OENO One
VINE AND WINE
OPEN ACCESS JOURNAL

Multivariate characterisation of Italian monovarietal red wines using MIR spectroscopy

Guiseppina.P. Parpinello[1], Arianna Ricci[1], Panagiotis Arapitsas[2], Andrea Curioni[3], Luigi. Moio[4], Susanna Rio Segade[5], Maurizio Ugliano[6] and Andrea Versari[1*]

[1]Department of Agricultural and Food Sciences, University of Bologna, Italy
[2]Fondazione Edmund Mach, Research and Innovation Centre, Department of Food Quality and Nutrition, San Michele all'Adige, Italy
[3]Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Italy
[4]Department of Agricultural Sciences, Division of Vine and Wine Sciences, University of Naples Federico II, Avellino, Italy
[5]Dipartimento di Scienze Agrarie, Forestali e Alimentari, Università degli Studi di Torino, Italy
[6]Department of Biotechnology, University of Verona, Italy

*Corresponding author: andrea.versari@unibo.it

This article is published in cooperation with the 11th OenoIVAS International Symposium, June 25–28 2019, Bordeaux, France.

ABSTRACT

Aim: The aim of this study was to investigate the application of mid-infrared (MIR) spectroscopy combined with multivariate analysis, to provide a rapid screening tool for discriminating among different Italian monovarietal red wines based on the relationship between grape variety and wine composition in particular phenolic compounds.

Methods and results: The MIR spectra (from 4000 to 700 cm^{-1}) of 110 monovarietal Italian red wines, vintage 2016, were collected and evaluated by selected multivariate data analyses, including principal component analysis (PCA), linear discriminant analysis (DA), support vector machine (SVM), and soft intelligent modelling of class analogy (SIMCA). Samples were collected directly from companies across different regions of Italy and included 11 grape varieties: Sangiovese, Nebbiolo, Aglianico, Nerello Mascalese, Primitivo, Raboso, Cannonau, Teroldego, Sagrantino, Montepulciano and Corvina.

PCA showed five wavelengths that mainly contributed to the PC_1, including a much-closed peak at 1043 cm^{-1}, which correspond to the C–O stretch absorption bands that are important regions for glycerol, whereas the ethanol peaks at around 1085 cm^{-1}. The band at 877 cm^{-1} are related to the C–C stretching vibration of organic molecules, whereas the asymmetric stretching for C–O in the aromatic –OH group of polyphenols is within spectral regions from 1050 to 1165 cm^{-1}. In particular, the (1175)–1100–1060 cm^{-1} vibrational bands are combination bands, involving C–O stretching and O–H deformation of phenolic rings. The 1166–1168 cm^{-1} peak is attributable to in-plane bending deformations of C–H and C–O groups of polyphenols, respectively, for which polymerisation may cause a slight peak shift due to the formation of H-bridges.

The best result was obtained with the SVM, which achieved an overall correct classification for up to 72.2% of the training set, and 44.4% for the validation set of wines, respectively. The Sangiovese wines (n=19) were split into two sub-groups (Sang-Romagna, n=12 and Sang-Tuscany, n=7) considering the indeterminacy of its origins, which is disputed between Romagna and Tuscany. Although the classification of three grape varieties was problematic (Nerello Mascalese, Raboso and Primitivo), the remaining wines were almost correctly assigned to their actual classes.

Conclusions: MIR spectroscopy coupled with chemometrics represents an interesting approach for the classification of monovarietal Italian red wines, which is important in quality control and authenticity monitoring.

Significance and impact of the study: Authenticity is a main issue in winemaking in terms of quality evaluation and adulteration, in particular for origin certified/protected wines, for which the added marketing value is related to the link of grape variety with the area of origin. This study is part of the D-wine project "The diversity of tannins in Italian red wines".

KEYWORDS

authenticity, D-Wine project, MIR, tannins, red wine

INTRODUCTION

With over 300 grape varieties, Italy has one of the richest ampelographic heritages worldwide. Among these grapes, several varieties are extremely rich in tannins (Mattivi *et al.*, 2002; Mattivi *et al.*, 2009). From a marketing point of view, the ability to associate each wine with specific sensory attributes (Gambuti *et al.,* 2013; García-Estévez *et al.*, 2017) is often a tool for the commercial success. Moreover, the assessment of wine authenticity is of great importance for consumers, producers and regulatory agencies to guarantee the wine labelling in terms of geographical origin and grape variety (Arvanitoyannis *et al.,* 1999; Versari *et al.,* 2014; Villano *et al.*, 2017).

As wine is a complex mixture of organic and inorganic compounds whose composition is affected by several variables (e.g. soil, climate, year, grape variety, winemaking practices), from a chemical point of view the authenticity of wine should rely on those parameters that are not affected during production or are difficult to forge. Thus, UV, visible (Vis), near infrared (NIR) and mid-infrared (MIR) spectroscopies with multivariate data analysis are suitable tools to ascertain wine composition, including phenolic compounds (Pique *et al.*, 2001; Fernández and Agosin, 2007; Bauer *et al.*, 2008; Jensen *et al.*, 2008; Laghi *et al.*, 2011; Dambergs *et al.*, 2012; Martelo-Vidal and Vázquez, 2014; Aleixandre-Tudo *et al.*, 2015; Saad *et al.*, 2016; Palade and Popa, 2018).

