Food and Bioprocess Technology (2023) 16:24–42
https://doi.org/10.1007/s11947-022-02836-x

REVIEW ARTICLE

Could QSOR Modelling and Machine Learning Techniques Be Useful to Predict Wine Aroma?

Virginia Cardoso Schwindt[1] · Mauricio M. Coletto[1] · Mónica F. Díaz[2,3] · Ignacio Ponzoni[4,5]

Received: 27 December 2021 / Accepted: 26 May 2022 / Published online: 2 July 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
Food informatics is having an increasing impact on the food industry and improving the quality of end products, as well as the efficiency of manufacturing processes. In the case of winemaking, a particular application of interest for food informatics is the sensory analysis of wines. This problem can benefit from the strong development that machine learning has achieved in recent decades. However, these data-driven techniques require accurate and sufficient information to generate models capable of predicting the sensory profile of wines. A review of the sensory analysis and volatile composition of wines is presented in this work, along with significant studies on the use of machine learning models to predict wine-related characteristics such as the antioxidant activity of polyphenols of wine and aroma compounds. In this sense, data from a sensory panel and analytical technology were gathered. This literature review reveals the lack of a homogeneous and sufficiently large database of sensory analysis related to the volatile composition of wines to develop machine learning models. However, among artificial intelligence approaches, the application of quantitative structure-odour relationship (QSOR) models is currently gaining importance. Recent studies show that it would be possible to predict quantitatively the sensory analysis of wines by QSOR models, using general volatile composition information. Therefore, the purpose of this review is to identify key aspects and guidelines for the development of this area.

Keywords Machine learning · QSOR · Volatile composition · Wine aroma

Introduction

The 'tools of the trade' in cheminformatics, an emerging discipline in which chemical data are collected, organized and analysed, are generally applied in the pharmaceutical field. They can also be applied to other types of chemical datasets, such as those containing food chemicals. The interest in the use of chemical information methodologies to address food-related challenges is growing and will continue to grow as the methods prove their usefulness, particularly in providing practical solutions to food industry challenges (Martínez & Medina-Franco, 2014).

Improving food quality is one of the main objectives of food science and technology through the development of

✉ Virginia Cardoso Schwindt
vcardoso@unrn.edu.ar

Mauricio M. Coletto
mmcoletto@unrn.edu.ar

Mónica F. Díaz
mdiaz@plapiqui.edu.ar

Ignacio Ponzoni
ip@cs.uns.edu.ar

[1] Universidad Nacional de Río Negro (UNRN), Centro de Investigación y Transferencia Río Negro (CONICET-UNRN), Villa Regina, Río Negro, Argentina

[2] Planta Piloto de Ingeniería Química (PLAPIQUI), Universidad Nacional del Sur (UNS) -Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Bahía Blanca, Argentina

[3] Departamento de Ingeniería Química, (DIQ-UNS), Bahía Blanca, Argentina

[4] Instituto de Ciencias e Ingeniería de la Computación (ICIC), Universidad Nacional del Sur (UNS) - Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Bahía Blanca, Argentina

[5] Departamento de Ciencias e Ingeniería de la Computación, (DCIC-UNS), Bahía Blanca, Argentina

new processes, food supplements and nutraceuticals. The application of artificial intelligence approaches is an emerging field of study for obtaining new food materials in the last years (Kakani et al., 2020; Sun et al., 2019; Misra et al., 2020; Camaréna, 2020). The behaviour of food chemical compounds can be modelled, and a wide number of changes can be carried out, such as discovering new enzymes and describing substrates, products or inhibitors. Kar et al. (2017) presented the currently available information on diverse groups of molecules with applications in agriculture and food science that have been subjected to artificial intelligence models. There is also enlisted the availability of agrochemical, food and flavour databases along with a list of software tools and online resources for in silico studies. According to Roy et al. (2015a, b), artificial intelligence approaches can be used to predict the activity of advanced active agrochemicals. Moreover, they have a huge role in the food industry. In silico studies constitute a dependable tool, among others, in the research of efficient antioxidant molecules with enhanced activity, allowing the identification of the fundamental molecular fragments responsible for the antioxidant propensity to several classes of chemicals and the response pharmacophore. A large number of these models have been published for the antioxidant activity in the last 10 years, as well as several cases of study of food protein-derived bioactive peptides. However, artificial intelligence studies in the areas of food flavour, taste and food supplements, which is the scope of this article, are not yet complete for full application in the food industries, and a lot of new research is still necessary (Roy et al., 2015a, b).

One of the most important quality factors of wine is aroma, and it is one of the key factors of consumer acceptance (Lockshin & Corsi, 2012; Rapp, 1999; Sáenz-Navajas et al., 2013). In many cases, the grape variety employed in the elaboration of a particular wine determines completely the aroma of that wine. This is due to the persistence of certain compounds present in the grape throughout the entire process of vinification (Gómez García-Carpintero et al., 2011). Wine aroma is defined by more than one thousand three hundred volatile compounds, including alcohols, esters, acids, aldehydes, isoprenoids, lactones and ketones, with a wide concentration range (Villamor & Ross, 2013). Differences in the aromatic profile of wines are determined by changes in the type, proportion and concentration of these volatile compounds (Atanasova et al., 2005).

The most widely used methodology for characterizing the aromatic profile of wine has been descriptive analysis (DA) with highly trained panels (Presa-Owens & Noble, 1995; Heymann & Noble, 1987; Noble et al., 1984b; Stone, 1992). Although this method allows to obtain detailed and reproducible results (Lawless & Heymann, 2010), to create and to maintain well-trained and calibrated sensory panels can be economically challenging and time consuming (Varela & Ares, 2012). Moreover, due to extensive training, highly trained assessors can perceive wine aroma differently from consumers (Fariña et al., 2015).

A promising method to overcome this situation is the application of sensor systems such as electronic tongues (E-tongues), which have the additional features of being fast and low cost. (Legin et al., 2002; Vlasov et al., 2002, 2005; Dias et al., 2017). This instrument was applied by Legin et al. (2003) to analyse fifty-six samples of the Italian red wines Barbera d'Asti and Gutturnio. The E-tongue was able to differentiate all wine samples of the same denomination and vintage, but from different vineyards. Moreover, it was possible to measure several parameters, such as total and volatile acidity, pH, and contents of ethanol, tartaric acid, sulphur dioxide, total polyphenols and glycerol with a precision within 12%. In addition, human sensory scores within a precision of about 13% for Barbera d'Asti wines and 8% for Gutturnio wines were predicted by the system.

Another instrument to help resolve some of the issues involved in the use of human sensory panels is the electronic nose (E-nose) (Di Rosa et al., 2020). According to the basis of their detection systems, these devices are classified into two categories: classical instruments, which are based on solid-state gas sensors, and new instruments, which are based on mass spectrometry (MS). Due to the interference caused by the high ethanol content on the solid-state gas sensors, MS-based instruments are more appropriate for the analysis of alcoholic beverages (Martí et al., 2005). Even though E-noses present several advantages over traditional methods, they have still some drawbacks: compounds that are not relevant to aroma cause noise in the results, the sensors may be poisoned during the analysis, and the sensor may give ambiguous responses. In addition, the water present in the sample headspace might interfere, as the sensors are reactive to the presence of water (Di Rosa et al., 2020).

In general, although E-noses and E-tongues are inspired by human senses, correlations with human perceptions are not easy to be established due to several factors. They detect chemical compounds that human systems cannot, and their perception intensity is not directly correlated with the concentration of a certain compound. Quite often, the compounds in a mixture have a synergistic effect and the resulting aroma or taste is not the addition of the smell or taste of the individual components. In addition, human taste also perceives the mouthfeel (astringency, heat, viscosity, etc.) or flavours that contribute to the perception as well (Rodríguez-Méndez et al., 2016). Both electronic instruments are presented in Table A2, in the supplementary material.

In summary, a great development of electronic instruments for food sensory analysis has been achieved so far. However, in most cases, they are not analytical instruments, and it is not possible to completely describe the organoleptic properties of food, particularly fermented beverages.

Therefore, nowadays, E-tongues and E-noses cannot replace their biological equivalents, but they can provide relevant supplementary information for describing the aroma and taste profile of food.

The use of quantitative structure-property/activity/odour relationship (QSPR/QSAR/QSOR) models is one of the artificial intelligence approaches that has increased in recent years. The design of these models relates a set of ‘descriptors’ (known variables related to the chemical structure of molecules) to a certain property/activity/odour (known as ‘target’ (Enciso et al., 2016)). It is necessary to identify which descriptors are most closely related to the target property to infer a QSPR/QSAR/QSOR model, known as the feature selection (FS) problem (Soto et al., 2009; Martínez et al., 2015). Then, a mathematical relationship between a set of these descriptors and the property/activity is established by fitting a training set, that is a group of molecules whose experimental property/activity value is known. In general, a QSPR/QSAR/QSOR model only requires a small number of descriptors to estimate the property under study (Cravero et al., 2019).

The mathematical relationship in QSPR/QSAR/QSOR models is defined as a function $Y = f(X)$, where $X = (x_1, x_2, \cdots, x_d)$ is a vector of molecular descriptors (features), Y is an experimental target property/activity, and d is the number of descriptors. From this database, the function f can be learned by using a supervised training method, such as random forest, supported vector machine, decision tree, neural networks, and random committee. Once f has been inferred, it may be applied to unseen molecules not covered by the training method. Thus, f can predict in silico the value of a property based on the analysis of data from other experiments. To assess these models, it is necessary to identify first which molecular descriptors are related to the property under study.

In the case of the aromatic profile of wines, the features are the volatile composition parameters of wines, and the properties to predict are the aromas present in wine. This means that the problem has multiple targets, since there are multiple properties to predict. Therefore, Y is a vector of targets to predict, $Y = (y_1, y_2, \cdots, y_t)$, and t is the number of targets.

Although it is well known that smell is determined by physiochemical properties of chemical compounds, there are multiple mathematical relationships between them that describe the underlying phenomenon. These relationships are called quantitative structure-odour relationship (QSOR), and so far it is not known how these properties or chemical structures affect odour quality (Licon et al., 2019). Nevertheless, a major progress has recently been made in the QSOR approach. Sharma et al. (2021a) developed a deep neural network (DNN) with physicochemical properties and molecular fingerprints (PPMF) and a convolution neural network (CNN) with chemical-structure images (IMG) to predict the smells of chemical compounds using their simplified molecular input line entry specification (SMILES) as molecular representations. They applied their QSOR model to an independent test set of chemical compounds and achieved a smell prediction accuracy of 97.3% and 98.3% from DNN + PPMF and CNN + IMG, respectively. Furthermore, Sharma et al. (2021b) presented OlfactionBase, a free, open-access web server where they gather together knowledge on many aspects of the olfaction mechanism in a single place that contains detailed information of components such as odours and odourants, among other aspects.

The scope of the present paper is threefold; first, to review the state of the art in the use of QSPR/QSAR/QSOR models to predict characteristics of wines; second, to review the publications on the aromatic profile and volatile composition of wines during the last decades, useful to develop a database for QSOR prediction of aromatic profile; and finally, to identify challenges and opportunities for the prediction of aromatic profiles of wines using QSOR models. To the best of our knowledge, this is the first review that addresses these goals of food informatics applied to wine.

Application of QSAR/QSPR Modelling in Wine-Related Studies

According to our research, there are few wine-related studies using QSAR/QSPR models. In a pioneer work, amino acid profiles had been used as a criterion for the authenticity of wines in different countries. Duchowicz et al. (2013) applied QSPR models to the aminograms obtained by high-performance liquid chromatography (HPLC) in their laboratories for Torrontés and Merlot wines. Their QSPR predictions for the amino acid profiles presented high concordance with the experimental data. In addition, their QSPR model showed other worthy applications such as the identification of each wine varietal, testing their authenticity, and estimating the concentrations of non-available experimental data in the amino acid profiles of Torrontés and Merlot wines. In addition, Kang et al. (2014) applied QSPR models between initial amino acids in Korean rice wine (makgeolli) mash and major aromatic compounds. According to their results, fusel alcohols and their acetate esters were positively correlated with the amino acid profile at the initial fermentation, but ethyl esters of medium-chain fatty acids were independent of amino acid profile.

