European Food Research and Technology
https://doi.org/10.1007/s00217-020-03568-y

ORIGINAL PAPER

Physico-chemical parameters complemented with aroma compounds fired up the varietal discrimination of wine using statistics

Ioannis K. Karabagias[1] · Dionysia Sykalia[1] · Alberto Mannu[2] · Anastasia V. Badeka[1]

Received: 16 May 2020 / Revised: 29 June 2020 / Accepted: 4 July 2020

Abstract
Wine comprises a beloved food and human companion since the early times of humans on earth. In this study, wine samples of different type (red, white, and rosé) and variety (Agiorgitiko, Augoustiatis, Cabernet Sauvignon, Syrah, Vlahiko, Assyrtiko, Chardonnay, Debina, Moschofilero, Vidiano, Syrah plus Mandilari, and Xinomavro) were subjected to physico-chemical and aroma compounds analyses, in an effort to characterize their identity and discriminate these samples according to variety using statistics. Results showed significant differences ($p < 0.05$) for wine samples of different variety in regard to the measured physico-chemical parameters (pH, electrical conductivity, total dissolved solids, salinity, L^*, a^*, b^*, and *Chroma**) and aroma compounds (alcohols, esters, phenolic compounds, pyran compounds, and terpenoids/norisoprenoids). Application of multivariate analysis of variance, linear discriminant analysis, and weighted least-squares regression analysis fired up the perfect varietal discrimination (~ 100%) of wine samples and modeling of results, contributing to new information in the literature about the identity of these wine varieties.

Keywords Wine · Physico-chemical indices · Aroma compounds · Varietal discrimination · Modeling

Introduction

Wine is the alcoholic drink traditionally made from fermented grapes. Different varieties of grapes and strains of yeasts produce different types of wine. These variations result from the complex interactions between the biochemical development of the grape, the reactions involved in fermentation, the grape's growing environment (terroir), and the production process. Wine has been produced for thousands of years and played an important role in the culture and religion of many civilizations [1].

Terroir is a concept that encompasses the varieties of grapes used, elevation and shape of the vineyard, type and chemistry of soil, climate, and seasonal conditions, along with the local yeast cultures [2]. The range of possible combinations of these factors can result in considerable differences among wines, influencing among others, the fermentation, finishing, and aging process. Many wineries use growing and production methods that preserve or accentuate the aroma and taste influences of their unique terroir. Aroma is one of the most important intrinsic factors that influence perceived wine quality and consumer acceptance [3].

In the literature, there are a considerable number of studies concerning the determination of wine quality and authenticity based on physico-chemical [4–6], phenolic composition [6], and aroma compounds analyses [6–11]. The most widespread used analytical technique for the characterization of the aroma profile of wine is gas chromatography/mass spectrometry, combined with new methodologies such as headspace solid-phase micro-extraction, and gas chromatography–olfactrometry [7, 11]. The aroma compounds of wine consist of (1) the primary aroma or aroma arising directly from the grapes and modifications during grape processing, (2) the secondary aroma or aroma produced by fermentation, or (3) the tertiary aroma or bouquet, which results from the transformation of the aroma during aging [12]. These aroma compounds are a complex mixture of alcohols, esters,

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00217-020-03568-y) contains supplementary material, which is available to authorized users.

✉ Ioannis K. Karabagias
ioanniskarabagias@gmail.com; ikaraba@cc.uoi.gr

[1] Department of Chemistry, Laboratory of Food Chemistry, University of Ioannina, 45110 Ioannina, Greece

[2] Department of Chemistry, University of Turin, Via Pietro Giuria, 7, I-10125 Turin, Italy

aldehydes, ketones, acids, terpenoids, norisoprenoids, phenols, sulfur compounds/thiols, etc., in varying amounts [8, 16].

Substantial differences in the proportions and the characteristic aroma notes of wine can be greatly influenced by both viticultural (climate, soil, water, cultivar, and grape-growing practices) and enological factors (condition of grapes, fermentation, and post-fermentation treatments). To date, 700–800 aroma compounds have been identified in wine [1, 13], which is a strong evidence for its complex composition.

On the other hand, for the complete characterization of the geographical or varietal origin of foods and wine, the implementation of statistical analysis is mandatory. Chemometrics is a powerful tool for researchers to allocate efficiently the investigated matrix in specific group (geographical origin or varietal origin) [17], and highlight among others, specific product status such as PDO (protected designation of origin) and PGI (protected geographical indication) and properties such as, i.e., flavor. Indeed, flavor differences are less desirable for producers of mass-market table wine or other cheaper wines, where consistency takes precedence. In that sense, the complete characterization and grouping of the aroma of a certified wine may generate the complete shield against the production of cheaper, adulterated, or mislabeled wines.

Based on the aforementioned, the aim of the present study was to characterize 12 wine varieties of product status (PDO or PGI) originating from grapes cultivated in different regional departments in Greece, on the basis of physico-chemical and aroma compounds analyses, and investigate whether these wine samples could be distinguished according to variety using statistics. An additional scope was the modeling of results, to monitor the effectiveness of the applied statistical models and to eliminate any drawbacks related to the different vintage of the wine samples, which were, however, of a certified status. To the best of our knowledge, this is the first report in the literature that studies together 12 different wine varieties using statistical modeling of data.

Experimental—materials and reagents

Wine samples

The total number of the studied wine samples was 60 (N=60). All samples analyzed were of Protected Designation of Origin (PDO) and Protected Geographical Indication (PGI). Samples were grouped according to type (red, white, and rosé) and according to variety (Agiorgitiko, Augoustiatis, Cabernet Sauvignon, Syrah, Vlahiko, Assyrtiko, Chardonnay, Debina, Moschofilero, Vidiano, Syrah and Mandilari, and Xinomavro). The group of dry red wines consisted of Agiorgitiko (PDO) (vintage of 2016), Augoustiatis (PGI) (vintage of 2016), Cabernet Sauvignon (PGI) (vintage of 2011), Syrah (PGI) (vintage of 2017), and Vlahiko (PGI) (vintage of 2016) varieties (N=25), which originated from the regions of Nemea (Peloponnese), Ilia (Peloponnese), Atalanti Valley (Fthiotida), Heraklion (Crete), and Zitsa (Ioannina, Epirus), respectively; The group of dry white wines consisted of Assyrtiko (vintage of 2017), Chardonnay (vintage of 2017), Debina (vintage of 2017), Moschofilero (vintage of 2018), and Vidiano (vintage of 2017) (N=25), which originated from the regions of Drama (Macedonia), Atalanti Valley (Fthiotida), Zitsa (Ioannina, Epirus), Arcadia (Peloponnese), and Heraklion (Crete), respectively; The group of dry rosé wines Syrah plus Mandilari (Vintage of 2016) (N=5) and Xinomavro (vintage of 2018) (N=5) originated from the regions of Heraklion (Crete) and Epanomi (Macedonia), respectively, and were used as the ''test/blind'' sample for the estimation of the efficacy of the statistical analysis discrimination model.

Determination of pH

The pH of wine samples was measured in 10% (w/v) aqueous wine solutions with distilled water [4], using a Delta OHM (model HD 3456.2 Padova, Italy) pH meter with a high precision (± 0.002 pH units). The instrument was calibrated with buffer solutions (pH = 4.0 and pH = 7.0) (HACH, UK). Prior the pH measurements (three replicates), the wine solution was vigorously shaken to remove the carbon dioxide.

Determination of electrical conductivity, total dissolved solids, and salinity

The electrical conductivity of wine samples was measured in a 20% (w/v) wine solution (free of carbon dioxide) in distilled water using a Delta OHM conductimeter (model HD 3456.2, Padova, Italy) at 18 ± 1 °C. The probe was calibrated automatically resorting to the 1413 μS/cm conductivity standard solution (Hannah Instruments, Inc., Woonsocket, USA). Temperature was measured by four-wire Pt 100 and two-wire Pt 1000 sensors by immersion. Similarly, salinity and total dissolved solids of a 20% (w/v) aqueous wine solution in distilled water were measured at 18 ± 1 °C using the aforementioned conductivity meter. Results were expressed as g/L and mg/L, respectively. Each sample was analyzed in triplicate.

Determination of colour

The colour parameters (L^*, a^*, b^*), were determined in pure wine samples. Colour parameter L^* corresponds to the degree of brightness; colour parameter a^* (positive values) corresponds to the degree of redness, and when a^* shows negative values to the degree of greenness; and colour parameter b^* corresponds to yellowness of colour (when positive) and to blueness of colour (when negative). The aforementioned chromaticity coordinates were measured using a Hunter Lab model DP-9000 optical sensor colorimeter (Hunter Associates Laboratory, Reston VA, USA). The sample consisting of 50 mL of wine was introduced in a glass Petri-dish and measurements ($n=5$) were carried out by manual rotation of the sample in 45° of viewing aperture. The colorimeter was calibrated with a white standard plate (YCIE = $L^*=83.87$, XCIE = $a^*=81.82$ and ZCIE = $b^*=99.59$) prior the measurements.

Chroma values (*Chroma**) were determined on the basis of a^* and b^* colour parameters using the following equation:

$$Chroma* = \left(a*^2 + b*^2\right)^{1/2} \quad [18]. \tag{1}$$

Determination of aroma compounds—headspace solid-phase micro-extraction coupled to gas chromatography/mass spectrometry (HS-SPME/GC–MS)

Isolation of aroma compounds

The extraction of aroma compounds found in the headspace of wine was done using a divinyl benzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) fiber of 50/30 µm purchased by Supelco (Bellefonte, PA, USA). Before analysis of samples, the fiber was conditioned and cleaned daily using the method of the "clean" program. More specifically, during the "cleaning" of the fiber, oven temperature was held at 80 °C for 0 min, and then increased to 260 °C at 10 °C/min (2 min hold). The inlet temperature was 270 °C. A split/splitless injection mode was followed with a ratio of 1:10. The auxiliary temperature was 280 °C and that of the MS source 230 °C. For the analysis of wine samples, the optimized conditions were as follows: 20 min equilibration time, 15 min sampling time, 5 mL sample volume, and 50 °C water bath temperature. The samples consisted of 5 mL of wine and were directly placed in 25 mL screw-cap vials equipped with PTFE/silicone septa. The vials were maintained at 50 °C in a water bath under continuous stirring at 600 rpm during the headspace extraction.