Several authors have focused on predicting wine authenticity using a combination of wine composition and multivariate data analysis. For instance, Edelmann *et al.* (2001) correctly discriminated more than 97% of Austrian red wines (n=38) – including the cultivars Cabernet Sauvignon, Merlot, Pinot Noir, Blaufränkisch (Lemberger), St. Laurent and Zweigelt – using MIR with soft independent modelling of class analogy (SIMCA) of the phenolic extracts. Cozzolino *et al.* (2003) correctly classified 100% of Australian Riesling wines (n=144) and up to 96% of Chardonnay wines (n=125) using Vis-NIR in combination with discriminant partial least squares (DPLS) regression. Around 86% of the Sauvignon Blanc wines (n=64) from Australia and New Zealand were correctly classified using the MIR spectrum with full cross-validation (leave-one-out) and partial least squares discriminant analysis (PLS-DA)

(Cozzolino *et al.*, 2011). Similarly, MIR coupled with linear discriminant analysis (LDA) with cross-validation could discriminate around 88% of five Marsala wines with different ageing times (Condurso *et al.*, 2018).

Louw *et al.* (2009) obtained a 98.3% correct separation of South African wines (n=496) – including Pinotage, Merlot, Cabernet Sauvignon, Shiraz, Chardonnay and Sauvignon Blanc – based on their MIR spectra and LDA. Similarly, Bevin *et al.* (2007) applied MIR for varietal classification of three red (n=119) and four white (n=72) Australian wines, and achieved a classification rate of more than 90% using LDA statistical analysis.

Furthermore, the combination of MIR spectroscopy and LDA made it possible to correctly classify more than 75% of the red and white wines sourced from organic (n=57) and non-organic (n=115) production systems from 13 growing regions in Australia (Cozzolino *et al.*, 2009). Recently, Basalekou *et al.* (2016) correctly classified four wines (n=88) made by two white (Vilana and Dafni) and two red grape varieties (Kotsifali and Mandilari) based on their phenolic content and colour parameters using the 1800–900 cm^{-1} MIR spectral region. Basalekou *et al.* (2017) further implemented the dataset of wines (n=154) with 82% correct classification using LDA and cross-validation.

As part of the D-wine project "The diversity of tannins in Italian red wines", this study aimed to investigate the application of MIR spectroscopy combined with multivariate analysis to provide a rapid screening tool for discriminating among different Italian monovarietal red wines.

MATERIALS AND METHODS

1. Samples

A total of 110 monovarietal red wines, vintage 2016, were collected directly from companies across different regions of Italy (Figure 1) and included 11 grape varieties: Sangiovese (n=19; seven from Tuscany and 12 from Romagna); Nebbiolo (Piemonte: n=11); Aglianico (Puglia: n=10); Nerello Mascalese (Sicilia: n=3); Primitivo (Puglia: n=11); Raboso (Veneto: n=10); Cannonau (Sardegna: n=9); Teroldego (Trentino: n=11); Sagrantino (Umbria: n=11); Montepulciano (Abruzzo: n=8); and Corvina (Veneto: n=7). All samples were monovarietal, i.e. 100% from a single grape variety, and were

FIGURE 1. Map of Italy with regions selected for sampling red wines.

not blended with wines from other regions. Therefore, although they do not represent commercially available products, these wines exhibit their respective varietal uniqueness and were selected based on their importance at regional level. The aim of this study was to gain insight on the composition of real wines in terms of authenticity of grape and wine, therefore each winery performed its original vinification protocols optimised for each specific grape variety.

2. Tannin assay

All wines were analysed for iron reactive tannins [tannins-Fe] according to the literature (Harbertson *et al.*, 2003) and as described elsewhere in detail (Versari *et al.*, 2007). Briefly, the wine tannins are precipitated with bovin serum albumin (BSA), then the pellet is dissolved in buffer and the tannins are determined by reaction with ferric chloride, yielding a coloured product quantified at 510 nm (UV–Vis spectrophotometer Cary 60, Agilent Technologies, Santa Clara, CA) and using (+)-catechin as calibration standard (mg/L CE) (Sigma, Milano, Italy).

3. Mid-infrared analysis

Infrared analysis was carried out using a diamond Attenuated Total Reflectance (ATR) Smart Orbit accessory (Thermo Optec), equipped

with a deuterated triglycine sulfate detector and a KBr window for measuring the medium infrared (MIR) region. The incident beam had a 45° geometry with respect to the diamond surface, yielding 25 internal reflections. The wines were analysed without any pretreatment or purification, simply by pouring a few drops of the samples over the ATR crystal. For each sample, the whole MIR spectra range from 4000 to 700 cm^{-1} was averaged over 128 consecutive scans with a resolution of 4 cm^{-1} (Figure 2). The samples were analysed in duplicate against an air background that was collected immediately prior to analysis, and the averaged spectrum was processed as a single sample spectrum for further multivariate analysis.

4. Data processing and multivariate analyses

The whole MIR spectra and several spectral regions were tested for chemometric elaboration, including the so-called "fingerprint region" from 1500 to 700 cm^{-1}. According to Louw *et al.* (2009), some spectral regions (5000–2970 cm^{-1}; 1716–1543 cm^{-1}) can be excluded prior to multivariate analysis to avoid strong absorption of water and spectral features that are little related to wine composition.

