Modelling wine astringency from its chemical composition using machine learning algorithms

Maria P. Sáenz-Navajas[1], Sara Ferrero-del-Teso[2], Miguel Romero[3], Dario Pascual[3],
David Diaz[3], Vicente Ferreira[1] and Purificación Fernández-Zurbano[2]

[1]Laboratorio de análisis del aroma y enología (LAAE. Department of Analytical Chemistry,
Universidad de Zaragoza, Instituto Agroalimentario de Aragón (IA2) (UNIZAR-CITA)
Associate unit to Instituto de Ciencias de la Vid y del Vino (ICVV) (UR-CSIC-GR) Zaragoza, Spain
[2]Instituto de Ciencias de la Vid y del Vino (Universidad de La Rioja-Consejo Superior de Investigaciones
Científicas-Gobierno de La Rioja. Department of Enology, Logroño, La Rioja, Spain
[3]SDG-Group, SDG-Consulting Spain, Logroño, Spain

Corresponding author: maria@saenz.info

**This article is published in cooperation with the International Conference in Wine Sciences,
Macrowine 2018, May 23-28 2018, Zaragoza (Spain)**

ABSTRACT

Aims: The present work aims to predict sensory astringency from wine chemical composition using machine learning algorithms.

Material and results: Moristel grapes from different vineblocks and at different stages of ripening were collected. Eleven different wines were produced in 75 L tanks in triplicate, and further sensory factors were described by the rate-all-that-apply method with a trained panel of participants. The polyphenolic composition was characterised in wines by measuring the concentration and activity of tannins using UHPLC-UV/VIS, the mean degree of polymerisation (mDP. and the composition of tannins using thiolysis followed by UHPLC-MS. Conventional oenological parameters were analysed using FTIR and UV-Vis. Machine learning was applied to build models for predicting a wines astringency from its chemical composition. The best model was obtained using the support vector regressor (radial kernel) algorithm presenting a root-mean-square error (RMSE) value of 0.190.

Conclusions: The main variables of the astringency model were the % of procyanidins constituting tannins and ethanol content, followed by other eight variables related to tannin structure and acidity.

Significance of the study: These results increase the knowledge of chemical variables related to the perception of wine astringency and provide tools to control and optimise grape and wine production stages to modulate astringency and maximise quality and the consumer appeal of wines.

KEYWORDS

wine, astringency, machine learning, modelling, tannin activity, sensory

INTRODUCTION

Consumption experience and thus wine appreciation is the result of interactions between the consumer and products properties (Charters and Pettigrew, 2007; Prescott *et al.*, 2002). Product-related factors refer to both intrinsic and extrinsic categories (Jover *et al.*, 2004). Intrinsic cues are related to flavour, while extrinsic cues refer to properties that are not physically part of the wine, such as bottle weight, bottling place, type of wine or appellation, etc. While consumers rely on both types of cues when selecting a product (D'Alessandro and Pecotich, 2013), there is a wide body of work focused on understanding the impact of extrinsic cues on wine appreciation. However, much less is known about the intrinsic cues I.e. flavour. A products flavour is the result of the interaction between sensory modalities including colour, aroma, taste and mouthfeel. In complex systems, the formation of mouthfeel is the least understood overall. This is especially true for astringency perception, which is reported to be mainly driven by alcohol content, polyphenolic compounds, their interaction with oral components (i.e. saliva, mucosa and oral receptors) and brain processing (Canon *et al.*, 2018). There is a lack of consensus in the scientific literature when defining the perceptual phenomenon and mechanisms driving astringency, however, the interactions between phenolics and proteins seem to be the most important driver (García-Estévez *et al.*, 2018). It is important to note that there is great interindividual variability in astringency perception, which has been attributed to differences in saliva composition and buccal microbiota among consumers (Lamy, 2018). This variability in salivary proteome is related to differences in the sensory sensitivity for astringency (Lamy *et al.*, 2017) and has shown to influence the acceptability of phenol-rich foods (Dinnella *et al.*, 2011; Masi *et al.*, 2015).

Tannins have been reported as important drivers of wine astringency, which is mainly understood as dryness perception. These molecules are constituted by catechin or epicatechin (procyanidins. epigallocatechin (prodelphinidins) or epicatechin gallate (gallocatechins) units linked through C4–C8 interflavanoid bonds. Differences in constitutive molecules, concentrations and mean degree of polymerisation (mDP) (Ma *et al.*, 2014) have been reported to have an impact on wine astringency.

Recently, tannin activity, measured as the enthalpy of interaction between polyphenols and an hydrophobic surface, has been hypothesised to correlate with wine astringency (Revelette *et al.*, 2014; Watrelot *et al.*, 2016). However, the correlation of sensory astringency with the different chemical variables that characterise phenolic compounds is far from clear. This difficulty in establishing consistent relationships between phenolic compounds and astringency has been attributed to: 1) a lack of enough chemical variability among the samples studied to induce significant sensory differences; 2) a lack of analytical relevant variables analysed in samples; or 3) the presence of cross-modal sensory interactions between aroma or taste properties with astringency and thus the capacity of other sensory modalities to modulate astringent perception (Watrelot *et al.*, 2016). Besides these limitations, in the present work it is hypothesised that relationships between phenolic chemical properties and astringency do not necessarily have to be linear, as most studies have tried to show.

