material). The production of these high-quality PDO wine vinegars in Spain is centered in Andalusia, each of them made from the corresponding protected wines (*Jerez, Montilla-Moriles and Condado de Huelva*), which provides singular and specific characteristics to each vinegar. In addition, the production of high-quality vinegars requires an ageing period in wooden butts. During the period of aging, some chemical modifications take place providing them with unique organoleptic properties and higher sensory quality (Morales, Tesfaye, García-Parrilla, Casas, & Troncoso, 2002). According to the sweetness, time and method of ageing ("*criaderas and solera"* and *"añada"* system*)*, different categories are considered within each Spanish PDO (Table 1).

The longer aging time is directly related to both the higher quality and production costs of these wine vinegars. This fact increases the final market price and makes the quality assurance and authentication of the Spanish PDOs wine vinegars an important issue. For this reason, objective analytical methodologies are required to guarantee the wine vinegar authenticity and fight against frauds. However, the most common analytical techniques used for the characterization and authentication of these vinegars rely on chromatographic methods that are often expensive and time-consuming (Aceña, Vera, Guasch, Busto, & Mestres, 2011; Cirlini, Caligiani, Palla, & Palla, 2011). Thus, in recent years there has been a growing interest in developing rapid,

inexpensive, non-destructive and direct methodologies based on non-targeted techniques for food authentication. Fluorescence spectroscopy has been increasingly applied as a competitive, high sensitivity, fast and non-destructive technique in food analysis (Karoui & Blecker, 2011). This spectroscopic technique has been more commonly used in wine (Airado-Rodŕiguez, Galeano-Díaz, Durán-Merás, & Wold, 2009; Azcarate et al., 2015), but rarely adopted for wine vinegar samples (Callejón et al., 2012) and hence, there is still scarce information about vinegar fluorescent components.

In this sense, wine vinegar is a very complex multi-component mixture of chemical compounds due to its traditional making procedure, the raw material used and the ageing period and method employed. Some of these chemical compounds are polyphenols, amino acids and vitamins (Airado-Rodríguez, Durán-Merás, Galeano-Díaz, & Wold, 2011), whose presence is related to the wine chemical basis. To handle this complexity, fluorescence multidimensional measurements, such as excitation-emission fluorescence spectroscopy, combined with adequate multi-way methods (Andersen & Bro, 2003; Sádecká & Tóthová, 2007) have been proven to be useful for characterization of complex food matrices (Callejón et al., 2012; Christensen, Becker, & Frederiksen, 2005; Elcoroaristizabal et al., 2016; Lenhardt, Bro, Zeković, Dramićanin, & Dramićanin, 2015). Measuring the emission spectra at different excitation wavelengths results into a bi-dimensional Excitation-Emission Matrix (EEM), which contains unique information of each measured sample. Therefore, a three dimensional array is obtained when all the samples are gathered together, so requiring an appropriate data processing for its interpretation.

An adequate multiway method, such as PARAllel FACtor Analysis (PARAFAC), can be used to decompose fluorescence EEMs into different independent groups of fluorescence components (fluorophores), as well as their relative concentration (scores) in each sample (Bro, 1997). The information provided by the resolved fluorophores has been successfully applied in food quality control, since it can reveal clearer insights into the relationships between the intrinsic food properties and the quality of the product. For instance, EEM-PARAFAC has been applied for monitoring the changes occurring during the storage and production of different food samples (Christensen, Becker, & Frederiksen, 2005; Elcoroaristizabal et al., 2016) and their characterization (Lenhardt et al., 2015; Tena, Aparicio, & García-González, 2012). Furthermore, the information obtained after EEM data decomposition by PARAFAC

84 modeling could be coupled with different classification methods in order to characterize
85 and classify different food products or detect fraudulent samples (Callejón et al., 2012).

86 There are numerous classification algorithms such as Partial Least Square
87 Discrimination Analysis (PLS-DA), K-Nearest Neighbors (KNN), Support Vector
88 Machines (SVM) and Soft Independent Modelling of Class Analogy (SIMCA) (Cover &
89 Hart, 1967; Vapnik, 1999; Wold, 1966; Wold, 1976). Among them, Partial Least
90 Squares-Discriminant Analysis (PLS-DA) and Support Vectors Machines (SVM) are
91 two of the most common used ones. PLS-DA is a supervised class-modelling method
92 used for building linear discriminant models (Nocairi, Qannari, Vigneau, & Bertrand,
93 2005), which has been successfully applied to a wide variety of food matrices for
94 classification purposes (Azcarate, Cantarelli, Pellerano, Marchevsky, & Camiña, 2013;
95 Lenhardt et al., 2015; Liu, He, & Wang, 2008). The main advantage of the PLS-DA
96 approach is the ability of handling highly collinear and noisy data. However, one of the
97 main issues is that PLS-DA models need a sufficient and balanced amount of samples
98 for each class; and sometimes it is difficult to acquire sufficient samples of some
99 classes, due to their cost of production or their non-availability in the market. Moreover,
100 classes that are not effectively separated linearly are common in food products.
101 Support Vector Machines (SVM) is an effective non-linear machine learning technique
102 suitable for both classification and regression analysis (Xu, Zomer, & Brereton, 2006).
103 In comparison to PLS-DA, the main advantage of SVM is its flexibility in modelling
104 complex classification problems that are non-linear. A common disadvantage of SVM is
105 the lack of transparency of the results, since there are no statistics such as scores and
106 loadings available for easy visualization.

107 Several researchers have tested the SVM's performance in different food
108 authentication problems obtaining better results than other traditional classification
109 methods. For instance, Acevedo et al. (2007) (Acevedo, Jiménez, Maldonado,
110 Domínguez, & Narváez, 2007) observed that SVM performed better than SIMCA, k-
111 NN, and PLS-DA for discrimination of wines according to their PDO, which also
112 enabled the selection of the most relevant UV-Vis wavelengths for samples
113 classification. In the same way Callejón et al. (2012) (Callejón et al., 2012) proved that
114 SVM in conjunction with excitation-emission fluorescence spectroscopy was a more
115 adequate methodology than PLS-DA for the classification of sherry vinegars according
116 to their ageing time. However, the aforementioned study was only focused on the
117 classification of a limited number of wine vinegar categories (aged vinegars) belonging
118 to one Spanish PDO (*"Vinagre de Jerez").*
119

In this context, the aim of this work was to investigate the feasibility of using excitation-emission fluorescence spectroscopy combined with several chemometric techniques for characterization and classification of the three Spanish PDOs wine vinegars and their commercialized categories. First, EEM data will be analyzed by PARAFAC in order to characterize spectroscopically and chemically different commercialized wine vinegar categories (aged and sweet) according to each Spanish PDO. Then, these results will be used to build reliable classification models able to differentiate between the wine vinegar categories corresponding to each Spanish PDO, and each PDO within the same wine vinegar category.

2. Materials and Methods

2.1 Wine vinegar samples

Seventy-nine wine vinegar samples from the three Spanish PDOs coming from several producers were analyzed in this study (Table 1): 30 *"Vinagre de Jerez"*, 18 *"Vinagre de Montilla-Moriles"* and 21 *"Vinagre de Condado de Huelva"* samples. Among the aged categories, these vinegars are aged by the traditional system called "*criaderas and solera*", except for the *"Vinagre de Condado de Huelva Añada"* which is aged by using the static aging system called *"añada"*. Regarding the "*Pedro Ximenez*" category, it should be highlighted that this sweet category differs from the aged category not only by the aging time but also by other factors such as their different production process. Thus, they are produced by the addition of must of raisined "Pedro Ximenez" grapes (in the case of "Vinagre de Montilla-Moriles") or the addition of "Pedro Ximenez" wine to the vinegar. All the samples were purchased from local wineries working in compliance with current regulations of each Spanish PDO. The samples were collected in triplicate and stored in amber vials at room temperature until the analysis.