To attempt the classification of wines based on the relationship between grape variety and wine composition, the raw MIR spectra were mean-
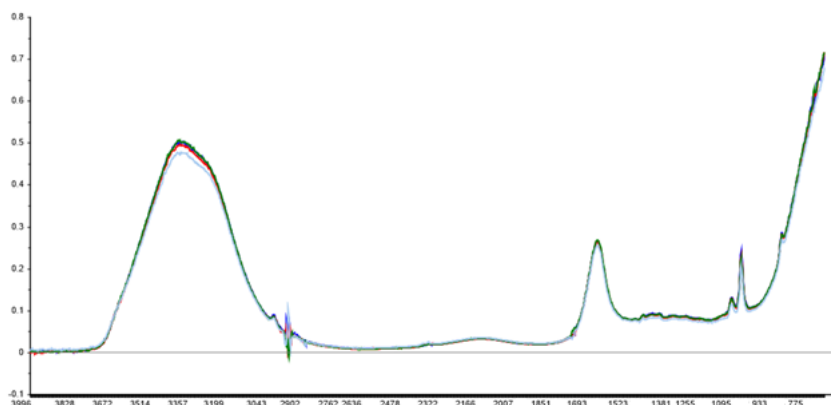
FIGURE 2. Example of MIR spectra of red wines.

centred and scaled to the same variance by standardisation throughout all multivariate statistical procedures (Unscrambler software version 10.3, Camo AS, Nedre Vollgate, Norway; XLSTAT 2018, Addinsoft, Paris, France), using principal component analysis (PCA), linear discriminant analysis (LDA), soft intelligent modelling of class analogy (SIMCA), and support vector machine (SVM). The selected multivariate tools have been previously used for classification of grapevine varieties (Yu *et al.,* 2017; Canuti *et al.*, 2018) and grape nectars authenticity (Whei Miaw *et al.*, 2018).

In particular, PCA is an unsupervised technique that reduces the dimensionality of the response matrix (i.e. MIR spectra) to a few new principal components (PCs) and was used to examine the hidden structure of the dataset, to determine correlations between observations and variables and to describe the overall variation in the data (Esbensen, 2002; Naes *et al*., 2002).

LDA, a supervised classification technique, was used to assign the wine samples according to variety, and the classification performance was evaluated by comparing the number of correctly assigned objects to their total number. LDA maximises between-group variance and performs best when there are fewer variables than samples and when the variables are orthogonal, i.e. uncorrelated (Naes *et al*., 2002). LDA classification matrices were developed using a full cross-validation (leave-one-out) method on the PCA sample scores for the five principal components (PC_{1-5}) that gave the highest level of separation (high variance) in the PCA models developed. PCA reduces the spectral data, thus allowing two important criteria to be satisfied:

(i) data is orthogonal and (ii) the number of variables is lower than the number of samples.

SIMCA is a supervised classification technique based on PCA, which exploits the similarity among samples within a class more than the difference between the classes (Ballabio and Todeschini, 2009). The SIMCA was modelled in two steps: (i) the PCA was carried out for each grape variety and the number of components was determined through cross-validation; (ii) the classification of new samples was carried out by means of a SIMCA model formed through PCA. Thus, the sample set was divided at random into calibration and validation sets, grouping 70% and 30% of wine samples, respectively. In particular, the modelling power (MP) is a measure of the influence of a variable over a given model. This measure has values between 0 and 1; the closer to 1, the more that variable is taken into account in the class model, the higher the influence of that variable, and the more relevant it is to that particular class. Usually, any variable with a MP higher than 0.3 is considered relevant in the model (Wold *et al.*, 1981). The discrimination power (DP) of a variable shows the ability of each variable to discriminate between two models, i.e. wines. Thus, a variable with a high DP (regarding two particular models) is very important for the differentiation between the two corresponding classes.

SVM is a supervised learning technique that uses an optimal separation hyperplane to separate samples in multi-class feature space. SVM can model cases of fewer samples and nonlinear relations (Bishop, 2006) and it does not require the use of PCA to reduce the dimensionality of the input matrix. The parameters of nonlinear
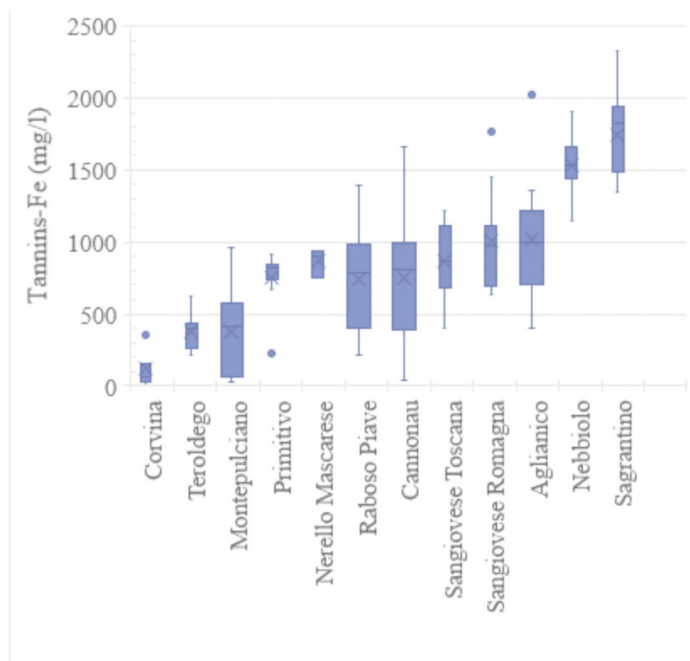
FIGURE 3. Tannin content in Italian red wines (n=110).

model based on radial basis function (RBF) kernel SVM were adjusted through some preliminary tests as follows: g value, 0.0005; C value, 0.5; Nu value, 0.5; cross-validation segment, 10.

