

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Relationships between skin flavonoid content and berry physical-mechanical properties in four red wine grape cultivars (*Vitis vinifera* L.)

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1567475> since 2017-05-19T09:32:50Z

Published version:

DOI:10.1016/j.scienta.2015.09.053

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)



UNIVERSITÀ DEGLI STUDI DI TORINO

This Accepted Author Manuscript (AAM) is copyrighted and published by Elsevier. It is posted here by agreement between Elsevier and the University of Turin. Changes resulting from the publishing process - such as editing, corrections, structural formatting, and other quality control mechanisms - may not be reflected in this version of the text. The definitive version of the text was subsequently published in:

Scientia Horticulturae 197 (2015) 272–279; <http://dx.doi.org/10.1016/j.scienta.2015.09.053>

You may download, copy and otherwise use the AAM for non-commercial purposes provided that your license is limited by the following restrictions:

- (1) You may use this AAM for non-commercial purposes only under the terms of the CC-BY-NC-ND license.
- (2) The integrity of the work and identification of the author, copyright owner, and publisher must be preserved in any copy.
- (3) You must attribute this AAM in the following format: Creative Commons BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>), <http://www.sciencedirect.com/science/article/pii/S0304423815302181>

Relationships between skin flavonoid content and berry physical-mechanical properties in four red wine grape cultivars (*Vitis vinifera* L.)

Luca Brillante^{1*}, Diego Tomasi¹, Federica Gaiotti¹, Simone Giacosa², Fabrizio Torchio², Susana R o Segade², Ren e Siret³, Imen Zouid³, Luca Rolle²

¹*CREA-Council for Agricultural Research and Economics, Viticulture Research Center, Via XXVIII Aprile 26, Conegliano (TV), Italy.*

²*University of Turin, Dipartimento di Scienze Agrarie, Forestali e Alimentari, Largo Paolo Braccini 2, 10095 Grugliasco (TO), Italy.*

³*LUNAM Universit , SFR 4207 QUASAV, UMT VINITERA, UPSP GRAPPE, Groupe ESA, 55 Rue Rabelais, Angers F-49007, France.*

*Corresponding author: brillanteluca@live.it

ABSTRACT

Flavonoids are a class of bioactive compounds extremely important in food and wine industry. The development of rapid methods for their quantification in grape berries is one of the modern challenges in viticulture and enology research. Total flavonoid (TF) amount changes during grape ripening and also berry physical-mechanical properties, as evaluated by instrumental texture analysis, change in the same period. In this work, TF and berry physical-mechanical parameters were linked together through predictive models. Models were developed for each of four red wine grape cultivars: Brancellao, Cabernet Franc, Menc a and Merenzao, and another one considered all cultivars together. These models reached high accuracy and allowed to predict TF in grape berries with a low error (RMSE from 0.15 ± 0.07 mg g⁻¹ to 0.35 ± 0.10 mg g⁻¹ in prediction, as evaluated by cross-validation).

Berry weight (BW) was the parameter having the largest influence on TF predictions, and also

was the only variable having part in all models. BW and chewiness had a similar behavior and when berry weight was excluded, chewiness was able to substitute its role in all models.

The other physical-mechanical characteristics displayed a different behavior across cultivars.

In conclusion, this work shows that it is possible to predict TF from physical-mechanical predictors in grape berries and that cultivar specific models reach higher accuracy for this purpose than the multi-cultivar model.

Keywords: Multivariate Adaptive Regression Splines (MARS); total flavonoids; berry weight; mechanical properties; red wine grape cultivars

ABBREVIATIONS

BW Berry Weight

Co Cohesiveness

Ch Chewiness

E_{sk} Skin Young's modulus

F_{sk} Skin break force

GCV General Cross Validation

R Resilience

Sp_{sk} Skin thickness

W_{sk} Skin break energy

TA Texture Analysis

TF Total Flavonoids Index

1. INTRODUCTION

Flavonoids are a class of bioactive compounds having nutraceutical properties. Recently, they have received a great attention because of their beneficial effects on human health. In particular,

their antioxidant (Lourenço et al., 2008) and hypocholesterolemic (Gonzalez et al., 2014) properties make these compounds the most known “French paradox” contributors (Sun et al., 2002). Flavonoids greatly affect the sensory and nutritional quality of fruits and vegetables (Harnly et al., 2006) and their derived products like wine (Ristic et al., 2010). Red wine contains more flavonoids than white one due to the quantity originally present in grapes (González-Neves et al., 2004), but in particular to differences in wine-making practices. The biosynthesis and concentration of grape flavonoids greatly depend on cultivar, vineyard practices, soil and climate (Koundouras et al., 2006). Quantitative and qualitative modifications of flavonoids and their easiness to be released during maceration occur during the ripening period (Zouid et al., 2010; Kuhn et al., 2013), and therefore the harvest date is an important parameter to determine the potential amount of flavonoids in wine. In this context, modern grape-growing practices are focused on the optimization of flavonoid content in grapes.