To identify the odourant molecules essential for lowering the odour threshold properties (OTP), Ojha and Roy (2018) modelled the OTP of various aroma components present in different types of wine using 2D and 3D descriptors by employing QSPR. First, they selected the most relevant descriptors; then, they developed a partial least squares (PLS) regression model; finally, they validated it considering its acceptability and predictivity to improve confidence in QSPR predictions. Furthermore, they verified the obtained results with the observations reported by Wang et al. (2017) and concluded that their developed model could be a valuable tool for a better understanding of the relationship between the aroma characteristics of different types of wines obtained under diverse manufacturing conditions and their aroma constituents.

Huangjiu is a traditional Chinese wine with special taste and flavour, but changes of aroma compounds during storage remain unclear. Feng et al. (2020) developed QSAR models to predict the flavour thresholds for alcohols, acids and esters in Huangjiu, obtaining higher accuracy for alcohols and acids. Their study is relevant because it provides valuable information to unveil the regulation of Huangjiu flavour on a molecular basis. However, they did not develop QSAR/QSPR models for sensory analysis of wines according to their volatile composition.

Some recent non-QSAR/QSPR studies in which researchers used machine learning algorithms to predict wine flavour and aroma are listed below.

Fuentes et al. (2020) used machine learning modelling strategies to predict the aroma profile of Australian Pinot noir wines (2008–2016 vintages). They determined aroma profile by gas chromatography (GC) and chemometric analysis and used weather and water management information from a Pinot noir vineyard as input data. They obtained high accuracy artificial neural network (ANN) models in the prediction of both parameters compared to the experimental values: aroma profile and chemometric analysis.

ANN was found to be the best predictive method for wine sensory quality grading as a function of the aroma chemistry in Sauvignon Blanc according to Zhu et al. (2021). The authors also detected a correlation between certain volatile compounds and wine sensory grading. They applied static headspace-gas chromatography-ion mobility spectrometry (SHS-GC-IMS) to wine aroma analysis and tested the quality grading prediction capability of six machine learning models. In addition, they identified a set of volatile compounds that have been seldom reported in the literature, such as methyl acetate, ethyl format and amyl acetate.

Garde-Cerdan et al. (2021) conducted a study to differentiate Tempranillo and Tempranillo blanco grapes and wines from the La Rioja region (Spain) using machine learning techniques. They determined nitrogen and phenolic compounds and volatile compounds using HPLC and gas chromatography–mass spectrometry (GC-MS), respectively. They then used a machine learning approach for both wine and grape discrimination. According to their results, some of the chemical compounds were useful parameters for both discriminations, whereas others were useful for only one discrimination approach.

Summerson et al. (2021) studied how smoke contamination exposure of grape vineyards affected the volatile aromatic compounds of Cabernet Sauvignon wines. Using the results of an E-nose, they developed two high-precision ANN-based models. One of them predicted smoke aroma intensity from sensory evaluation, and the other predicted the volatile aromatic compounds present in the wine. Nevertheless, in none of these studies, the authors developed a QSOR model for wine aroma.

Finally, Table A6 shows a summary of the computational methods used for QSPR modelling in the works cited in this section. In some cases, the mentioned methods are used for feature selection steps and, in other cases, are used for building regression and classification models. As can be observed, the machine learning approaches were applied in the most recent publications.

Available Data for Aromatic Profile Studies in Wines

A fundamental aspect to address any problem using machine learning techniques is the availability of sufficient quantity and quality of data for effective model training. In particular, for the aromatic profiles of wines, there are challenges due to the scarcity and heterogeneity of most of the available databases.

Wine data sources are classified in the present work as red wines and white wines. At the same time, each category is divided into three sections: studies with sensory analysis and without volatile characterization, studies with aromatic profile and total volatile composition but without sensory analysis, and studies with sensory analysis and volatile characterization. Data sources are also classified in Table 1 and Table A1 (included in supplementary materials); several extraction methods for analytical volatile compounds determination are presented in Table A3; the sensory analyses used in the studied works are described in Table A4; finally mathematical methods for data processing and analysis are shown in Table A5. In addition, the countries of origin of the wines mentioned in the present work are shown in Fig. 1 and the main grape varities are represented in a word cloud in Fig. 2.

Red Wines

Studies with Sensory Analysis and Without Volatile Characterization

Malolactic fermentation (MLF) is encouraged in wine for the purposes of deacidification, converting malic to lactic acid, and increasing aroma and flavour complexity as well as biological stability (Amerine et al., 1982; Davis et al., 1985; Kunkee, 1974). Furthermore, several studies have demonstrated that MLF significantly improves wine flavour. Gámbaro et al. (2001) reported a descriptive sensory analysis on Tannat wine (1999 vintage), with

Table 1 Analysed wine grapes and applied methodologies in each reviewed article

Reference	Grape wines		Sensory analysis: data analysis and reported results									Volatile composition determination			
	Red	White	PCA	CA	MF	GM	ANOVA	QDA	GC-O	HCA	Other methods	GC-FID	GC-MS	HPLC	Other methods
McDaniel et al. (1987)	x			x							x				
Gámbaro et al. (2001)	x		x												
Gámbaro et al. (2003)	x		x								x				
Varela and Gámbaro (2006)	x		x				x				x				
Goldner and Zamora (2007)	x		x				x								
Tao et al. (2009)	x				x										
Campo et al. (2010)	x		x	x						x					
Dimitrov et al. (2018)	x											x			
Dimitrov et al. (2019)	x											x			
Boido et al. (2003)	x											x	x		
Fanzone et al. (2022)	x													x	x
del Barrio Galán et al. (2022)	x												x		x
Miranda-Lopez et al. (1992)	x								x		x	x	x		
Cliff and Dever (1996)	x	x	x									x			
Aznar et al. (2003)	x										x	x	x		
Cullere et al. (2004)	x								x		x	x	x		
Gürbüz et al. (2006)	x								x		x		x		
Escudero et al. (2007)	x				x				x			x	x		
Goldner et al. (2009)	x		x				x								
Gómez García-Carpintero et al. (2011)	x										x	x	x		
Costello et al. (2012)	x						x						x		
Vilanova et al. (2012)	x					x		x					x		
Fariña et al. (2015)	x									x		x	x		
Sánchez-Palomo et al. (2017)	x						x	x					x		
Longo et al. (2020)	x						x						x		x
Denat et al. (2021)	x						x					x			x
Rodriguez et al. (1990)		x					x	x			x			x	
Campo et al. (2008)		x		x						x					
Versini et al. (1994)		x										x	x		
Zhang et al. (2022)		x											x	x	x
Ubeda et al. (2022)		x											x	x	x
Gonzalez-Viñas et al. (1996)		x										x			
González-Viñas et al. (1998)		x	x												
Guth (1997)		x							x				x	x	x

Table 1 (continued)

Reference	Grape wines		Sensory analysis: data analysis and reported results									Volatile composition determination			
	Red	White	PCA	CA	MF	GM	ANOVA	QDA	GC-O	HCA	Other methods	GC-FID	GC-MS	HPLC	Other methods
Guth (1998)		×							×						×
Campo et al. (2005)		×			×		×		×		×	×	×		
Campo et al. (2006)		×			×				×			×	×		
Peťka et al. (2006)		×			×		×		×		×	×			
Sánchez-Palomo et al. (2007)		×										×	×		
Verzera et al. (2008)		×					×								
Muñoz González et al. (2011)		×	×				×				×	×	×		
Vilanova et al. (2013)		×				×	×	×						×	
Ayestarán et al. (2019)		×				×								×	×
Naranjo et al. (2021)		×				×	×						×		×
Sancho-Galán et al. (2022)		×									×	×	×		×

samples from commercial and experimental wines. Panellists were trained to recognize twenty-seven tertiary tier descriptors from the Wine Aroma Wheel proposed by Noble et al. (1987). They analysed their data using principal component analysis (PCA), and found significant differences between MLF and non-MLF wines. MLF resulted in changes in the intensity of various aroma descriptors of Tannat wines: decrease of the secondary descriptors 'berry fruit' and 'fresh vegetative' and of the related tertiary descriptors 'blackcurrant', 'apricot', 'cut green grass' and 'green pepper'.

McDaniel et al. (1987) studied six samples of Pinot noir wine (1981 vintage) with different MLF treatments by applying DA. Panellists were trained to recognize thirty-three tier descriptors from the Wine Aroma Wheel proposed by Noble et al. (1984a), which was divided into groups of terms that describe similar aroma characters. The main groups, primary tier terms, were further divided into specific characters (secondary and tertiary tier terms). Panellists used a 9-point scale to rate the intensity of each wine aroma: (1) none; (9) extreme. They carried out a balanced incomplete block (BIB) experimental design, and data were analysed as presented by Gacula and Singh (1984) for each panellist and the panel as a whole. They found significant differences in twenty of the aroma descriptors, showing that the strain of malolactic bacteria selected for use can affect aroma perception.

Gámbaro et al. (2003) evaluated by DA the aroma properties of fourteen commercial red wines from the Uruguayan market (five Tannat, five Cabernet Sauvignon and four Merlot). The aim was to characterize the aromatic profile of Tannat wines from a sensory approach and compare it with the ones of Merlot and Cabernet Sauvignon from the same Uruguayan regions, to establish which sensory aroma characteristics are unique to the Uruguayan Tannat. They applied generalized Procrustes analysis (GPA) to differentiate among the three groups of samples and obtain sensory attributes that were responsible for these differences. The samples of Tannat were differentiated from the samples of Cabernet and Merlot, and their aroma profiles were characterized by the secondary and tertiary descriptors 'blackcurrant', 'prune', 'oak', 'liquorice' and 'yeasty'.

Accordingly, Varela and Gámbaro (2006) evaluated by sensory DA thirteen Uruguayan Tannat wines. Furthermore, they correlated the quantitative sensory data with the quality assessment obtained from a panel of wine consumers. Aroma profile was determined for each wine by a panel of twenty-two members, using a 9-point structured scale: (1) threshold, (9) very intense. Panellists were trained to recognize the thirty tertiary tier descriptors selected from the Wine Aroma Wheel proposed by Noble et al. (1987). Quality evaluation was performed by an amateur tasting group of thirty wine consumers, using a 9-point structured quality scale: (1) very bad, (9) excellent. PCA and cluster analysis were used to evaluate panel performance, and ANOVA was performed on the sensory panel and the quality panel. According to their results, an increase in the 'dried fruit', 'phenolic' and 'berry' aromas resulted in higher quality scores, but high intensities of 'yeasty', 'burned' and 'earthy' aromas were not desirable in Tannat wines.

Goldner and Zamora (2007) studied the aroma profile of fifty-six Argentine Malbec wines (vintage 2004) from seven regions using DA. Sensory DA was performed by a panel of ten not-sighted assessors, which allowed them an evaluation of the wines only by taste and smell, without the influence of visual attributes. Then, ANOVA was carried out to determine the aroma attributes that were significantly different among wines from different regions, and PCA was used to explore the relationship between aroma attributes and regions. They found significant differences between aroma descriptors and the viticultural regions studied.

Tao et al. (2009) performed a study by sensory analysis of eight vintages (1998–2005) of Cabernet Sauvignon dry red wine from Changli County (China). Panellists scored the intensity of each wine aroma using a 5-point scale: (0) not detected; (1) weak, hardly recognizable note; (2) clear, but weak; (3) clear but not an intense note; (4) intense note. The description of the aroma of each wine was expressed by modified frequency (MF), calculated with the formula proposed by Dravnieks (1985). According to their findings, 32 aromatic descriptors were relevant to aroma characters of the sampled wines, and these wines were characterized by 'blackcurrant', 'green pepper', 'smoke', 'redcurrant', 'cut hay', 'vanilla', 'bilberry', and 'cinnamon' aromas.