In this context, the present work aimed to firstly generate wines with enough phenolic variability able to induce astringency differences and then model astringency by applying different machine learning algorithms.

The first hypothesis was that grapes from the same variety (Moristel in this case), harvested at different maturity levels and processed with a similar winemaking strategy would yield wines with maximal variability in phenolic composition and most probably astringency, and with minimal aroma variability (as observed among wines from different varieties and winemaking processes). This would reduce the presence of cross-modal interactions. The second hypothesis was that sensory–chemical relationships between phenolics and astringency do not necessarily have to be linear.

MATERIALS AND METHODS

1. Samples and winemaking

Eleven different wines were produced in 2017 with Moristel red grapes (*Vitis vinifera L.*) at the Pirineos winery (Barbastro, Spain). Four different vineyard blocks were selected based on historical data related to maturity evolution as measured by Dyostem® (Vivelys, France). According to commercial information, this tool monitors sugar loading and changes in the colour

of the fruit to determine the maturity level of grapes (i.e. polyphenolic and technological maturities related to sugar and acidity levels) and optimal harvest date. Based on this approach, one block did not experience any significant maturity evolution in a four-week period, and thus it was harvested only once. In contrast, the other three blocks were harvested at three or four different points, each separated by one or two weeks. According to the commercial system, the second point of maturity was the optimal point to harvest; therefore the fruit was harvested one week before and one and/or two weeks after this optimal point, to ensure grapes with different maturity levels and thus, *a priori*, maximal variability in chemical composition and most probably in astringency.

Grapes were manually harvested from the Somontano region (Huesca, Spain). For each of the 11 selected vineblocks and/or maturity points, 150 kg of fruit was harvested and processed with an automatic crusher/destemmer. Each lot was divided into three stainless steel tanks (75 L capacity. sulphur dioxide was adjusted to 50 mg L^{-1} and the fruit was further inoculated with commercial yeasts (Lalvin ICV D254, Lallemand) at 106 cells mL^{-1}. Alcoholic fermentations (FOH) took place on skins for an average of 10 days. Once FOH was finished, lactic bacteria (Lalvin VP41) were inoculated at a rate of 9 mg L^{-1}. Wines were bottled around 3 months after FOH (free SO_2 adjusted to 30 mg L^{-1}). Glass bottles (750 mL capacity) were sealed with natural cork closures.

2. Chemical characterisation of wines

2.1. Conventional oenological analysis

The total polyphenol index (TPI) was estimated as absorbance at 280 nm (Ribéreau-Gayon, 1970) and colour intensity (CI) as the sum of absorbance at 420, 520 and 620 nm (Glories, 1984) For TPI determination, the absorbance at 280 nm of samples diluted 1:100 in deionised water was measured in 1-cm quartz cuvettes. For CI, absorbance of undiluted samples was measured in 2-mm crystal cuvettes. Reducing sugars, ethanol content, pH, malic and lactic acid, as well as titratable and volatile acidities, were analysed using infrared spectrometry with Fourier Transformation with a WineScanTM FT 120 (FOSS®, Barcelona, Spain. which was previously calibrated with the official OIV methods.

2.2. Analysis of anthocyanin-derived pigments

Determination of monomeric (MP. small polymeric pigments (SPP) and large polymeric pigments (LPP) in wines and fractions was carried out as described elsewhere (Harbertson *et al.*, 2003). MPs were the group of compounds bleachable with bisulphite, while SPP and LPP were resistant to bisulphite bleaching. SPP did not precipitate with ovalbumin, in contrast with LPP, which did. Levels of MP, SPP and LPP were expressed as absorbance at 520 nm.

2.3. Characterisation of tannins

Acid-catalysed degradation in the presence of toluene-α-thiol was performed according to the method described by Labarde *et al.* (1999) but with some modifications as described by Gonzalo-Diago *et al.* (2013). Quantification was performed in the negative mode from the extracted ion chromatogram (EIC) for flavan-3-ols and in the positive mode for malvidin-3-*O*-glucosyde. The area under the peaks of malvidin-3-*O*-glucosyde and flavan-3-ol monomers (terminal units) before and after thiolysis, as well as toluene-α-thiol adducts (extension units) released from the depoly-merisation reaction were integrated.

Calibration curves were established with malvidin-3-*O*-glucosyde, (+)-catechin, (−)-epicatechin, (−)-epicatechin-3-*O*-gallate and (−)-epigallocatechin. In the absence of the standards of the thiol derivatives and considering that the thiolytic derivatives were shown to have similar response factors as the correspondent monomeric units, their concentrations were calculated from the respective monomer calibration curves. The mean degree of polymerisation (mDP) was calculated as the ratio of total units (extension + terminal) to terminal units (calculated as the difference before and after thiolysis). The % of tannins linked to malvidin-3-*O*-glucosyde (%T-M) was calculated as the molar ratio of malvidin-3-*O*-glucosyde linked to tannins (calculated as the difference before and after thiolysis) to the sum of the total units of terminal malvidin-3-*O*-glucosyde and extension + terminal units of (+)-catechin, (−)-epicatechin, (−)-epicatechin-3-*O*-gallate and (−)-epigallocatechin (i.e. total units of tannins). The % of procyanidins (%PC) was calculated as the ratio of total units (extension and terminal) of catechin and epicatechin to total units of tannins. The % of prodelphinidins

(%PD) and galloylated (%G) units was calculated as the ratio of total units of PD and G to the total units of tannins, respectively.