Within each PDO, a different number of samples were collected for the established categories (aged and sweet) according to the production/sale rates of each category during the last years (2014-2015). In these years, the general trend for the three Spanish PDOs reveals a higher production of the categories with less ageing time due to market trends. For instance, *"Vinagre de Jerez"* (JCR) represented approximately 60% of total sales in the PDO *"Vinagre de Jerez"*, while sales of *"Vinagre de Jerez Reserva"* (JRE) and *"Vinagre de Jerez Gran Reserva"* (JGR) accounted for 40% and 1% of the total, respectively. Similarly, *"Vinagre de Condado de Huelva"* (CSC) category had the highest sales growth up 38% of the total. Meanwhile among the aged

categories of *"Vinagre de Condado de Huelva"* PDO, the most commercialized vinegar categories were, in decreasing order: *"Solera"* (CSO), *"Reserva"* (CRE) and *"Añada"* (CAN). In the same way, *"Vinagre de Montilla-Moriles Crianza"* (MCR) was the most commercialized one of the *"Vinagre de Montilla-Moriles"* PDO due to the recent incorporation to the Spanish PDOs.

2.2 Fluorescence analysis

Fluorescence measurements were recorded using a Varian Cary-Eclipse fluorescence spectrophotometer (Varian Iberica, Madrid, Spain), equipped with two Czerny-Turner monochromators, and a Xenon discharge lamp pulsed at 80 Hz with a half peak height of ~2 µs (peak power equivalent to 75 kW). A high-performance R298 photomultiplier tube detector was used for collection of the fluorescence spectra. Wine vinegar samples were directly analysed without sample pre-treatment by pipetting them into 3.5 mL quartz cuvettes before measurement. Standard quartz cells (Hellma Analytics, Müllheim, Germany) of 1 cm path length were used to carry out the measurements in a Peltier thermostatted cuvette holder (25.00 ± 0.05 ◦C). The spectrometer was interfaced to a computer with Cary-Eclipse software for spectral acquisition and exportation.

The fluorescence Excitation-Emission Matrices (EEMs) were obtained by varying the excitation wavelength (λ_{ex}) ranging between 250 and 700 nm (every 5 nm), and recording the emission spectra (λ_{em}) from 300 to 800 (every 2 nm). For these measurements, excitation and emission slits were both set at 5 nm, and the scan rate was fixed to 1200 nm min^{-1}. The system was wavelength calibrated every day by means of the water Raman peak to account for a possible wavelength drift of the instrument. EEMs were registered by triplicate for each sample and preprocessed in order to avoid noisy and non-informative areas by selecting shorter spectral ranges (λ_{ex} from 250 to 680 nm, and λ_{em} from 310 to 800 nm).

2.3 Software and data analysis

EEM data analysis was performed by using the PLS_Toolbox 7.9.5 (Eigenvector Research Inc., Wenatchee, WA) working under Matlab v.8.5.0 environment (The Mathworks Inc., Natick, MA). Before the analysis, EEMs data were corrected for – Rayleigh and Raman scattering (Elcoroaristizabal, Bro, García, & Alonso, 2015) – by removing and replacing the scattering areas with interpolated values by using the FLUCUT function included in the PLS_Toolbox. FLUCUT Removes Rayleigh scattering (and possibly Raman) by inserting NaN and 0 values in Excitation-Emission Matrices

(EEMs) where the Rayleigh bands are. Alternatively, FLUCUT may also be used to generate weights that can be used for deweighting (instead of eliminating) these regions.

2.3.1 Parallel Factor Analysis (PARAFAC)

PARAllel FACtor models were performed on the corrected EEM data in order to extract the relevant information and develop models for: (1) different wine vinegar categories belonging to the same Spanish PDO, and (2) similar wine vinegar categories belonging to different Spanish PDOs.

Before modelling, the EEM landscapes corresponding to the same Spanish PDO (1) were rearranged into a three-dimensional structure (**X**) of size (3 replicated samples x λ_{em} x λ_{ex}): 90 x 246 x 87 for the PDO *"Vinagre de Jerez"*; 54 x 246 x 87 for the PDO *"Vinagre de Montilla-Moriles"*, and 63 x 246 x 87 for the PDO *"Vinagre de Condado de Huelva"*. In a similar way, the EEM landscapes corresponding to similar wine vinegar categories but different Spanish PDOs (2) were organized into a three-way array (**X**) of size (3 replicated samples x λ_{em} x λ_{ex}): 78 x 246 x 87 for *"Crianza"*; 66 x 246 x 87 for *"Reserva"* category, and 27 x 246 x 87 for *"Pedro Ximenez"* categories. No PARAFAC analysis was carried out for the *"Gran Reserva"* category due to the limited number of samples.

Then, each three-way dataset (**X**) was decomposed by PARAFAC (Bro, 1998). The proper number of factors for each model was determined by using the CORe CONsistency DIAgnostic test (CORCONDIA) (Bro & Kiers, 2003), the percentage of variance explained by the model and the visual inspection of the recovered spectral profiles and residuals. Non-negative constraints for all modes (concentrations and both spectral profiles) were applied to obtain meaningful chemical solutions.

2.3.2 Classification methods

Partial Least Squares-Discriminant Analysis (PLS-DA) (Nocairi et al., 2005) and Support Vectors Machines (SVM) (Vapnik, 1999) algorithms were used to build classification models for discrimination the wine vinegar category within each Spanish PDO. On the one hand, PLS-DA is a classification method based on partial least squares regression (PLS) that transforms the data into a set of linear latent variables for predicting the dependent or class variable, making models that allow the maximum separation among classes. The class variable forms a so-called dummy matrix that indicates whether a sample belongs to a certain class or category. In our study, three

different wine vinegar categories were considered in each Spanish PDO, therefore, the dimensions of each dummy matrix was 3x3.

On the other hand, SVM is a relative new chemometric tool based on the statistical learning theory (SLT). It is a supervised learning method that searchers for the optimal separating hyperplane between the different data classes by maximizing the distance between the hyperplane and the closest samples of the training set (the support vectors), keeping the classification error as low as possible (Xu et al., 2006). Only two parameters need to be tuned in SVM, including C (cost) and the kernel parameter γ in Gaussian kernel function. C is a tuning parameter, which weights in-sample classification errors and controls the generalization ability of an SVM. Moreover, within the different kernel functions, an appropriate γ parameter is related to a stable generalization performance. Furthermore, this method does not need a large number of samples to be trained, it is not affected by the presence of outliers and it has been successfully applied to solve a variety of practical classification problems (Acevedo et al., 2007; Liu et al., 2008; Xu et al., 2006).

As the results obtained in our study for the classification of wine vinegar categories within each PDO showed that SVM developed better classification models, only Support Vectors Machines (SVM) was used to build classification models for distinguishing the Spanish PDOs in each similar wine vinegar category (*"Crianza"*, *"Reserva"* and *"Pedro Ximenez"*). For both approaches, scores of each sample from PARAFAC models were used, 95% confidence intervals were considered for the classification models and vinegar samples were randomly divided into two groups.

The first group of samples (training set), comprising the 75% of samples, was used for calibration and internal validation of the models by means of a venetian blinds cross-validation procedure. For discrimination of the wine vinegar category within each Spanish PDO, this dataset (samples analysed in triplicate) was formed by 63 *("Vinagre de Jerez"*), 33 (*"Vinagre de Montilla-Moriles"*), and 39 (*"Vinagre de Condado de Huelva"*) samples. Meanwhile, this dataset consisted of 54 (*"Crianza"*), 48 (*"Reserva"*), and 18 (*"Pedro Ximenez"*) samples, in order to distinguish the Spanish PDO corresponding to each wine vinegar category.