Conventional descriptive analysis (conventional DA) has been widely used for sensory profiling of a variety of food products and involves the evaluation of both the qualitative and quantitative sensory characteristics of products by a trained panel. However, although this tool is generally well adapted when applied to simple products, it is less suited to profile complex products, especially when dealing with odour (Lawless, 1999). This could be due to the difficulties of humans to discriminate odour qualities in a mixture (Laing, 1991; Laing & Glemarec, 1992; Marshall et al., 2006) or the limited capacity of humans to reliably differentiate concentration and/or intensity levels in a mixture (Engen & Pfaffmann, 1959). Campo et al. (2010) compared the odour properties of twelve Pinot noir wines (2005 vintage) described by two independent panels that performed, respectively, an intensity-based method (conventional DA) and a citation frequency-based method (Campo et al., 2008). Data from conventional DA and the frequency of citation method were analysed by PCA and correspondence analysis (CA), respectively. Their main objective was to compare both methods when applied to the description of wine samples, based on three criteria: similarity of sensory maps, panel monitoring and practical aspects of each technique. They found that the citation frequency-based method can represent a plausible alternative to conventional DA when a detailed description of a complex aroma product such as wine is required.

On the other hand, the effect of blending Cencibel (Tempranillo) grapes with other grape varieties cultivated in La Mancha region (Spain) (Rojal, Moravia Dulce, and Tortosí) on the wine aroma profile was studied by Sánchez-Palomo et al. (2018). They compared the aroma profile of monovarietal wines with bivarietal ones by means of a sensory panel integrated by fifteen trained judges. The authors applyied ANOVA followed by Student-Newman-Keulstoghether on data aroma descriptors in order to differenciate among the means of chemical data. Finally, they employed PCA for determining the aroma terms of the wines. In sum, they found that wines obtained by blending presented an increased aroma intensity, and consequently an improved aroma complexity.

Studies with Aromatic Profile and Total Volatile Composition but Without Sensory Analysis

Dimitrov et al. (2018, 2019) studied the aromatic profile and total volatile composition of six types of red wines (2017 vintages) from the Central Northern region of Bulgaria, produced by the grape varieties Rubin, Storgozia, Bouquet, Trapezitsa, Kaylashky Rubin and Pinot noir. They obtained the aromatic profiles of the wines by GC with flame ionization detector (GC-FID) analysis, identifying twenty-four volatile compounds: nine esters, eight higher alcohols, one aldehyde and five terpene alcohols. In addition, two succesive vintages of young Tannat (1999–2000), a typical red wine from Uruguay, were characterized by Boido et al. (2003). Using GC-FID and GC-MS, they identified and quantitatively determined fifty-one volatile components, including alcohols, esters, carbonyl compounds, acids, terpenes and norisoprenoids. Studies by Gámbaro et al. (2001) indicate that some of these volatile components can make a sensory contribution, although to understand the impact of these individual compounds that can alter the aroma and flavour of Tannat wine, Boido et al. (2003) suggested that future research on this variety should involve sensory studies on individual constituents within the group of released flavourants and the application of the gas chromatography-olfactometry (GC-O) technique.

Fanzone et al. (2022) evaluated the combined effect of microwave-assisted extraction (MW) application with stem additions in different conditions, before fermentation, on the chemical composition of Bonarda Argentinian wines. The results shown these strategies modified the chromatic characteristics and phenolic composition, enhanced the colour stability and changed the volatile and polysaccharide profile of these wines. Volatile compounds from the wines were extracted using headspace solid phase microextraction

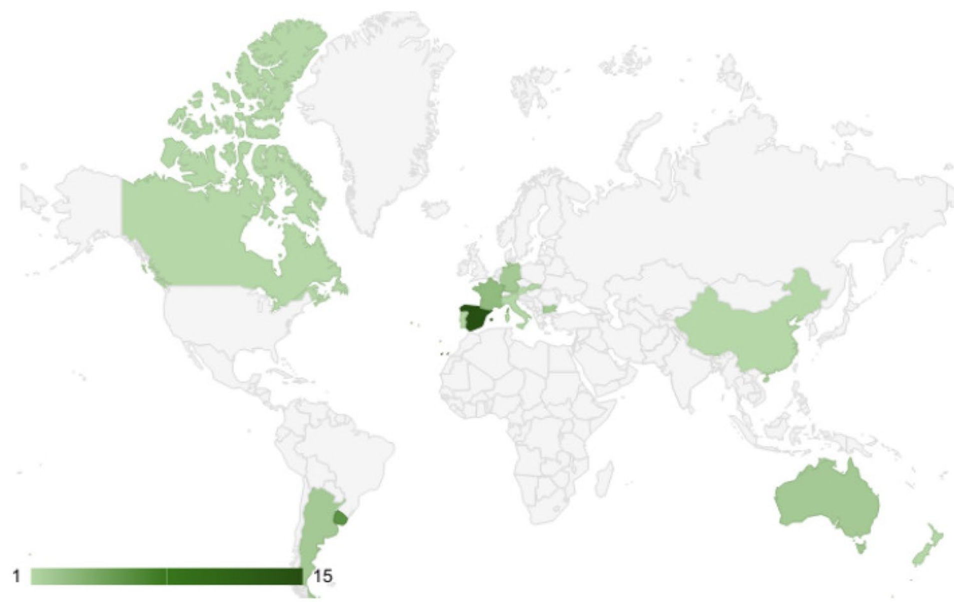

Fig. 1 Country of origin of the wines classified in this article

(HS-SPME) and analysis was performed on high-performance liquid chromatograph equipped with a diode array detector, a quaternary pump, and an autosampler (HPLC-DAD/ESI-MS). del Barrio Galán et al. (2022) studied the volatile profile of the red wines from different Spanish Protected Designations of Origin, which are very closely geographically, and/or categories, to differentiate them. Fifty-three volatile compounds were identified and quantified using the headspace solid-phase micro-extraction technique and gas chromatography-mass spectrometry analysis (HS-SPME-GC/MS).

Studies with Sensory Analysis and Volatile Characterization

The odour profiles of Pinot noir wines from the 1987 and 1988 vintages, and made with grapes harvested at different degree of maturation, were analysed by Miranda-Lopez et al. (1992) using the Osme method. In addition, the results were completed with a quantitive analysis employing GC-MS. In summary, they found that wines at the same level of maturity in different vintages had more differences than similarities and that the odour character varied notably depending on the climatic conditions in a given harvest season.

Cliff and Dever (1996) used sensory and compositional analyses to profile fourteen Pinot noir wines (1992–1993 vintages) produced by ten wineries from British Columbia. They performed the following analyses: titratable acidity, pH, absorbance, phenol and alcohol content, and the sensory profile was evaluated by ten expert wine judges. In addition, they studied the relationships between the wine attributes and samples and classified the wines into groups by performing PCA on the mean scores. As a result, they found good agreement between compositional and sensory analyses.

Aznar et al. (2003) studied fifty-seven Spanish aged red wines from seven different Spanish Denominations of Origin, and developed partial least squares regression (PLSR) models to predict some of the wine aroma nuances from its chemical composition. The sensory panel consisted of fifty-one judges, divided into five committees of eight to twelve professionals. Each of the judges judge tasted up to thirty-six different samples; the measure of intensity of a given sensory term was the frequency with which an aromatic descriptor was used to define a given wine. Chemical composition was obtained by GC-FID and GC-MS, analyzing sixty-nine odourants. Finally, the authors developed models for eighteen sensory terms and twenty-seven chemicals or groups of chemicals obtaining satisfactory models for the most important aromatic descriptors: 'wood-vanillin-cinnamon', 'animal-leather-phenolic', 'toasted-coffee', 'old wood-reduction', 'vegetal-pepper', 'raisin-flowery', 'sweet-candy-cacao', 'fruity' and 'berry fruit'. These models confirmed complex multivariate relationships between chemicals and odours.

Using quantitative GC-O and techniques of quantitative chemical analysis, Cullere et al. (2004) studied the aroma of six premium quality Spanish red wines, revealing the presence of eighty-five aromatic notes in which seventy-eight aroma compounds were identified. The GC-O study was carried out by eight trained judges, who were asked to measure the overall intensity of each odour using a 0–3 scale, being half values allowed, using a GC equipped with a FID and a sniffing port. Furthermore, microextraction and GC-FID analysis, and solid phase extraction (SPE) and GC-ion trap MS analysis were employed for quantitative analysis. The authors found that the number of components present in concentrations above their threshold value was quite substantial, reporting also the presence of 1-nonen-3-one (temptatively) and 2-acetylpyrazinesare in wine aroma for the first time. Moreover, using time-intensity GC-O and GC-MS, Gürbüz et al. (2006) analysed the volatiles of Merlot and Cabernet Sauvignon wines produced in California and Australia, noting the presence of seventy-four aroma active compounds. With a GC equipped with a FID and a sniffing port, they employed two experienced olfactory assessors, and each assessor sniffed at least three times each sample. They estimated the overall wine aroma by grouping the aroma active compounds into nine odour categories based on similar odour descriptors, using the Wine Aroma Wheel of Noble et al. (1987). Then, by GC-MS, they identified sixty-six volatiles GC-MS, twenty-eight esters and nineteen minor alcohols. According to their results, the most relevant compounds were ethanol, ethyl octanoate, ethyl decanoate, ethyl acetate, 3-methyl-1-butanol, ethyl hexanoate, diethyl succinate, and 2-phenylethanol. Although there were distinct quantitative differences between Merlot and Cabernet wines, the relative aroma category profiles of the four wines were similar, characterized by high 'fruity', 'caramel', 'green' and 'earthy' aroma.

Using sensory descriptive analysis, quantitative GC-O and chemical quantitative analysis, Escudero et al. (2007) studied the aroma profile of five premium red wines. The sensory panel that performed the sensory descriptive analysis of the wines, consisting of nine judges, determined the nine aroma terms that were best suited for further descriptive analysis of the selected wines: 'raisin', 'berry fruit', 'veggie', 'phenolic', 'toasted', 'woody', 'alcohol', 'sweet' and 'reduction'. They scored the intensity of each attribute using a 7-point scale: (0) non-detected; (1) weak, hardly recognizable note; (2) clear but not intense note; (3) intense note, being half values allowed. Another panel consisting of twelve assessors performed the sensory evaluation of samples spiked with aroma compounds. The data processed were a mixture of intensity and frequency of detection (MF) according to Dravnieks (1985). The quantitative chemical analysis involved liquid-liquid microextraction (LLM) and GC-FID analyses, and SPE and GC-ion trap-MS analyses. They obtained in most cases satisfactory agreement between GC-O and quantitative data, identifiyng forty-five odourants and at least thirty-seven odourants at concentrations above their odour threshold. The most relevant findings were confirmed by sensory analysis.

Goldner et al. (2009) studied how ethanol level affected the aroma attributes and volatile compounds in twenty-three samples of Malbec wines. Volatile compounds were analysed by SPME and GC-MS, and sensory DA was performed by a panel of ten blind assessors using 9-point intensity scales. ANOVA was carried out to determine the characteristic aroma attributes and volatile compounds in wines with different levels of alcohol. Then, using PLS, they confirmed the influence of alcohol level on the presence of volatile compounds and the perception of aroma descriptors in wine.