The % of malvidin-3-*O*-glucosyde linked to tannins (%M-T) was calculated as the molar ratio of the malvidin linked to flavanols (difference of malvidin before and after thiolysis) to total malvidin (before and after thiolysis).

Concentration and activity of tannins were estimated using HPLC-UV-Vis following the method proposed by Revelette *et al.* (2014). Tannin activity is related to the thermodynamics of interaction between tannins and a hydrophobic surface (polystyrene divinyl-benzene HPLC column) as discussed elsewhere (Barak and Kennedy, 2013; Revelette *et al.*, 2014).

3. Sensory characterisation of wines

3.1. Participants and procedure

The 33 wines produced (11 different wines in triplicate) were evaluated sensory characterised in February and March 2018 by 17 participants at the Instituto de Ciencias de la Vid y del Vino (ICVV) and Universidad de La Rioja (Spain). The participants were mainly oenology students and oenologists (11 women and five men, age range 22–34 years, average age 28) recruited on the basis of interest and availability and were not paid for their participation. They attended a total of 13 sessions spread over four weeks, comprising nine training sessions (90 min each and taking place at 12 p.m.) and four sessions to describe the wines studied (eight or nine samples per session). The participants worked in two subgroups and followed the same guidelines. The first session was devoted to generating aroma terms that differed among samples. During the following training sessions, reference standards (prepared at Universidad de Zaragoza) were presented for the 12 selected aroma terms and six for taste and mouthfeel terms. For in-mouth terms, solutions were prepared containing different concentrations of table sugar (0–7 g L^{-1}) for testing sweetness, tartaric acid (0–3 g L^{-1}) for acidity, quinine sulphate (0–40 mg L^{-1}) for bitterness and potassium, aluminium sulphate (0–5 g L^{-1}) for astringency, absolute alcohol (0–15% v/v) for alcoholic sensation and carboxymethylcellulose (0-1.5 g L^{-1}) for viscosity. During a typical training session, the participants were presented with references illustrating the different aroma, taste and mouthfeel terms. Next, between two and four wines were first individually described and then the ratings were discussed until the participants achieved consensus. The wines were described in the last four sessions: participants were asked to taste the wines and rate the intensity on a 7-point scale (1 = very low; 7 = very high) using only those terms (out of 18 available terms) that applied to the sample, according to rate-all-that-apply (RATA) methodology (Ares *et al.*, 2014). Terms that did not apply to the wine were allocated a value of 0 when collecting data. To avoid bias due to order of presentation, terms in the list appeared in different and randomised orders for each assessor. The use of a sip (rinsing solutions: water and 1 g L^{-1} pectin solution) and spit protocol between each sample was imposed as described elsewhere (Colonna *et al.*, 2004). Participants tasted samples in a sequential monadic manner: 20-mL samples were served in dark wine glasses labelled with random three-digit codes and covered with plastic Petri dishes according to a random arrangement that was different for each participant. Samples were served at room temperature and evaluated in a ventilated and air-conditioned tasting room at around 20 °C.

3.2. Data analysis

Only the data for astringency are reported here.

The discriminability potential of chemical variables among wines was calculated as the rate between maximal and minimum level (max/min) for each variable.

Two-way ANOVAs (participants as random and wines as fixed factors) were calculated for the term astringency. Next, a pair-wise comparison test (Fischer test) was applied (5% risk) using XLSTAT (2015).

The first step for modelling was to search for the best simple models, then the best learning algorithms were merged to obtain better predictive performance. This technique is known as model ensemble.. Machine learning algorithms were boosted by SDG using the DataRobot Platform. Therefore, dozens of independent challenger models were developed and validated by cross-validation. Model accuracy was evaluated by root-mean-square error (RMSE), i.e. differences between astringency scores predicted by a model and the scores observed. A robust model k-fold cross-

TABLE 1. Chemical variables analysed in the 33 wine samples in the study. Range of occurrence, median, rate of maximal and minimal level, Pearson correlation coefficients (r) between sensory astringency and chemical variables.

	Range	Median	Max/min	r (astringency)
pH	3.1–3.4	3.3	1.1	0.24
volatile acidity (VA) (g L^{-1})	0.3–0.5	0.4	2.1	0.59
total acidity (TA) (g L^{-1})	5.3–7.0	6.0	1.3	0.02
reducing sugars (RS) (g L^{-1})	1.6–2.7	2.3	1.7	0.34
malic acid (MA) (g L^{-1})	0.2–0.8	0.4	5.5	0.16
lactic acid (LA) (g L^{-1})	0.0–0.5	0.3	70.8	-0.45
ethanol content (% v/v)	**12.3–15.8**	**13.2**	**1.3**	**0.71***
colour intensity (CI) (a.u.)	**4.8–15.3**	**11.8**	**3.2**	**0.65***
total polyphenol index (TPI) (a.u.)	**22–53**	**43**	**2.4**	**0.72***
tannin activity (TA) (-KJ mol^{-1})	**853–2751**	**1422**	**3.2**	**0.77***
tannin concentration (TC) (g L^{-1})	1993–4188	2714	2.1	0.51
monomeric pigments (MP) (a.u.)	**0.4–1.0**	**0.8**	**2.9**	**0.65***
small polymeric pigments (SPP) (a.u.)	0.2–0.5	0.4	2.8	0.32
large polymeric pigments (LPP) (a.u.)	0.1–0.4	0.2	4.6	0.25
mean degree of polymerisation (mDP)	0.1–2.8	1.3	28	0.51
% of procyanidins in tannins (%PC)	51–85	69	1.7	-0.62
% of prodelphinidins in tannins (%PD)	**0.6–7.3**	**2.9**	**11.4**	**0.66***
% galloylated tannins (%G)	0.2–10.1	1.9	54.9	-0.40
% of tannins linked to malvidin (%T-M)	10.7–39.5	23.8	3.7	0.61
% of malvidin linked to tannins (%M-T)	11.1–32.2	20.6	2.8	0.34