The second group with the remaining samples (test set) was used as external independent dataset to evaluate the discriminative power of the models (external validation). This dataset was formed by 25% of the samples, and consisted of (samples analyzed in triplicate): 21 (*"Vinagre de Jerez")*, 15 (*"Vinagre de Montilla-Moriles")*, and 18 (*"Vinagre de Condado de Huelva"*) samples, in order to discriminate the wine

vinegar category within each Spanish PDO. For differentiating the Spanish PDO in each wine vinegar category, this dataset was formed by (samples analyzed in triplicate): 24 (*"Crianza"*), 18 (*"Reserva"*), and 9 (*"Pedro Ximenez"*) samples. Afterwards, the samples belonging to each dataset were randomly selected making sure that in both datasets at least one sample, with the corresponding replicates, of each category/PDO was included. Samples belonging to *"Gran Reserva"* and *"Añada"* category in each PDO (JGR, MGR and CAN) with a low number of samples (≤ 2) were not used. Further information about the number of samples used in each case for calibration and external validation can be found in Table II (Supplementary Material).

The statistical assessment of the quality of both classification models was carried out by means of the comparison of the sensitivity, specificity and classification error of calibration (CAL), cross-validation (CV) and prediction (PRED) parameters (Margraf, Santos, de Andrade, van Ruth, & Granato, 2016) according to Eqs. (5) and (6):

$$\text{Sensitivity (\%)} = [TP / (TP+FN)] *100\% \quad (5)$$

$$\text{Specificity (\%)} = [TN / (TN+FP)] *100\% \quad (6)$$

whereby TP and TN represent the number of samples correctly classified as their real class (e.g. the number of JCR samples predicted as JCR and the number of MCR samples predicted as MCR samples, respectively). On the other hand, FP and FN represent the number of samples misclassified (e.g. the JCR samples assigned to MCR class and MCR samples assigned to JCR class, respectively).

3. Results and discussion

3.1 Fluorescence landscapes of the Spanish PDO wine vinegars

Typical fluorescence landscapes of several samples belonging to the different wine vinegar categories of each Spanish PDO are shown in **Figure 1** (after removing and replacing the first and second order Rayleigh scattering). As it can be observed, the shape of the EEM spectra varies within the same Spanish PDO, which allows us to confirm *a priori* differentiation according to the wine vinegar category (aged or sweet).

A visual assessment of the fluorescence features of the aged categories points out a general tendency for the spectral maxima to be shifted towards longer excitation and emission wavelengths with the aging of the vinegars. Furthermore, similar fluorescence maxima were observed for the different Spanish PDOs wine vinegars according to the aging period. In general, vinegars with a minimum of 6 months of aging (JCR, MCR

and CSO) show their maximum peaks at 370/450 nm for both excitation/emission wavelengths ($\lambda_{ex}/\lambda_{em}$), whereas the maximum peaks corresponding to the *"Reserva"* category (JRE, MRE and CRE) appear at higher wavelengths, around 370-470 nm of λ_{ex} and 470-550 nm of λ_{em}. Finally, the most aged *"Gran Reserva"* category samples (JGR, MGR and CAN) show their maxima at 470-500 nm of excitation and 550-600 nm of emission wavelengths, following the general observed trends. The spectral features of the wine vinegars without aging period (CSC), which shows maximum peaks at the shortest wavelengths (around 370/440 nm $\lambda_{ex}/\lambda_{em}$), also confirm this tendency. Interestingly, some samples of the "*Reserva"* category (e.g. *"Vinagre de Montilla-Moriles"* PDO, *"Reserva"* sample *"MRE"* in Figure 1) show two different maximum peaks probably due to the broader aging period, which can vary from 24 to 120 months (*"Vinagre de Jerez"* PDO and *"Vinagre de Montilla-Moriles"* PDO) or even to longer periods (*"Vinagre de Condado de Huelva"* PDO) (Table 1). These observed spectral features are probably related to the different chemical complexity of the aged categories. In fact, a similar fluorescence trend with the aging of wine samples was observed by Airado-Rodriguez et.al (2011), whose fluorescence landscapes showed a tendency to increase the emission at longer wavelengths with the ageing of the wine samples, due an increase in concentration of fluorescence substances (Airado-Rodríguez et al., 2011).

In contrast, the fluorescence landscapes of the sweet vinegar categories, denoted as "*Pedro Ximenez*" (JPX and MPX) show a highly intense fluorescent area between 550-570 nm and 600-650 nm of excitation and emission wavelengths, respectively. Indeed, these sweet vinegars show their excitation and emission maxima even at longer wavelengths than the ones corresponding to the aged categories. This phenomenon could be explained by the different production and composition of the sweet vinegars in comparison with the aged categories. Thus, the sweet vinegars are produced with the addition of *"Pedro Ximenez"* Sherry wine in the case of "*Vinagre de Jerez*" PDO (containing at least 60 g/L of reducing material from this wine) (Council Regulation (EC) No 510/2006), or adding must of raisined *"Pedro Ximenez"* grapes during the maturing process for the "*Vinagre de Montilla-Moriles*" PDO (Council Regulation (EC) No 510/2006). These sweet vinegars have a high carbohydrate content (glucose and fructose) and other compounds (brown pigments and volatile compounds) produced by a Maillard reaction of the carbohydrates and free amino acids (Casale, Sáiz Abajo, González Sáiz, Pizarro, & Forina, 2006), which may be responsible for the observed fluorescence at longer excitation and emission wavelengths.

From these observations, it is clear that the fluorescence landscapes of these Spanish PDOs vinegars contain several fluorophores that are highly overlapped in both excitation and emission spectra. In this sense, further decomposition of EEM spectra by PARAFAC will help to clarify the potential fluorophores present in each vinegar category.

3.2 Potential fluorophores of the Spanish PDO wine vinegars

Three individual PARAFAC models were built in order to extract the excitation and emission profiles of the main fluorophores present in the Spanish PDOs vinegars (as described in section 2.3.1). The optimum number of factors for each PARAFAC model was selected comparing the quality parameters of the models built for an increasing number of factors (ranging from one to seven). Specifically, the best PARAFAC models obtained for each Spanish PDO were 5-factor PARAFAC models for the *"Vinagre de Jerez"* and "*Vinagre de Montilla-Moriles"* PDOs, and a 4-factor PARAFAC model for the *"Vinagre de Condado de Huelva"* PDO. The obtained models were enough robust, explaining more than 99% of the variance with a core consistency over zero (Table III Supplementary Material), and represented the underlying chemical spectra of the fluorophores present in these vinegars. Figure 2 includes the PARAFAC loadings (excitation and emission spectra) of the extracted factors present in each Spanish PDO vinegar, whose corresponding fluorescence emission and excitation maxima are listed in Table IV (Supplementary Material). The fluorescent loading patterns of the modelled factors can be matched to fluorophores described in the literature (Airado-Rodríguez et al., 2011; Dufour, Letort, Laguet, Lebecque, & Serra, 2006; Elcoroaristizabal et al., 2016). However, it is important to note that vinegar contains a wide variety of naturally occurring fluorescent compounds, being each emission/excitation profiles a sum of related fluorescent molecules and not only to a single one (Airado-Rodríguez et al., 2011). The difference in these modelled factors is probably related to the different chemical composition of these vinegars as a consequence of the different raw materials (wines) and production and ageing processes for each Spanish PDO. This is corroborated by the variation in the score values of the fluorophores modelled for each Spanish PDO according to the vinegar category (Table V Supplementary Material).