Gómez García-Carpintero et al. (2011) presented results from the first experiment performed on the free and bound aroma compounds of Moravia Agria wines, a minority grape variety cultivated in the Castilla La Mancha region. They studied five consecutive vintages (2004–2008) of these wines using GC-MS analysis and sensory analysis performed by a trained panel of fifteen experienced wine testers to determine the influence of grape variety on the aroma of wine. Quantitative analysis was carried out using a GC-FID for major volatiles and SPE to isolate free and glycosidically bound aroma compounds for subsequent analysis by GC-MS. Over these five consecutive harvests, they identified and quantified ninety-two free aroma compounds and sixty-seven bound aroma compounds and classified the odour activity values (OAVs) for the different compounds into seven odourants. DA was applied to evaluate the sensory profile of Moravia Agria wines, considering eight olfactive attributes as the ones that best described the aroma characteristics of Moravia Agria wines. To rate the intensity of each attribute, the judges used a unstructured scale, in which the left-hand end of the scale was 'attribute not perceptible' and the right-hand end was 'attribute strongly perceptible'. They found that although some aromatic series such as 'floral', 'green' and 'fresh' were the minor aroma categories according to GC-MS analysis, they were among the most characteristic attributes in the sensory profile of Moravia Agria wines. In addition, even though the aromatic series including 'pungent', 'chemical', 'fatty' and 'dry' aromas constitute a major aroma category (according to quantitative chemical analysis), these attributes were not detected in the sensory flavour profile by the authors. This was attributed to several factors that, when combined, may alter the intensity of the descriptors, masking the descriptors of some aromatic series and increasing the intensity of other odour descriptors. Finally, their results are in agreement with the results obtained by Gürbüz et al. (2006) in red wines made from Merlot and Cabernet Sauvignon grape varieties.

Costello et al. (2012) studied the effects on the chemical and sensory impacts of MLF in Cabernet Sauvignon wine. MLF was conducted on two styles of Cabernet Sauvignon wine, each of them divided into four equal volumes, three of which were inoculated with three different commercial *Oenococcus oeni* strains, and the fourth used as a non-MLF control. Both Cabernet Sauvignon wine sets were studied separately by a sensory panel of eleven participants (set 1) and thirteen participants (set 2) trained by the same panel leader. The panel selected twenty-four sensory attributes to distinguish the wines: appearance (two attributes), aroma (twelve attributes) and palate (ten attributes). For each attribute, an ANOVA analysis was conducted, useful to evaluate the effect of the MLF treatment and fermentation replicate nested within the MLF treatment, controlling for the effect of presentation replicate and judge. Seven attributes were affected in wines from set 1 ('dark fruit aroma', 'savoury aroma', 'overall fruit flavour', 'green flavour', 'savoury flavour', 'viscosity' and 'bitterness') and six in wines from set 2 ('colour intensity', 'overall fruit flavour', 'hotness', 'coarseness', 'astringency' and 'bitterness'). Two of them were common to both sets of wines: 'overall fruit flavour' and 'bitterness'. They quantified twenty-eight fermentation-derived volatile compounds in each wine using SPME, stable isotope dilution analysis and GC-MS. Then, they conducted PLS regression analysis to model the sensory data as a dependent variable with the compositional data. This analysis revealed a strong correlation between important chemical components and sensory attributes, including overall 'fruit flavour' and 'dark fruit aroma'.

The volatile composition and aroma of five red wines were studied during three consecutive vintages (2007–2009) by Vilanova et al. (2012). Brancellao, Mencía, Merenzao, Mouratón and Sousón, traditional red cultivars grown in Galicia, were characterized using GC-MS and sensory anaysis. In addition, volatile and sensory results were analysed using PLSR. Sensory analysis was carried out by a panel of eight judges using the Quantitative Descriptive Analysis (QDA) methodology (Lawless & Heymann, 2010). Panellists scored the intensity of each aromatic attribute using a 9-point scale. The frequency, intensity and geometric mean (%GM) of the different descriptors were calculated for each wine, and ANOVA was performed on the attribute intensity scores. They identified twenty out of fifty-one volatile compounds at concentrations higher than their corresponding odour thresholds, which contributed to the final wine aroma, and analysed the relationships between volatile composition and aromatic descriptors by applying PLSR. They obtained a satisfactory model for the prediction of four important aroma descriptors in these wines: 'aroma quality', 'aroma intensity', 'herbaceous' and 'red fruit'.

The aroma profiles of ten Uruguayan Tannat wines were characterized by Fariña et al. (2015). Volatile composition was studied by GC-MS, identifying and quantifying sixty-two volatile compounds, being alcohols and esters the most present. Sensory characterization was performed by a panel of thirty wine professionals using projective mapping, and these data were analysed using multiple factor analysis (MFA). The aroma descriptors used by the panellists in the description phase were qualitatively analysed and grouped according to the Wine Aroma Wheel of Noble et al. (1987). PLSR was used to study the relationship between volatile composition and sensory characterization. They found that the most important sensory descriptors, namely 'woody', 'earthy', 'phenolic', 'sulfur', 'chemical' and 'microbiological', were related to volatile composition. Their aroma profile results were in agreement with those reported by Varela and Gámbaro (2006).

Sánchez-Palomo et al. (2017) characterized by chemical and sensory analysis four consecutive vintages (2013–2016) of the aroma of Malbec red wines from La Mancha. Volatile composition was determined by SPE and GC-MS: they identified seventy-nine aroma compounds. Sensory analysis was carried out by a panel of ten experts that selected the nine aroma attributes that best describe the main characteristics of these wines and the differences among them. Panellists rated the intensity of each attribute using a unstructured scale, in which the left-hand end of the scale was 'none/weak' and the right-hand end was 'strong'. QDA was conducted to identify the sensory aroma profile, and ANOVA was performed for each judge and each attribute. They found that the sensory aroma of these wines was characterized by 'red fruit', 'clove', 'caramel', 'liquorice', 'leather', 'tobacco' and 'coffee' aroma descriptors. Their results were consistent with the ones obtained by Gómez García-Carpintero et al. (2011), Gürbüz et al. (2006) for red wines made from the Moravia Agria, Merlot and Cabernet Sauvignon grape varieties.

Longo et al. (2020) studied colour components, volatile and sensory attributes of 15 Australian Pinot noir wines, in order to discriminate them according to their region of origin. A HS-SPME (headspace solid-phase micro-extraction) GC-MS method was used to quantify 28 fermentative compounds and all of the datasets were analysed by ANOVA. Sensory analysis was performed by a panel of 11 judges using a Pivot© profile sensory method (Thuillier et al., 2015), a frequency-based descriptive method based on free description. The sensory panel generated 53 descriptors: 10 appearance, 22 aroma and 21 palate, but only 14 attributes (four appearance, and five for both aroma and palate) were included in the CA. They found that the region of origin is a strong driver of aroma typicity of wine. For these wines, ethyl decanoate, ethyl 2-methylpropanoate, ethyl 2-methylbutanoate, and decanoic acid appeared to significantly contribute to the distinctiveness of the producing regions, and the most relevant sensory attributes were 'red fruits', 'floral' and 'oaky' aromas, 'acidic', 'astringent', 'complex' and 'soft' palate descriptors.

Denat et al. (2021) studied the effects after fermentation and after ageing on the sensory and chemical aroma profiles of Tempranillo wine from Rioja, Spain (vintage 2019). They also used two *S. cerevisiae* strains. For twelve samples, volatile composition was determined by SPE and GC-FID, and aroma compounds were arranged into aroma vectors. Aroma vectors, according to Ferreira et al. (2022), are groups of aroma compounds which share chemical and sensory characteristics. They identified 17 aroma vectors. They did two sensory studies: a sorting task carried out by twenty judges without previous training, and a descriptive analysis by flash profile with a trained panel of fourteen judges. They found that the strain of yeast introduces significant and consistent differences. Not only sensory profile but also 11 of the aroma vector profiles were affected. Particularly, wines made with one one of the strains contained higher levels of ethyl esters, acetic acid, cinnamates and ethyl acetate (related to black and fresh fruit), and lower levels of linear fatty acids, β-damascenone, acetaldehyde, higher alcohols and lactones (related to white and compote fruits).

Finally, a compilation of considered key compounds in wine aroma was presented by Francis and Newton (2005), who provided an overview of the relationships between wine sensory properties and volatile chemical composition obtained by GC-O for rosé, young red and aged red wines, and also reported their odour threshold.

White Wines

Studies with Sensory Analysis and Without Volatile Characterization

Sauvageot and Vivier (1987) also studied the differences between MLF and no-MLF Chardonnay wines (1991–1992 vintages). MLF wines were perceived higher in 'hazelnut', 'fresh bread' and 'dried fruit' aromas, whereas non-MLF wines kept specific aromas such as apple and grapefruit-orange. Also for Chardonnay wine and according to Gámbaro et al. (2001), after MLF, Laurent et al. (1994) found a significant decrease of the attributes 'berry fruit' and 'floral', and an increase of the aromatic descriptors 'earthy' and 'buttery'. The increase of the 'buttery' character after MLF in Chardonnay wine was also reported by Rodriguez et al. (1990), after an experienced panel of seven judges measured four aromatic characters: 'fruitiness', 'butteriness', 'cheesiness' and 'sauerkraut'. Panellists also determined the acidity level of each sample. QDA measurements were made and analysed using ANOVA and Tukey's honestly significant difference (HSD) test (Stone & Sidel, 2004). Campo et al. (2008) studied twenty-three commercial monovarietal young white wines (2004 vintage) produced in Spanish regions from different grape varieties (Albariño, Chardonnay, Godello, Macabeo, Sauvignon Blanc, Verdejo, Treixadura, Xarello and Palomino Fina). A panel of thirty-two judges sorted the wine samples into groups on the basis of odour similarity, followed by sensory DA based on citation frequencies for each sample. DA was analysed by CA and HCA. They found that only Verdejo and Sauvignon Blanc varietals were similarly perceived and described by the panel, being tropical fruit the major descriptor and that no other clear grouping was observed for the rest of grape varieties.

Nanou et al. (2020), in turn, studied sensory aroma profiles of white wines of the indigenous Greek grape varieties Assyrtiko, Malagousia, Moschofilero and Roditis. They employed a panel integrated by twenty-three trained members and applied a frequency of attribute citation method for identifying the aromatic descriptors, besides different descriptive analysis for testing panel performance, experiment reproducibility, and so on. They found banana and vanila profiles for Roditis variety; lemon, grapefruit, and citrus blossom character as descriptors for Assyrtiko variety; and floral odour profile, particularlly, rose, jasmine or more citrus blossom-like for Moschofilero variety. In addition, Malagousia and Assyrtiko presented some common aromatic descriptors such as mushroom and earthy, whereas Malagousia and Moschofilero shared floral and citrus notes.

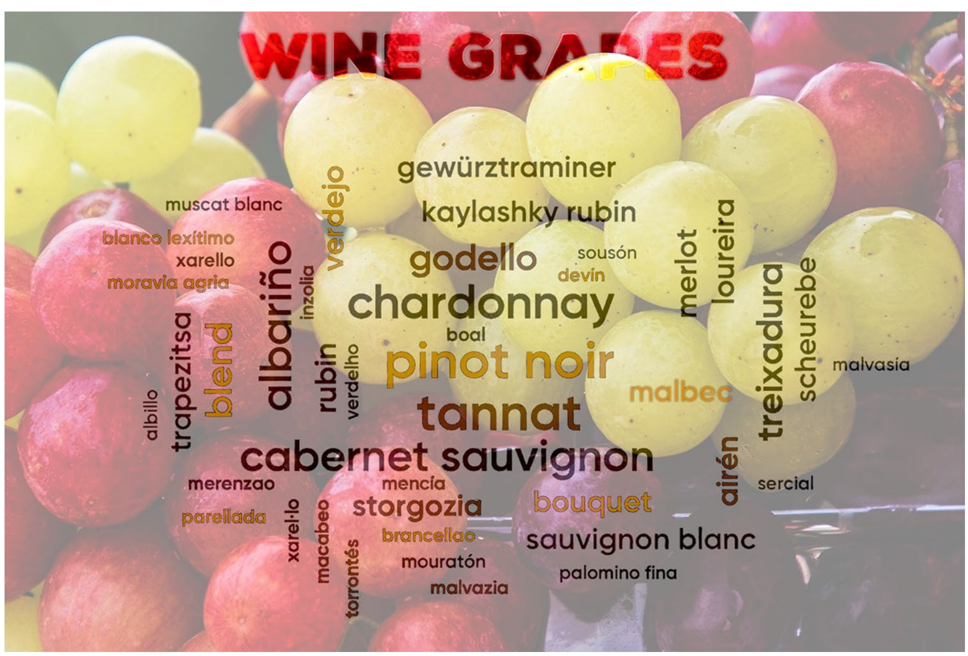

Fig. 2 Grapes from which the wines classified in this article come from

Studies with Aromatic Profile and Total Volatile Composition but Without a Panel

Three aroma categories, volatile compounds (except for monoterpenols), monoterpenols and bound forms were analysed in depth by Versini et al. (1994) using GC-MS and GC-FID for wines from the three most interesting Galician white grape varieties (Albariño, Loureira and Godello). Data about young white wines were also compiled by Francis and Newton (2005) in their review.