Chemical variables with significant lineal correlation with astringency are marked in bold. *Significant linear correlation (P<0.05)

validation framework to test the out-of-sample stability of each model was employed. In addition to the cross-validation partitioning, a holdout sample was calculated to further test out-of-sample model performance and ensure that overfitting did not occur. Therefore, 18% of the training data was set aside as a holdout dataset. This dataset was used to verify that the final model performs well on data that has not been examined throughout the training process. For further model validation, the remainder of the data was divided into five cross-validation partitions (selected by random sampling). For the best models, five-fold cross-validation training and scoring was completed. Then, the mean score of the complete model cross-validation was calculated across all folds.

The best model for astringency consisted of applying a smooth ridit transformation followed by the calculation of the support vector regressor (SVR; radial kernel). The ridit (or score for a variable) transformation can be interpreted as an adjusted percentile score and extends the Bross ridit method (Bross, 1958) by applying the method to numerical values and normalising the score such that the mean calculated for the reference population will always be 0 and the score will be in the interval [-1,1]. The SVR is a generalised version of support vector machine classifiers that uses a special loss function to convert a regression problem into a classification problem. Support vector machines (SVMs) are an extremely robust machine learning model and are very efficient in high-dimensional spaces. In addition, a "kernel" function was used, which allows for a non- linear transformation of the data before fitting the SVM. These kernel functions can be a very useful way to transform a non-linear problem into a linear domain.

The permutated impact of variables on the models is calculated by observing the effect on model scores when altering the input data of a given variable. The algorithm employed normalises the scores so that the values of chemical variables included in the model are
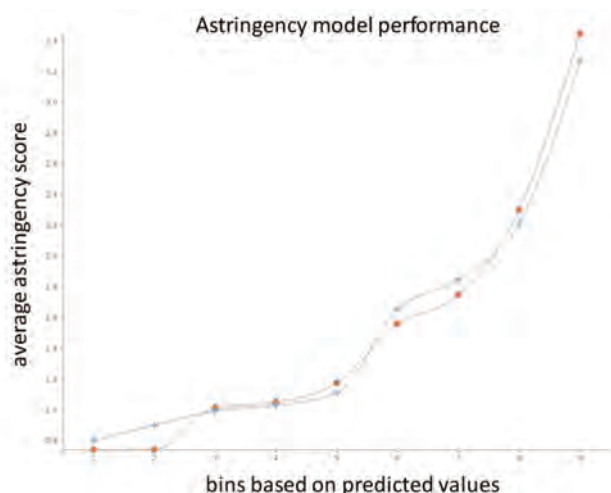
FIGURE 1. Lift charts for astringency model. Lines in orange and blue displays the average actual and predicted scores for a given bin.

normalised to the most important variable (%PC).

RESULTS AND DISCUSSION

The first objective of the present work was to produce Moristel wines with important chemical variabilities, focusing on parameters typically known to be related to astringency perception. Results show that for the 20 chemical variables analysed, highly significant differences (P<0.0001 in all cases) were observed among the 33 wines. Table 1 shows ranges and median values of the parameters measured. These data show important chemical variability among wines, with particular importance placed on the differences, and thus the discriminability potential (measured as the rate max/min) among wines for lactic acid (max/min = 70.8. % of galloylation of tannins (max/min = 55. mean degree of polymerisation of tannins (max/min = 41) and % of prodelphinidins constituting tannins (gallocatechins and epigallocatechins) (max/min = 11).

It is important to note that bitterness and astringency do not present significant linear correlations (r=0.40, P>0.1. which confirms that the participants were not confused and were able to differentiate both sensations (Lea and Arnold, 1978). Astringency scores ranged between 0 and 4 (6 being the maximum possible score) and significant differences among wines (F=15.13; P<0.0001) were observed. These data confirm our first hypothesis related to the strategy followed (selection of grapes from different vineblocks at different maturity points) to generate wines with different chemical compositions, inducing sensory differences in astringency. Table 1 shows that astringency scores present significant (P<0.05) positive linear correlations with six of the 20 chemical variables studied: tannin activity, TPI, ethanol content, % of prodelphinidins in tannins, colour intensity and monomeric pigments.