Accurate determination of flavonoids originally present in grape and of their potential extractability is therefore of extreme importance in the wine industry. There are several analytical procedures to determine flavonoids in fruits, which are based on the use of spectrophotometry, chromatography and mass spectrometry (Ignat et al., 2011 for a generic review and Lorrain et al., 2013 for the case of grape and wine). While these methods have proven to be accurate, they are time consuming or require great initial and/or utilization costs. Therefore, in the last ten years, researchers have tried to develop rapid and cheap methods to obtain a faster screening of flavonoids in grapes. The greatest advances in this field have been made with the use of Near InfraRed (NIR) spectroscopy (Ferrer-Gallego et al., 2011; Rolle et al., 2012; Cozzolino, 2015 among others). In this work, we the potentialities of another different approach based on instrumental Texture Analysis (TA). This analytical technique was already successfully applied

to assess the flavonoid content in grape seed acoustic parameters of seeds highly correlated with extractable flavonoid (Torchio et al., 2012). However, at our knowledge, no study was performed in order to evaluate the relationship among flavonoid concentration in grape skin. Berries accumulate flavonoids in the skin during the ripening phase (Kuhn et al., 2013), and at the same time change their physical-mechanical properties (Le Moigne et al., 2008; R o Segade et al., 2011a). Therefore, it interesting to try to link these two phenomena. Berry becomes softer and more cohesive, whereas the berry skin is more elastic as ripening progresses (Le Moigne et al., 2008; R o Segade et al., 2011a). Although these physical modifications can be evaluated by sensory panels to support harvesting decisions, TA generally permits to reduce the variability associated with the subjectivity of sensory analysis through quantitative and objective measures. TA has already shown to be effective for an accurate evaluation of physical-mechanical characteristics in grapes (Letaief et al., 2008a, Battista et al., 2015) and has been suggested for the evaluation of extractable anthocyanins (R o Segade et al., 2011b,c).

This work aims to extend the use of TA to the prediction of total flavonoids in grape skin. Furthermore, a deeper knowledge of the relationships between the physical-mechanical properties and chemical composition of grapes will also address new approaches in vineyard management to improve those properties related to high quality berries. In this article, predictive models for the evaluation of total flavonoids in four red wine grape cultivars developed over a dataset composed of 480 berries. The contribution of each physical-mechanical predictor to the performance of the model described and differences among cultivars underlined.

2. MATERIALS AND METHODS

2.1 Grape sampling

Grape samples of *Vitis vinifera* L. cv Mencía, Brancellao and Merenzao were harvested in various vineyards located in different sub-zones of the Ribeira Sacra Denomination of Origin (Galicia, Northwest Spain). The samples of Cabernet Franc were collected from different vineyards in the Loire Valley (France). For each cultivar, approximately 1000 grape berries were randomly picked with attached short pedicels. The berries were visually inspected, and those with damaged skins were discarded. For each cultivar, a set of 120 berries was randomly selected (with a total of 480 berries) to determine physical-mechanical properties and total flavonoid content. Before the analysis the weight of each berry was determined using a technical balance (Gibertini E1700, Modena, Italy).

2.2 Instrumental texture analysis

Analyses were made with a Universal Testing Machine (UTM) TAxT2i texture analyzer (Stable Micro Systems, Godalming, Surrey, UK) equipped with a HDP/90 platform and a 5 kg load cell. All data were acquired at 400 Hz, and the mechanical properties were calculated using the Texture Expert Exceed software package (Stable Micro Systems). Each grape berry was singularly weighed with an analytical laboratory balance, Radwag 109 AS 220/X, Radwag, Radom, Poland, and a texture profile analysis (TPA) non-destructive mechanical test was then performed (Letaief et al., 2008a). Each whole berry was individually compressed in the equatorial position using a SMS P/35 flat cylindrical probe (Stable Micro Systems) under 25% deformation, with a waiting time between the two bites of 2 s and a test speed of 1 mm s⁻¹. The following mechanical properties of whole berry were determined: hardness (N, as H), cohesiveness (adimensional, as Co), gumminess (N, as G), springiness (mm, as S), chewiness (mJ, as Ch) and resilience (adimensional, as R). A puncture test was then carried out on the same berries using a SMS P/2N needle probe (Stable Micro Systems), a test speed of 1 mm s⁻¹ and a

penetration depth of 3 mm (Letaief et al., 2008b). Each berry was individually punctured in the lateral face, and the following three skin mechanical parameters were measured: skin break force (N, as F_{sk}), skin break energy (mJ, as W_{sk}) and skin resistance to the axial deformation ($N\text{ mm}^{-1}$, as E_{sk}). The use of a needle probe minimizes the possible interference from the pulp firmness. Finally, skin thickness (μm , as $S_{p_{sk}}$) was individually measured in each berry by a compression test after the skin was manually removed from the pulp using a laboratory spatula (Río Segade et al., 2011c). This test was carried out using a SMS P/2 flat cylindrical probe and a test speed of 0.2 mm s^{-1} .