Ubeda et al. (2022) subjected a Chilean Sauvignon Blanc wine to a maturation during 6 months by using four different types of vessels, and subsequently wines were bottled using three different closures (natural cork, synthetic cork, and screwcaps). The volatile compound profiles of the wine samples were recorded by SPME-GC-MS throughout vessel maturation as well as after the bottle storage period. They conclude the selection of vessels and closures during wine maturation could be employed as a tool by winemakers to modulate wine features and could be a helpful strategy to mitigate the side effects of uncontrolled elements such as climate change.

Zhang et al. (2022) investigated the effects of triple mixed culture of *T. delbrueckii*, *H. vineae*, and *S. cerevisiae* with different inoculation ratios on the basic wine parameters, aroma profiles, biogenic amines and phenolic compounds in Petit Manseng wines. The results collectively indicated that co-inoculation of *T. delbrueckii* and *H. vineae* is an effective method to make up the species shortages and further improve the overall quality of wines. The technique used for identification of volatile compounds is micro-extraction coupled with gas chromatography-mass spectrometry (HS-SPME-GC-MS). The phenolic compounds were separated and analysed using high-performance liquid chromatography triple-quadrupole tandem mass spectrometry (HPLC-QqQ-MS/MS).

Studies with Sensory Analysis and Volatile Characterization

Cliff and Dever (1996) studied Chardonnay wines (1988–1993 vintages) produced by wineries from British Columbia by sensory and compositional analyses. For each sample, titratable acidity, pH, absorbance, phenol and alcohol content were determined, and sensory profile was performed by a panel of ten expert wine judges. Data were analysed by performing PCA on the mean scores of sixteen samples of wines to assess the relationship between attributes and wines, and they found good agreement between compositional and sensory analyses. They also used discriminant analyses on twenty-five samples of wine to determine the linear combination of variables that distinguished between vintages and wineries, and obtained that discriminant analysis effectively distinguished between the vintages: in general, the 1992–1993 vintages were more 'fruity' than the 1988–1991 vintages. In addition, they found specific winery 'styles': some wineries produced 'yellow', 'oaky', 'buttery' and 'astringent' wines, whereas others produced wines that were more 'fruity' and 'floral'.

Guth (1997) and (1998) studied the aroma profile of two wine varieties: Gewürztraminer and Scheurebe. Volatile composition was obtained by GC-MS and HPLC. Sensory profile was determined by a panel of six assessors using GC-O and a 3-point scale: (0) none, (1) weak, (2) medium, (3) strong. In addition, Guth (1998) studied the influence of barrel ageing on the flavour of Gewürztraminer wine. They found the presence of 4-Mercapto-4-methylpentan-2-one only in Schreube wines, whereas *cis*-rose oxide was one of the most relevant odourants present only in Gewürztraminer wines.

To study the effect of in-bottle storage time on the organoleptic characteristics of wines, González-Viñas et al. (1998) performed a sensory analysis of the aroma attributes of twenty-six young Airén white wines (1991–1994 vintages), stored under usual commercial conditions for 6, 18, 30 and 42 months. Another aim of their work was to evaluate the shelf life of wines in the absence of careful storage conditions. A panel of twelve assessors selected seven attributes that best suited to the differences between wines stored in a bottle for different time periods. Then, six aroma attributes were evaluated for each wine sample: 'fresh-citric', 'floral', 'apple', 'spicy-green pepper', 'banana' and 'sweet-raisin-prune'. The judges rated the intensity of each attribute using a unstructured scale, in which the left-hand end of the scale was 'attribute not perceptible' and the right-hand end was 'attribute strongly perceptible'. PCA was applied to this data: according to their results, 'fresh-citric' aroma disappeared after 18 months, and 'spicy-green pepper' and 'sweet-raisin-prune' aromas became noticeable to consumers. This study is related with an earlier study by Gonzalez-Viñas et al. (1996), in which four consecutive vintages (1990–1993) of Airén wines were analysed using GC-FID for major and minor volatile determination. They found significant changes in composition within 12 months of storage in bottles as well as in subsequent months.

Using GC-O, Campo et al. (2005) developed a PLSR model to predict wine sensory properties of six monovarietal young white wines (2001 vintage). Monovarietal wines were selected from the following grapes: Albariño, Godello, Malvasía, Parellada, Treixadura and Verdejo. The volatile composition of wines was determined by GC-FID and GC-MS, and a panel of eight judges scored the ten aroma terms selected for descriptive analysis using a 4-point scale: (0) not detected; (1) weak, hardly recognizable note; (2) clear but not intense note; (3) intense note. Data were processed using MF (Dravnieks, 1985). Then, DA data were analysed by CA, χ^2 analysis and ANOVA, in which the considered factors were judges and wine varieties. Finally, PLSR models were validated by sensory analysis performed by a panel composed of twelve judges, confirming most of the sensory descriptor predictions described by the model. Then, by sensory analysis, GC-O and GC-MS, Campo et al. (2006) studied the aroma profile of Malvazia, Boal, Verdelho and Sercial monovarietal wines, four of the most emblematic grape varieties from Madeira, Portugal. The sensory panel, composed of eight judges, determined the seven aroma terms that best described the selected wines for further DA: 'dried fruit', 'candy', 'lacquer', 'nutty', 'maderized', 'toasty' and 'spicy'. The intensity of each attribute was scored using a 7-point scale: (0) non-detected; (1) weak, hardly recognizable note; (2) clear but not intense note; (3) intense note, being half values allowed. The data processed were a mixture of intensity and frequency of detection (MF) according to Dravnieks (1985). Quantitative analysis was carried out using GC-FID, SPE and GC-ion trap-MS analyses. Furthermore, to identify the odourants specifically related to the process of elaboration of Madeira wines, GC-O profiles of the four wines were compared to the GC-O profiles obtained from three young white monovarietal wines elaborated with Malvasia, Boal and Verdello, respectively. Their GC-MS results confirmed most of the results of the GC-O study, leading to the conclusion that GC-O is a useful tool for the detection of presence of active odourants in wine.

Peťka et al. (2006) studied the aroma profile of a Slovakian white wine made from Devín grapes, using GC-O profile and sensory characterization performed by a panel of judges. GC-O sniffings were carried out by a panel of eight judges, and the profile revealed that this wine aroma resembles a mixture of Gewürztraminer, Sauvignon Blanc and Muscat. A panel of thirteen experts evaluated the orthonasal, retronasal and residual wine aroma, using a list of twenty-five aromatic descriptors previously agreed and a 7-point scale: (0) non-detected; (1) weak, hardly recognizable note; (2) clear but not very intense odour; (3) extremely strong odour, being half values allowed. Data were processed using adjusted frequencies (MF) according to Dravnieks (1985) combined with ANOVA and GPA. They found that aromatic profile was primarily 'Muscat' and intense 'fruity', 'sweet', and 'herbaceous' notes.

Using GC-FID and GC-MS, Sánchez-Palomo et al. (2007) studied the volatile composition of must and wines from Muscat 'à petits grains' and Albillo grape cultivars harvested at different degree of maturation (2001–2002 vintages). They found higher ester and fatty acid concentrations in wines with a low degree of maturity, but less terpene compounds and benzene derivatives than in wines from more mature grapes. Although they do not provide information about the panel composition, the judges performed a sensory profile for each wine. Wines obtained from Albillo varieties had 'citric' and 'fruity' aromas with 'floral' notes, whereas wines made from Muscat 'à petits grains' had a 'muscat' aroma, and 'fresh' and 'fruity' odour.

The volatile composition and sensory profile of wines made from Inzolia grape, which is one of the most widespread native white grapes in Sicily, Italy, were characterized by Verzera et al. (2008). Twelve samples from 2006 vintage were studied. They identified fifty-six volatile compounds and alcohols, esters, terpenes and fatty acids using HS-SPME/GC-MS. Sensory characterization was performed by a panel of ten judges, who evaluated eight descriptors selected by the panellists as the most representative of these wines. Aromatic attributes were defined according to the Wine Aroma Wheel proposed by Noble et al. (1987). Sensory attributes were quantified using a 9-point intensity scale and analysed using ANOVA. Wine samples were described mainly with the descriptors 'fruity', 'banana', 'rippened apple', 'floral', 'acid' and 'pungent', which was in agreement with volatile composition.

Muñoz González et al. (2011) studied the volatile and sensory characterization of twenty-five Xarel·lo monovarietal white wines (2005–2008 vintages), a representative wine from Penedés region (Catalonia, Spain). By GC-FID and HS-SPME-GC-MS, they found fifty-nine aroma compounds in the wine samples, although some of them were not detected in every sample. Only twenty-five volatile compounds were found in more than 90% of the samples. DA was performed by a panel composed by twelve experts, who selected the sixteen best suited descriptors to characterize these wines and used a 10-point scale to rate their presence in each sample, being (0) descriptor not perceived and (9) highest intensity the bottom and upper bounds, respectively. In addition, they applied several statistical methods for the data analysis (ANOVA and Scheffe test, PCA, cluster analysis) and found considerable differences in the sensory attributes and volatile composition between aged and younger wines. Finally, the relationship between the sensory attributes and volatile composition was predicted by applying PLS.

In addition, the influence of geographical location on volatile composition and perceived flavour of Sauvignon Blanc wines from Marlborough (New Zealand), Sancerre, Loire and Saint Bris (France), and Styria (Austria) was studied by Green et al. (2011). Sensory analyses of eighteen commercial Sauvignon Blanc wines (2006–2007 vintages) were performed by a panel of nineteen judges using eleven experimenter-provided flavour descriptors and analysed using HCA. Chemical analyses were determined using HS-SPME, SPME-GC-MS and SPME-GC-MS/MS. Their results indicated that wines from New Zealand were dominated by 'green' characteristics ('green capsicum', 'herbaceous', 'grassy' and 'leafy'), Austrian wines were perceived as 'fruity' ('tropical' and 'stonefruit'), and French wines received relatively high intensity ratings for 'mineral-smoky-flinty' aromas.

Using GC-MS and QDA, Vilanova et al. (2013) studied the aroma and volatile composition of wines from five white grape cultivars from northwestern Spain (Loureira, Blanco lexítimo, Torrontés, Treixadura and Albariño) during three consecutive vintages (2007–2009). The sensory evaluation of the wines was performed by eight wine tasters using DA, who scored the intensity of each attribute using a 9-point scale. The frequency, intensity and GM of each of the thirty-six descriptors were calculated for each wine. They also performed ANOVA on the individual intensity scores of attributes and found that the effect of wine was significant for the terms 'gold color', 'odor intensity', 'floral', 'herbaceous', and 'ripe fruit' aromas, 'balanced', 'sweetness' and 'acidity', 'body and global value'. This means that these ten terms were useful for characterizing differences among the five cultivars. They identified twenty out of forty-six volatile compounds at concentrations higher than their corresponding odour thresholds, thus contributing to the final wine aroma, and analysed the relationships between volatile composition and aromatic descriptors by applying of PLSR. They obtained a satisfactory model for the prediction of three important aroma descriptors in this set of wines, 'floral', 'herbaceous' and 'ripe fruit' aromas, and aroma intensity from instrumental analysis data.