RESULTS

1. Tannin composition

Results of tannin composition from the whole dataset are summarised in Figure 3. As expected, the tannin content of red wines showed great variability (range, 6–2327 mg/L CE) due to the varietal diversity, with a minor right-skewed distribution (data not shown), meaning a large proportion of samples was found on the left side of the distribution compared to the Gaussian distribution. The sampling strategy effectively selected red wines with a wide range of tannins and therefore the dataset is also suitable for further study on astringency.

2. Principal component analysis

The raw spectra within the fingerprint region (i.e. 1500–700 cm^{-1}) were analysed without rotation to improve the interpretation of X-loadings. The original X-matrix, when analysed by PCA, was reduced to five principal components (PC$_{1-5}$) globally explaining the 98% of total variability, and these five PC scores were later used for further LDA analysis. The PCA plot scattered the samples in two dimensions and therefore the data are suitable to be independently treated. Although the PC$_1$ was responsible for 78% of variance there was a lack of visual grouping according to grape variety on the first two PCs (PC$_2$= 12% of variance explained) (data not shown). The X-loading plot showed five wavelengths (1043, 877, 1083, 1164 and 1178 cm^{-1}) that mainly contributed to the PC$_1$.

3. LDA

LDA was used in the first attempt to classify the red wines according to their grape variety. Due to the limited classification rate achieved (Table 1), the SVM approach was then used to improve the classification of the wines.

4. SIMCA

SIMCA was modelled on the fingerprint spectra region (i.e. MIR spectra 1500–700 cm^{-1}) to counterbalance the limited sample size and the number of independent variables. Although SIMCA can work with as few as ten samples per class, the Nerello Mascalese wines were not considered for SIMCA due to limited samples available (n=3). SIMCA does not provide a single plot for looking at all the groups concurrently as it uses different PC models for each group. While SIMCA showed limited

TABLE 1. Confusion matrix of Italian red wines using LDA.

| Actual/predicted	Sang-T	Sang-R	Nebb	Agl	Ner	Prim	Rab	Cann	Terol	Sagr	Montep	Corv	**Total actual (n)**	**Correct (%)**
Sang-Toscana	**3**	0	0	1	1	0	0	1	0	0	0	1	7	42.9
Sang-Romagna	2	**0**	0	1	0	0	3	3	0	0	1	2	12	0.0
Nebbiolo	0	0	**1**	3	0	0	0	5	1	0	1	0	11	9.1
Aglianico	0	0	0	**5**	0	1	1	1	0	0	0	2	10	50.0
Nerello	0	0	0	0	**2**	0	1	0	0	0	0	0	3	66.7
Primitivo	0	0	0	2	0	**5**	1	0	0	3	0	0	11	45.5
Raboso	1	0	0	1	2	0	**4**	2	0	0	0	0	10	40.0
Cannonau	1	1	1	0	0	1	1	**2**	0	1	0	0	8	25.0
Teroldego	0	2	1	0	0	1	1	2	**1**	2	1	0	11	9.1
Sagrantino	0	0	0	1	1	5	0	1	0	**3**	0	0	11	27.3
Montepulciano	1	0	2	2	1	1	0	0	0	0	**1**	0	8	12.5
Corvina	1	0	0	1	4	0	0	0	0	0	0	**0**	6	0.0
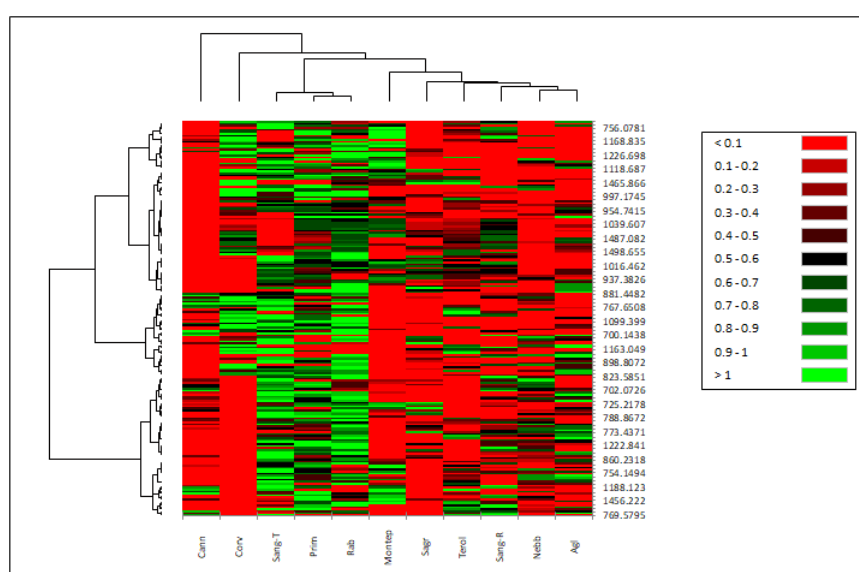
FIGURE 4. SIMCA modelling power of monovarietal red wines from selected Italian grape cultivars analysed by MIR.

performance to model the identification of red wines due to the complexity of the dataset, it provides useful information to visualise the important factors in terms of modelling power (MP) and discrimination power (DP). The heat plot of MP for the selected red wines showed the more relevant variables (i.e. MIR wave numbers) for each particular class of wine (Figure 4). The variables with higher influence for each class model are coloured in green and show that some wines/cultivars (e.g., Sangiovese Toscana, Primitivo, and Raboso) have many variables that are potentially relevant for helping each principal component to model variation in the data.