2.3 Chemical analysis

Once finalized the skin thickness test, each berry skin was individually immersed into 5 mL of a hydroalcoholic buffer solution of pH 3.2 containing 5 g L^{-1} tartaric acid, 2 g L^{-1} sodium metabisulphite and 12% v/v ethanol (Di Stefano and Cravero, 1991). Each skin was then homogenized at 8000 rpm for 1 min using a Ultraturrax T25 high-speed homogenizer (IKA Labortechnik, Staufen, Germany), and centrifuged in a PK 131 centrifuge (ALC International, MI, Italy) for 15 min at $3000\times g$ at $20\text{ }^{\circ}\text{C}$. The supernatant was used for analysis after dilution with an ethanolic solution of HCl (70:30:1, ethanol:water:HCl, v/v) (Di Stefano and Cravero, 1991). Total flavonoid index (TF) was determined by a spectrophotometric method using a UV-1601PC spectrophotometer (Shimadzu Scientific Instruments Inc., Columbia, MD, USA) and expressed as mg g^{-1} of (+)-catechin (Di Stefano and Cravero, 1991; Torchio et al., 2012).

2.4 Chemometrics

Before starting the statistical analysis for the 480 berries, 160 unique samples were obtained by averaging three berries at a time, randomly selected without replacement inside the same cultivar.

For each cultivar, a model was then fitted in order to predict TF as a function of TA predictors. For model fitting purposes, the Multivariate Adaptive Regression Splines (MARS) algorithm (Friedman, 1991; Hastie et al., 2009) was used. This technique builds a linear first-order spline for each predictor, which then added to a linear regression model instead of the original predictor. The algorithm isolates portions of data where the relationship between the response and a predictor is constant and can be approximated by a linear function. These linear relationships are then joined at knotpoints that are defined as values of the predictor where the relationship with the outcome varies.

Mathematically, MARS defines a series of piecewise linear basis functions of the form $(x - c)_+$ and $(c - x)_+$. The symbol “+” means positive part, so:

$$(x - c)_+ = \begin{cases} c - x, & x < c \\ 0, & x \geq c \end{cases} \quad (1)$$

$$(c - x)_+ = \begin{cases} x - c, & x > c \\ 0, & x \leq c \end{cases} \quad (2)$$

where c is the value of the parameter at the knotpoint. The spline function is then added to the model in a linear form, so:

$$y = \beta_0 + \beta_1(x - c)_+ + \beta_2(c - x)_+ + \varepsilon \quad (3)$$

where y is the outcome, x is the predictor, c is the knotpoint value, β_0 is the intercept, β_1 and β_2 are the coefficients, ε is the error and the symbol “+” means positive part. The procedure is iterative; for all predictors, every data point is evaluated as a possible knotpoint, and the linear relationship and the related error are calculated. The data point that allows the lowest error, once used as knotpoint, is retained in the final model. The utility of simple linear terms (i.e. no spline)

is also evaluated. Once the full set of splines has been created, the algorithm sequentially removes terms and functions that do not contribute significantly to the root mean square error (RMSE) reduction. To determine the contribution of each term and function, MARS uses the general cross validation (GCV) statistic. This procedure also estimates how much the error rate is decreased by including a predictor in the model, allowing the classification of predictors in order of importance. The algorithm does not add terms if these do not further reduce the GCV error. The MARS model allows the use of interactions between predictors, and also, in this case, splines are created. In addition to the internal GCV statistic, which is in general optimistic, we tuned the MARS model for the number of terms and the degree of interaction using 25 repeats of 10-fold cross-validation; the RMSE was used as a metric. The same resampling technique was used to train the model and to estimate its performance on new data. The MARS model is sensible to correlated predictors. Therefore, before fitting the model, all available predictors were filtered in order to maintain the maximum correlation between predictors lower than 0.7.

Statistical analysis, data preparation and management were performed with the R GNU statistical software environment (R v3.1.2 R Core Team 2012), using the earth package (Milborrow, 2014) to fit the MARS models.

3. RESULTS

3.1 Grape composition variability

As described in the previous section 2.4, available predictors were filtered for collinearity before continuing the analysis. This led to select berry weight (BW), W_{sk} , E_{sk} , Sp_{sk} , Co, Ch and R as predictors of TF. In the whole dataset, some of them were correlated (Ch with BW and E_{sk}). When cultivars were considered independently, these predictors were correlated in just half of the groups (in Merenzao and Brancellao). All of these three predictors (E_{sk} , Ch and BW) were

therefore retained for the following exploratory analysis, and their contemporary use was justified when developing models for cultivars where they were not correlated (Cabernet Franc and Mencía). In the current work, we also tried to discriminate between the role of BW against that of TA predictors. Indeed, while the role of BW on grape flavonoid content is well known (Barbagallo, et al 2011) other physical-mechanical predictors have been introduced just recently (Letaief, et al. 2008) and their effect under active study (Zsófi, Zs. et al. 2014). Therefore in the following section (section 3.2), models fitted with and without BW. For those cultivars where Ch was correlated with BW and E_{sk} , Ch was excluded to allow the determination of the role of BW. The exclusion of Ch solved the problem of correlations among predictors for Brancellao and Merenzao grape varieties.