Gas chromatography–mass spectrometry unit and analysis conditions

The GC unit used in the study for the gas chromatography/mass spectrometry analysis of wine samples was an Agilent 7890A model coupled to an MS detector (Agilent 5975). The capillary column used in the analysis was DB-5MS (cross linked 5% PH ME siloxane) (60 m × 320 µm i.d., × 1 µm film thickness). Helium served as the carrier gas (purity 99.999%), at a flow rate of 1.5 mL/min. The MS source and the injector were maintained at 230 °C and 270 °C, respectively, whereas, during the analysis, the oven temperature was held at 80 °C for 0 min, and then increased to 120 °C at 4 °C/min (3 min hold), then to 240 °C at 8 °C/min for 2 min, and finally increased to 260 °C at 8 °C/min for 1 min. Electron impact mass spectra were recorded at 50–550 mass range and the ionization energy was 70 eV, whereas a split/splitless injection mode was applied with a ratio of 1:10. To avoid any source of contamination, blank runs were carried during the analysis of wine samples of different variety.

Identification of aroma compounds of wine

The identification of aroma compounds was achieved using the Wiley 7, NIST 2005 mass spectral library. For the calculation of Kovats indices, a mixture of *n*-alkanes (C_8–C_{20}) 40 mg/L each in *n*-hexane was supplied by Supelco (Bellefonte, PA, USA) and the retention time of standards was determined according to the GC–MS methodology discussed above. Aroma compounds having ≥ 80% similarity with Wiley mass spectral library were tentatively identified using the GC–MS spectra. The method of identification was based on the combination of mass spectral data found in the Wiley 7 NIST 2005 mass spectral library and data of Kovats index values that were determined for each volatile compound and then compared with those included in the Wiley MS library. Data were expressed as % of the total measured area in the total ion chromatogram (% Area pct).

Statistical analysis

The basic step was to create the group of wine samples according to type (red, white, and rosé) and according to variety. The columns indicated the variety of wine, whereas the rows comprised either the physico-chemical parameters or the aroma compounds. Thereafter, multivariate analysis of variance (MANOVA) was applied to the data set considering the varieties of wine samples as the fixed factors and the investigated parameters as the dependent variables. The Pillai's Trace and Wilks' Lambda indices were both computed to explore a possible significant effect of the aforementioned variables on wine variety. MANOVA indicated the significant parameters that could be used for the discrimination of wine samples. Linear discriminant analysis (LDA) was then applied using only the selected/significant variables to explore the possibility of differentiating wine samples according to variety. For the LDA analysis, variety was taken as the grouping variable, whereas the determined physico-chemical parameters or aroma compounds were taken as the independent variables. The discriminant analysis assumes that the predictor variables are normally distributed within each class, i.e., the data from each group follow a multivariate normal distribution. Moreover, data were centered and scaled prior to the application of LDA. For the estimation of the classification ability of samples according to the initial group of origin (variety), the original and cross-validation methods were used [19]. Data were further evaluated using weighted least-squares (WLS) regression analysis, also known as weighted linear regression. WLS is a generalization of ordinary least-squares and linear regression, in which the error covariance matrix is allowed to be different from an identity matrix. WLS is also a specialization of the generalized least squares in which the above matrix is diagonal. In statistics, linear regression is a linear approach for modeling the relationship between a scalar response (dependent variable) and one or more explanatory variables (independent variables) [20]. Correlations among the measured variables were obtained using Pearson correlation coefficient (R) at the confidence level $p < 0.05$. The higher the positive values of the Pearson's R ($-1 \leq R \leq +1$), the stronger is the positive correlation among the measured parameters. Statistical analysis was carried out using the SPSS version 26 (IBM, Armonk, NY, USA) for Windows.

Results and discussion

Physico-chemical parameters of dry red, white, and rosé wines

Table 1 shows the physico-chemical parameters of dry red, white, and rosé wines according to variety. In addition, there are given the Fisher's function values representing the F distribution of data, and the significant level of confidence for the differences among the physico-chemical parameters measured in relation to wine variety. Significant differences ($p < 0.05$) were observed for the measured physico-chemical parameters with respect to the wine variety. What is worth mentioning is that significant differences ($p < 0.05$) were obtained between the determined physico-chemical parameters, when dry red or white wine samples were tested individually according to variety (Tables S1 and S3). In this work, lower acidity values were obtained for dry white wines, followed by rosé and red dry wines. In a previous study dealing with red and white wines from the Canary Islands, the authors [4] reported pH values ranging between 3.60 ± 0.18 for red wines, whereas those of the white wines were 3.40 ± 0.31. Such values are in agreement with the present results concerning the studied dry red and white wines, while cover the range of pH values for the dry rosé wine varieties of the present study. Electrical conductivity shows the ability of a material to conduct the electric current, giving at the same time information about the "metallic character" of the material/matrix. In general, the dry red wines showed higher electrical conductivity values than those of the white or rosé varieties showing statistically significant differences ($p < 0.05$) (Table 1). The highest values were obtained for Syrah wine variety, followed by Augoustiatis wine variety. However, and in the case of the studied dry white or rosé wine varieties, significant differences ($p < 0.05$) were also observed, indicating the potential of electrical conductivity as a physico-chemical indicator of the varietal differentiation of wine.

Concerning the total dissolved solids, a similar finding was recorded. The dry red wine varieties showed significantly higher ($p < 0.05$) total dissolved solid content values compared to dry white or rosé wine varieties. Given that total dissolved solids comprise a measure of the dissolved content of all inorganic and organic substances present in a liquid matrix, in molecular, ionized, or micro-granular suspended form, the respective determinations in wine may exhaustively give information about the variety of interest. The higher values obtained in dry red wines, for example, may give partial information about the presence and binding of tannins to proteins, amino acids, minerals, or alkaloids of dry red wine varieties. The solid content of wine (among other factors such as yeast strain, yeast growth, ethanol production, fermentation temperature, must pH, aeration, grape variety, grape maturity, and skin contact time) may also affect its aroma, as it was reported that affects, i.e., the alcohol production [13, 21].

Significant variations ($p < 0.05$) were also obtained for the salinity content values among the different wine varieties. Salinity may reflect the salt content of wine, giving secondary information about the area of grapes cultivation, grape variety, etc. The higher salinity content values were obtained for the dry red wine varieties followed by the respective dry rosé and white wine varieties. Concerning the chromaticity parameters, significant differences were also observed among the different wine varieties. This was expected given the primary differences observed by the eye during the analysis of different types of wine samples (red, white, and rosé). However, significant differences ($p < 0.05$) were observed in all cases when wine samples of the same type were grouped according to variety (Tables S1 and S3). Another practical application of the determination of colour in wine is that the obtained differences may be given in numbers and not just with a physical observation, giving the opportunity to analysts for statistical handling of data and the correct classification of wine samples according to type, variety, geographical origin, etc. The lightest wines (higher L^* values) were those of Moschofilero and Vidiano varieties, whereas the most reddish wine was that of Agiorgitiko variety. The present results concerning the Agiorgitiko wine variety are in agreement with a recent study dealing with the characterization and classification of four different dry red

Table 1 Physico-chemical parameters of wine samples according to variety and statistical indices

Physico-chemical parameters	Agiorgitiko	Augoustiatis	Cabernet Sauvignon	Syrah	Vlahiko	Assyrtiko	Chardonnay	Debina	Moschofilero	Vidiano	Syrah + Mandilari	Xinomavro	Wilks' Lambda	F	*p*
Effective acidity (pH)	3.57 ± 0.01	3.60 ± 0.02	3.59 ± 0.02	3.59 ± 0.03	3.50 ± 0.02	3.16 ± 0.01	3.49 ± 0.02	3.34 ± 0.01	3.23 ± 0.01	3.36 ± 0.03	3.38 ± 0.02	3.19 ± 0.01	0.010	433.791	0.000
Electrical conductivity (μS/cm)	763 ± 2	839 ± 2	817 ± 1	864 ± 1	679 ± 2	577 ± 5	669 ± 2	588 ± 1	606 ± 1	629 ± 2	632 ± 1	683 ± 1	0.000	14,273.32	0.000
Total dissolved solids (mg/L)	381 ± 1	420 ± 1	408 ± 1	432 ± 1	341 ± 3	290 ± 2	335 ± 1	294 ± 1	303 ± 1	313 ± 2	316 ± 0	341 ± 1	0.001	6,088.77	0.000
Salinity (g/L)	0.44 ± 0.01	0.49 ± 0.00	0.47 ± 0.01	0.50 ± 0.00	0.40 ± 0.01	0.34 ± 0.00	0.39 ± 0.00	0.34 ± 0.00	0.36 ± 0.00	0.35 ± 0.01	0.36 ± 0.00	0.39 ± 0.00	0.002	2,295.92	0.000
*L**	17.42 ± 0.24	12.75 ± 0.21	12.61 ± 0.15	6.71 ± 0.10	29.78 ± 0.22	78.35 ± 0.19	78.42 ± 0.09	77.77 ± 0.11	78.90 ± 0.04	78.89 ± 0.13	62.71 ± 0.23	74.36 ± 0.08	0.000	187,052.02	0.000
*a**	44.78 ± 1.07	40.55 ± 0.81	40.50 ± 0.95	33.10 ± 0.65	40.87 ± 1.14	−4.13 ± 0.16	−5.54 ± 0.09	−3.85 ± 0.14	−3.66 ± 0.17	−4.85 ± 0.08	14.19 ± 0.42	2.47 ± 0.35	0.001	5,910.34	0.000
*b**	23.10 ± 0.19	18.59 ± 0.46	18.80 ± 0.30	8.43 ± 0.20	31.24 ± 0.43	6.45 ± 0.06	9.77 ± 0.08	7.88 ± 0.09	3.47 ± 0.07	7.82 ± 0.13	23.51 ± 0.27	9.72 ± 0.07	0.001	6,724.45	0.000
*Chroma**	50.39 ± 0.86	44.61 ± 0.57	44.66 ± 0.75	34.16 ± 0.64	51.44 ± 0.99	7.66 ± 0.13	11.23 ± 0.11	8.77 ± 0.07	5.05 ± 0.14	9.21 ± 0.12	27.46 ± 0.26	10.03 ± 0.13	0.001	6,558.14	0.000