Our second hypothesis was that sensory and chemical composition do not necessarily follow a linear correlation (i.e. the higher/lower concentration the higher astringency). Thus, astringency scores were modelled from the 20 chemical variables using machine learning algorithms. A highly satisfactory model was obtained. Residual error (measured through root-mean-square deviation, RMSE, by full cross-validation) was 0.19. The best algorithm was SVR (radial kernel). Figure 1 shows the lift chart, which confirms model performance and thus its capability to predict astringency. Model performance was evaluated by calculating possible predictions partitioned into subsegments, deciles or bin. For each bin, average predicted astringency scores (blue line) were compared to average actual values (orange value). Both predicted and actual scores were very closely projected and lines consistently increased, both being indicators of satisfactory model performance and the accuracy of the model.

Figure 2 shows the permutated impact of chemical variables on the model. Half of the chemical variables (10/20) were included in the model, with the most important of these the % of

procyanidins constituting tannins (%PC) and ethanol content (1.0 and 0.98 of normalised impact, respectively). The next important variables, all with normalised impact factors lower than 0.5, were TPI, % of prodelphinidins constituting tannins (%PD. mean degree of polymerisation of tannins (mDP. volatile acidity, tannin activity, % of tannins linked to malvidin (%T-M) and % of galloylated (%G) flavanols constituting tannins.

These results confirm the importance of tannin structure, ethanol content and acids (measured as titratable acids or total acidity and volatile acidity) on the modulation of astringency perception, which is not surprising given the many publications that mention these variables as important drivers of this sensory perception in wine (Ma *et al.*, 2014; Sáenz-Navajas *et al.*, 2012; Soares *et al.*, 2017; Watrelot *et al.*, 2016). However, most of the existing literature tries to establish linear correlations between astringency and chemical composition (i.e. higher/lower levels of a component generate higher astringency. which could be the main source of contradictory results reported when establishing sensory–chemical relationships. In the present work, different tendencies in astringency perception can be observed depending on the levels of a given chemical variable (Figure 3 and supplementary material). Figure 3a shows the partial dependence plot of astringency and % procyanidins (%PC) in tannin structures. Three main tendencies could be identified depending on the %PC: 1) a steep negative linear relationship for %PC< 68%; 2) a moderate positive linear relationship for the 68–76% range, and 3) no change in astringency

associated with different %PC for %PC>76%. Interestingly, the % of total (extension + terminal) catechins and epicatechins (%PC) presents a significant correlation with the % of epicatechin units (r = 0.80; P < 0.001). Thus, this modulation of astringency with %PC could be attributed to changes in the stereochemistry of tannins related to the % of epicatechins in procyanidins. Thus, at low levels of epicatechin subunits, astringency decreases with increases in terminal and extension epicatechins in tannins. However, at intermediate levels (68–76%) of epicatechin units in tannins, increasing epicatechins generate higher astringency, but at higher levels no effect is observed. Results observed for intermediate levels (68–76%) are in agreement with results reported in the literature (Quijada-Morín *et al.*, 2012) in which higher astringency is observed for tannins with higher proportions of epicatechin than catechin subunits. However, in the present non-linear model, two further trends could be identified depending on the % levels of PC (one with negative effect for low levels of PC and other with no effect for higher levels). These results could be explained in terms of structural/ conformational differences of tannins with different structural properties that have different site-specific bindings with tannin (De Freitas and Mateus, 2001; Thorngate and Noble, 1995). Further research in this topic should be carried out to find a plausible explanation.

Figure 3b shows a general positive correlation of astringency and ethanol content with two main tendencies: 1) <14.4% of ethanol (moderate positive linear correlation) and 2) >14.4% of ethanol (steep positive linear correlation). This is
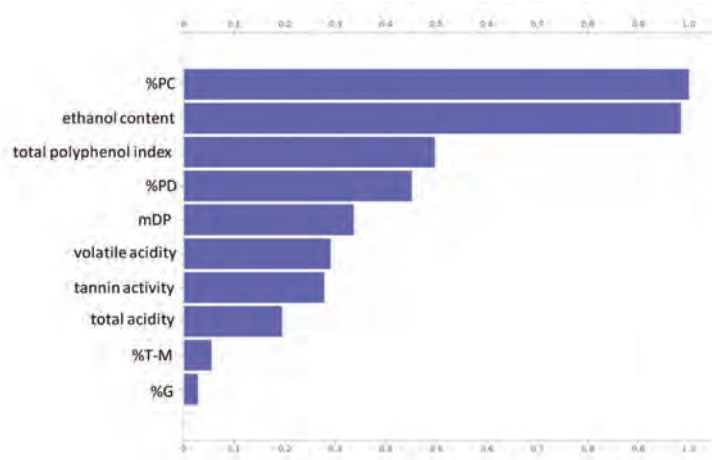

FIGURE 2. Permutated impact factor of chemical variables included in the astringency model.

well in accordance with data reported for commercial wines containing ethanol levels of 13–17% (v/v) (Sáenz-Navajas *et al.*, 2010; Sáenz-Navajas *et al.*, 2012; Watrelot *et al.*, 2016) but contradicts studies carried out with model wines at typical wine ethanol levels of 11–15% (Fontoin *et al.*, 2008; Vidal *et al.*, 2004). These studies report a decrease in astringency with ethanol content, which has been attributed to a decrease of the interaction power between tannins and proteins from 10% of ethanol (hydrophobic + hydrogen-bond interactions) to 15% (hydrogen-bond interactions) (McRae *et al.*, 2015). Thus, the positive correlation between ethanol content and astringency observed in the present work could be attributed to an indirect relationship with phenolic content. Grapes harvested at earlier stages present lower levels of extractable polyphenols but also lower sugar content, yielding wines with lower ethanol levels and polyphenolic concentration and resulting in lower astringency. However, it cannot be ruled out that ethanol can induce astringency-related sensations by mechanisms other than polyphenol-protein interactions. This would be supported by an important number of papers that have established positive relationships between ethanol content and astringency perception (Sáenz-Navajas *et al.*, 2010; Sáenz-Navajas *et al.*, 2012; Watrelot *et al.*, 2016). Additional investigation is needed to understand the relationship between ethanol and astringent sensations.