The first factor (F1, blue line in Figure 2) of the PDO *"Vinagre de Jerez"* has a maximum excitation centered at 465 nm and a maximum emission around 535 nm. According to Airado-Rodríguez et al. (2011), this factor could be related to vitamin B2 and its principal forms such as riboflavin, Flavin mononucleotide (FMN), and Flavin adenine dinucleotide (FAD). In contrast, F1 appears at lower wavelengths, specifically

365 at 375/460 nm and 370/470 nm (λ_{ex}/ λ_{em}), for the *"Vinagre de Montilla-Moriles"* and
366 *"Vinagre de Condado de Huelva"* PDOs, respectively. In these PDOs, and taking into
367 account these wavelengths, F1 could be due to the presence of coumarins, tannins
368 and other unknown fluorescent compounds originating from wooden casks (Tóthová &
369 Sádecká, 2008), as well as phenols and flavonols, usually in abundance in these
370 vinegars and naturally presented in wines (Airado-Rodríguez et al., 2011; Sádecká &
371 Tóthová, 2007).

372 The second factor (F2, red line in Figure 2) has a similar profile for the three PDOs
373 with an excitation and emission maxima between 400-420 nm and 480-505 nm
374 respectively. 5-Hydroxymethylfurfural (HMF), which has been determined in vinegars
375 as a product being formed during the Maillard reaction (García Parrilla, Heredia, &
376 Troncoso, 1999), could match with the wavelengths of F2 according to Zhu et al., 2009
377 (Zhu, Ji, Eum, & Zude, 2009). Furthermore, caramel, which is frequently added to
378 vinegars as a colorant, showed a maximum excitation/emission wavelength at 390-
379 410/482-498 nm according to Sádecká et al., 2009 (Sádecká, Tóthová, & Májek, 2009),
380 and Tóthová et al., 2008 (Tóthová & Sádecká, 2008). Its presence in these vinegars
381 could be also related to this second factor.

382 The third factor (F3, yellow line in Figure 2) of *"Vinagre de Jerez"* PDO is a peak
383 centered at 500 nm (λ_{ex}) and 580 nm (λ_{em}). This component could be associated to
384 brown pigments produced by some acetic bacteria strains present in vinegar (Polo &
385 Sanchez-Luengo, 1991) since they showed similar excitation/emission wavelengths.
386 However, in *"Vinagre de Montilla-Moriles"* PDO, this factor F3 shows excitation and
387 emission maxima around 470/550 nm, which agrees with the presence of vitamin B2
388 and its principal forms (Lenhardt et al., 2015). Finally, the third factor of *"Condado de*
389 *Huelva"* PDO is centered at 300/425 nm (λ_{ex}/ λ_{em}), and these wavelengths could be
390 associated with the phenolic compounds present in these vinegars (Rodríguez-
391 Delgado, Malovaná, Pérez, Borges, & García Montelongo, 2001).

392 The fourth factor (F4, purple line in Figure 2) has 340/420 nm and 350/440 nm of
393 excitation and emission maxima for the *"Vinagre de Montilla-Moriles*" and "*Vinagre de*
394 *Jerez"* PDOs. According to the literature, the excitation/emission wavelengths of this
395 factor could be related to phenolic compounds, the best known fluorescent molecules
396 naturally present in wine that differ in accordance to the grape variety and the vinegar
397 ageing. This group of compounds includes phenolic acids and phenolic aldehydes, as
398 well as oxidation and Maillard reaction products (present due to browning processes
399 and oxidative mechanisms taking place during ageing and storage), which have shown

maximum excitation/emission wavelengths around 330/420 nm (Airado-Rodríguez et al., 2011; Azcarate et al., 2015; Callejón et al., 2012; Dufour et al., 2006; Elcoroaristizabal et al., 2016; Sádecká & Tóthová, 2007). In contrast, in *"Condado de Huelva"* PDO this F4 presents its excitation and emission maxima around 485/560 nm (λ_{ex}/ λ_{em}), and it could be associated with the aforementioned vitamin B2.

Finally, a fifth factor (F5, green line in Figure 2) appears only for the *"Vinagre de Montilla-Moriles*" and *"Vinagre de Jerez"* PDOs, showing excitation and emission maxima at 530/605 nm and 585/655 nm (λ_{ex}/ λ_{em}), respectively. There are not reported fluorophores matching exactly with this emission/excitation profile. However, it seems to be related to the special characteristics of the category *"Pedro Ximenez"* for which a higher mean values of this factor was detected in this category (Table V Supplementary Material). The absence of this factor in the *"Vinagre Condado de Huelva"* model also confirms this hypothesis since this sweet category is not registered in this PDO.

3.3 Vinegar category classification within each Spanish PDO

Two different approaches, Partial Least Squares Discriminant Analysis (PLS-DA) and Support Vector Machines (SVM), were used for the development of classification models of Spanish PDO vinegars according to their category. In all cases, the best PLS-DA models were obtained using two latent variables (LVs). This optimum number of LVs was chosen based on the Root Mean Square Error of Cross-Validation (RMSECV), the ROC curves and de variance captured. On the other hand, the optimal parameters for the optimization of SVM models, $\log_{10}(C)$ and $\log_{10}(\gamma)$, were found to be 2 and between -2 and -0.5, respectively. The statistical assessment of the performance of both classification models was carried out by calculating and comparing different classifiers (described in section 2.3.2) such as sensitivity, specificity and classification error of calibration (CAL), cross-validation (CV) and prediction (PRED). These statistical results are shown in Table 2.

Regarding the PLS-DA models, high sensitivity and specificity values were obtained for the sweet category (JPX and MPX) of *"Vinagre de Jerez"* and *"Vinagre de Montilla-Moriles"* PDOs. The classification errors of prediction obtained for these categories (100% of the samples correctly classified), demonstrated that these samples can be successfully separated from the rest of classes. This is probably due to their different chemical and fluorescence spectral features, since this sweet category emitted at the longest wavelengths (Figure 1). In a similar way, concerning the non-

aged category (CSC) of *"Vinagre de Condado de Huelva"* PDO, the different fluorescence profile observed in the landscapes and the higher mean values of F1 and F3 with respect to the rest of categories (Table V Supplementary Material), could be related to the lowest errors of classification obtained for this category (25.0% of the samples misclassified). However, unsatisfactory results were obtained for the rest of aged categories within each Spanish PDO: JCR, JRE, MCR, CSO and CRE. The low sensitivity and specificity values (mainly under 65.0%) and high classification errors obtained in terms of prediction (between 25.0 and 63.0%), confirm the difficulty in correctly classifying these aged vinegar categories (≥ 6 and ≥ 24 months) by using linear classification models. This could be related to the similar score values followed by the modelled factors of these categories within each Spanish PDOs (Table V Supplementary Material). Among them, the ≥ 6 months aged samples (JCR, MCR and CSO) were the worst classified ones in all the Spanish PDO vinegar models, showing classification errors until 63.0%. This may be explained due to these wine vinegars are aged over a wide range of time (from 6 to 24 months). Thus, those samples aged until 24 months are expected to be spectroscopically and chemically quite similar to the vinegars of the following category (≥ 24 month). Similar results were obtained by Callejón et al., (2012).