MANOVA in comparison of values at the confidence level $p < 0.05$. The reported results are triplicates of five independent wine samples according to variety. F: value of the F distribution; *p*: probability wine varieties (Agiorgitiko, Xinomavro, Mavrotragano, and Fociano). The reported values for Agiorgitiko variety were: $L^* = 19.53 \pm 3.48$, $a^* = 45.97 \pm 1.12$, and $b^* = 24.68 \pm 3.97$ [6]. The *Chroma** values, as a mathematical transformation of the a^* and b^* values, may also give important information about the colour intensity of wine according to variety, as this parameter represents the yellow and reddish components in the chromaticity space [18]. The highest *Chroma** values were obtained for the dry red wine varieties, and, in particular, for Vlahiko and Agiorgitiko wine samples.

Aroma compounds of dry red, white, and rosé wines: internal comparison between the samples

Table 2 shows the aroma compounds identified in wine samples according to variety and the respective statistical indices. The proportion of each aroma compound is given in percent (%) of the total mass area (% area pct) and presented with the average ± standard deviation. The results of HS-SPME/GC–MS are expressed as relative percentage, which could be used as a rapid tool for the internal comparison between the different wine samples, providing qualitative standards for the presence or absence of specific aroma compounds in Greek wines. In total, 18 aroma compounds were tentatively identified belonging to alcohols, esters, phenolic compounds, pyran compounds, and terpenoids/norisoprenoids. Figure 1 shows a typical gas chromatogram of the dry red wine variety ''Vlahiko'' from Zitsa (regional unit of Ioannina) indicating with numbers some key aroma compounds.

The most abundant class of aroma compounds was the alcohols due to the high proportions of ethanol in all the studied wine varieties. However, significant differences ($p < 0.05$) were also observed in the proportions of all the other aroma compounds according to variety. The second most abundant class was the ethyl esters of the respective fatty acids.

Ethanol (a yeast-derived volatile compound) was identified in higher proportions in Chardonnay wines followed by those of Agiorgitiko, Syrah, Debina, and Moschofilero. As it can be observed, there is a different trend among the studied dry red or white wine varieties. On the contrary, the Xinomavro variety (dry rosé wine) had the lowest ethanol proportions, while the blend of Syrah (70%) and Mandilari (30%) had ethanol proportions close to the Cabernet Sauvignon, Vlahiko, and Assyrtiko wine varieties. Ethanol has been reported to be the major volatile component in wine and somewhat "masks" the developed aroma, since it affects the solubility of other aroma compounds. It has been reported in the literature that a higher concentration of ethanol affects the aroma intensity of wine [8, 13]. On the other hand, some other higher chain alcohols may give a characteristic pungent odor. In our case, these were 3-methyl-1-butanol which gives a whiskey, malt, and burnt aroma; 2-methyl-1-butanol which gives a lemon and orange aroma; and 2-methyl-1-propanol which gives the characteristic wine, solvent, and bitter aroma [8]. The aroma compounds 3-methyl-1-butanol, 2-methyl-1-butanol, and 2-methyl-1-propanol had the higher proportions in dry red wines. 3-Methyl-1-butanol had the highest proportions in Augoustiatis, Cabernet Sauvignon, and Vlahiko wine varieties. Similarly, 2-methyl-1-butanol had the highest proportions in Cabernet Sauvignon, Augoustiatis, and Syrah wine varieties. Finally, 2-methyl-1-propanol had the highest proportions in Augoustiatis, Syrah, and Vlahiko wine varieties. Another alcohol which contributed to the aroma of the studied wine varieties was hexanol. Hexanol had the highest proportions in the studied dry red wines, and especially in Vlahiko variety, which was aged for 12 months in oak barrels. This volatile compound has been reported to give a resin and green aroma note in wine [8, 9]. Another important class of odorants in wine is the ethyl esters which give a fruity aroma note [13].

Of the esters identified, the most dominants were octanoic acid ethyl ester, acetic acid ethyl ester, decanoic acid ethyl ester, and hexanoic acid ethyl ester, accompanied by smaller proportions of others, which also contribute to wine aroma [8]. Octanoic acid ethyl ester had the highest proportions in dry rosé and white wines. More specifically, Xinomavro followed by Vidiano and Assyrtiko varieties had the highest proportions of octanoic acid ethyl ester. Octanoic acid ethyl ester has been reported previously contributing to the aroma of Spanish and Uruguayan premium red aged wines [8], and in Chinese Chardonnay dry white wines [9]. On the contrary, acetic acid ethyl ester had higher proportions in dry red wines, and in particular, in Augoustiatis and Cabernet Sauvignon varieties. Of the dry white wines, acetic acid ethyl ester had considerable proportions in Assyrtiko variety. Decanoic acid ethyl ester had the highest proportions in dry rosé wines, and especially in the Xinomavro variety, followed by the blend of Syrah plus Mandilari, and that of Debina variety. The presence of decanoic acid ethyl ester either in dry red or white wines is quite common [8, 9], in agreement with the present results. Hexanoic acid ethyl ester had the highest proportions in dry white wines, followed by dry rosé and red wines, respectively. Vidiano and Assyrtiko dry white wines had the higher proportions (Table 2). Apart from the general fruity note that hexanoic acid ethyl ester gives to wine, it has been additionally reported to give an apple peel flavor [10]. Hexanoic acid ethyl ester has been reported previously to contribute to the aroma of Chinese dry red and white wine varieties such as Cabernet Sauvignon and Chardonnay [9, 10]. 3-Methyl-1-butanol-acetate (or isoamyl acetate) formed from isoamyl alcohol and acetic acid was identified in higher proportions in dry red wines, and more specifically, in Agiorgitiko variety. This compound has been reported to give a banana-like aroma in either dry

Table 2 Aroma compounds of wine samples (% area pct) according to variety and statistical indices

RT (min)	Aroma compounds (RIexp)	Agiorgitiko	Augoustiatis	Cabernet Sauvignon	Syrah	Vlahiko	Assyrtiko	Chardonnay	Debina	Moschofilero	Vidiano	Syrah + Mandilari	Xinomavro	Wilks' Lambda	F	*p*
	Alcohols															
4.53	Ethanol (< 800)	63.95 ± 4.04	61.35 ± 5.25	60.23 ± 3.56	63.63 ± 2.11	61.13 ± 4.91	58.63 ± 4.36	66.52 ± 6.46	63.59 ± 6.39	63.14 ± 5.43	56.18 ± 1.30	59.08 ± 5.53	50.52 ± 4.66	0.520	4.022	0.000
5.85	2-Methyl-1-propanol (< 800)	1.35 ± 0.20	1.59 ± 0.16	1.27 ± 0.72	1.51 ± 0.16	1.47 ± 0.16	1.15 ± 0.15	1.40 ± 0.30	0.89 ± 0.23	0.82 ± 0.21	0.78 ± 0.09	0.72 ± 0.42	0.58 ± 0.35	0.404	6.434	0.000
7.59	3-Methyl-1-butanol (< 800)	9.68 ± 0.77	10.80 ± 1.33	10.74 ± 2.80	9.74 ± 1.07	10.37 ± 1.57	4.40 ± 0.80	5.89 ± 0.51	7.26 ± 0.94	5.40 ± 1.19	4.89 ± 0.23	6.16 ± 0.97	3.22 ± 0.71	0.152	24.326	0.000
7.69	2-Methyl-1-butanol (< 800)	3.27 ± 0.44	4.00 ± 0.41	4.07 ± 0.96	3.89 ± 0.27	3.52 ± 0.66	1.50 ± 0.30	2.05 ± 0.06	2.29 ± 0.44	1.61 ± 0.33	1.28 ± 0.08	1.61 ± 0.42	0.99 ± 0.10	0.113	34.119	0.000
11.17	Hexanol (< 800)	0.15 ± 0.02	0.13 ± 0.07	0.22 ± 0.12	0.18 ± 0.10	0.44 ± 0.08	ni	0.05 ± 0.07	0.09 ± 0.08	0.16 ± 0.05	0.16 ± 0.05	0.16 ± 0.06	ni	0.242	13.666	0.000
	Sum	78.40	77.87	76.53	78.95	76.93	65.68	75.91	74.12	71.13	63.29	67.73	55.31			
	Esters															
5.66	Acetic acid ethyl ester (< 800)	3.40 ± 0.29	5.66 ± 0.65	4.96 ± 0.50	3.36 ± 0.20	4.40 ± 0.65	4.77 ± 0.70	3.16 ± 0.16	2.92 ± 0.32	1.49 ± 0.17	2.98 ± 0.17	2.55 ± 0.28	1.52 ± 0.09	0.078	51.439	0.000
9.06	Butanoic acid ethyl ester (800)	ni	ni	ni	0.06 ± 0.06	0.08 ± 0.05	0.08 ± 0.08	0.05 ± 0.07	0.02 ± 0.04	0.03 ± 0.07	0.15 ± 0.02	0.12 ± 0.07	0.08 ± 0.05	0.479	4.737	0.000
10.51	2-Methylbutanoic acid ethyl ester (800)	ni	ni	0.14 ± 0.08	0.02 ± 0.04	0.08 ± 0.08	ni	ni	ni	ni	ni	ni	ni	0.34	8.472	0.000
10.62	3-Methylbutanoic acid ethyl ester (800)	0.02 ± 0.05	ni	0.17 ± 0.10	0.12 ± 0.07	0.16 ± 0.02	0.02 ± 0.05	0.02 ± 0.04	0.07 ± 0.06	0.01 ± 0.03	0.07 ± 0.05	0.09 ± 0.04	0.05 ± 0.04	0.42	6.026	0.000
11.35	3-Methyl-1-butanol acetate (< 800)	1.41 ± 0.15	0.67 ± 0.20	0.75 ± 0.10	0.97 ± 0.09	0.83 ± 0.14	0.40 ± 0.04	0.47 ± 0.24	0.43 ± 0.11	0.29 ± 0.04	0.55 ± 0.07	0.65 ± 0.14	0.40 ± 0.05	0.132	28.764	0.000
16.08	Hexanoic acid ethyl ester (945)	1.85 ± 0.28	1.33 ± 0.68	1.48 ± 0.27	1.40 ± 0.24	1.91 ± 0.28	4.10 ± 0.59	1.98 ± 1.11	2.17 ± 1.02	3.21 ± 1.24	4.43 ± 0.16	3.33 ± 0.45	2.54 ± 1.47	0.323	9.136	0.000
21.81	Butanedioic acid diethyl ester (1208)	ni	0.07 ± 0.09	0.44 ± 0.36	0.02 ± 0.05	0.12 ± 0.14	0.02 ± 0.04	ni	0.05 ± 0.07	ni	ni	0.05 ± 0.08	0.06 ± 0.09	0.476	4.811	0.000
22.20	Octanoic acid ethyl ester (1229)	6.02 ± 1.05	4.88 ± 3.09	4.28 ± 1.52	5.94 ± 0.49	7.22 ± 0.49	15.73 ± 1.47	9.98 ± 3.69	9.87 ± 4.86	10.81 ± 7.99	19.42 ± 0.63	15.44 ± 1.65	23.36 ± 3.75	0.205	16.927	0.000