Figure 1 shows descriptive statistics for the whole dataset. The mean BW for all data together was 1.92 g. Merenzao was the cultivar with the heaviest berries in average (2.32 g), followed by Mencía (2.26 g), then Brancellao (1.74 g) and finally Cabernet Franc with the lightest berries (1.35 g). The average W_{sk} was higher for Cabernet Franc (0.56 mJ) than for Mencía (0.41 mJ), Merenzao (0.35 mJ) and Brancellao (0.26 mJ), while the mean W_{sk} for all cultivars was 0.39 mJ. Regarding Sp_{sk} , Cabernet Franc and Mencía showed the highest average values (287 and 285 μm), followed by Merenzao (250 μm) and Brancellao (234 μm). The Spanish cultivars had a very similar average E_{sk} (approx. 0.29 N mm^{-1}), but value was higher than that of Cabernet Franc (0.23 N mm^{-1}). A similar trend was observed for Ch, for which Cabernet Franc showed the lowest average values (2.82 mJ), while Merenzao had the highest (6.55 mJ). However, the average Ch of Merenzao berries was quite different from that of Brancellao (5.29 mJ), but not from that of Mencía (6.12 mJ). The mean Ch for all data together was 5.2 mJ. Lower variability was observed between cultivars for Co and R, and this last variable showed limited range of

variation even inside the same cultivar. The mean Co for the whole dataset was 0.72 with a variance of 0.001, while the mean R was 0.37 with a variance of 0.001. Conversely, TF greatly varied across cultivars and allowed to well differentiate them. Cabernet Franc was the cultivar with the highest average TF (4.30 mg g^{-1}), followed by Mencía (3.59 mg g^{-1}), Brancellao (2.60 mg g^{-1}) and Merenzao (1.88 mg g^{-1}).

In the whole dataset, BW and Ch were the parameters better related to TF, having respectively a *r*-value of -0.63 (BW) and -0.65 (Ch). This significant and negative correlation between BW and TF was observed in all cultivars, and *r* was -0.61 for Mencía, -0.64 for Merenzao, -0.76 for Brancellao, while the highest value of -0.89 was recorded for Cabernet Franc. It appears that if the correlation is significant in the whole dataset, it is generally more scattered than when cultivars are considered independently, as shown in figure 2. This figure plots BW versus TF for all data; the color is mapped to the cultivar, and the size of the points to Ch. Figure 2 also shows the negative correlation between Ch and TF in all cultivars. In opposition to what seen for BW, the relationship of Ch with TF was higher when the whole dataset was considered than when it was splitted by cultivar, with the exception of Cabernet Franc (*r*= -0.66). In detail, the last relationship was not significant for Mencía, while *r* was -0.52 for Merenzao and -0.48 for Brancellao.

From the exploratory analysis, it appears that there are strong differences in the behavior and the variability of the TA predictors in all cultivars. Therefore, in the following section, a different model was developed for each cultivar to predict TF, and another one for all cultivars together. Results of all models and the role of the predictors across cultivars will then be compared.

3.2 Model fitting

First linear regression were used to fit a function relating TF to BW alone, without texture parameters. On the train set, the linear regression for Brancellao had a R^2 value of 0.57 and a RMSE of 0.17 mg g^{-1} , for Cabernet a R^2 of 0.78 and a RMSE of 0.36 mg g^{-1} , for Mencia a R^2 of 0.36 and a RMSE of 0.33 mg g^{-1} , for Merenzao a R^2 of 0.39 and a RMSE of 0.17 mg g^{-1} .

To improve these results, the MARS technique was used for model fitting, and texture analysis parameter were included as predictors. This technique was chosen because, in the exploratory analysis, several predictors (BW is a noticeable example) were found to be well related to the outcome in an approximately linear way. Linear regression could be used in these cases, but MARS has been preferred because not all predictors showed linear relationships. MARS is advantageous in these situations, being able to model non-linear relationships. With respect to linear regression, MARS also allows to test predictors and the presence of interactions in an extensive fashion, because of the built-in backward feature selection routine. This avoids the use of predictors and terms that do not significantly reduce the error in the prediction of unseen data as evaluated by cross-validation (GCV more specifically).

A MARS model was therefore fitted for each cultivar in the dataset. All models were tuned for the number of terms and degree of interaction, and performances on future data were evaluated using 25 repetitions of ten-fold cross validation. Once tuned, the number of terms and degree of interaction varied across all optimized models. The number of terms was 5 for Brancellao, 3 for Cabernet Franc, 7 for Mencia and 4 for Merenzao. Except for Mencia where the degree of interaction was 1, which indicates absence of interactions, models for the other cultivars used up to 2 degrees of interaction among predictors for Merenzao and Brancellao, and up to 3 for Cabernet Franc. The model built on the comprehensive dataset, containing all data from all

cultivars, was composed of 7 terms and 2 degrees of interaction.