The DP, i.e. the variables most capable of discriminating between two models, showed two main groups: (i) Corvina, Primitivo, Montepulciano, Teroldego and Sagrantino; and (ii) Cannonau, Sangiovese, Raboso, Aglianico and Nebbiolo. The former has variables with a high DP (with regard to two particular models), and therefore it is important for the differentiation between the two corresponding classes (Figure 5). For practical reasons, the Sangiovese wine from Tuscany was selected as the reference class due to the great importance of this variety at national level.

5. SVM

SVM was further attempted to improve the class modelling of wines. Like PCA, the SVM method
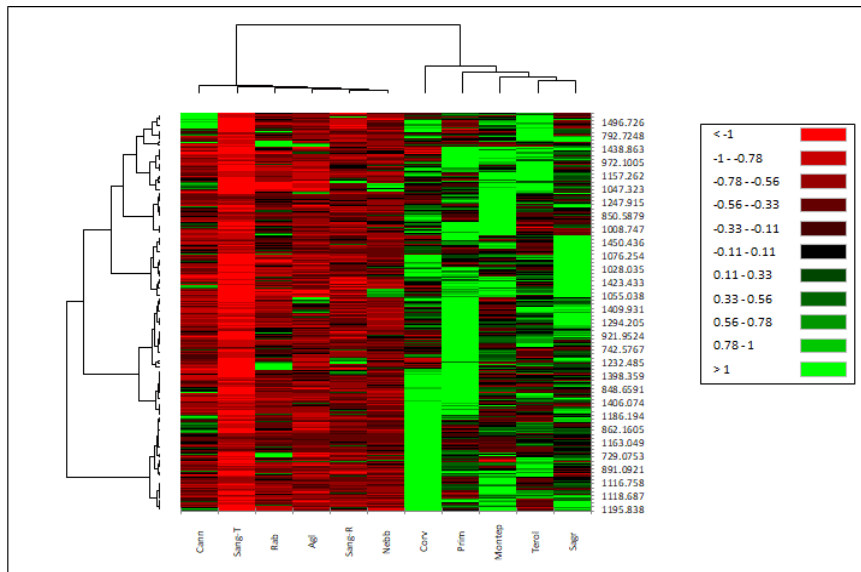
FIGURE 5. SIMCA discrimination power of monovarietal red wines from selected Italian grape cultivars analysed by MIR.

TABLE 2. Confusion matrix of Italian red wines using SVM with an overall correctly classified value of 72.2%.

	i	ii	iii	iv	v	vi	vii	viii	ix	x	xii	xii	Total (n)	Correct (n)	Correct (%)
i. Sang-T	**6**	0	0	0	0	0	1	0	0	0	0	0	7	6	85.7
ii. Sang-R	0	**11**	1	0	0	0	0	0	0	0	0	0	12	11	91.7
iii. Nebbiolo	0	0	**11**	0	0	0	0	0	0	0	0	0	11	11	100.0
iv. Aglianico	0	1	1	**7**	0	0	0	0	0	0	0	0	10	7	70.0
v. Nerello	2	1	0	0	**0**	0	0	0	0	0	0	0	3	0	0.0
vi. Primitivo	0	0	5	3	0	**3**	0	0	0	0	0	0	11	3	27.3
vii. Cannona	0	0	0	0	0	1	**7**	0	0	0	0	0	8	7	87.5
viii. Teroldeg	0	0	0	0	0	0	0	**10**	0	0	0	1	11	10	90.9
ix. Sagrantin	0	0	2	0	0	0	0	0	**8**	0	0	0	10	8	80.0
x. Montepulc	0	0	0	0	0	0	0	1	0	**7**	1	0	9	7	77.8
xi. Corvina	0	0	0	0	0	0	0	0	0	0	**6**	0	6	6	100.0
xii. Raboso	0	4	2	2	0	0	0	0	0	0	0	**2**	10	2	20.0

used the fingerprint spectra region (i.e. MIR spectra 1500–700 cm^{-1}), and the cross-validation approach due to the small number of samples. The results obtained with the SVM classification are encouraging, with 72.2% of overall correct classification for training set (Table 2), and 44.4% for the validation set of wines.

DISCUSSION

The D-wine project "The diversity of tannins in Italian red wines" focuses on polyphenols due to the great contribution of this class of compounds on red wine colour, mouthfeel and aroma longevity. The tannin content of Italian red wines is consistent with the literature (Harbertson *et al.*, 2008) in which a range from 30 to 1895 mg/L catechin equivalents (CE) was reported for five grape varieties: Pinot noir (n = 261 wines), Syrah (n = 266), Merlot (n = 197), Zinfandel (n = 182), and Cabernet Sauvignon (n = 364).

As a matter of fact, the Italian red wines with high tannin content (e.g. Sagrantino, Nebbiolo, and Aglianico) are more suitable for ageing, whereas the other wines with low tannin concentration (e.g. Corvina, Teroldego, and Montepulciano) seem more appropriate for young wines. Thus, there are several well-known practices that can modify the 'natural' tannin content of red wine, including the addition of exogenous tannins, the aging in barriques and the blending. It is worth noting out that all the wines sampled in this study were obtained without contact with any wood source (e.g. barriques, staves, chips, oenological tannins, etc.), therefore the current findings disclosed the 'natural' tannin content of Italian monovarietal red wines.