For the rest of chemical variables included in the model, three different tendencies were globally identified (see supplementary material).

The first trend is observed for the mean degree of polymerisation. For low mDP values (up to approximately 1.4) astringency increases with mDP, while for higher values astringency decreases. This result is well in accordance with the positive linear relationships observed between astringency and DP with low molecular tannins by Peleg *et al.* (1999). Interestingly, Chira *et al.* (2009) also found significant (P = 0.04) positive correlations, but only for skin tannins in one of the two years studied. This sample set presented an average mDP of 21 (range 4–49. which is far out of the range for the wines studied here. This lack of significance for the rest of sample sets (year 2016 and seed tannins of 2006 and 2007) could be attributed to the presence of different relationships depending on the level of mDP as observed in the present work. The effect of the size of tannins (measured through the mDP) on astringency could be explained in terms of tannin hydrophobicity. Thus, even if higher tannin polymerisation can bring more hydrophobic parts, and thus higher astringency (due to higher tannin-protein interactions. this relationship is thought not to be linear and is attributed to conformational arrangements and aggregation processes (Ma *et al.*, 2014).

The second trend is observed for total acidity, volatile acidity, % of prodelphinidins (%PD. total polyphenol index (TPI) and tannin activity (measured as the enthalpy of interaction of
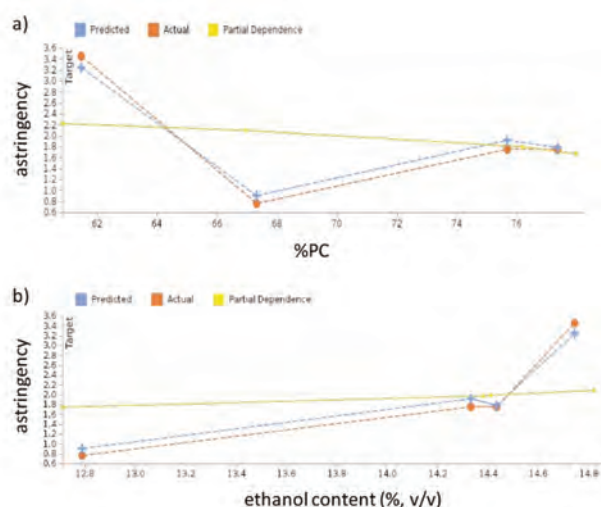

FIGURE 3. Partial dependence plot of astringency with a) %PC and b) ethanol content.

tannins with a hydrophobic surface). These present positive linear relationships with astringency, with this relationship more pronounced at higher values of the corresponding chemical variable and astringency perception. It has been shown that the effect of acidity on astringency is attributed to changes in pH. Thus, for similar pH values, changes in acidity do not have significant effects on astringency (Fontoin *et al.*, 2008). This increase of astringency is attributed to the presence of more phenolate forms and an augmentation of charged molecules, susceptible to participating in protein binding (Siebert and Euzen, 2008). Concerning the positive relationship observed between astringency and %PD, it is interesting to note that this is more important at higher levels (range 1.8–3.2) of astringency. This result is in apparent contradiction with other studies, which have shown in simple model solutions that procyanidins present a faster and stronger interaction with salivary proteins than prodelphinidins (Ferrer-Gallego *et al.*, 2015). At present it is difficult to explain such a relationship, because it is likely that astringency differences related to polyphenolic structures are the result of conformational differences among tannins that cannot yet be measured in such a complex mixture such as wine. To this regard, the measure of tannin activity by HPLC seems to be a promising index. Tannin activity is a parameter that measures the enthalpy of interaction of tannins with a hydrophobic surface. It appeared as an interesting measure of tannin affinity to proteins and thus of wine astringency (Revelette *et al.*, 2014). However, until now no direct relationship with tannins could be established that was attributed to the presence of strong interactions (with polysaccharides or aroma perception) appearing in wines with very different chemical and sensory spaces (i.e. different varieties, winemaking processes, origins, etc) (Watrelot *et al.*, 2016). Thus, working with the same grape variety from a similar origin and processed with the same winemaking protocol could have helped establish such interesting linear relationships between tannin activity and sensory astringency (i.e. drying sensation). It is interesting to note that this is the first time a relationship of this chemical variable with sensory perception has been established.

The third trend is related to the % of galloylated tannins (%G) and tannins linked to malvidin (%T-M. which show a similar relationship with astringency as % PC. Thus, two segments are observed: one for low levels of the chemical variable with a negative linear relationship with astringency and a second for high values with a positive linear correlation. As explained above, the constitutive units of tannins as well as their polymerisation degree play an important role in tannin conformational structures, which drives tannin hydrophobicity, tannin-protein interactions and thus perceived astringency.