In contrast, higher sensitivity and specificity levels were obtained for all Spanish PDO vinegars using SVM models (Table 2). The optimal parameters for the optimization of SVM models ,$\log_{10}(C)$ and $\log_{10}(\gamma)$, were optimized in the traditional way by using an independent test set (Cristianini & Shawe-Taylor, 2000). Between 92% and 100% of the samples were correctly classified in all categories. Even more, all samples belonging to the Spanish PDO *"Vinagre de Montilla-Moriles"* were perfectly classified. Further information about the misclassified category samples within each Spanish PDO is summarized in the confusion matrices shown in Table VI (Supplementary Material). These results also point out that SVM does not need a large number of samples to make a good model, as occurs in our study with some categories, and further, it is not affected by the presence of outliers. These results demonstrated that this methodology could be successfully used for the authentication of the vinegar category belonging to each Spanish PDO.

3.4 Spanish PDO classification within similar vinegar categories (*"Crianza"*, *"Reserva"* and *"Pedro Ximenez"*)

In this classification task, the objective was to classify the samples by their PDO for a single category. Three PARAFAC models were built according to the vinegar categories under study, i.e. *"Crianza"*, *"Reserva"* and *"Pedro Ximenez"*, in order to discriminate their corresponding Spanish PDO. In a similar way to the previous sections, the best PARAFAC models obtained for each vinegar category were selected comparing the quality parameters of the models, which are shown in Table VII (Supplementary Material). In this case, a 4-factor PARAFAC model was obtained for "*Crianza*", while a 5-factor PARAFAC model was built for "*Reserva*", and a 3-factor PARAFAC model was constructed for "*Pedro Ximenez*" category. The obtained models explained more than 97.0% of the variance with a core consistency over zero. The related PARAFAC loadings (excitation and emission spectra) obtained for the models corresponding to each vinegar category are illustrated in Figure 3.

The maxima wavelengths (λ_{ex} and λ_{em}) of the different factors obtained for the "*Crianza*", "*Reserva*" and "*Pedro Ximenez*" PARAFAC models match with the different fluorophores described in detail in section 3.2. As shown in **Figure 3**, the *"Crianza"* category is characterized by factors with excitation and emission ranges between 340-500 nm and 430-580 nm, respectively. These factors are mainly related to fluorescent compounds naturally presented in high concentration in wine such as phenols, flavonols and vitamins as previously described (section 3.2). These compounds have a higher contribution in this category (Table V Supplementary Material) since the less aged wine vinegars retain more compounds coming from the raw materials. Regarding the *"Reserva"* samples, a higher number of factors were required to model this category, i.e., more fluorescent compounds with a wide range of excitation/emission spectra were needed to describe its underlying chemical composition. In this case, the longer aging period undergone by vinegars of the "*Reserva"* category plays a crucial role in this chemical complexity. There is an enrichment of these vinegars with more phenolic compounds (released by the wood barrels) and oxidation products (derived from the development of certain chemical reactions among vinegar components), whose concentration levels have been proven to increase during the aging process (Callejón, Morales, Silva Ferreira, & Troncoso, 2008). The wavelengths of the factors modelled for "*Pedro Ximenez*" category are associated to fluorophores emitting at the highest excitation and emission wavelengths (upper than 475 and 550 nm, respectively). However, more information, not available in the literature, is needed to identify these fluorophores.

For all the categories under study, the relative values of these factors (scores) vary as a function of the Spanish PDO, which highlights that the composition of the vinegar

categories depends also on the raw material used (wine) and on the different production methods to which the vinegars have been subjected in each PDO. Thus, these particular characteristics reflected by the scores provide the chance to discriminate the Spanish PDO corresponding to similar wine vinegar categories. In this case, related to the proven higher ability of prediction previously obtained (Table 2), only SVM classification models were built. The parameters for the optimization of the SVM models, $\log_{10}(C)$ and $\log_{10}(\gamma)$, were found to be 2 and between -1 and 0, respectively. Table 3 summarizes the statistical results (sensitivity, specificity and errors) of the performance of the SVM models**.**

Regarding the "*Crianza*" category, high sensitivity and specificity values were obtained for these samples according to their origin (Spanish PDO) with classification errors of prediction lower than 3.5% (Table 3). These results demonstrate that it is possible to successful differentiate the Spanish PDO of *"Crianza"* vinegars according to their fluorescent composition that is highly related to the raw material (wine) used. Furthermore, all the samples belonging to the "*Pedro Ximenez*" category were correctly classified in the "*Vinagre de Jerez*" and "*Vinagre de Montilla-Moriles*" PDOs. The high levels of sensitivity and specificity and the good classification rates obtained were explained by the different production process employed by each Spanish PDO. "*Reserva*" was the worst classified category according to the PDOs, showing sensitivity and specificity values higher than 70% and predicted errors lower than 15%. In the case of samples belonging to *"Vinagre de Montilla-Moriles*", 100% of the *"Reserva"* vinegars were correctly classified, and only some samples were misclassified between *"Vinagre de Jerez"* and *"Vinagre de Condado de Huelva"* PDOs (Table VIII Supplementary Material). These results are considered acceptable considering the high variability of these samples due to the wide range of ageing periods that are reflected by their complex chemical composition (Casale et al., 2006; García Parrilla et al., 1999).

4. Conclusions

The analytical methodology proposed in this study, namely fluorescence excitation–emission spectroscopy coupled to PARAFAC modeling and SVM classification method, has demonstrated to be able to characterize and classify the three Spanish PDOs wine vinegars according to their "Protected Denomination of Origin" as well as their categories (aged and sweet). As a simple preliminary characterization, a visual assessment of the fluorescence Excitation-Emission Matrices (EEMs) of the aged categories pointed out similarities in the fluorescence landscapes

for the three Spanish PDOs wine vinegars: the spectral maxima were shifted towards longer wavelengths with the aging of these vinegars. Moreover, the sweet category *"Pedro Ximenez"* showed its excitation and emission maxima even at longer wavelengths than the aged categories, probably due to the different production process to which these vinegars are subjected. PARAFAC was carried out to spectroscopically and chemically characterize the different wine vinegars. It gave information about the potential fluorescent compounds present in the wine vinegars as well as their contribution in each Spanish PDO and category. These dissimilar spectroscopic and chemical features allowed us their differentiation according to their category and origin (Spanish PDO) using suitable classification methods. The feasibility of SVM methodology to classify the different categories of wine vinegars within each PDO was successfully demonstrated. The built SVM classification models proved a higher ability of prediction (between 92% and 100% correctly classified samples) than PLS-DA models, especially for classifying aged vinegar categories with similar spectroscopic characteristics. Furthermore, SVM models were also able to differentiate the Spanish PDOs even for similar vinegar categories due to their spectral differences.

The advantages of this methodology, e.g. fast, non-destructive and non- sample preparation, would allow implementing this method as an alternative tool for PDO regulatory councils and producers to be implemented in routine analysis. It could be applied for assessing the authenticity of the Spanish PDO and the vinegar category. Finally, it is expected that further information about the specific ageing periods to which these vinegars are subjected will improve the performance of some classification models.

5. Acknowledgments

The authors want to thank "Consejería de Economía, Innovación y Ciencia" of the "Junta de Andalucía" for the financial support provided through the project P12-AGR-1601, and the Spanish Regulatory Councils of the wine vinegar PDOs for their invaluable help with the acquisition of the samples for this study. In addition, authors want to thanks "Ministerio de Educación, Cultura y Deporte" (MECD, Spanish Government) for the FPU pre-doctoral fellowship and UPV/EHU (University of the Basque Country) for the grant received from the program *''Contratación de Doctores Recientes"*.