Ayestarán et al. (2019) studied the effect of maceration on the volatile composition and aroma characteristics of Tempranillo Blanco wines during three consecutive vintages (2014–2016). Sensory analysis was performed by an experienced panel using a total of nine aromatic descriptors. In addition, the GM of each one was calculated, and GC-ion trap and CG-MS were employed to determine the volatile composition. They analysed the relationships between volatile compounds and sensory analysis using PLS, concluding that while wines with carbonic maceration process were more aromatic and had 'ripe fruit' descriptors, conventionally made wines were associated with 'citrus', 'tropical' and 'seed fruit' aromatic descriptors.

Naranjo et al. (2021) studied the volatile composition and aroma characteristics of Maturana Blanca wines (vintage 2018), an autochthonous minor variety of grape from Rioja, Spain. Volatile composition was determined by GC-MS after liquid-liquid extraction (LLE), quantifying 33 volatile compounds. The sensory evaluation of the wines was carried out by thirteen wine tasters. The sensory panel selected a consensual group of descriptors, and the GM of each one of the descriptors was calculated. They found that Maturana Blanca wines were influenced the most by acetates, ethyl esters, 2-phenylethanol and γ-decalactone, which are related to 'fruity' and 'floral' aromas and 'spicy' notes. Sancho-Galán et al. (2022) also studied an autochthonous cultivar from Spain, specifically from Cadiz, and the influence of the grape over-ripening in the production of white wines for a particular grape. They characterized the volatile composition and sensory profile of wines made from Palomino Fino. Volatile composition was determined by GC-FID (for major volatile compounds) and GC-MS after SPE (for free minor volatile compounds). Sensory analysis was performed by a 20-member panel, using an 11-point scale, with 0 points representing the lowest score and 10 points representing the highest score of the aroma attributes selected. They found that grape over-ripening implies modifications in the volatile composition of wines, and that there are differences between grapes that have been over-ripened naturally or in the sun versus those that have been over-ripened in a climatic chamber under controlled conditions. The wines that were made with overripe grapes were dominated by fruity and floral notes.

Country of origin, most relevant aromas and volatile compounds of wines of the studied works are detailed in Supplementary Information.

Conclusions

A review of the available datasets about wine features, chemical composition and aromatic profiles was performed in the present paper. In addition, the state of the art on the use of machine learning–based models to predict characteristics of wines was carried out. As a first conclusion, as far as we have found, none of the reviewed works presented a study on the relationship between the quantitative sensory analysis of wines and their volatile composition by applying the QSOR modelling approach. Additionally, the homogeneous datasets of wines available in the literature is scarce, and those that can be found contain a reduced amount of information. This makes the development of predictive models based on supervised machine learning approaches considerably complicated.

In this regard, there are several requirements and guidelines that a dataset must meet in order to allow the generation of reliable models. The first one is the homogeneity in the processes and standards used for data acquisition. We have seen a great diversity of criteria both in the definition of the descriptive variables of the wines, such as their chemical compositions, and in the characterization of the variables to be predicted, whether they are aromatic descriptors or other types of objective variables. This heterogeneity in the data makes it impossible to integrate most of the datasets that are available in the literature. Another important guideline is that the variety and quantity of wines represented in the datasets should be large enough to guarantee a wide domain of applicability of the QSOR models, because models derived from small datasets are going to have a very limited scope and practical use in the industry, even when they have been properly trained.

Nevertheless, as we mentioned in the "Introduction", in 2021, a key progress for developing QSOR modelling approach was published by Sharma et al. They applied their QSOR model to an independent test set of chemical compounds and achieved high accuracy on smell predictions. Furthermore, the same research group published OlfactionBase, a free, open-access web server in which the knowledge about many aspects of the olfaction mechanism is gathered, containing detailed information of components such as odours and odourants, among other aspects. Using this new database, we think that the aromatic profiles of wines can be predicted transitively. In other words, whereas traditional technologies determine the composition of wine in terms of its chemical compounds, aromatic descriptors and their intensity might be inferred by machine learning models. This is the most relevant finding of this review, which can give future directions to the community working in food informatics applied to the wine industry.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11947-022-02836-x.

Country of origin, most relevant aromas and volatile compounds of wines of the studied works are presented as supplementary material to the present work.

Acknowledgements This project is partially supported by PICT 2019-03350, ANPCyT, Argentina. The authors also acknowledge Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Universidad Nacional de Río Negro (UNRN) and Universidad Nacional del Sur (UNS) for the facilities employed for materializing this work.

Author Contribution Not applicable.

Availability of Data and Material Not applicable.

Code Availability Not applicable.

Declarations

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Competing Interests The authors declare no competing interests.

References

Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician, 46*(3), 175–185.

Amerine, M., Berg, H. W., & Cruess, W. V. (1982). Chapter 16 - Bacteria in wine. In M. Amerine & R. Kunkee (Eds.) *Technology of wine making* (4th ed., pp 557–581). Avi Publishing, Westport, Connecticut. https://www.sciencedirect.com/science/article/pii/B9780080918310500102

Atanasova, B., Thomas-Danguin, T., Chabanet, C., et al. (2005). Perceptual interactions in odour mixtures: Odour quality in binary mixtures of woody and fruity wine odorants. *Chemical Senses, 30,* 209–17. https://doi.org/10.1093/chemse/bji016

Ayestarán, B., Martínez-Lapuente, L., Guadalupe, Z., et al. (2019). Effect of the winemaking process on the volatile composition and aromatic profile of Tempranillo Blanco wines. *Food Chemistry, 276*, 187–194. https://doi.org/10.1016/j.foodchem.2018.10.013. https://www.sciencedirect.com/science/article/pii/S0308814618317783

Aznar, M., Lopez, R., Cacho, J., et al. (2003). Prediction of aged red wine aroma properties from aroma chemical composition. Partial least squares regression models. *Journal of Agricultural and Food Chemistry, 51,* 2700–7. https://doi.org/10.1021/jf026115z

Boido, E., Lloret, A., Medina Rland, K., et al. (2003). Aroma composition of Vitis vinifera cv. Tannat: The typical red wine from Uruguay. *Journal of Agricultural and Food Chemistry, 51,* 5408–13. https://doi.org/10.1021/jf030087i

Brereton, R. G., & Lloyd, G. R. (2014). Partial least squares discriminant analysis: Taking the magic away. *Journal of Chemometrics, 28*(4), 213–225. https://doi.org/10.1002/cem.2609

Camaréna, S. (2020). Artificial intelligence in the design of the transitions to sustainable food systems. *Journal of Cleaner Production, 271*, 122–574. https://doi.org/10.1016/j.jclepro.2020.122574. https://www.sciencedirect.com/science/article/pii/S0959652620326214

Campo, E., Ferreira, V., Escudero, A., et al. (2005). Prediction of the wine sensory properties related to grape variety from dynamic-headspace gas chromatography-olfactometry data. *Journal of Agricultural and Food Chemistry, 53,* 5682–90. https://doi.org/10.1021/jf047870a

Campo, E., Ferreira, V., Escudero, A., et al. (2006). Quantitative gas chromatography-olfactometry and chemical quantitative study of the aroma of four Madeira wines. *Analytica Chimica Acta, 563,* 180–187. https://doi.org/10.1016/j.aca.2005.10.035

Campo, E., Do, B., Ferreira, V., et al. (2008). Aroma properties of young Spanish monovarietal white wines: A study using sorting task, list of terms and frequency of citation. *Australian Journal of Grape and Wine Research, 14*(2), 104–115. https://doi.org/10.1111/j.1755-0238.2008.00010.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1755-0238.2008.00010.x

Campo, E., Ballester, J., Langlois, J., et al. (2010). Comparison of conventional descriptive analysis and a citation frequency-based descriptive method for odor profiling: An application to burgundy pinot noir wines. *Food Quality and Preference, 21,* 44–55. https://doi.org/10.1016/j.foodqual.2009.08.001

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp 785–794)

Cliff, M., & Dever, M. (1996). Sensory and compositional profiles of British Columbia Chardonnay and pinot noir wines. *Food Research International, 29,* 317–323. https://doi.org/10.1016/0963-9969(96)83271-6

Cortes, C., & Vapnik, V. (1995). Support-vector networks. In *Machine Learning* (pp 273–297)

Costello, P., Francis, I., & Bartowsky, E. (2012). Variations in the effect of malolactic fermentation on the chemical and sensory properties of Cabernet Sauvignon wine: interactive influences of *Oenococcus oeni* strain and wine matrix composition. *Australian Journal of Grape and Wine Research, 18*(3), 287–301. https://doi.org/10.1111/j.1755-0238.2012.00196.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1755-0238.2012.00196.x

Cravero, F., Schustik, S. A., Martínez, M. J., et al. (2019). Computer-aided design of polymeric materials: Computational study for characterization of databases for prediction of mechanical properties under polydispersity. *Chemometrics and Intelligent Laboratory Systems, 191*, 65–72. https://doi.org/10.1016/j.chemolab.2019.06.006. https://www.sciencedirect.com/science/article/pii/S0169743919301509

Cullere, L., Escudero, A., Cacho, J., et al. (2004). Gas chromatography-olfactometry and chemical quantitative study of the aroma of six premium quality Spanish aged red wines. *Journal of Agricultural and Food Chemistry, 52,* 1653–60. https://doi.org/10.1021/jf0350820

Davis, C., Wibowo, D., Eschenbruch, R., et al. (1985). Practical implications of malolactic fermentation: A review. *American Journal of Enology and Viticulture, 36,* 290–301.

del Barrio Galán, R., Bueno-Herrera, M., de la Cuesta, P. L., et al. (2022). Volatile composition of Spanish red wines: Effect of origin and aging time. *European Food Research and Technology*, 1–14

Denat, M., Pérez, D., Heras, J., et al. (2021). Impact of two yeast strains on Tempranillo red wine aroma profiles throughout accelerated ageing. *OENO One, 55*, 181–195. https://doi.org/10.20870/oeno-one.2021.55.4.4732

Di Rosa, A. R., Leone, F., & Chiofalo, V. (2020). Electronic noses and tongues. In Y. Pico (Ed.) *Chemical analysis of food* (2nd ed., chap 7, pp. 353–389). Academic Press. https://doi.org/10.1016/B978-0-12-813266-1.00007-3. https://www.sciencedirect.com/science/article/pii/B9780128132661000073

Dias, L., Meirinho, S. G., Veloso, A., et al. (2017). Electronic tongues and aptasensors. In L. Rodrigues, & M. Mota (Eds.) *Bioinspired materials for medical applications*. Woodhead Publishing.

Dimitrov, D., Haygarov, V., & Yoncheva, T. (2018). Aromatic profile of red wines from grapevine varieties Rubin, Storgozia, Bouquet, Trapezitsa, Kaylashky Rubin and Pinot Noir, cultivated in the region of central Northern Bulgaria. *Journal of Microbiology, Biotechnology and Food Sciences, 8,* 885–889. https://doi.org/10.15414/jmbfs.2018-19.8.3.885-889

Dimitrov, D., Haygarov, V., & Yoncheva, T. (2019). Aromatic characterization and total volatile composition of red wines from the region of Central Northern Bulgaria. *Ukrainian Journal of Food Science, 7,* 207–216. https://doi.org/10.24263/2310-1008-2019-7-2-5

Dravnieks, A. (1985). *Atlas of odor character profiles*. Philadelphia, PA: American Society for Testing and Materials.

Duchowicz, P. R., Castro, E. A., Fernández, F. M., et al. (2005). A new search algorithm for QSPR/QSAR theories: Normal boiling points of some organic molecules. *Chemical Physics Letters, 412*(4–6), 376–380.