Regarding PCA, the lack of grouping according to grape variety origin based on the first two PCs can be due to several variables, including the high number of grape varieties (n=11), the lack of equality in group sizes and the limited sample size. Bevin *et al.* (2008), using MIR spectra of wines, failed to discriminate among wines from four white grape varieties (Chardonnay, Riesling, Sauvignon Blanc and Viognier) with the first two PCs that explained 94% of variation in the spectra. Similarly, the red wines (Cabernet Sauvignon, Shiraz, and Merlot) showed little grouping using the first two PCs ($PC_1 = 69\%$ and $PC_2 = 20\%$). Regarding the wavelengths that mainly contributed to the PC_1, Zhang *et al.* (2010) showed that the much-closed peak at 1043 cm^{-1} corresponds to the C–O stretch absorption bands, which are important regions for glycerol, whereas the ethanol peaks at about 1085 cm^{-1} (Shurvell, 2001). The band at 877 cm^{-1} would be related to C–C stretching vibration of organic molecules. Some authors located the asymmetric and symmetric stretching for C–O in the aromatic –OH group of hydrolyzable tannins, with the spectral regions 1050 to 1165 cm^{-1} (Pantoja-Castroa and González-Rodríguez, 2011). The band close to peak 1176 cm^{-1} was found to be a typical feature of C–O stretching of hydrolisable tannins (Agatonovic-Kustrin *et al.,* 2013). According to Jensen *et al.* (2008) two MIR regions (1485–1425 and 1060–995 cm^{-1}) were likely particularly important for tannin quantification.

It is clear that for optimum performance LDA needs a balanced design (i.e. a similar number of objects in various classes), and the relevant information should be in the mean of the data (not in the variance). Moreover, LDA is a parametric method that provides inferior results for nonlinear problems, so SVM was used to address non-linearly separable cases by applying the kernel approach with RBF. The Sangiovese wines (n=19) were split into two sub-groups (Sang-Romagna n=12 and Sang-Tuscany n=7) considering the indeterminacy of its origins, disputed between Romagna and Tuscany. Although the SVM classification of three grape varieties was problematic (Nerello Mascalese, Raboso and Primitivo), the remaining wines were correctly assigned up to 100% (Table 2). These accuracy values are comparable to those obtained for grapevine varietal classification using near infrared (NIR) spectroscopy with cross-validation and SVM (Yu *et al.,* 2017).

SIMCA may not provide satisfactory results if sample distribution in variable space is not uniform (Di Egidio *et al.*, 2011). Moreover, the high dimensionality of wine identification requires the selection of variables to ensure the model's performance with fewer variables. SIMCA creates an individual submodel for each class based on a supervised pattern recognition approach, and therefore the variables selected can differ from the global PCA. Thus, the MP of SIMCA outlined the important variables (i.e. wavenumbers) for model variation, whereas the discriminatory power related the contribution of the variables to the identification of wines in the data set. Deleting variables with both a low MP and a low DP may sometimes help in improving the classification. However, the current wine dataset is problematic due to the high number of classes and low number of samples, and therefore it was a challenge to select (few) features to advance the performance of the model and to simplify the analysis of the results. In future, the SIMCA approach could be useful to refine the model once further wines are sampled.

For this reason, the SVM with nonlinear kernel was attempted considering that the method does not need a large number of samples to be trained and is not affected by the presence of outliers. The superior results of SVM confirmed its ability to classify samples when linear functions are not adequate to achieve complete class separation. Acevedo *et al.* (2007) discriminated Spanish red and white wines – with classification rates above 96% – according to their denomination of origin by ultraviolet (UV)-

visible spectrophotometric techniques combined with SVM.

In conclusion, according to the literature the MIR spectrum of wine can be used to discriminate the varietal origin of wines, and this preliminary study challenged for the first time a large number of grape varieties (11). Although the number of red wines under investigation was limited (110 samples), to provide a definitive classification of each and every variety based on their MIR spectra, the current findings showed the occurrence of a peculiar MIR pattern for some Italian grape cultivars. This needs further study to disclose the effect of additional variables, such as vintage. In particular, Nebbiolo, Corvina, Teroldego, Sangiovese Romagna and Sagrantino wines were, to a large extent, classified using the current SVM approach.

Further analysis is in progress to provide full information about the distinctive physico-chemical and sensory characteristics of the selected monovarietal red wines.

REFERENCES

Acevedo F.J., Jiménez J., Maldonado S., Domínguez E., and Narváez, A., 2007. Classification of wines produced in specific regions by UV-visible spectroscopy combined with support vector machines. *J. Agric. Food Chem*. 55 (17), 6842–6849. doi:10.1021/jf070634q

Agatonovic-Kustrin S., Morton D.W., and Yusof A.P.M., 2013. The use of Fourier transform infrared (FTIR) spectroscopy and artificial neural networks (ANNs) to assess wine quality. *Mod. Chem. Appl.* 1 (4), 1–8, doi:10.4172/2329-6798.1000110

Aleixandre-Tudo J.L., Nieuwoudt H., Aleixandre J.L., and du Toit W.J., 2015. Robust Ultraviolet−Visible (UV−Vis) Partial Least-Squares (PLS) models for tannin quantification in red wine. *J. Agric. Food Chem*. 63, 1088−1098. doi:10.1021/ jf503412t

Arvanitoyannis I.S., Katsota M.N., Psarra E.P., Soufleros E.H., and Kallithraka S., 1999. Application of quality control methods for assessing wine authenticity: Use of multivariate analysis (chemometrics). *Trends Food Sci. Technol.* 10, 321−336. doi:10.1016/S0924-2244(99)00053-9

Ballabio D. and Todeschini R., 2009. Multivariate Classification for Qualitative Analysis. In: S. Da-Wen (Ed.), *Infrared Spectroscopy for Food Quality Analysis and Control*. Amsterdam: Elsevier; pp. 83−104.