3.2.1 Models including BW

Figure 3 shows the results for all cultivars on the train set. These results are improved with respect to the use of linear regression with BW alone as predictor, except for Cabernet. In this cultivar results are comparable because even the MARS approach uses BW alone as predictor and does not include texture analysis parameters in the model. In the cross-validation procedure, which simulates performances on future and unseen data, the best R^2 -values were obtained for Cabernet Franc ($R^2 = 0.83 \pm 0.17$, Fig.3), probably because this cultivar had the highest variance in TF (0.58 mg g^{-1}). R^2 -values were 0.73 ± 0.25 for the model developed for Brancellao, followed by 0.65 ± 0.31 for Merenzao and finally 0.56 ± 0.33 for Mencía. The RMSE values were lower for those cultivars where the variance of TF was also lower. The lowest RMSE corresponded to Brancellao (RMSE = 0.15 ± 0.06) and Merenzao (RMSE = 0.15 ± 0.10), followed by Mencía (RMSE = 0.32 ± 0.12) and Cabernet Franc (RMSE = 0.35 ± 0.10). Because the TF values for these cultivars were not on the same scale, the comparison of RMSE does not make sense. Therefore, normalization is required dividing RMSE by the mean of the outcome for all cultivars to calculate the RMSE coefficient of variation (CoV). This parameter estimates performances as percentage of error over the mean. The lowest CoV computed for prediction RMSE, as evaluated by cross-validation, were obtained for Brancellao ($0.058 \pm 0.028\%$), Merenzao ($0.080 \pm 0.053\%$), Cabernet Franc ($0.081 \pm 0.023\%$) and Mencía ($0.089 \pm 0.033\%$)

The tuned model built for the whole dataset had a R^2 -value of 0.70 ± 0.14 in cross-validation and RMSE of 0.57 ± 0.11 , which corresponded to CoV(RMSE) of $0.184 \pm 0.035\%$.

Because MARS models are fairly simple, being just additive linear splines, they could be directly

applied to infer TF in grape berries. However, the validity of these models is limited to the same cultivar and TA data in the ranges observed here. The formulas for all models will be given hereafter (terms are rounded to 2 digits, unrounded coefficients are available from the corresponding author). Model formulas will also allow an easier description of terms different from BW, which will be reported in the following section 3.3.

Brancellao:

$$TF = 2.7 - 0.20(BW - 0.33)_+ + 0.27(0.33 - BW)_+ - 0.15(0.20 - E_{sk})_+ + 0.09(W_{sk} - 0.50)_+ (E_{sk} - 0.20)_+ \quad (4)$$

Cabernet Franc:

$$TF = 3.90 - 0.54(BW - 0.43)_+ + 0.80(0.43 - BW)_+ \quad (5)$$

Mencía:

$$TF = 4.46 - 0.37(BW - 0.43)_+ - 0.45(0.28 - E_{sk})_+ - 0.35(Sp_{sk} - 0.37)_+ - 0.49(0.55 - Sp_{sk})_+ - 0.22(0.16 - Co)_+ + 0.79(-0.53 - Ch)_+ \quad (6)$$

Merenzao:

$$TF = 1.85 + 0.38(-0.67 - BW)_+ + 0.69(BW - 0.67)_+ (Co - 0.63)_+ - 0.10(BW - 0.67)_+ (R - 0.15)_+ \quad (7)$$

where TF is the amount of total flavonoids (mg g^{-1}), BW is berry weight (g), W_{sk} is skin break energy (mJ), E_{sk} is skin resistance to the axial deformation (N mm^{-1}), Sp_{sk} is skin thickness (μm), Co is berry cohesiveness (adimensional), Ch is berry chewiness (mJ) and R is berry resilience (adimensional).

3.2.2 Models without BW

BW was excluded from the global physical-mechanical predictors to test the effect of TA predictors alone. Models were refitted and re-tuned on the same set of re-samples previously used to build the model including BW. This will permit to compare the two versions of the model. Once tuned as described in the section 2.4, the best model for Brancellao and Merenzao used 3 terms and no interactions, that for Cabernet Franc used 2 terms and 2 degrees of interaction, and that for Mencía used 7 terms and no interactions. All cultivars together were better modeled by using 7 terms and 3 degrees of interaction. In the cross-validation procedure, Brancellao had $CV(RMSE) = 0.195 \pm 0.069 \text{mg/kg}$ and $R^2 = 0.56 \pm 0.33$, Cabernet Franc had $CV(RMSE) = 0.615 \pm 0.204 \text{mg/kg}$ and $R^2 = 0.50 \pm 0.34$, Mencía had $CV(RMSE) = 0.349 \pm 0.116 \text{mg/kg}$ and $R^2 = 0.51 \pm 0.34$, and Merenzao had $CV(RMSE) = 0.146 \pm 0.092 \text{mg/kg}$ and $R^2 = 0.68 \pm 0.29$. For models without BW, performances changed and generally decreased when compared with those of models including BW. The greatest reduction was observed for Cabernet Franc, while no reduction was observed at all for Merenzao.