Duchowicz, P. R., Giraudo, M. A., Castro, E. A., et al. (2013). Amino acid profiles and quantitative structure–property relationship models as markers for Merlot and Torrontés wines. *Food Chemistry, 140*(1), 210–216. https://doi.org/10.1016/j.foodchem.2013.02.064. https://www.sciencedirect.com/science/article/pii/S0308814613002240

Enciso, M., Meftahi, N., Walker, M. L., et al. (2016). Bioppsy: An open-source platform for QSAR/QSPR analysis. *PLOS ONE, 11*(11), 1–11. https://doi.org/10.1371/journal.pone.0166298

Engen, T., & Pfaffmann, C. (1959). Absolute judgments of odor intensity. *Journal of Experimental Psychology, 58*(1), 23–6.

Escudero, A., Campo, E., Fariña, L., et al. (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. *Journal of Agricultural and Food Chemistry, 55,* 4501–10. https://doi.org/10.1021/jf0636418

Fanzone, M., Coronado, I., Sari, S., et al. (2022). Microwave-assisted maceration and stems addition in Bonarda grapes: Effects on wine chemical composition over two vintages. *Food Research International, 156*, 111–169. https://doi.org/10.1016/j.foodres.2022.111169. https://www.sciencedirect.com/science/article/pii/S0963996922002265

Fariña, L., Villar, V., Ares, G., et al. (2015). Volatile composition and aroma profile of Uruguayan Tannat wines. *Food Research International*. https://doi.org/10.1016/j.foodres.2014.12.029

Feng, T., Hu, Z., Chen, L., et al. (2020). Quantitative structure-activity relationships (QSAR) of aroma compounds in different aged Huangjiu. *Journal of Food Science*. https://doi.org/10.1111/1750-3841.15421

Ferreira, V., Blanco, A., & Sáenz-Navajas, M. P. (2022). Wine aroma vectors and sensory attributes. pp 3–39. https://doi.org/10.1016/B978-0-08-102067-8.00008-7

Forina, M., Casolino, C., & Pizarro Millan, C. (1999). Iterative predictor weighting (IPW) PLS: A technique for the elimination of useless predictors in regression problems. *Journal of Chemometrics, 13*(2), 165–184. https://doi.org/10.1002/(SICI)1099-128X(199903/04)13:2<165::AID-CEM535>3.0.CO;2-Y

Francis, I., & Newton, J. (2005). Determining wine aroma from compositional data. *Australian Journal of Grape and Wine Research, 11,* 114–126. https://doi.org/10.1111/j.1755-0238.2005.tb00283.x

Fuentes, S., Tongson, E., Torrico, D. D., et al. (2020). Modeling Pinot noir aroma profiles based on weather and water management information using machine learning algorithms: A vertical vintage analysis using artificial intelligence. *Foods, 9*(1). https://doi.org/10.3390/foods9010033. https://www.mdpi.com/2304-8158/9/1/33

Gacula, M. C., & Singh, J. (1984). 5 - incomplete block experimental designs. In M. C. Gacula, & J. Singh (Eds.) *Statistical Methods in Food and Consumer Research. Food Science and Technology* (pp. 141–175). Academic Press, San Diego. https://doi.org/10.1016/B978-0-08-091831-0.50010-2. https://www.sciencedirect.com/science/article/pii/B9780080918310500102

Gámbaro, A., Boido, E., Zlotejablko, A., et al. (2001). Effect of malolactic fermentation on the aroma properties of Tannat wine. *Australian Journal of Grape and Wine Research, 7*(1), 27–32. https://doi.org/10.1111/j.1755-0238.2001.tb00190.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1755-0238.2001.tb00190.x

Gámbaro, A., Varela, P., Boido, E., et al. (2003). Aroma characterization of commercial red wines of Uruguay. *Journal of Sensory Studies, 18,* 353–366. https://doi.org/10.1111/j.1745-459X.2003.tb00394.x

Garde-Cerdan, T., Costa, N., Rubio-Bretón, P., et al. (2021). The most important parameters to differentiate Tempranillo and Tempranillo Blanco grapes and wines through machine learning. *Food Analytical Methods*. https://doi.org/10.1007/s12161-021-02049-6

Goldner, M. C., & Zamora, M. C. (2007). Sensory characterization of *Vitis vinifera* cv. Malbec wines from seven viticulture regions of Argentina. *Journal of Sensory Studies, 22*(5), 520–532. https://doi.org/10.1111/j.1745-459X.2007.00123.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1745-459X.2007.00123.x

Goldner, M. C., Zamora, M. C., Di Leo Lira, P., et al. (2009). Effect of ethanol level in the perception of aroma attributes and the detection of volatile compounds in red wine. *Journal of Sensory Studies, 24*(2), 243–257. https://doi.org/10.1111/j.1745-459X.2009.00208.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1745-459X.2009.00208.x

Gómez García-Carpintero, E., Sánchez-Palomo, E., Gómez Gallego, M., et al. (2011). Volatile and sensory characterization of red wines from cv. Moravia Agria minority grape variety cultivated in La Mancha region over five consecutive vintages. *Food Research International, 44,* 1549–1560. https://doi.org/10.1016/j.foodres.2011.04.022

González-Viñas, M., Pérez-Coello, M., Cabezudo, M., et al. (1998). Sensory analysis of aroma attributes of young Airen white wines during storage in the bottle. *Journal of Food Quality, 21,* 285–297. https://doi.org/10.1111/j.1745-4557.1998.tb00523.x

Gonzalez-Viñas, M., Perez-Coello, M., Salvador, M., et al. (1996). Changes in gas-chromatographic volatiles of young Airen wines during bottle storage. *Food Chemistry, 56*(4), 399–403. https://doi.org/10.1016/0308-8146(95)00207-3. https://www.sciencedirect.com/science/article/pii/0308814695002073

Green, J., Parr, W., Breitmeyer, J., et al. (2011). Sensory and chemical characterisation of Sauvignon blanc wine: Influence of source of origin. *Food Research International, 44,* 2788–2797. https://doi.org/10.1016/j.foodres.2011.06.005

Gürbüz, O., Rouseff, J., & Rouseff, R. (2006). Comparison of aroma volatiles in commercial Merlot and Cabernet Sauvignon wines using gas chromatography-olfactometry and gas chromatography-mass spectrometry. *Journal of Agricultural and Food Chemistry, 54,* 3990–6. https://doi.org/10.1021/jf053278p

Guth, H. (1997). Identification of character impact odorants of different white wine varieties. *Journal of Agricultural and Food Chemistry, 45*(8), 3022–3026. https://doi.org/10.1021/jf9608433

Guth, H. (1998). Comparison of different white wine varieties in odor profiles by instrumental analysis and sensory studies. *714*, 39–52. https://doi.org/10.1021/bk-1998-0714.ch004

Heymann, H., & Noble, A. (1987). Descriptive analysis of commercial Cabernet Sauvignon wines from California. *American Journal of Enology and Viticulture, 38,* 41–44.

Hocking, R. R. (1976). A biometrics invited paper. The analysis and selection of variables in linear regression. *Biometrics*, 1–49

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology, 24,* 498–520.

Kakani, V., Nguyen, V. H., Kumar, B. P., et al. (2020). A critical review on computer vision and artificial intelligence in food industry. *Journal of Agriculture and Food Research, 2*, 100,033. https://doi.org/10.1016/j.jafr.2020.100033. https://www.sciencedirect.com/science/article/pii/S2666154320300144

Kang, B., Lee, J., & Park, H. (2014). Qualitative and quantitative prediction of volatile compounds from initial amino acid profiles in Korean rice wine (Makgeolli) model. *Journal of Food Science*. https://doi.org/10.1111/1750-3841.12489

Kar, S., Roy, K., & Leszczynski, J. (2017). *On applications of QSARs in food and agricultural sciences: History and critical review of recent developments* (pp. 203–302). https://doi.org/10.1007/978-3-319-56850-8_7

Kunkee, R. E. (1974). *Malo-lactic fermentation and winemaking* (chap 7, pp. 151–170). https://doi.org/10.1021/ba-1974-0137.ch007. https://pubs.acs.org/doi/abs/10.1021/ba-1974-0137.ch007

Laing, D. G. (1991). Characteristics of the human sense of smell when processing odor mixtures (pp. 241–259). Springer Berlin Heidelberg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-76223-9_12

Laing, D., & Glemarec, A. (1992). Selective attention and the perceptual analysis of odor mixtures. *Physiology & Behavior, 52*(6), 1047–1053. https://doi.org/10.1016/0031-9384(92)90458-E. https://www.sciencedirect.com/science/article/pii/003193849290458E

Laurent, M., Henick-Kling, T., & Acree, T. (1994). Changes in the aroma and odor of chardonnay wine due to malolactic fermentation. *Fermentation Viticulture and Enology Science, 49,* 3–10.

Lawless, H. T. (1999). Descriptive analysis of complex odors: Reality, model or illusion? *Food Quality and Preference, 10*(4), 325–332. https://doi.org/10.1016/S0950-3293(98)00052-4. https://www.sciencedirect.com/science/article/pii/S0950329398000524

Lawless, H., & Heymann, H. (2010). Sensory evaluation of food: Principles and practice (vol. 5999). https://doi.org/10.1007/978-1-4419-6488-5

Legin, A., Rudnitskaya, A., Lvova, L., et al. (2003). Evaluation of Italian wine by the electric tongue: Recognition, quantitative analysis and correlation with human sensors perception. *Analytica Chimica Acta, 484,* 33–44. https://doi.org/10.1016/S0003-2670(03)00301-5

Legin, A., Rudnitskaya, A., & Vlasov, Y. (2002). Electronic tongues: Sensors, systems, applications. *Sensors Update, 10,* 143–188. https://doi.org/10.1002/1616-8984(200201)10:1<143::AID-SEUP143>3.0.CO;2-Q

Licon, C. C., Bosc, G., Sabri, M., et al. (2019). Chemical features mining provides new descriptive structure-odor relationships. *PLOS Computational Biology, 15*(4), 1–21. https://doi.org/10.1371/journal.pcbi.1006945

Lockshin, L., & Corsi, A. (2012). Consumer behaviour for wine 2.0: A review since 2003 and future directions. *Wine Economics and Policy, 1,* 2–23. https://doi.org/10.1016/j.wep.2012.11.003

Longo, R., Pearson, W., Merry, A., et al. (2020). Preliminary study of Australian Pinot noir wines by colour and volatile analyses, and the Pivot© Profile method using wine professionals. *Foods, 9*(9). https://doi.org/10.3390/foods9091142. https://www.mdpi.com/2304-8158/9/9/1142

Marshall, K., Laing, D. G., Jinks, A., et al. (2006). The capacity of humans to identify components in complex odor-taste mixtures. *Chemical Senses, 31*(6), 539–545. https://doi.org/10.1093/chemse/bjj058. https://academic.oup.com/chemse/article-pdf/31/6/539/17426412/bjj058.pdf

Martí, P., Boqué, R., Busto, O., et al. (2005). Electronic noses in the quality control of alcoholic beverages. *Trends in Analytical Chemistry - TrAC, 24,* 57–66. https://doi.org/10.1016/j.trac.2004.09.006

Martínez, K., & Medina-Franco, J. (2014). Foodinformatics: Applications of chemical information to food chemistry. https://doi.org/10.1007/978-3-319-10226-9

Martínez, M., Ponzoni, I., Diaz, M., et al. (2015). Visual analytics in cheminformatics: User-supervised descriptor selection for QSAR methods. *Journal of Cheminformatics, 7,* 1. https://doi.org/10.1186/s13321-015-0092-4

McDaniel, M., Henderson, L. A., Watson Jr. B. T., et al. (1987). Sensory panel training and screening for descriptive analysis of the aroma of Pinot noir wine fermented by several strains of Malolatic bacteria. *Journal of Sensory Studies, 2*(3), 149–167. https://doi.org/10.1111/j.1745-459X.1987.tb00414.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1745-459X.1987.tb00414.x

Miranda-Lopez, R., Libbey, L., Watson, B., et al. (1992). Odor analysis of Pinot noir wines from grapes of different maturities by a gas chromatography-olfactometry technique (OSME). *Journal of Food Science, 57,* 985–993. https://doi.org/10.1111/j.1365-2621.1992.tb14339.x