Basalekou M., Strataridaaki A., Pappas C., Tarantilis P.A., Kotseridis Y., and Kallithraka S., 2016. Authenticity Determination of Greek-Cretan Mono-Varietal White and Red Wines Based on their Phenolic Content Using Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy and Chemometrics. *Curr. Res. Nutr. Food Sci.* 1 (SI. 1), 54–62. doi:10.12944/CRNFSJ.4.Special-Issue-October.08

Basalekou M., Pappas C., Tarantilis P.A., Kotseridis Y., and Kallithraka S., 2017. Wine authentication with Fourier Transform Infrared Spectroscopy: a feasibility study on variety, type of barrel wood and ageing time classification. *Int. J. Food Sci. Technol.* 52, 1307–1313. doi:10.1111/ijfs.13424

Bauer R., Nieuwoudt H., Bauer F.F., Kossmann J., Koch K.R., and Esbensen K.H., 2008. FTIR spectroscopy for grape and wine analysis. *Anal. Chem.* 80 (5), 1371–1379. doi:10.1021/ac086051c

Bevin C.J., Dambergs R.G., Fergusson A.J., and Cozzolino D., 2008. Varietal discrimination of Australian wines by means of mid-infrared spectroscopy and multivariate analysis. *Anal. Chim. Acta*, 621, 19–23 doi:10.1016/j.aca.2007.10.042

Bishop C.M., 2006. Pattern recognition and machine learning (1st ed.). New York: Springer.

Canuti V., Puccioni S., Storchi P., Zanoni B., Picchi M., and Bertuccioli M., 2018. Enological eligibility of grape clones based on the SIMCA method: the case of the Sangiovese cultivar from Tuscany. *Ital. J. Food Sci.* 30, 1, 184–199. doi:10.14674/IJFS-1020

Condurso C., Cincotta F., Tripodi G.L., and Verzera A., 2018. Characterization and ageing monitoring of Marsala dessert wines by a rapid FTIR-ATR method coupled with multivariate analysis. *Eur. Food Res. Technol.* 244, 1073–1081. doi:10.1007/s00217-017-3025-9

Cozzolino D., Smyth H.E., and Gishen M., 2003. Feasibility study on the use of visible and near-infrared spectroscopy to discriminate between white wine of different varietal origin. *J. Agric. Food Chem*. 52, 7701–7711. doi:10.1021/jf034959s

Cozzolino D., Holdstock M., Dambergs R.G., Cynkar W.U., and Smith P.A., 2009. Mid infrared spectroscopy and multivariate analysis: A tool to discriminate between organic and non-organic wines grown in Australia. *Food Chem.* 116, 761–765. doi:10.1016/j.foodchem.2009.03.022

Cozzolino D., Cynkar W.U., Shah N., and Smith P.A., 2011. Can spectroscopy geographically classify Sauvignon Blanc wines from Australia and New

Zealand? *Food Chem.* 126, 673–678. doi:10.1016/j.foodchem.2010.11.005

Di Egidio V., Oliveri P., Woodcock T., and Downey G., 2011) Confirmation of brand identity in foods by near infrared transflectance spectroscopy using classification and class-modelling chemometric techniques — The example of a Belgian beer. *Food Res. Int.* 44, 544–549. doi:10.1016/j.foodres.2010.11.021

Edelmann A., Diewok J., Schuster K.C. and Lendl B., 2001. Rapid method for the discrimination of red wine cultivars based on Mid-Infrared spectroscopy of phenolic wine extracts. *J. Agric. Food Chem.* 49, 1139–1145. doi:10.1021/jf001196p

Esbensen K., 2002. *Multivariate Data Analysis in Practice* (4th ed.). CAMO Process AS: Oslo, Norway; pp. 19–72.

Fernández C. and Agosin E., 2007. Quantitative analysis of red wine tannins using Fourier-Transform Mid-Infrared Spectrometry. *J. Agric. Food Chem.* 55, 7294–7300. doi:10.1021/jf071193d

Gambuti A., Rinaldi A., Ugliano M. and Moio L., 2013. Evolution of phe-nolic compounds and astringency during aging of red wine: Effect of oxy-gen exposure before and after bottling. *J. Agric. Food Chem.* 61, 1618–1627. doi:10.1021/jf302822b

García-Estévez I., Pérez-Gregorio R., Soares S., Mateus N., and de Freitas V., 2017. Oenological perspective of red wine astringency. *OENO One* 51, 3. doi:10.20870/oeno-one.2017.51.2.1816

Harbertson J.F., Picciotto E.A., and Adams D.O., 2003. Measurement of polymeric pigments in grape berry extracts and wines using a protein precipitation assay combined with bisulfite bleaching. *Am. J. Enol. Vitic.* 54 (4), 301–306.

Harbertson J.F., Hodgins R.E., Thurston L.N., Schaffer L.J., Reid M.S., Landon J.L., Ross C.F., and Adams D.O., 2008. Variability of tannin concentration in red wines. *Am. J. Enol. Vitic.* 59 (2), 210–214.