Table 2 (continued)

RT (min)	Aroma compounds (RIexp)	Agiorgitiko	Augoustiatis	Cabernet Sauvignon	Syrah	Vlahiko	Assyrtiko	Chardonnay	Debina	Moschofilero	Vidiano	Syrah + Mandilari	Xinomavro	Wilks' Lambda	F	*p*
26.82	Decanoic acid ethyl ester (1488)	0.17 ± 0.14	0.53 ± 0.40	0.19 ± 0.17	0.63 ± 0.29	0.70 ± 0.12	1.35 ± 0.55	0.81 ± 0.40	2.42 ± 1.34	2.41 ± 1.34	1.95 ± 0.06	4.88 ± 0.97	8.96 ± 1.85	0.089	44.48	0.000
	Sum	12.87	13.14	12.41	12.52	15.50	26.47	16.47	17.95	18.25	29.55	27.11	36.97			
	Phenolic compounds															
20.82	2-Phenylethanol (1159)	ni	0.04 ± 0.09	0.37 ± 0.22	0.05 ± 0.10	0.19 ± 0.19	ni	ni	0.07 ± 0.10	ni	ni	ni	ni	0.390	6.813	0.000
	Sum	ni	0.04	0.37	0.05	0.19	ni	ni	0.07	ni	ni	ni	ni			
	Pyran compounds															
21.36	Nerol oxide (1186)	ni	ni	Ni	ni	ni	ni	ni	ni	0.08 ± 0.02	ni	ni	ni	0.063	64.723	0.000
	Sum	ni	ni	Ni	ni	ni	ni	ni	ni	0.08	ni	ni	ni			
	Terpenoids/norisoprenoids															
17.82	dl-Limonene (1019)	ni	ni	Ni	ni	ni	ni	ni	Ni	0.01 ± 0.02	ni	ni	ni	0.629	2.579	*0.012*
25.02	Vitispirane (1382)	0.16 ± 0.04	0.18 ± 0.10	0.21 ± 0.02	0.21 ± 0.03	0.33 ± 0.03	0.54 ± 0.07	0.30 ± 0.46	0.12 ± 0.07	0.11 ± 0.02	0.21 ± 0.03	0.13 ± 0.06	0.19 ± 0.02	0.548	3.592	*0.001*
	Sum	0.16	0.18	0.21	0.21	0.33	0.54	0.30	0.12	0.11	0.21	0.13	0.19			

RT retention time, *RIexp* experimental retention time indices values calculated on the basis of the retention times of the standard alkane solution. MANOVA in comparison of values at the confidence level $p < 0.05$. The reported results are duplicates of five independent wine samples according to variety. *Sum* Total average % mass area. F: value of the F distribution, *p* probability, *ni* not identified. The non-identified values were treated as zeros for chemometrics

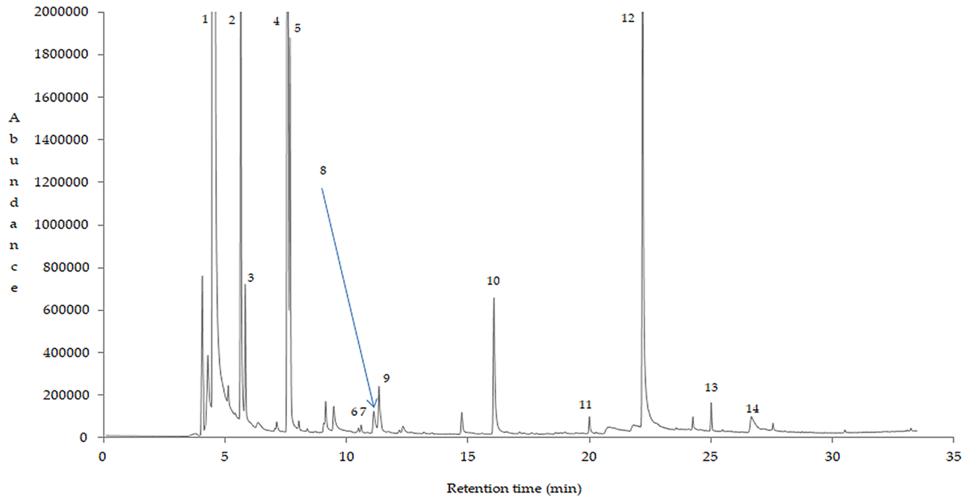


Fig. 1 A typical gas chromatogram of the dry red wine variety "Vlahiko" from the region of Zitsa. 1: Ethanol, 2: Acetic acid ethyl ester, 3:2-Methyl-1-propanol, 4:3-Methyl-1-butanol, 5: 2-Methyl-1-butanol, 6:2-Methylbutanoic acid ethyl ester, 7: 3-Methylbutanic acid ethyl ester 8:3-Methyl-1-butanol acetate, 9: Hexanoic acid ethyl ester, 10: 2-Phenylethanol, 11: Butanedioic acid diethyl ester, 12: Octanoic acid ethyl ester, 13: Vitispirane, 14: Decanoic acid ethyl ester

Table 3 Pearson's correlation coefficient (*R*) between the average values of total dissolved solids content (mg/L) and sum of average percent (%area pct) of alcohols and esters in relation to dry red, white, and rosé wine varieties

Wine variety	Total dissolved solids (mg/L) Vs. alcohols (%area pct)	Total dissolved solids (mg/L) Vs. esters (%area pct)
Dry red	$R=0.458$, $p=0.438$	$R=-0.853$, $p=0.066$
Dry white	$R=0.362$, $p=0.550$	$R=-0.279$, $p=0.649$
Rosé	$R=-1.000$, $p=0.01$	$R=1.000$, $p=0.01$

Significant correlations at the confidence level $p<0.05$. *Vs.* versus. Pearson's correlation coefficient r takes values between −1 and +1

red or white wine varieties [8–10]. Butanedioic acid diethyl ester, the diethyl ester of succinic acid, was identified in higher proportions in dry red wines, and especially, in the Cabernet Sauvignon variety. It has been reported previously contributing to the aroma of Spanish and Uruguayan dry red wines [8].

Before going any further, and considering that alcohols and esters were the dominant aroma compounds in the studied wine samples, Pearson's correlation coefficient was investigated between the average values of total dissolved solids content (mg/L) and the sum of average percent (% area pct) of alcohols and esters in relation to dry red, white, and rosé wine varieties (Table 3). The correlations indicated in Table 3 may be of particular interest to wine makers and producers in terms of the preparation of a wine with a specific aroma after having considered the variations in the total dissolved solids during vinification. Especially for the rosé wines which are the outcome of a mixed vinification procedure of red and white grapes, it is quiet important.