6. References

Aceña, L., Vera, L., Guasch, J., Busto, O., & Mestres, M. (2011). Chemical

577 characterization of commercial sherry vinegar aroma by headspace solid-phase
578 microextraction and gas chromatography-olfactometry. *Journal of Agricultural and*
579 *Food Chemistry*, *59*(8), 4062–4070. http://doi.org/10.1021/jf104763u

580 Acevedo, F. J., Jiménez, J., Maldonado, S., Domínguez, E., & Narváez, A. (2007).
581 Classification of wines produced in specific regions by UV-visible spectroscopy
582 combined with support vector machines. *Journal of Agricultural and Food*
583 *Chemistry*, *55*(17), 6842–6849. http://doi.org/10.1021/jf070634q

584 Airado-Rodríguez, D., Durán-Merás, I., Galeano-Díaz, T., & Wold, J. P. (2011). Front-
585 face fluorescence spectroscopy: A new tool for control in the wine industry.
586 *Journal of Food Composition and Analysis*, *24*(2), 257–264.
587 http://doi.org/10.1016/j.jfca.2010.10.005

588 Airado-Rodríguez, D., Galeano-Ḋiaz, T., Durán-Merás, I., & Wold, J. P. (2009).
589 Usefulness of fluorescence excitation-emission matrices in combination with
590 parafac, as fingerprints of red wines. *Journal of Agricultural and Food Chemistry*,
591 *57*(5), 1711–1720. http://doi.org/10.1021/jf8033623

592 Andersen, C. M., & Bro, R. (2003). Practical aspects of PARAFAC modeling of
593 fluorescence excitation-emission data. *Journal of Chemometrics*, *17*(4), 200–215.
594 http://doi.org/10.1002/cem.790

595 Azcarate, S. M., Araújo, A. De, Alcaraz, M. R., Ugulino, M. C., Araújo, D., Camiña, J.
596 M., & Goicoechea, H. C. (2015). Modeling excitation- emission fluorescence
597 matrices with pattern recognition algorithms for classification of Argentine white
598 wines according, *184*, 214–219.

599 Azcarate, S. M., Cantarelli, M. A., Pellerano, R. G., Marchevsky, E. J., & Camiña, J. M.
600 (2013). Classification of argentinean sauvignon blanc wines by UV spectroscopy
601 and chemometric methods. *Journal of Food Science*, *78*(3), 432–436.
602 http://doi.org/10.1111/1750-3841.12060

603 Bro, R. (1997). PARAFAC. Tutorial and applications. *Chemometrics and Intelligent*
604 *Laboratory Systems*, *38*(2), 149–171. http://doi.org/10.1016/S0169-
605 7439(97)00032-4

606 Bro, R. (1998). Multi-way analysis in the food industry. Models, algorithms, and
607 applications., (1998).

608 Bro, R., & Kiers, H. A. L. (2003). A new efficient method for determining the number of
609 components in PARAFAC models. *Journal of Chemometrics*, *17*(5), 274–286.

http://doi.org/10.1002/cem.801

Callejón, R. M., Amigo, J. M., Pairo, E., Garmón, S., Ocaña, J. A., & Morales, M. L. (2012). Classification of Sherry vinegars by combining multidimensional fluorescence, parafac and different classification approaches. *Talanta*, *88*, 456–462. http://doi.org/10.1016/j.talanta.2011.11.014

Callejón, R. M., Morales, M. L., Silva Ferreira, A. C., & Troncoso, A. M. (2008). Defining the typical aroma of Sherry vinegar: Sensory and chemical approach. *Journal of Agricultural and Food Chemistry*, *56*(17), 8086–8095. http://doi.org/10.1021/jf800903n

Casale, M., Sáiz Abajo, M. J., González Sáiz, J. M., Pizarro, C., & Forina, M. (2006). Study of the aging and oxidation processes of vinegar samples from different origins during storage by near-infrared spectroscopy. *Analytica Chimica Acta*, *557*(1–2), 360–366. http://doi.org/10.1016/j.aca.2005.10.063

Christensen, J., Becker, E. M., & Frederiksen, C. S. (2005). Fluorescence spectroscopy and PARAFAC in the analysis of yogurt. *Chemometrics and Intelligent Laboratory Systems*, *75*(2), 201–208. http://doi.org/10.1016/j.chemolab.2004.07.007

Christiani, N., & Shawe-Taylor, J. (2000). An Introduction to Support Vector Machines. *Cambridge Univ. Press, Cambridge, UK.*

Cirlini, M., Caligiani, a., Palla, L., & Palla, G. (2011). HS-SPME/GC-MS and chemometrics for the classification of Balsamic Vinegars of Modena of different maturation and ageing. *Food Chemistry*, *124*(4), 1678–1683. http://doi.org/10.1016/j.foodchem.2010.07.065

Council Regulation (EC) No 510/2006 of 20 March 2006 on the protection of geographical indications and designations of origin for agricultural products and foodstuffs.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, *13*, 21 – 27.

Dufour, É., Letort, A., Laguet, A., Lebecque, A., & Serra, J. N. (2006). Investigation of variety, typicality and vintage of French and German wines using front-face fluorescence spectroscopy. *Analytica Chimica Acta*, *563*(1–2 SPEC. ISS.), 292–299. http://doi.org/10.1016/j.aca.2005.11.005

Elcoroaristizabal, S., Bro, R., García, J. A., & Alonso, L. (2015). PARAFAC models of

643 fluorescence data with scattering: A comparative study. *Chemometrics and*
644 *Intelligent Laboratory Systems*, *142*, 124–130.
645 http://doi.org/10.1016/j.chemolab.2015.01.017

646 Elcoroaristizabal, S., Callejón, R. M., Amigo, J. M., Ocaña-González, J. A., Morales, M.
647 L., & Ubeda, C. (2016). Fluorescence excitation-emission matrix spectroscopy as
648 a tool for determining quality of sparkling wines. *Food Chemistry*, *206*, 284–290.
649 http://doi.org/10.1016/j.foodchem.2016.03.037

650 García Parrilla, M. C., Heredia, F. J., & Troncoso, A. M. (1999). Sherry wine vinegars:
651 Phenolic composition changes during aging. *Food Research International*, *32*(6),
652 433–440. http://doi.org/10.1016/S0963-9969(99)00105-2

653 Joint FAO/WHO Food Standards Programme.Report No: FAO-ESN-CX/PFV-98/8.
654 Codex Committee on Processed Fruits and Vegetables. Sess 19. Whashington,
655 DC, 16-18, March 1998. FAO, R. Conversion of European Regional Standard for
656 Vinegar into World-wide Standard. (1998).

657 Karoui, R., & Blecker, C. (2011). Fluorescence Spectroscopy Measurement for Quality
658 Assessment of Food Systems-a Review. *Food and Bioprocess Technology*, *4*(3),
659 364–386. http://doi.org/10.1007/s11947-010-0370-0

660 Karoui, R., Dufour, É., & De Baerdemaeker, J. (2006). Front face fluorescence
661 spectroscopy coupled with chemometric tools for monitoring the oxidation of semi-
662 hard cheeses throughout ripening. *Food Chemistry*, *101*(3), 1305–1314.
663 http://doi.org/10.1016/j.foodchem.2006.01.028

664 Lenhardt, L., Bro, R., Zeković, I., Dramićanin, T., & Dramićanin, M. D. (2015).
665 Fluorescence spectroscopy coupled with PARAFAC and PLS DA for
666 characterization and classification of honey. *Food Chemistry*, *175*, 284–291.
667 http://doi.org/10.1016/j.foodchem.2014.11.162

668 Liu, F., He, Y., & Wang, L. (2008). Determination of effective wavelengths for
669 discrimination of fruit vinegars using near infrared spectroscopy and multivariate
670 analysis. *Analytica Chimica Acta*, *615*(1), 10–17.
671 http://doi.org/10.1016/j.aca.2008.03.030