Misra, N. N., Dixit, Y., Al-Mallahi, A, et al. (2020). IoT, big data and artificial intelligence in agriculture and food industry. *IEEE Internet of Things Journal, 1*. https://doi.org/10.1109/JIOT.2020.2998584

Muñoz González, C., Moreno-Arribas, M. V., Martín-Álvarez, P., et al. (2011). Volatile and sensory characterization of Xarel.lo white wines. *Flavour and Fragrance Journal, 26,* 153–161. https://doi.org/10.1002/ffj.2038

Nanou, E., Mavridou, E., Milienos, F. S., et al. (2020). Odor characterization of white wines produced from indigenous Greek grape varieties using the frequency of attribute citation method with trained assessors. *Foods*. https://doi.org/10.3390/foods9101396

Naranjo, A., Martínez-Lapuente, L., Ayestarán, B., et al. (2021). Aromatic and sensory characterization of Maturana Blanca wines made with different technologies. *Beverages, 7*(1). https://doi.org/10.3390/beverages7010010. https://www.mdpi.com/2306-5710/7/1/10

Noble, A., Arnold, R., Masuda, B., et al. (1984a). Progress towards a standardized system of wine aroma terminology. *American Journal of Enology and Viticulture, 35*

Noble, A., Arnold, R., Buechsenstein, J., et al. (1987). Modification of a standardized system of wine aroma terminology. *American Journal of Enology and Viticulture, 38*

Noble, A., Williams, A., & Langron, S. (1984b). Descriptive analysis and quality ratings of 1976 wines from four Bordeaux communes. *Journal of the Science of Food and Agriculture, 35,* 88–98. https://doi.org/10.1002/jsfa.2740350115

Ojha, P., & Roy, K. (2018). Chemometric modeling of odor threshold property of diverse aroma components of wine. *RSC Advances, 8,* 4750–4760. https://doi.org/10.1039/C7RA12295K

Peťka, J., Ferreira, V., González-Viñas, M., et al. (2006). Sensory and chemical characterization of the aroma of a white wine made with Devín grapes. *Journal of Agricultural and Food Chemistry, 54,* 909–15. https://doi.org/10.1021/jf0518397

Presa-Owens, C., & Noble, A. (1995). Descriptive analysis of three white wine varieties from Penedès. *American Journal of Enology and Viticulture, 46,* 5–9.

Rapp, A. (1999). Volatile flavour of wine: Correlation between instrumental analysis and sensory perception. *Die Nahrung, 42,* 351–63. https://doi.org/10.1002/(SICI)1521-3803(199812)42:063.3.CO;2-U

Rodriguez, S., Amberg, E., Thornton, R., et al. (1990). Malolactic fermentation in Chardonnay: Growth and sensory effects of commercial strains of Leuconostoc oenos. *Journal of Applied Microbiology, 68,* 139–144. https://doi.org/10.1111/j.1365-2672.1990.tb02558.x

Rodríguez-Méndez, M., De Saja, J., Gonzalez-Anton, R., et al. (2016). Electronic noses and tongues in wine industry. *Frontiers in Bioengineering and Biotechnology*. https://doi.org/10.3389/fbioe.2016.00081

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological review, 65*(6), 386.

Roy, K., Kar, S., & Das, R. (2015a). A primer on QSAR/QSPR modeling: Fundamental concepts. https://doi.org/10.1007/978-3-319-17281-1

Roy, K., Kar, S., & Das, R. (2015b). Understanding the basics of QSAR for applications in pharmaceutical sciences and risk assessment.

Sáenz-Navajas, M. P., Ballester, J., Pêcher, C., et al. (2013). Sensory drivers of intrinsic quality of red wines. Effect of culture and level of expertise. *Food Research International, 54,* 1506–1518. https://doi.org/10.1016/j.foodres.2013.09.048

Sánchez-Palomo, E., Cañas, P. M. I., Delgado, J. A. B., et al. (2018). Sensory characterization of wines obtained by blending Cencibel grapes and minority grape varieties cultivated in La Mancha region. *Journal of Food Quality, 2018,* 1–8.

Sánchez-Palomo, E., González-Viñas, M., Hidalgo, M. C., et al. (2007). Aroma potential of Albillo wines and effect of skin-contact treatment. *Food Chemistry, 103,* 631–640. https://doi.org/10.1016/j.foodchem.2006.08.033

Sánchez-Palomo, E., Trujillo, M., García Ruiz, A., et al. (2017). Aroma profile of Malbec red wines from La Mancha region: Chemical and sensory characterization. *Food Research International, 100*, 201-208. https://doi.org/10.1016/j.foodres.2017.06.036. https://www.sciencedirect.com/science/article/pii/S0963996917302909

Sancho-Galán, P., Amores, A., Palacios, V., et al. (2022). Volatile composition and sensory characterisation of dry white wines made with overripe grapes by means of two different techniques. *Foods, 11,* 509. https://doi.org/10.3390/foods11040509

Sauvageot, F., & Vivier, P. (1987). Effects of malolactic fermentation on sensory properties of four burgundy wines. *American Journal of Enology and Viticulture, 48*(2), 187–192. https://www.ajevonline.org/content/48/2/187

Sharma, A., Kumar., R., Ranjta, S., et al. (2021a). Smiles to smell: Decoding the structure-odor relationship of chemical compounds using the deep neural network approach. *Journal of Chemical Information and Modeling, 61*(2), 676–688. https://doi.org/10.1021/acs.jcim.0c01288. pMID: 33449694.

Sharma, A., Saha, B. K., Kumar, R., et al. (2021b). OlfactionBase: A repository to explore odors, odorants, olfactory receptors and odorant-receptor interactions. *Nucleic Acids Research*. https://doi.org/10.1093/nar/gkab763. gkab763. https://academic.oup.com/nar/advance-article-pdf/doi/10.1093/nar/gkab763/40215129/gkab763.pdf

Soto, A., Cecchini, R., Vazquez, G., et al. (2009). Multi-objective feature selection in QSAR using a machine learning approach. *QSAR & Combinatorial Science, 28*(11–12), 1509–1523. https://doi.org/10.1002/qsar.200960053. https://onlinelibrary.wiley.com/doi/abs/10.1002/qsar.200960053

Stone, H. (1992). *Quantitative descriptive analysis (QDA)* (pp. 15–21). ASTM International, West Conshohocken, PA. https://doi.org/10.1520/MNL10523M

Stone, H., & Sidel J. L. (2004). 6 - descriptive analysis. In H. Stone, & J. L. Sidel (Eds.) *Sensory evaluation practices* (3rd ed., pp. 201–245). Food Science and Technology, Academic Press, San Diego. https://doi.org/10.1016/B978-012672690-9/50010-X. https://www.sciencedirect.com/science/article/pii/B978012672690950010X

Summerson, V., Gonzalez Viejo, C., Pang, A., et al. (2021). Assessment of volatile aromatic compounds in smoke tainted Cabernet Sauvignon wines using a low-cost e-nose and machine learning modelling. *Molecules, 26*(16). https://doi.org/10.3390/molecules26165108. https://www.mdpi.com/1420-3049/26/16/5108

Sun, Q., Zhang, M., Mujumdar, A. S. (2019). Recent developments of artificial intelligence in drying of fresh food: A review. *Critical Reviews in Food Science and Nutrition 59*(14), 2258–2275. https://doi.org/10.1080/10408398.2018.1446900. pMID: 29493285.

Tao, Y. S., Yq, Liu, & Li, H. (2009). Sensory characters of Cabernet Sauvignon dry red wine from Changli county (China). *Food Chemistry - FOOD CHEM, 114,* 565–569. https://doi.org/10.1016/j.foodchem.2008.09.087

Thuillier, B., Valentin, D., Marchal, R., et al. (2015). Pivot© profile: A new descriptive method based on free description. *Food Quality and Preference, 42,* 66–77. https://doi.org/10.1016/j.foodqual.2015.01.012

Ubeda, C., Peña-Neira, Á, Cortiella, M. G. (2022). Combined effects of the vessel type and bottle closure during Chilean Sauvignon Blanc wine storage over its volatile profile. *Food Research International, 156*, 111–178. https://doi.org/10.1016/j.foodres.2022.111178. https://www.sciencedirect.com/science/article/pii/S0963996922002356

Varela, P., & Ares, G. (2012). Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Research International, 48,* 893–908. https://doi.org/10.1016/j.foodres.2012.06.037

Varela, P., & Gámbaro, A. (2006). Sensory descriptive analysis of Uruguayan Tannat wine: Correlation to quality assessment. *Journal of Sensory Studies, 21,* 203–217. https://doi.org/10.1111/j.1745-459X.2006.00061.x

Versini, G., Orriols, I., & Serra, A. (1994). Aroma components of Galician Albariño, Loureira and Godello wines. *Vitis, 33,* 165–170.

Verzera, A., Ziino, M., Scacco, A., et al. (2008). Volatile compound and sensory analysis for the characterization of an Italian white wine from "inzolia" grapes. *Food Analytical Methods, 1,* 144–151. https://doi.org/10.1007/s12161-008-9027-2

Vilanova, M., Campo, E., Escudero, A., et al. (2012). Volatile composition and sensory properties of *Vitis vinifera* red cultivars from North West Spain: Correlation between sensory and instrumental analysis. *Analytica Chimica Acta, 720,* 104–11. https://doi.org/10.1016/j.aca.2012.01.026

Vilanova, M., Escudero, A., Graña, M., et al. (2013). Volatile composition and sensory properties of north west Spain white wines. *Food Research International, 54*(1), 562-568. https://doi.org/10.1016/j.foodres.2013.07.036. https://www.sciencedirect.com/science/article/pii/S096399691300402X

Villamor, R., & Ross, C. (2013). Wine matrix compounds affect perception of wine aromas. *Annual Review of Food Science and Technology, 4,* 1–20. https://doi.org/10.1146/annurev-food-030212-182707

Vlasov, Y., Legin, A., & Rudnitskaya, A. (2002). Electronic tongues and their analytical application. *Analytical and Bioanalytical Chemistry, 373,* 136–46. https://doi.org/10.1007/s00216-002-1310-2

Vlasov, Y., Legin, A., Rudnitskaya, A., et al. (2005). Nonspecific sensor arrays ("electronic tongue") for chemical analysis of liquids

(IUPAC technical report). *Pure and Applied Chemistry, 77*(11), 1965–1983. https://doi.org/10.1351/pac200577111965

Wang, X. J., Tao, Y. S., Wu, Y., et al. (2017). Aroma compounds and characteristics of noble-rot wines of chardonnay grapes artificially botrytized in the vineyard. *Food Chemistry, 226*, 41–50. https://doi.org/10.1016/j.foodchem.2017.01.007. https://www.sciencedirect.com/science/article/pii/S0308814617300079

Zhang, B., Liu, H., Xue, J., et al. (2022). Use of *Torulaspora delbrueckii* and *Hanseniaspora vineae* co-fermentation with *Saccharomyces cerevisiae* to improve aroma profiles and safety quality of Petit Manseng wines. *LWT, 161*, 113–360. https://doi.org/10.1016/j.lwt.2022.113360. https://www.sciencedirect.com/science/article/pii/S002364382200295X

Zhao, N., Mio, W., & Liu, X. (2011). A hybrid PCA-LDA model for dimension reduction. In *The 2011 International Joint Conference on Neural Networks* (pp. 2184–2190). IEEE.

Zhu, W., Benkwitz, F., & Kilmartin, P. A. (2021). Volatile-based prediction of Sauvignon Blanc quality gradings with static headspace-gas chromatography-ion mobility spectrometry (SHS-GC-IMS) and interpretable machine learning techniques. *Journal of Agricultural and Food Chemistry, 69*(10), 3255–3265. https://doi.org/10.1021/acs.jafc.0c07899. pMID: 33661647.

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