Of the phenolic volatiles, the major representative in the studied wine varieties was 2-phenylethanol. 2-Phenylethanol had the highest proportions in dry red wines, and in particular, in the Cabernet Sauvignon variety. 2-Phenylethanol has been identified in the extract of rose, carnation, hyacinth, Aleppo pine, orange blossom, ylang–ylang, geranium, neroli, and champaca plant species, whereas it has been reported to be an antimicrobial compound and an auto-antibiotic, produced by the fungus *Candida albicans* [22]. It has been also reported in the literature that 2-phenylethanol can be produced by bio-transformation from L-phenylalanine using the immobilized yeast *Saccharomyces cerevisiae*, comprising, thus, a yeast-derived volatile compound [23]. Therefore, its presence in dry red wine may have a complex role: it may give a rose and honey-like odor [13], and, at the same time, may enrich the antimicrobial agents of wine, in combination with ethanol, and give probably a trend about the phenylalanine content of dry red wine or the fermentation practices followed. This is a future prospective that may be sometime approved.

Terpenoids or norisoprenoids have been also reported as aroma indicators of wine [8, 10]. These compounds are a diverse class of aroma compounds that contribute to the varietal character of many wines, especially those of the Riesling variety [24]. In the present study, dl-limonene (1-methyl-4-(prop-1-en-2-yl)cyclohex-1-ene) and vitispirane (2,6,6-trimethyl-10-methylidene-1-oxaspiro[4.5]dec-8-ene) were the major terpenoids/norisoprenoids that were identified. Dl-limonene was identified only in dry white wine of the Moschofilero variety, giving a citrus-like aroma note.

Similarly, nerol oxide (3,6-dihydro-4-methyl-2-(2-methyl-1-propenyl)-2*H*-pyran), the product of oxidation of the monoterpene nerol, was identified only in the Moschofilero dry white wine variety. This compound may give a floral, lemongrass, green, and sweet odor. It has been reported previously contributing to the aroma of Australian Pinot Noir wines [25]. On the other hand, vitispirane is an aroma compound that is commonly found in alcoholic beverages, as a constituent of the juice of wine grape (*Vitis vinifera*).

Vitispirane has an aroma that has been described as floral, fruity, woody, reminiscent of eucalyptus, rose, and honey-like [26]. Vitispirane has been reported previously to contribute to the aroma of Riesling and Pinot Noir wines [24, 25]. Vitispirane was identified in all the studied wine varieties showing the highest proportions in the Assyrtiko dry white wine variety and the red dry wine variety of Vlahiko.

Considering the aforementioned differences in the volatile compounds identified in the studied wine samples, the observed differences are mainly linked to the technology of vilification, which, in turn, is linked to the secondary metabolites of grape, and, on the other hand, are linked to the different variety of grape cultivated in a specific region [14–16].

Varietal discrimination of dry red, white, and rosé wines using physico-chemical parameters and statistics

Part A: varietal discrimination of dry red wines

The qualitative criteria of multivariate analysis of variance such as Pillai's Trace = 3.524 (F = 14.812, df = 32, p = 0.000) and Wilks' Lambda = 0.000 (F = 380.608, df = 32, p = 0.000) showed that there was a statistically significant effect ($p < 0.05$) of the dry red wine variety on the measured physico-chemical parameters (Table S1). Therefore, these significant parameters were subjected to LDA. Results showed that four canonical discriminant functions of the following qualitative criteria were formed: Wilks' Lambda = 0.000 (X^2 = 355.523, df = 32, p = 0.000) for the first function; Wilks' Lambda = 0.000 (X^2 = 184.581, df = 21, p = 0.000) for the second function; Wilks' Lambda = 0.010 (X^2 = 81.437, df = 12, p = 0.000) for the third function; and Wilks' Lambda = 0.452 (X^2 = 13.898, df = 5, p = 0.016) for the fourth function. The first discriminant function accounted for 97.7% of the total variance, and had the highest eigenvalue (17,467.525) and canonical correlation (1.000). The second discriminant function had a lower eigenvalue (361.824) and canonical correlation (0.999), while accounted for 2% of the total variance. The third discriminant function had a lower eigenvalue (46.436) and canonical correlation (0.989) accounting for 0.3% of the total variance. Finally, the fourth discriminant function had the lowest eigenvalue (1.213) and canonical correlation (0.740) as it explained the zero tolerance (0% of total variance). All discriminant functions accounted for 100% of the total variance. Figure 2a shows the clear discrimination of the five dry red wine varieties. The classification rate was 100% using the original and 100% using the cross-validation method. As it can be observed, a perfect classification rate was obtained for all the dry red wine varieties. Supplementary Table 2 (Table S2) shows the perfect allocation of samples according to the initial group of wine variety. Figure 2 also shows the group centroid values. In particular, the group centroid values represent the unstandardized canonical discriminant functions evaluated at group means [19]. Each centroid has two numbers which reflect the coordinates. The abscissa is the first discriminant function and the ordinate is the second. The group centroid values were: (− 27.198, 18.324), (68.683, 8.051), (56.205, 12.634), (119.206, − 27.449), and (− 216.896, − 11.560) for Agiorgitiko, Augoustiatis, Cabernet Sauvignon, Syrah, and Vlahiko varieties. The classification function coefficients for the building of the discriminant function concerning the dry red wines using physico-chemical parameters and LDA are given in Supplementary Table 3 (Table S3).

Part B: varietal discrimination of dry white wines

Similarly, Pillai's Trace = 3.904 (F = 81.600, df = 32, p = 0.000) and Wilks' Lambda = 0.000 (F = 302.578, df = 32, p = 0.000) showed that there was a statistically significant effect ($p < 0.05$) of dry white wine variety on the measured physico-chemical parameters (Table S4). Thus, these significant parameters were subjected to LDA. Results showed that four canonical discriminant functions of the following qualitative criteria were formed: Wilks' Lambda = 0.000 (X^2 = 340.784, df = 32, p = 0.000) for the first function; Wilks' Lambda = 0.000 (X^2 = 211.201, df = 21, p = 0.000) for the second function; Wilks' Lambda = 0.002 (X^2 = 107.976, df = 12, p = 0.000) for the third function; and Wilks' Lambda = 0.053 (X^2 = 51.557, df = 5, p = 0.016) for the fourth function. The first discriminant function accounted for 80.2% of the total variance and had the highest eigenvalue (1642.812) and canonical correlation (1.000). The second discriminant function had a lower eigenvalue (363.512) and canonical correlation (0.999), while accounted for 17.7% of the total variance. The third discriminant function had an even lower eigenvalue (24.127) and canonical correlation (0.980) accounting for 1.2% of the total variance. Finally, the fourth discriminant function had the lowest eigenvalue (18.032) and canonical correlation (0.973), while explained only the 0.9% of the total variance. All discriminant functions accounted for 100% of the total variance. Figure 2b shows a very clear discrimination of the five dry white wine varieties. The classification rate was 100% using the original and 96% using the cross-validation method. The classification rate was 100% for Assyrtiko, Chardonnay, Debina, and Moschofilero varieties, whereas the respective classification rate of Vidiano variety was 80%. Supplementary Table 5 (Table S5) shows the allocation of samples according to the initial group of white wine variety. The group centroid values were: (19.257, − 19.795), (− 54.175 16.445), (− 9.104, − 19.329), (55.280, 20.311), and (− 11.258, 2.367) for Assyrtiko, Chardonnay, Debina, Moschofilero, and Vidiano varieties. The classification function coefficients for the

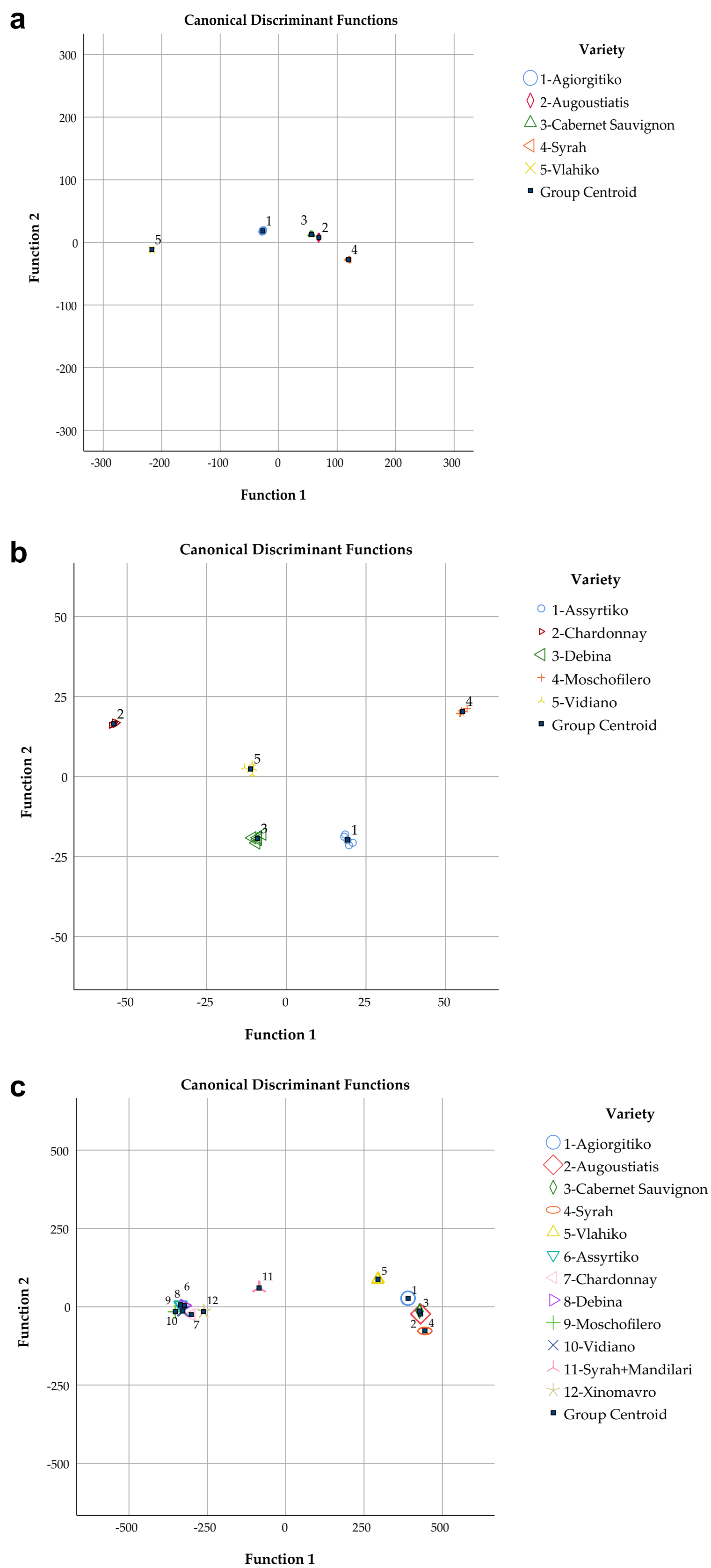


Fig. 2 **a** Varietal discrimination of dry red wines using the measured physico-chemical parameters and LDA. **b** Varietal discrimination of dry white wines using the measured physico-chemical parameters and LDA. **c** Varietal discrimination of dry red, rosé, and white wines using the measured physico-chemical parameters and LDA

building of the discriminant function concerning the dry white wines using physico-chemical parameters and LDA are given in Supplementary Table 6 (Table S6).