672 Margraf, T., Santos, É. N. T., de Andrade, E. F., van Ruth, S. M., & Granato, D. (2016).
673 Effects of geographical origin, variety and farming system on the chemical
674 markers and in vitro antioxidant capacity of Brazilian purple grape juices. Food
675 Research International, 82, 145–155. http://doi.org/10.1016/j.foodres.2016.02.003

Morales, M. L., Tesfaye, W., García-Parrilla, M. C., Casas, J. a, & Troncoso, A. M. (2002). Evolution of the aroma profile of sherry wine vinegars during an experimental aging in wood. *Journal of Agricultural and Food Chemistry*, *50*(11), 3173–8. http://doi.org/10.1021/jf011313w

Nocairi, H., Qannari, E. M., Vigneau, E., & Bertrand, D. (2005). Discrimination on latent components with respect to patterns. Application to multicollinear data. *Computational Statistics and Data Analysis*, *48*(1), 139–147. http://doi.org/10.1016/j.csda.2003.09.008

Polo M.C., & Sanchez-Luengo, A. (1991). El vinagre de vino. In *in: C. Llaguno, M.C. Polo (Eds.), El vinagre de vino, C.S.I.C. Espana, Madrid* (pp. 23–47).

Rodríguez-Delgado, M. A., Malovaná, S., Pérez, J. P., Borges, T., & García Montelongo, F. J. (2001). Separation of phenolic compounds by high-performance liquid chromatography with absorbance and fluorimetric detection. *Journal of Chromatography A*, *912*(2), 249–257. http://doi.org/10.1016/S0021-9673(01)00598-2

Sádecká, J., & Tóthová, J. (2007). Fluorescence Spectroscopy and Chemometrics in the Food Classification− a Review. *Czech J. Food Sci. Vol*, *25*(4), 159–173. Retrieved from http://www.journals.uzpi.cz/publicFiles/00302.pdf

Sádecká, J., Tóthová, J., & Májek, P. (2009). Classification of brandies and wine distillates using front face fluorescence spectroscopy. *Food Chemistry*, *117*(3), 491–498. http://doi.org/10.1016/j.foodchem.2009.04.053

Tena, N., Aparicio, R., & García-González, D. L. (2012). Chemical changes of thermoxidized virgin olive oil determined by excitation-emission fluorescence spectroscopy (EEFS). *Food Research International*, *45*(1), 103–108. http://doi.org/10.1016/j.foodres.2011.10.015

Tóthová, J., & Sádecká, Ľ. Ž. J. (2008). Characterization and Classification of Distilled Drinks Using Total Luminescence and Synchronous Fluorescence Spectroscopy. *Acta Chimica Slovaca*, *1*(1), 265–275.

Vapnik, V. (1999). The Nature of Statistical Learning Theory . In *New York : Springer-Verlag*.

Wold, H. (1966). Estimation of principal components and related models by iterative least squares . In *In: Multivariate Analysis (Krishnaiah PR , ed.). New York : Academic Press .*

Wold, S. (1976). Pattern recognition by means of disjoint principal components models . *Pattern Recognition*, *8*, 127 – 139.

Xu, Y., Zomer, S., & Brereton, R. G. (2006). Support Vector Machines: A Recent Method for Classification in Chemometrics. *Critical Reviews in Analytical Chemistry*, *36*(3–4), 177–188. http://doi.org/10.1080/10408340600969486

Zhu, D., Ji, B., Eum, H. L., & Zude, M. (2009). Evaluation of the non-enzymatic browning in thermally processed apple juice by front-face fluorescence spectroscopy. *Food Chemistry*, *113*(1), 272–279. http://doi.org/10.1016/j.foodchem.2008.07.009

FIGURE CAPTIONS

Figure 1. *Excitation-Emission fluorescence landscapes obtained for different categories of Spanish PDO vinegars. The color scale of fluorescence intensity (in arbitrary units) varies from dark blue (lowest signal intensity) to yellow (highest signal intensity).The acronyms for the different wine vinegar categories are defined in Table 1.*

Figure 2. *Excitation and Emission spectra (PARAFAC loadings) of the main fluorophores present in the vinegars of the three Spanish PDOs.*

Figure 3. *Excitation and Emission spectra (PARAFAC loadings) of the main fluorophores present in different vinegars categories of the Spanish PDOs.*

731 **Table 1**. Wine vinegar samples analyzed according to the Spanish PDOs

Protected Designation of Origin (PDO)	***Vinegar Category***	***Category Name***	***Acronym***	***Aging time (months)***	***Number of samples****
"Vinagre de Jerez"	Aged	"Vinagre de Jerez"	JCR	≥6	13
		"Reserva"	JRE	≥24	11
		"Gran Reserva"	JGR	≥ 120	2
	Sweet	"Pedro Ximenez"	JPX	-	4
"Vinagre de Montilla-Moriles"	Aged	"Crianza"	MCR	≥6	7
		"Reserva"	MRE	≥24	4
		"Gran Reserva"	MGR	≥120	2
	Sweet	"Pedro Ximenez"	MPX	-	5
"Vinagre de Condado de Huelva"	Non – aged	"Vinagre Condado de Huelva"	CSC	0	6
	Aged	"Viejo Solera"	CSO	≥6	6
		"Viejo Reserva"	CRE	≥24	7
		"Viejo Añada"	CAN	≥36	2

732

733 *Note: each sample- corresponds to different producers

734

735 **Table 2.** Sensitivity, specificity and classification errors (%) obtained for SVM and PLS-
736 DA classification models corresponding to the vinegar category of each Spanish PDO.