Jensen J.S., Egebo M., and Meyer A.S., 2008. Identification of spectral regions for the quantification of red wine tannins with Fourier Transform Mid-Infrared spectroscopy. *J. Agric. Food Chem.* 56 (10), 3493–3499. doi: 10.1021/jf703573f

Laghi L., Versari A., Parpinello G.P., Nakaji D.Y., and Boulton R.B., 2011. FTIR spectroscopy and direct orthogonal signal correction preprocessing applied to selected phenolic compounds in red wines. *Food Anal. Methods* 4 (4), 619–625. doi:10.1007/s12161-011-9240-2

Louw L., Roux K., Tredoux A., Tomic O., Naes T., Nieuwoudt E.H., and van Rensburg P., 2009. Characterization of selected South African young cultivar wines using FTMIR spectroscopy, gas chromatography and multivariate data analysis. *J.*

Agric. Food Chem. 57 (7), 2623–2632. doi:10.1021/jf8037456

Martelo-Vidal M.J. and Vázquez M., 2014. Determination of polyphenolic compounds of red wines by UV–VIS–NIR spectroscopy and chemometrics tools. *Food Chem.* 158, 28–34. doi:10.1016/j.foodchem.2014.02.080

Mattivi F., Zulian C., Nicolini G., and Valenti L., 2002. Wine, Biodiversity, Technology, and Antioxidants. *Ann. N.Y. Acad. Sci.* 957, 37–56. doi:10.1111/j.1749-6632.2002.tb02904.x

Mattivi F., Vrhovšek U., Masuero D. and Trainotti D., 2009. Differences in the amount and structure of extractable skin and seed tannins amongst red grape varieties. *Aust. J. Grape Wine Res.* 15, 27–35 doi:10.1111/j.1755-0238.2008.00027.x

Naes T., Isaksson T., Fearn T., and Davies T., 2002. *A user-friendly guide to multivariate calibration and classification.* Chichester: NIR Publications.

Palade L.M. and Popa M.E., 2018. Polyphenol fingerprinting approaches in wine traceability and authenticity: Assessment and implications of red wines. *Beverages* 4, 75, 1–18. doi:10.3390/beverages4040075

Pantoja-Castroa M.A. and González-Rodríguez H., 2011. Study by infrared spectroscopy and thermogravimetric analysis of tannins and tannic acid. *Rev. Latinoam. Quím.* 39, 107–112.

Pique D., Cattenoz T., and Corrieu G., 2001. Classification of red wines analysed by middle infrared spectroscopy of dry extract according to their geographical origin. *J. Int. Sci. Vigne Vin* 35, 3, 165–170. doi:10.20870/oeno-one.2001.35.3.1703

Saad R., Bouveresse D.J., Locquet N., and Rutledge D.N., 2016. Using pH variations to improve the discrimination of wines by 3D front face fluorescence spectroscopy associated to Independent Components Analysis. *Talanta* 153, 278–284. doi:10.1016/j.talanta.2016.03.023

Shurvell H.F., 2001. *Handbook of vibrational spectroscopy.* New York: John Wiley & Sons, Ed.

Versari A., Parpinello G.P., and Mattioli A.U., 2007. Characterization of polymeric pigments and color components of commercial red wines by using selected UV-Vis spectrophotometric methods. *South Afr. J. Enol. Vitic.* 28 (1), 6–10. doi:10.21548/28-1-1453

Versari A., Laurie V.F., Ricci A., Laghi L., and Parpinello G.P., 2014. Progress in authentication, typification and traceability of grapes and wines by chemometric approaches. *Food Res. Int.* 60, 2–18 doi:10.1016/j.foodres.2014.02.007

Villano C., Lisanti M.T., Gambuti A., Vecchio R., Moio L., Frusciante L., Aversano R., and Carputo D., 2017. Wine varietal authentication based on phenolics, volatiles and DNA markers: State of the

art, perspectives and drawbacks. *Food Control* 80, 1–10 doi:10.1016/j.foodcont.2017.04.020

Whei Miaw C.S., Martins Sena M., Carvalho de Souza S.V., Callao M.P., and Ruisanchez I., 2018. Detection of adulterants in grape nectars by attenuated total reflectance Fourier-transform mid-infrared spectroscopy and multivariate classification strategies. *Food Chem.* 266, 254–261. doi10.1016/j.foodchem.2018.06.006

Wold S., Albano Chr., Blomquist G., Coomans D., Dunn III. W.J., Edlund U., Eliasson B., Hellberg S., Johansson E., Norden B., Sjastrem M., Saderstrem B., and Wold H., 1981. Pattern recognition by means of disjoint principal component models (SIMCA). *In*: Hsskuldson A., *et al.* (ed.) *Proc. Symp. Anv. Stat.*

NEUCC, RECAU and RECKU. Copenhagen: Denmark; pp. 475–487.

Yu J., Zhan J., and Huang W., 2017. Identification of wine according to grape variety using Near-Infrared Spectroscopy based on Radial Basis Function Neural Networks and Least-Squares Support Vector Machines. *Food Anal. Methods* 10, 3306–3311. doi:10.1007/s12161-017-0887-1

Zhang Y.-L., Chen Y.-B., Lei Y., Zhou Q., Sun S.-Q., and Nod I., 2010. Discrimination of different red wine by Fourier-transform infrared and two-dimensional infrared correlation spectroscopy. *J. Mol. Struct.* 974, 144–150. doi:10.1016/j.molstruc.2010.03.021