Part C: estimation of the model efficacy—varietal discrimination of dry red, white, and rosé wines

To further investigate the efficiency of the discrimination model obtained for the varietal classification of dry red and white wines, the dry rosé wine samples were introduced in the analysis. Pillai's Trace = 6.589 ($F = 20.385$, df = 88, $p = 0.000$) and Wilks' Lambda = 0.000 ($F = 831.495$, df = 88, $p = 0.000$) showed that there was a statistically significant effect ($p < 0.05$) of the wine variety on the measured physico-chemical parameters (Table S7). Afterwards, LDA was performed. Results showed that eight canonical discriminant functions were formed: Wilks' Lambda = 0.000 ($X^2 = 1791.523$, df = 88, $p = 0.000$) for the first; Wilks' Lambda = 0.000 ($X^2 = 1208.220$, df = 70, $p = 0.000$) for the second; Wilks' Lambda = 0.000 ($X^2 = 833.553$, df = 54, $p = 0.000$) for the third; Wilks' Lambda = 0.000 ($X^2 = 554.294$, df = 40, $p = 0.000$) for the fourth; Wilks' Lambda = 0.002 ($X^2 = 301.733$, df = 28, $p = 0.000$) for the fifth; Wilks' Lambda = 0.024 ($X^2 = 181.884$, df = 18, $p = 0.000$) for the sixth; Wilks' Lambda = 0.198 ($X^2 = 79.284$, df = 10, $p = 0.000$) for the seventh; and Wilks' Lambda = 0.991 ($X^2 = 0.442$, df = 4, $p = 0.979$) for the eighth discriminant function. The eighth discriminant function, however, was not significant ($p > 0.05$). The first discriminant function accounted for 98.3% of the total variance and had the highest eigenvalue (147,876.484) and canonical correlation (1.000). The second discriminant function had a lower eigenvalue (2091.822) and canonical correlation (1.000), while accounted for 1.4% of the total variance. The third discriminant function had an even lower eigenvalue (297.617) and canonical correlation (0.998) accounting for 0.2% of the total variance. Finally, the fourth discriminant function had the lowest eigenvalue (172.176) and canonical correlation (0.997), while explained only the 0.1% of the total variance. The first four discriminant functions accounted for 100% of the total variance (therefore, no data are given for the rest discriminant functions). Figure 2c shows a perfect discrimination of the 12 dry white wine varieties. The classification rate was 100% using the original and 100% using the cross-validation method. The use of the two different dry rosé wine varieties did not affect the discrimination ability of the developed model, using the specific physico-chemical parameters. Supplementary Table 8 (Table S8) shows the allocation of samples according to the initial group of white wine variety. The group centroid values were: (390.435, 26.988), (430.777, − 23.036), (427.063, − 13.755), (444.205, − 77.074), (294.422, 88.018), (− 335.540, 5.744), (− 301.446, − 25.595), (− 322.943, 3.781), (− 352.280, − 15.945), (− 328.210, − 13.723), (− 84.859, 59.758), and (− 261.625, − 15.158) for Agiorgitiko, Augoustiatis, Cabernet Sauvignon, Syrah, Vlahiko, Assyrtiko, Chardonnay, Debina, Moschofilero, Vidiano, Syrah plus Mandilari, and Xinomavro wine varieties. Concerning the xy clustering in Fig. 2c of the dry rosé wines more closely to the whites wines, rather than those of dry red wines, it is probably owed to the different vintage and differences in the geographical area of grape cultivation. The classification function coefficients for the building of the discriminant function concerning the dry red, white, and rosé wines using physico-chemical parameters and LDA are given in Supplementary Table 9 (Table S9). The classification results were further evaluated using WLS regression modeling, and the consideration of coefficients of correlation *R*-squared, adjusted *R*-squared, and multiple *R*-squared. At this point, it should be stressed that R^2 is a statistic that will give information about the goodness of fit of a model. In regression analysis, the R^2 coefficient of determination is a statistical measure of how well the regression predictions approximate the real data points. An R^2 of 1 indicates that the regression predictions perfectly fit the data. The implementation of the adjusted *R*-squared statistic is almost the same as R^2, but it penalizes the statistic as extra variables are included in the model. Finally, the term multiple R^2 equals the square of the Pearson correlation coefficient between the observed and modeled (predicted) data values of the dependent variable [19]. Considering the above the WLS regression analysis model had the following characteristics multiple $R^2 = 0.967$, $R^2 = 0.936$, and adjusted-$R^2 = 0.926$, standard error of the estimate 3.119, and power value of − 1.50 (this ranges from − 2.0 to 2.0), at the confidence level $p = 0.000$.

Varietal discrimination of dry red, rosé, and white wines using aroma compounds and statistics

Part A: varietal discrimination of dry red wines

The next effort was oriented on the use of aroma compounds to investigate whether these could also fire up a satisfactory discrimination of dry red wine samples according to variety. Pillai's Trace = 3.837 ($F = 11.804$, df = 64, $p = 0.000$) and Wilks' Lambda = 0.000 ($F = 12.314$, df = 64, $p = 0.000$) showed that there was a statistically significant effect ($p < 0.05$) of the dry red wine variety on the proportions of ten aroma compounds (Table S10). In particular, the aroma compounds that showed significant differences among the different dry red wine varieties were acetic acid ethyl ester, butanoic acid ethyl ester, 2-methylbutanoic acid ethyl ester, 3-methylbutanoic acid ethyl ester, hexanol, 3-methyl-1-butanol-acetate, benzeneethanol, butanodioic acid ethyl ester, vitispirane, and decanoic acid ethyl ester. Therefore, these were subjected to LDA.

Results showed that four canonical discriminant functions of the following qualitative criteria were formed: Wilks' Lambda = 0.000 (X^2 = 168.424, df = 44, p = 0.000) for the first function; Wilks' Lambda = 0.002 (X^2 = 96.618, df = 30, p = 0.000) for the second function; Wilks' Lambda = 0.034 (X^2 = 53.969, df = 18, p = 0.000) for the third function; and Wilks' Lambda = 0.355 (X^2 = 16.571, df = 8, p = 0.035) for the fourth function. The first discriminant function accounted for 78.2% of the total variance, and had the highest eigenvalue (87.934) and canonical correlation (0.994). The second discriminant function had a lower eigenvalue (13.376) and canonical correlation (0.965), while accounted for 11.9% of the total variance. The third discriminant function had a lower eigenvalue (9.354) and canonical correlation (0.950) accounting for 8.3% of the total variance. Finally, the fourth discriminant function had the lowest eigenvalue (1.817) and canonical correlation (0.803), whereas explained the 1.6% of the total variance. All discriminant functions accounted for 100% of the total variance. Figure 3a shows the clear discrimination of the five dry red wine varieties. The classification rate was 100% using the original and 96% using the cross-validation method. A perfect classification rate was obtained for Agiorgitiko, Augoustiatis, Cabernet Sauvignon, and Vlahiko varieties. The classification rate of Syrah variety was 80% (Table S11). The respective group centroid values were: (13.595, − 1.400), (− 9.915, − 4.947), (− 3.870, 4.851), (5.273, − 0.351), and (− 5.082, 1.847) for Agiorgitiko, Augoustiatis, Cabernet Sauvignon, Syrah, and Vlahiko varieties. The classification function coefficients for the building of the discriminant function concerning the dry red wines using aroma compounds and LDA are given in Supplementary Table 12 (Table S12).