CONCLUSIONS

The present work has successfully modelled the perception of wine astringency from its chemical composition by applying machine learning approaches. This strategy has explained non-linear relationships by means of the SVR (radial kernel) algorithm, which showed a very low residual error between actual and predicted astringency scores. This and the fact that sensory perception is distinctly non-linear show the necessity of considering non-linear models to explain sensory precepts from chemical composition.

The main drivers of the astringency model were related to ethanol content (potentially elicited by a mechanism different from polyphenol-protein interactions. acidity (related to pH variations. as well as to effects of chemical variables linked to tannin structure, such as 1) the constitutive subunits of tannins (%PC, %PD, %G and %T-M. 2) tannin activity measured as the enthalpy of interaction with a hydrophobic surface, and 3) the size of tannins measured by the mean degree of polymerisation.

The results presented here increase understanding of astringency perception and provide wine producers with objective tools to help them control and optimise grape and wine production stages for further modulating astringency and thus maximising the quality and consumer appeal of their wines.

REFERENCES

Ares G., Bruzzone Fernanda, Vidal L., Silva Cadena R., Giménez A., Pineau B., Hunter D., Paisley A.G. and Jaeger S.R., 2014. Evaluation of a rating-based variant of check-all-that-apply questions: Rate-all-that-apply (RATA), *Food Quality and Preference*, 36, 87-95. doi:10.1016/j.foodqual.2014.03.006

Bross I.D.J., 1958. *How to Use Ridit Analysis, Biometrics*, 14 (1) 18-38. doi:10.2307/2527727

Canon F., Neiers F. and Guichard E., 2018. Saliva and Flavor Perception: Perspectives, *Journal of Agricultural and Food Chemistry*, 66 (30) 7873-79. doi:10.1021/acs.jafc.8b01998

Colonna A.E., Adams D.O. and Noble A.C., 2004. Comparison of procedures for reducing astringency carry-over effects in evaluation of red wines, *Australian Journal of Grape and Wine Research*, 10 (1) 26-31. doi:10.1111/j.1755-0238.2004.tb00005.x

Charters S. and Pettigrew S., 2007. The dimensions of wine quality, *Food Quality and Preference*, 18 (7) 997-1007. doi:10.1016/j.foodqual.2007.04.003

Chira K., Schmauch G., Saucier C., Fabre S. and Teissedre P.-L., 2009. Grape variety effect on proanthocyanidin composition and sensory perception of skin and seed tannin extracts from Bordeaux wine grapes (Cabernet-Sauvignon and Merlot) for two consecutive vintages (2006 and 2007), *Journal of Agricultural and Food Chemistry*, 57 (2) 545-53. doi:10.1021/jf802301g

D'Alessandro S. and Pecotich A., 2013. Evaluation of wine by expert and novice consumers in the presence of variations in quality, brand and country of origin cues, *Food Quality and Preference*, 28 (1) 287-303. doi:10.1016/j.foodqual.2012.10.002

De Freitas V. and Mateus N., 2001. Structural features of procyanidin interactions with salivary proteins, *Journal of Agricultural and Food Chemistry*, 49 (2) 940-45. doi:10.1021/jf000981z

Dinnella C., Recchia A., Tuorila H. and Monteleone E., 2011. Individual astringency responsiveness affects the acceptance of phenol-rich foods, *Appetite*, 56 (3) 633-42. doi:10.1016/j.appet.2011.02.017

Ferrer-Gallego R., Ferrer-Gallego N., Quijada-Morín N, F. Brás N., Gomes P., de Freitas V., Rivas-Gonzalo J.C. and Teresa Escribano-Bailón M.T., 2015. Characterization of sensory properties of flavanols-A molecular dynamic approach, *Chemical Senses*, 40 (6) 381-90. doi:10.1093/chemse/bjv018

Fontoin H., Saucier C., Teissedre P.-L. and Glories Y., 2008. Effect of pH, ethanol and acidity on astringency and bitterness of grape seed tannin oligomers in model wine solution, *Food Quality and Preference*, 19 (3) 286-91. doi:10.1016/j.foodqual.2007.08.004

García-Estévez I., Ramos-Pineda A.M. and Escribano-Bailón M.T., 2018. Interactions between wine phenolic compounds and human saliva in astringency perception, *Food and Function*, 9 (3) 1294-309. doi:10.1039/C7FO02030A

Glories Y., 1984. La couleur des vins rouges. Mesure, origine et interprétation. Partie I. , *Connaissance de la vigne et du vin*, 18, 195-217. doi:10.20870/oeno-one.1984.18.4.1744

Gonzalo-Diago A., Dizy M. and Fernandez-Zurbano P., 2013. Taste and Mouthfeel Properties of Red Wines Proanthocyanidins and Their Relation to the Chemical Composition, *Journal of Agricultural and Food Chemistry*, 61 (37) 8861-70. doi:10.1021/jf401041q

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, *American Journal of Enology and Viticulture*, 54 (4) 301-06.