3.3 Model description

In order to evaluate the weight of each predictor in each model, we computed the variable importance by taking advantage of the implicit backward feature selection of MARS models. Variable importance is computed on the basis of the reduction in the GCV statistics corresponding to the addition of each term for each predictor in the model. The total reduction, once the final number of basic functions is determined, is used as variable importance metric. When a predictor is never used, its influence in the model is obviously equal to 0 whereas predictors with the highest influence have a value of 100. Data are shown in Table 1 for best-performance models, thus including BW. Other ways of computing the variable importance in a

MARS model are available, specifically other metrics can be used instead of the GCV used here, such as the reduction in the Residual Sum of Square. This last metric was also tested, but the results were not different. The topological order of all predictors did not change at all, and the scaled value changed very slightly (1 or 2 digits).

Table 1 shows that BW is the predictor allowing the strongest reduction in the GCV error for all cultivars taken both singularly and all together. Therefore, BW is the predictor with the largest influence in all models, however this can also be a spurious effect due to the expression of TF by weight. In Cabernet Franc, where the linear relation between TF and BW is the strongest (see section 3.1), the other terms do not improve the relation and do not contribute to further reduce the error. For this cultivar, BW is the only predictor used in the model and is the only variable that has some influence in all models. With the exception of Cabernet Franc, all TA predictors contribute at least in one model. All predictors enter the model for the multicultivar dataset, except Ch that was previously excluded from this global dataset because of its correlation with BW. Generally, the stronger influence of BW involves better model performances (see also figure 3 and section 3.2). Other predictors can substitute the role of BW in the models and, when the relation between BW and TF is more scattered, the overall performances of the model decrease.

Considering variables different from BW, Co contributes to the GCV error reduction for Mencía and Merenzao, E_{sk} for Brancellao and Mencía, while all other predictors have some importance just for one single cultivar model. In the whole dataset, all used predictors (Ch was previously excluded) enter the model and therefore contribute to the reduction in the GCV error. Co enters the model for Merenzao (eq. 7) (and also that for the whole dataset, data not-shown) in interaction with BW; this predictor singularly contributes to the model only for Mencía (eq. 6). In

this last model, interactions are not present at all. Ch has been excluded from three of the five models because it was correlated with BW. Ch is present in the models for Mencía and Cabernet Franc but it contributes only in the first, while as already shown it is useless for Cabernet Franc (Table 1). E_{sk} enters the models individually for Brancellao and Mencía (eqs. 4 and 6, respectively) and also in interaction with W_{sk} in this first model (Brancellao, eq. 4), while it is used exclusively in interaction with BW in the model for all cultivars together. W_{sk} is used without interactions in the global model. In the model for Merenzao, R is used in interaction with BW and contributes equally to the model for all cultivars. Sp_{sk} is the second variable in order of importance for Mencía (Table 1) and is used in interaction with R in the global model.

When BW is excluded as predictor (data not shown) and Ch is included in all models (see section 3.2.2), Ch becomes the predictor with the highest influence for all cultivars except for Mencía, where it is Co. The predictor influence generally increases in relative, except when interactions with BW are present, but the order of predictor importance across all models does not change. In some models, where specific predictors are used exclusively in interaction with BW, they are no longer used once BW is excluded. Specifically, w_{sk} is no longer needed in the model for Brancellao as well as Co in the model for Merenzao, while the role of R in this last model stays important but its weight is reduced.

DISCUSSION

Texture analysis was already used to develop rapid methods for the evaluation of total phenolic content and phenol extractability in grape seeds with a good accuracy (Rolle et al., 2012), and of anthocyanin extractability in grape skins (Rolle et al., 2008; Ríó Segade et al., 2011b,c). The present work shows that the mechanical properties can also be successfully used as predictors in

models for the estimation of TF in skin grape berries.

Two kinds of models were developed; the first was single-cultivar based, while the second was a model built with all data available from all cultivars together. Single-cultivar models outperformed here the multi-cultivar model because, with some exceptions, physical-mechanical predictors are related to TF amount in a cultivar-specific way. Variables appearing well related to TF for a given cultivar were useless for models developed for other cultivars (Table 1). It is difficult to explain with certitude why this happens. It can be hypothesized a different genetic control of berry physical modifications with ripening and/or differences among cultivars in the non-linear accumulation of TF during ripening (Braidot et al., 2008). Predictors able to influence TF in the same way for all cultivars considered in the model are rare, as occurred in this study with BW and Ch. However, it is possible that other TA variables, well related to BW or Ch and therefore eliminated here from the predictor set because of collinearity problems, could be good candidate as multi-cultivar predictors (berry springiness is an example).

Single-cultivar models were very effective and achieved good results with other techniques used for the development of rapid methods to determine TF in grapes berries, skins, seeds and grape homogenates. Examples can be found in Near InfraRed (NIR) spectroscopy. Several studies (Cozzolino et al., 2008; Ferrer-Gallego et al., 2011; Torchio et al., 2013 among others) have reached very high performances with models for single cultivars, while models where different cultivars have been analyzed together are not very frequent in the literature.