737

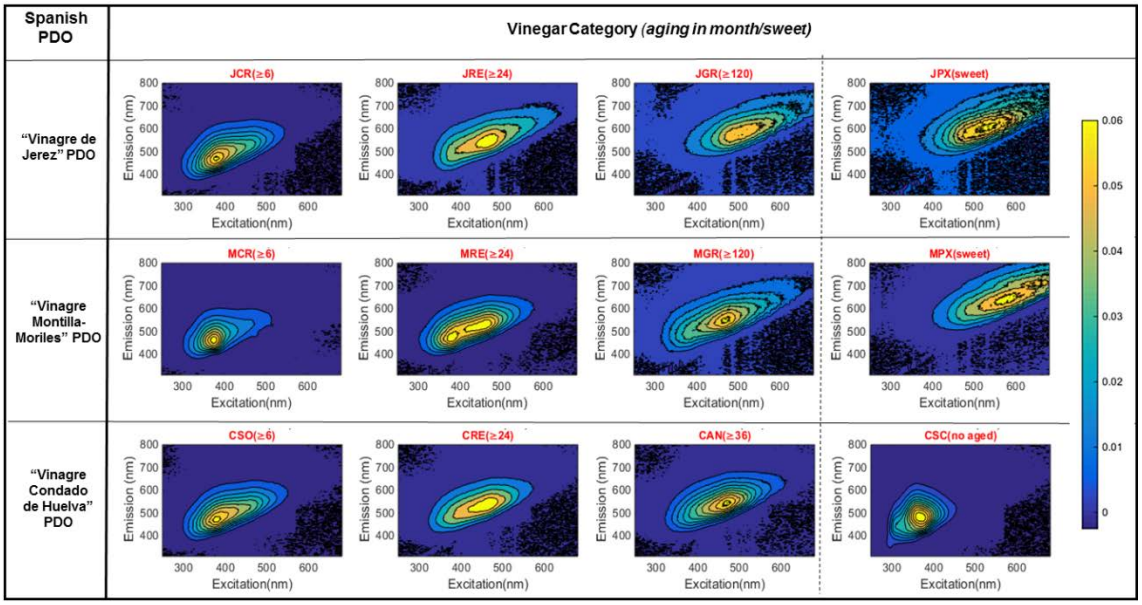
Spanish PDO	**"*Vinagre de Jerez*"**						**"*Vinagre de Montilla-Moriles*"**						**"*Vinagre de Condado de Huelva*"**					
Classification model	SVM			*PLS-DA*			SVM			*PLS-DA*			SVM			*PLS-DA*		
Category	JCR	JRE	JPX	JCR	JRE	JPX	MCR	MRE	MPX	MCR	MRE	MPX	CSC	CSO	CRE	CSC	CSO	CRE
Sensitivity CAL	92.9	92.3	100.0	*38.7*	*100.0*	*100.0*	100.0	100.0	100.0	*80.0*	*100.0*	*100.0*	83.3	100.0	100.0	*75.0*	*100.0*	*100.0*
Sensitivity CV	100.0	100.0	100.0	*38.7*	*100.0*	*100.0*	100.0	100.0	100.0	*80.0*	*66.7*	*100.0*	100.0	100.0	100.0	*75.0*	*100.0*	*100.0*
Sensitivity PRED	92.9	92.3	100.0	*10.0*	*66.7*	*100.0*	100.0	100.0	100.0	*50.0*	*100.0*	*100.0*	83.3	100.0	100.0	*50.0*	*100.0*	*100.0*
Specificity CAL	100.0	94.4	96.3	*90.9*	*32.5*	*94.5*	100.0	100.0	100.0	*83.3*	*75.0*	*100.0*	100.0	92.3	100.0	*100.0*	*33.3*	*62.5*
Specificity CV	100.0	100.0	100.0	*90.9*	*25.0*	*94.5*	100.0	100.0	100.0	*83.3*	*62.5*	*95.8*	100.0	100.0	100.0	*100.0*	*25.9*	*62.5*
Specificity PRED	100.0	94.4	96.3	*75.0*	*18.2*	*100.0*	100.0	100.0	100.0	*100.0*	*100.0*	*100.0*	100.0	91.7	100.0	*100.0*	*0.0*	*25.0*
Class. Error CAL	0.0	0.0	0.0	*35.1*	*33.7*	*2.7*	0.0	0.0	0.0	*18.3*	*12.5*	*0.0*	0.0	0.0	0.0	*12.5*	*33.3*	*18.7*
Class. Error CV	0.0	0.0	0.0	*35.1*	*37.5*	*2.7*	0.0	0.0	0.0	*18.3*	*35.4*	*2.1*	0.0	0.0	0.0	*12.5*	*37.0*	*18.7*
Class. Error PRED	3.5	6.6	1.8	*62.5*	*57.5*	*0.0*	0.0	0.0	0.0	*25.0*	*0.0*	*0.0*	8.3	4.1	0.0	*25.0*	*50.0*	*37.5*

738

739

Table 3. Sensitivity, specificity and classification errors (%) obtained for Spanish PDO classification models in similar wine vinegar categories.

Category	**CR**			**RE**			PX	
Classification model	SVM			SVM			SVM	
Spanish PDO	JCR	MCR	CSO	JRE	MRE	CRE	JPX	MPX
Sensitivity CAL	92.9	100.0	100.0	100.0	100.0	71.4	100.0	100.0
Sensitivity CV	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Sensitivity PRED	92.9	100.0	100.0	100.0	100.0	71.4	100.0	100.0
Specificity CAL	100.0	95.0	100.0	81.8	100.0	100.0	100.0	100.0
Specificity CV	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Specificity PRED	1.00.0	95.0	100.0	1.00.0	1.00.0	0.71.4	1.00.0	1.00.0
Class. Error CAL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Class. Error CV	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Class. Error PRED	3.5	2.5	0.0	9.1	0.0	14.3	0.0	0.0


Spanish PDO
Vinegar Category (aging in month/sweet)
"Vinagre de Jerez" PDO
"Vinagre Montilla-Moriles" PDO
"Vinagre Condado de Huelva" PDO
JCR(≥6)
JRE(≥24)
JGR(≥120)
JPX(sweet)
MCR(≥6)
MRE(≥24)
MGR(≥120)
MPX(sweet)
CSO(≥6)
CRE(≥24)
CAN(≥36)
CSC(no aged)
Emission (nm)
Excitation(nm)
0.06
0.05
0.04
0.03
0.02
0.01
0

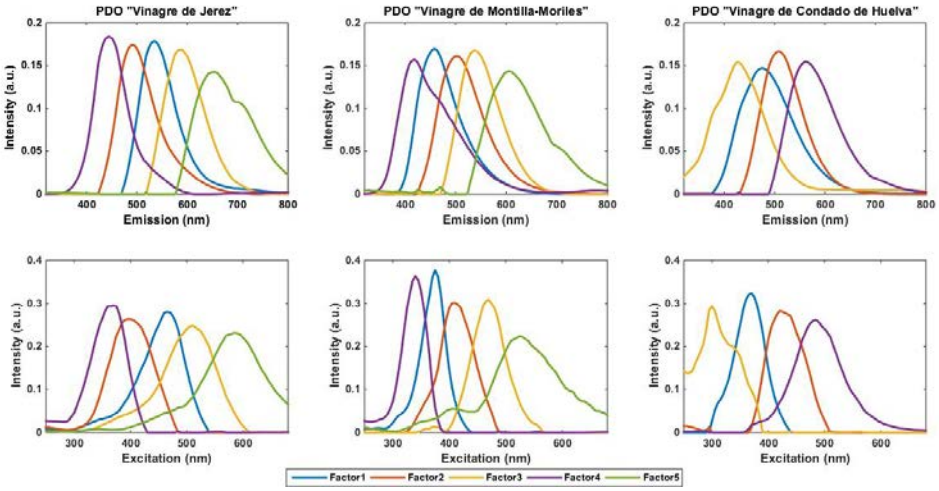
764

765

766

767

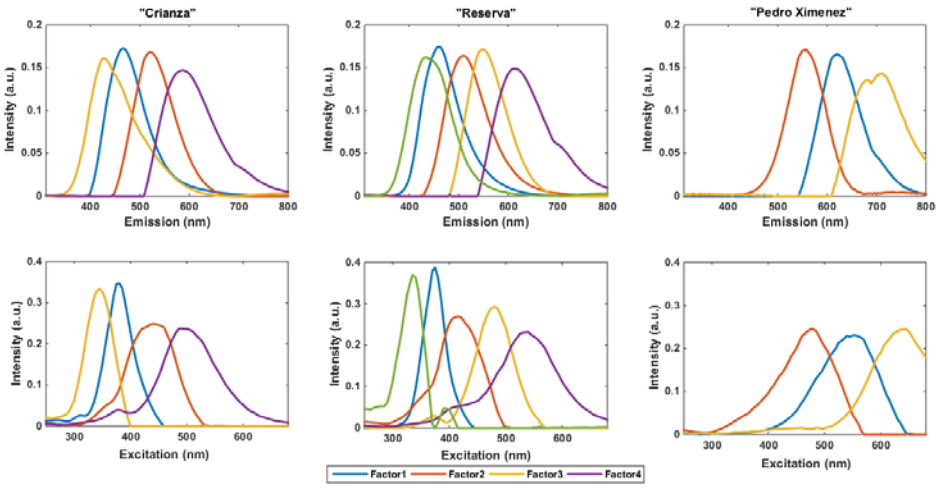
768



769

770

771


"Crianza"
"Reserva"
"Pedro Ximenez"
Intensity (a.u.)
Emission (nm)
Excitation (nm)
Factor1
Factor2
Factor3
Factor4

772