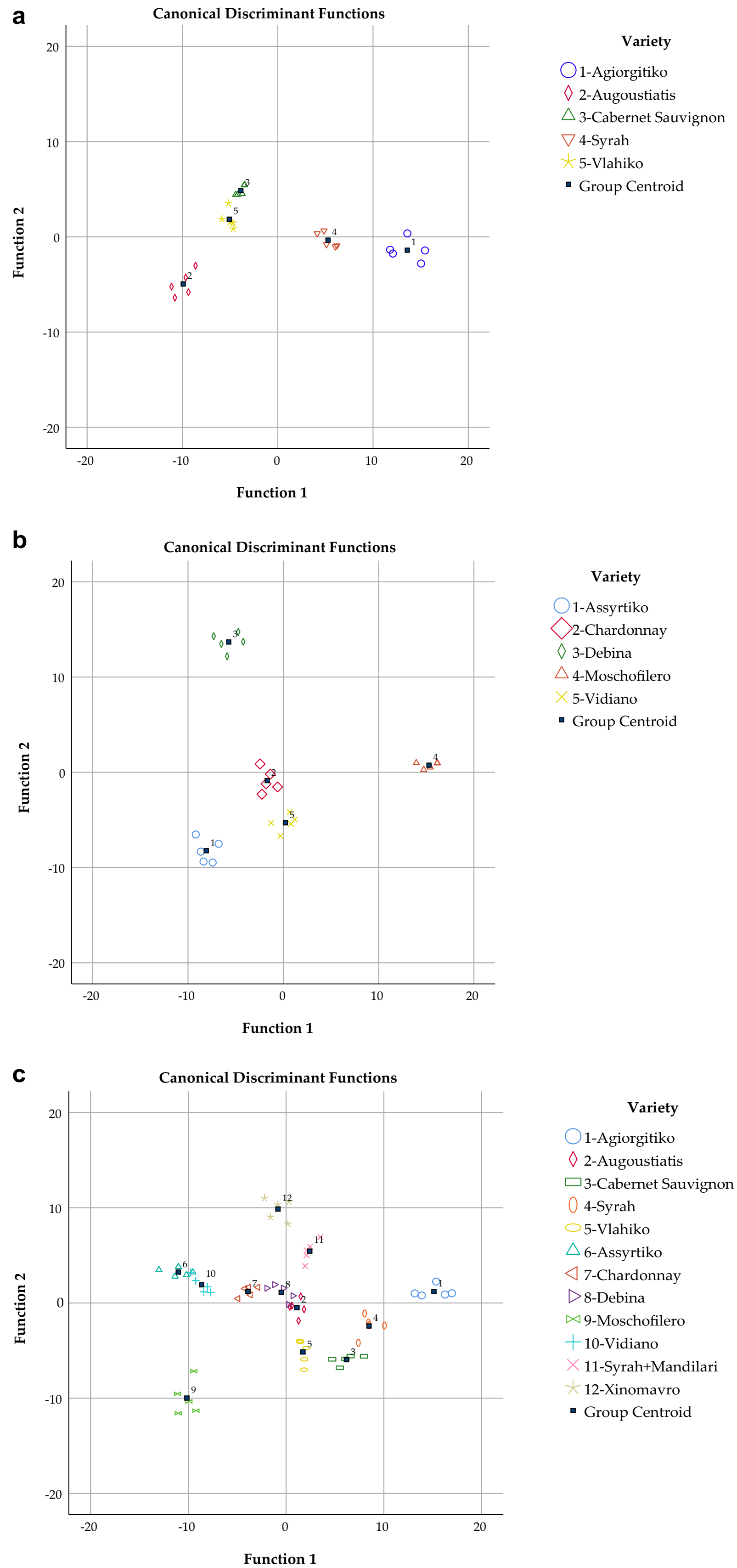
Part B: varietal discrimination of dry white wines

Similarly, Pillai's Trace = 3.829 (F = 9.248, df = 68, p = 0.000) and Wilks' Lambda = 0.000 (F = 12.393, df = 68, p = 0.000) showed that there was a statistically significant effect ($p < 0.05$) of the dry white wine variety on the proportions of aroma compounds. In particular, 12 aroma compounds with the exception of the non-significant ($p > 0.05$) 3-methylbutanoic acid ethyl ester, 3-methyl-1-butanol acetate, benzene ethanol, dl-limonene, and butanedioic acid diethyl ester showed significant variations ($p < 0.05$) in their proportions according to the dry white wine variety (Table S13). It should also be stressed that the aroma compound 2-methylbutanoic acid ethyl ester was excluded from the analysis, given that in the dry white wine varieties was not identified. Afterwards, the 12 significant aroma compounds were subjected to LDA. Results showed that four canonical discriminant functions of the following qualitative criteria were formed: Wilks' Lambda = 0.000 (X^2 = 190.511, df = 48, p = 0.000) for the first function; Wilks' Lambda = 0.000 (X^2 = 121.710, df = 33, p = 0.000) for the second function; Wilks' Lambda = 0.028 (X^2 = 55.392, df = 20, p = 0.000) for the third function; and Wilks' Lambda = 0.200 (X^2 = 24.964, df = 9, p = 0.003) for the fourth function. The first discriminant function accounted for 50.7% of the total variance and had the highest eigenvalue (83.670) and canonical correlation (0.994). The second discriminant function had a lower eigenvalue (71.139) and canonical correlation (0.993), while accounted for 43.1% of the total variance. The third discriminant function had a much lower eigenvalue (6.121) and canonical correlation (0.927) accounting for 3.7% of the total variance. Finally, the fourth discriminant function had the lowest eigenvalue (4.006) and canonical correlation (0.895), while explained only the 2.4% of the total variance. The four discriminant functions accounted for 100% of the total variance. As in the case of dry red wines, the classification rate was 100% using the original and 96% using the cross-validation method. The classification rate was 100% for Assyrtiko, Debina, Moschofilero, and Vidiano varieties, whereas the respective classification rate of the Chardonnay variety was 80% (Fig. 3b). Supplementary Table 14 (Table S14) shows the allocation of samples according to the initial group of dry white wine variety. The group centroid values were: (− 8.086, − 8.245), (− 1.691, − 0.873), (− 5.729, 13.679), (15.283, 0.746), and (0.223, − 5.306) for Assyrtiko, Chardonnay, Debina, Moschofilero, and Vidiano varieties, respectively. The classification function coefficients for the building of the discriminant function concerning the dry white wines using aroma compounds and LDA are given in Supplementary Table 15 (Table S15).

Part C: estimation of the model efficacy—varietal discrimination of dry red, white, and rosé wines

As in the case of physico-chemical parameter analysis, the rosé wine samples were introduced in the analysis. Pillai's Trace = 7.824 (F = 5.611, df = 198, p = 0.000) and Wilks' Lambda = 0.000 (F = 13.436, df = 198, p = 0.000) showed that there was a statistically significant effect ($p < 0.05$) of the wine variety on the proportions of aroma compounds (Table S16). Following the analysis, LDA was performed. Results showed that eleven canonical discriminant functions were formed: Wilks' Lambda = 0.000 (X^2 = 929.866, df = 198, p = 0.000) for the first; Wilks' Lambda = 0.000 (X^2 = 742.831, df = 170, p = 0.000) for the second; Wilks' Lambda = 0.000 (X^2 = 588.182, df = 144, p = 0.000) for the third; Wilks' Lambda = 0.000 (X^2 = 473.760, df = 120, p = 0.000) for the fourth; Wilks' Lambda = 0.001 (X^2 = 316.918, df = 98, p = 0.000) for the fifth; Wilks' Lambda = 0.005 (X^2 = 229.866, df = 78, p = 0.000) for the sixth; Wilks' Lambda = 0.031 (X^2 = 153.075, df = 60, p = 0.000) for the seventh; Wilks' Lambda = 0.125

Fig. 3 **a** Varietal discrimination of dry red wines using ten aroma compounds and LDA. **b** Varietal discrimination of dry white wines using 12 aroma compounds and LDA. **c** Varietal discrimination of dry red, white, and rosé wines using 18 aroma compounds and LDA



($X^2 = 91.502$, df = 44, $p = 0.000$) for the eighth; Wilks' Lambda = 0.333 ($X^2 = 48.339$, df = 30, $p = 0.018$) for the ninth; Wilks' Lambda = 0.547 ($X^2 = 26.558$, df = 18, $p = 0.088$) for the tenth; and Wilks' Lambda = 0.837 ($X^2 = 7.837$, df = 8, $p = 0.450$) for the eleventh discriminant function. The tenth and eleventh discriminant functions, however, were not significant ($p > 0.05$). All the identified aroma compounds were found significant ($p < 0.05$) for the varietal discrimination of wine samples. The first discriminant function accounted for 42.4% of the total variance, and had the highest eigenvalue (69.161) and canonical correlation (0.993). The second discriminant function had a lower eigenvalue (32.608) and canonical correlation (0.985), while accounted for 20% of the total variance. The third discriminant function had a slightly lower eigenvalue (30.231) and canonical correlation (0.984) accounting for 18.5% of the total variance. The fourth discriminant function had an even lower eigenvalue (14.236) and canonical correlation (0.967), while explained the 8.7% of the total variance. The fifth discriminant function had an even lower eigenvalue (6.232) and canonical correlation (0.928), while explained the 3.8% of the total variance. The sixth discriminant function had an even lower eigenvalue (4.727) and canonical correlation (0.909), while explained the 2.9% of the total variance. The seventh discriminant function had an even lower eigenvalue (3.053) and canonical correlation (0.868), while explained the 1.9% of the total variance. The eighth discriminant function had an even lower eigenvalue (1.667) and canonical correlation (0.791), while explained the 1% of the total variance. The ninth discriminant function had a much lower eigenvalue (0.641) and canonical correlation (0.625), while explained the 0.4% of the total variance. The tenth discriminant function had an even lower eigenvalue (0.530) and canonical correlation (0.589), while explained the 0.3% of the total variance. Finally, the eleventh discriminant function had the lowest eigenvalue (0.195) and canonical correlation (0.404), while explained only the 0.1% of the total variance. The eleven discriminant functions accounted for 100% of the total variance. Figure 3c shows a very clear discrimination of the 12 white wine varieties. The classification rate was 100% using the original and 96.7% using the cross-validation method, very satisfactory and almost perfect, considering the numerous studied wine varieties. As in the case of the physico-chemical parameters, the use of the dry rosé wine varieties did not affect the discrimination ability of the developed model. The 10 of the 12 wine varieties had a perfect classification rate (100%) using the cross-validation method. Only the Chardonnay and Xinomavro varieties had a respective classification rate of 80%. Supplementary Table 17 (Table S17) shows the allocation of samples according to the initial group of wine variety. The group centroid values were: (15.122, 1.189), (1.116, − 0.500), (6.173, − 5.955), (8.478, − 2.414), (1.726, − 5.148), (− 11.021, 3.244), (− 3.900, 1.224), (− 0.497, 1.126), (− 10.147, − 9.983), (− 8.638, 1.908), (2.422, 5.443), and (− 0.833, 9.867) for Agiorgitiko, Augoustiatis, Cabernet Sauvignon, Syrah, Vlahiko, Assyrtiko, Chardonnay, Debina, Moschofilero, Vidiano, Syrah plus Mandilari, and Xinomavro wine varieties. The classification function coefficients for the building of the discriminant function concerning the dry red, white, and rosé wines using aroma compounds and LDA are given in Supplementary Table 18 (Table S18).