In this work, we also highlight that BW is the most important physical-mechanical predictor of TF, at least for berries close to maturity as those considered here. In this case, TF is generally

determined to help assess the optimal harvest date because strongly related to wine phenolic composition (Cagnasso et al., 2008). . The relationship between BW and TF has already been studied in the literature and is well known in viticulture (Barbagallo et al., 2011 is an example). However, with the exception of Cabernet Franc, which has the lighter berries, BW alone is not able to well explain the amount of TF in grape berries. Other predictors are very useful to reduce the overall error in predictions and to increase model accuracy. On the contrary to BW, the role of other physical-mechanical predictors in determining the amount of TF in grape berries has not been yet extensively studied. E_{sk} has been found to be related to cellular maturity index (EA%) as predictor of anthocyanin extractability (Río Segade et al., 2011b), and also Sp_{sk} has been successfully used for the same purpose (Río Segade et al., 2011c). It is worthwhile to note that Sp_{sk} acts in a controversial way in the model for Mencía (eq. 6), where the relationship with TF has an approximate bell-shape. In this cultivar, an initial increase in Sp_{sk} corresponds to an increase in TF (ascendant), then at the knotpoint the relation becomes negative (descendant). Furthermore, other study showed that total extractable TF is significantly related to E_{sk} , although the correlation was weak (Rolle et al., 2011).

Ch is the most important predictor after BW, to which is also related in some of the cultivars studied (Brancellao and Merenzao). Berry cohesiveness is the product of berry hardness with berry cohesiveness and berry springiness, and is the measurement of the energy necessary to chew a solid food until it is ready for swallowing (Letaief et al., 2008a). The decrease in Ch has been observed with increased maturity of grapes (Río Segade et al., 2013; Zouid et al., 2010) and continues during post-harvest senescence due to the decrease in berry consistency (Deng et al., 2005).

CONCLUSIONS

This work shows that the influence of physical-mechanical predictors determining the amount of TF in grape berries varied across cultivars, even if some predictors such as Ch showed a stable behavior. This is probably due to a partial relation of this parameter with BW, which was the physical parameter better related to the TF amount. Therefore, a successful linkage between physical-mechanical predictors and TF amount in grape berries requires the development of a cultivar specific model as occurred in this work, where models for four different cultivars were developed. The error for the determination of TF by applying these models was low even on future data, as estimated by cross-validation, and allows use to monitor the TF through grape ripening in practice. These models are simple formulas that can be easily implemented and applied. Furthermore, the number of predictors needed was low, and their acquisition is fast, increasing the overall rapidity of this method. The developed models can be further improved by the use of a large number of data-points (more wine grape cultivars, growing zones and vintages), and also by extending the content range of TF used to build the models. This will increase the reliability and probably also the accuracy of the developed models.

REFERENCES

- Barbagallo, M.G., Guidoni, S., Hunter, J.J., 2011. Berry size and qualitative characteristics of *Vitis vinifera* L. cv. Syrah. South African J. Enol. Vitic. 32, 129–136.
- Battista, F., Tomasi, D., Porro, D., Caicci, F., Giacosa, S., Rolle, L., 2015. Winegrape berry skin thickness determination: comparison between histological observations and texture analysis determination. Ital. J. Food Sci. 27, 136-141.
- Braidot, E., Zancani, M., Petrusa, E., Peresson, C., Bertolini, A., Patui, S., Macrì, F., Vianello,

- A., 2008. Transport and accumulation of flavonoids in grapevine (*Vitis vinifera* L.). *Plant Signal. Behav.* 3, 626–632.
- Cagnasso E., Rolle L., Caudana A., Gerbi V. (2008) - Relationship between grape phenolic maturity and red wine phenolic composition. *Ital. J. Food Sci.* 20, 365-380.
- Cozzolino D., Cynkar, W., Damberg, R., Mercurio, M., Smith, P., 2008, Measurement of condensed tannins and dry matter in red grape homogenates using near infrared spectroscopy and partial least squares. *J. Agric. Food Chem.* 56, 7631-7636.
- Cozzolino, D., 2015. The role of visible and infrared spectroscopy combined with chemometrics to measure phenolic compounds in grape and wine samples. *Molecules* 20, 726–737.
- Deng, Y., Wu, Y., Li, Y., 2005. Effects of high O₂ levels on post-harvest quality and shelf life of table grapes during long-term storage. *Eur. Food Res. Technol.* 221, 392–397.
- Di Stefano, R., Cravero, M.C., 1991. Metodi per lo studio dei polifenoli dell' uva. *Riv. Vitic. Enol.* 44, 37–45.
- Ferrer-Gallego, R., Hernández-Hierro, Rivas-Gonzalo, J., Escribano-Bailón, M.T., 2011. Determination of phenolic compounds of grape skins during ripening by NIR spectroscopy. *LWT-Food Sci. Technol.* 44, 847–853.
- Friedman, J.H., 1991. Multivariate adaptive regression splines. *Ann. Stat.* 19 (1), 1–67.
- Gonzalez, J., Donoso, W., Sandoval, N., Reyes, M., Gonzalez, P., Gajardo, M., Morales, E., Neira, A., Razmilic, I., Yuri, J.A., Moore-carrasco, R., 2014. Apple peel supplemented diet reduces parameters of metabolic syndrome and atherogenic progression in ApoE *-/-* Mice. *Evidence-Based Complementary and Alternative Medicine*, vol. 2015, Article ID 918384, 10 pages, 2015. doi:10.1155/2015/918384