Jover A.J.V., Montes F.J.L. and Fuentes M.D.F., 2004. Measuring perceptions of quality in food products: the case of red wine, *Food Quality and Preference*, 15 (5) 453-69. doi:10.1016/j.foodqual.2003.08.002

Labarbe B., Cheynier V., Brossaud F., Souquet J.-M. and Moutounet M., 1999. Quantitative fractionation of grape proanthocyanidins according to their degree of polymerization, *Journal of Agricultural and Food Chemistry*, 47 (7) 2719-23. doi:10.1021/jf990029q

Lamy E., 2018. Salivary proteomics in ingestive behaviour research: Advances, potentialities and limitations, *Journal of Integrated OMICS*, 8 (1) 14-18. doi:10.5584/jiomics.v8i1.231

Lamy E., Mowe M., Pinheiro C., Rodrigues L., Lopes O. and Capela e Silva F., 2017. Oral food perception and polyphenol-rich foods acceptance - the importance of knowing individuals saliva characte-ristics for promoting consumption, *Journal of the International Society of Antioxidants*, 3 (2) 1-4. Doi:10.18143/JISANH_v3i4_1338

Lea A.G.H. and Arnold G.M. 1978. The phenolics of ciders: Bitterness and astringency, *Journal of the Science of Food and Agriculture*, 29 (5) 478-83. doi:10.1002/jsfa.2740290512

Ma W., Guo A., Zhang Y., Wang H. Liu Y. and Li H., 2014. A review on astringency and bitterness perception of tannins in wine, Trends in Food Science *and Technology*, 40 (1) 6-19. doi:10.1016/j.tifs.2014.08.001

Masi C., Dinnella C., Monteleone E. and Prescott J., 2015. The impact of individual variations in taste sensitivity on coffee perceptions and preferences, *Physiology and Behavior*, 138, 219-26. doi:10.1016/j.physbeh.2014.10.031

McRae J.M., Ziora Z.M., Kassara S., Cooper M.A. and Smith P.A., 2015. Ethanol concentration influences the mechanisms of wine tannin interactions with poly(l -proline) in model wine, *Journal of Agricultural and Food Chemistry*, 63 (17) 4345-52. doi:10.1021/acs.jafc.5b00758

Peleg H., Gacon K., Schlich P. and Noble A.C., 1999. Bitterness and astringency of flavan-3-ol monomers, dimers and trimers, *Journal of the Science of Food and Agriculture*, 79 (8) 1123-28. doi:10.1002/(SICI)1097-0010(199906)79:8<1123::AID-JSFA336>3.0.CO;2-D

Prescott J., Young O., O'Neil L. and Stevens R., 2002. Motives for food choice: a comparison of consumers from Japan, Taiwan, Malaysia and New Zealand, Food Quality and Preference, 13 (7-8) 489-95. doi:10.1016/S0950-3293(02)00010-1

Quijada-Morín N., Regueiro J., Simal-Gándara J., Tomás E., Rivas-Gonzalo J. and Escribano-Bailón M.T., 2012. Relationship between the Sensory-Determined Astringency and the Flavanolic Composition of Red Wines, *Journal of Agricultural and Food Chemistry*, 60 (50) 12355-61. doi:10.1021/jf3044346

Revelette M.R., Barak J.A. and Kennedy J.A., 2014. High-performance liquid chromatography determination of red wine tannin stickiness, *Journal of Agricultural and Food Chemistry*, 62 (28) 6626-31. doi:10.1021/jf501666z

Ribéreau-Gayon P., 1970. Le dosage des composés phénoliques totaux dans les vins rouges, *Chemist-Analyst*, 52, 627-31.

Sáenz-Navajas M.P., Tao Y.S., Dizy M., Ferreira V. and Fernández-Zurbano P., 2010. Relationship between nonvolatile composition and sensory properties of premium Spanish red wines and their correlation to quality perception, *Journal of Agricultural and Food Chemistry*, 58 (23) 12407-16. doi:10.1021/jf102546f

Sáenz-Navajas M.P., Avizcuri J.M., Ferreira V. and Fernández-Zurbano P., 2012. Insights on the chemical basis of the astringency of Spanish red wines, *Food Chemistry*, 134 (3) 1484-93. doi:10.1016/j.foodchem.2012.03.060

Siebert K.J. and Euzen C., 2008. The relationship between expectorant ph and astringency perception, *Journal of Sensory Studies*, 23 (2) 222-33. doi:10.1111/j.1745-459X.2008.00151.x

Thorngate J.H. and Noble A.C., 1995. Sensory evaluation of bitterness and astringency of 3R(−)-epicatechin and 3S(+)-catechin, *Journal of the Science of Food and Agriculture*, 67 (4) 531-35. doi:10.1002/jsfa.2740670416

Vidal S., Courcoux Philippe, Leigh Francis, Kwiatkowski M., Gawel R., Williams P., Waters E. and Cheynier V., 2004. Use of an experimental design approach for evaluation of key wine components on mouth-feel perception, *Food Quality and Preference*, 15 (3) 209-17. doi:10.1016/S0950-3293(03)00059-4

Watrelot A.A., Byrnes N.K., Heymann H. and Kennedy J.A., 2016. Understanding the Relationship between Red Wine Matrix, Tannin Activity, and Sensory Properties, Journal of Agricultural and Food Chemistry, 64 (47) 9116-23. doi:10.1021/acs.jafc.6b03767