Similar to the testing of the physico-chemical parameters as modeled predictors for the varietal discrimination of the 12 wine varieties, WLS regression modeling was implemented to the set of aroma compounds. The WLS regression analysis model had the following characteristics: multiple $R^2 = 0.988$, $R^2 = 0.977$, and adjusted-$R^2 = 0.967$, standard error of the estimate 0.284, and power value of 1.50, at the confidence level $p = 0.000$. Even though the varietal discrimination results of LDA, based on the cross-validation method, after subjection to analysis of the aroma compounds of dry rosé wine samples were slightly lower than those of the physico-chemical parameters (96.7% vs. 100%), the WLS regression analysis model gave better results for modeling the set of data of aroma compounds. This finding indicates the potential use of aroma compounds for the identical characterization and authentication of wine.

Conclusions

Results of the present study highlighted the potential use of physico-chemical parameters and aroma compounds for the characterization and discrimination of wine samples of different grape variety, in combination with statistical analysis. For the first time in the literature, there were studied together 12 different Greek wine varieties of PDO and PGI status. The study may be important then for different sectors. At first, it brings knowledge on the domestic wine industry/producers for these specific wine varieties and covers the demand of consumers for certified wine having identical characteristics, and secondly supports the literature by providing data for the physico-chemical parameters and aroma compounds of less known Greek wine varieties that are sold in the international market.

Acknowledgements The authors are grateful to the following wineries/estates/domaine: Zoinos Winery, Zitsa, Ioannina, Epirus; Ktima Gerovassiliou, Epanomi, Macedonia; Merkouri Estate, Korakochori, Ilia, Peloponnese; Diamantakis Winery, Heraklion, Crete; Skouras Domaine, Arcadia, Peloponnese; Hatzimichalis Domaine, Atalanti Valley, Fthiotida, Central Greece, for the donation of wine samples.

Author contributions Conceptualization, IKK and DS; methodology, IKK; software, AVB; validation, IKK, and AVB; formal analysis, IKK, DS, and AM; investigation, IKK, DS, AM; resources, AVB; data curation, IKK, and DS; writing—original draft preparation, IKK; writing—review and editing, IKK; visualization, IKK, and DS; supervision, IKK; project administration, IKK; funding acquisition, AVB. All authors have read and agreed to the published version of the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Compliance with ethics requirements This article does not contain any studies with human or animal subjects.

Informed consent Not applicable.

References

1. Berger RF (2007) Flavours and fragrances, chemistry, bioprocessing and sustainability. Springer, Berlin, pp 1–649
2. Fraga H, Malheiro AC, Moutinho-Pereira J, Cardoso RM, Soares PMM, Cancela JJ, Pinto JG, Santos JA et al (2014) Integrated analysis of climate, soil, topography and vegetative growth in Iberian viticultural regions. PLoS ONE 9(9):e108078. https://doi.org/10.1371/journal.pone.0108078
3. King ES, Kievit RL, Curtin C, Swiegers JH, Pretorius IS et al (2010) The effect of multiple yeasts co-inoculations on Sauvignon Blanc wine aroma composition, sensory properties and consumer preference. Food Chem 122:618–626. https://doi.org/10.1016/j.foodchem.2010.03.021
4. Darias-Martín J, Socas-Hernández A, Díaz-Romero C, Díaz-Díaz E (2003) Comparative study of methods for determination of titrable acidity in wine. J Food Comp Anal 16:555–562. https://doi.org/10.1016/S0889-1575(03)00032-2
5. Zhu L, Wang L, Yang W, Guo D (2016) Physicochemical data mining of Msalais, a traditional local wine in Southern Xinjiang of China. Int J Food Prop 19(11):2385–2395. https://doi.org/10.1080/10942912.2015.1033549
6. Karimali D, Kosma I, Badeka A (2020) Varietal classification of red wine samples from four native Greek grape varieties based on volatile compound analysis, color parameters and phenolic composition. Eur Food Res Technol 246:41–53. https://doi.org/10.1007/s00217-019-03398-7
7. Lopez R, Aznar M, Cacho JF, Ferreira V (2002) Quantitative determination of minor and trace volatile compounds in wine by solid-phase extraction and gas chromatography with mass spectrometric detection. J Chromatogr A 966:167–177. https://doi.org/10.1016/s0021-9673(02)00696-9
8. Escudero A, Campo E, Farina L, Cacho J, Ferreira V (2007) Analytical characterization of the aroma of five premium red wines. Insights into the role of odor families and the concept of fruitiness of wines. J Agric Food Chem 55:4501–4510. https://doi.org/10.1021/jf0636418
9. Li H, Tao YS, Wang H (2008) Impact odorants of Chardonnay dry white wine. Eur Food Res Technol 227:287–292. https://doi.org/10.1007/s00217-007-0722-9
10. Tao Y, Li H, Wang H, Zhang L (2008) Volatile compounds of young Cabernet Sauvignon red wine from County (China). J Food Comp Anal 21:689–694. https://doi.org/10.1016/j.jfca.2008.05.007
11. Munoz-Gonzalez C, Rodriguez-Bencomo JJ, Moreno-Arribas MV, Pozo-Bayon MA (2011) Beyond the characterization of wine aroma compounds: looking for analytical approaches in trying to understand aroma perception during wine consumption. Anal Bioanal Che 401:1497–1512. https://doi.org/10.1007/s00216-011-5078-0
12. Rapp A, Mandery H (1986) Wine aroma. Experientia 42:873–884. https://doi.org/10.1007/BF01941764
13. Villamor RR, Ross CF (2013) Wine matrix compounds affect perception of wine aromas. Annu Rev Food Sci Technol 4:1–20. https://doi.org/10.1146/annurev-food-030212-182707
14. Schüttler A, Friedel M, Jung R, Rauhut D, Darriet P (2015) Characterizing aromatic typicality of Riesling wines: merging volatile compositional and sensory aspects. Food Res Int 69:26–37. https://doi.org/10.1016/j.foodres.2014.12.010
15. Ruiz J, Kiene F, Belda I, Fracassetti D, Marquina D, Navascués E, Calderón F, Benito A, Rauhut D, Santos A, Benito S (2019) Effects on varietal aromas during wine making: a review of the impact of varietal aromas on the flavor of wine. Appl Microbiol Biotechnol 103(18):7425–7450. https://doi.org/10.1007/s00253-019-10008-9
16. Chen L, Capone DL, Nicholson EL, Jeffery DW (2019) Investigation of intraregional variation, grape amino acids, and pre-fermentation freezing on varietal thiols and their precursors for *Vitis vinifera* Sauvignon blanc. Food Chem 295:637–645. https://doi.org/10.1016/j.foodchem.2019.05.126
17. Granato D, Putnik P, Bursać Kovačević D et al (2018) Trends in chemometrics: food authentication, microbiology, and effects of processing. Compr Rev Food Sci Food Saf 17(3):663–677. https://doi.org/10.1111/1541-4337.12341
18. Mclellan MR, Lind LR, Kime RW (1995) Hue angle determinations and statistical analysis for multi quadrant HUNTER L, a, b data. J Food Qual 18:235–240. https://doi.org/10.1111/j.1745-4557.1995.tb00377.x
19. Karabagias VK, Karabagias IK, Louppis BA, Kontominas MG, Papastephanou C (2019) Valorization of prickly pear juice geographical origin based on mineral and volatile compound contents using LDA. Foods 8:123. https://doi.org/10.3390/foods8040123
20. Montgomery DC, Peck EA, Vining GG (2012) Introduction to linear regression analysis, 5th edn. Wiley, New York, pp 188–193
21. Fleet GH (1990) Growth of yeasts during wine fermentations. J Wine Res 1(3):211–223. https://doi.org/10.1080/09571269008717877
22. Lingappa BT, Prasad M, Lingappa Y, Hunt DF, Biemann K (1969) Phenethyl alcohol and tryptophol: Autoantibiotics produced by the fungus *Candida albicans*. Sci 163(3863):192–194. https://doi.org/10.1126/science.163.3863.192
23. Eshkol N, Sendovski M, Bahalul M, Katz-Ezov T, Kashi Y, Fishman A (2009) Production of 2-phenylethanol from L-phenylalanine by a stress tolerant *Saccharomyces cerevisiae* strain. J Appl Microbiol 106(2):534–542. https://doi.org/10.1111/j.1365-2672.2008.04023.x
24. Mendes-Pinto MM (2009) Carotenoid breakdown products the—norisoprenoids—in wine aroma. Arch Biochem Biophys 483:236–245. https://doi.org/10.1016/j.abb.2009.01.008
25. Song J, Smart RE, Damberg RG, Sparrow AM, Wells RB, Wang H, Qian MC (2014) Pinot Noir wine composition from different vine vigour zones classified by remote imaging technology. Food Chem 153:52–59. https://doi.org/10.1016/j.foodchem.2013.12.037
26. Eggers NJ, Bohna K, Dooley B (2006) Determination of vitispirane in wines by stable isotope dilution assay. Am J Enol Vitic 57(2):226–232

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.