- González-Neves, G., Charamelo, D., Balado, J., Barreiro, L., Bochicchio, R., Gatto, G., Gil, G., Tessore, A., Carbonneau, A., Moutounet, M., 2004. Phenolic potential of Tannat, Cabernet-Sauvignon and Merlot grapes and their correspondence with wine composition. *Anal. Chim. Acta* 513, 191–196.
- Harnly, J.M., Doherty, R.F., Beecher, G.R., Holden, J.M., Haytowitz, D.B., Bhagwat, S., Gebhardt, S., 2006. Flavonoid content of U.S. fruits, vegetables, and nuts. *J. Agric. Food Chem.* 54, 9966–9977
- Hastie, T., Tibshiranini, R., Friedman, J.H., 2009. *The Element of Statistical Learning*, second ed. Springer, Netherlands, 745 pp.
- Ignat, I., Volf, I., Popa, V.I., 2011. A critical review of methods for characterisation of polyphenolic compounds in fruits and vegetables. *Food Chem.* 126, 1821–1835.
- Koundouras, S., Marinos, V., Gkoulioti, A., Kotseridis, Y., Van Leeuwen, C., 2006. Influence of vineyard location and vine water status on fruit maturation of non-irrigated cv. Agiorgitiko (*Vitis vinifera* L.). Effects on wine phenolic and aroma components. *J. Agric. Food Chem.* 54, 5077–5086.
- Kuhn, N., Guan, L., Dai, Z. W., Wu, B. H., Lauvergeat, V., Gomès, E., Li, S. H., Godoy, F., Arce-Johnson, P., Delrot, S., 2013. Berry ripening: Recently heard through the grapevine, *J. Exp. Bot.*, 65, 4543-4559.
- Le Moigne, M., Maury, C., Bertrand, D., Jourjon, F., 2008. Sensory and instrumental characterisation of Cabernet Franc grapes according to ripening stages and growing location. *Food Qual. Pref.* 19, 220–231.
- Letaief, H., Rolle, L., Gerbi, V., 2008a. Mechanical behavior of wine grapes under compression

- tests. *Am. J. Enol. Vitic.* 59, 323–329.
- Letaief, H., Rolle, L., Zeppa, G., Gerbi, V., 2008b. Assessment of grape skin hardness by a puncture test. *J. Sci. Food Agric.* 88, 1567–1575.
- Lorrain, B., Ky, I., Pechamat, L., Teissedre, P.L., 2013. Evolution of analysis of polyphenols from grapes, wines, and extracts. *Molecules* 18, 1076–1086.
- Lourenço, F., Gago, B., Barbosa, R.M., De Freitas, V., Laranjinha, J., 2008. LDL isolated from plasma-loaded red wine procyanidins resist lipid oxidation and tocopherol depletion. *J. Agric. Food Chem.* 56, 3798–3804.
- Milborrow, S., 2014. *Earth: Multivariate Adaptive Regression Spline Models*. R package version 3.2-7. <http://CRAN.R-project.org/package=earth>
- R Core Team, 2014. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
- Río Segade, S., Giacosa, S., Torchio, F., de Palma, L., Novello, V., Gerbi, V., Rolle, L., 2013. Impact of different advanced ripening stages on berry texture properties of ‘Red Globe’ and ‘Crimson Seedless’ table grape cultivars (*Vitis vinifera* L.). *Sci. Hortic.* 160, 313–319.
- Río Segade, S., Orriols, I., Giacosa, S., Rolle, L., 2011a. Instrumental texture analysis parameters as winegrapes varietal markers and ripeness predictors. *Int. J. Food Prop.* 14, 1318–1329.
- Río Segade, S., Soto Vázquez, E., Orriols, I., Giacosa, S., Rolle, L., 2011b. Possible use of texture characteristics of winegrapes as markers for zoning and their relationship with anthocyanin extractability index. *Int J Food Sci Tech*, 46, 386–394.
- Río Segade, S., Giacosa, S., Gerbi, V., Rolle, L., 2011c. Berry skin thickness as main texture parameter to predict anthocyanin extractability in winegrapes. *LWT- Food Sci. and*

- Technol., 44, 392–398.
- Ristic, R., Bindon, K., Francis, L.I., Herderich, M.J., Iland, P.G., 2010. Flavonoids and C13-norisoprenoids in *Vitis vinifera* L. cv. Shiraz: Relationships between grape and wine composition, wine colour and wine sensory properties. *Aust. J. Grape Wine Res.* 16, 369–388.
- Rolle, L., Río Segade, S., Torchio, F., Giacosa, S., Cagnasso, E., Marengo, F., Gerbi, V., 2011. Influence of grape density and harvest date on changes in phenolic composition, phenol extractability indices, and instrumental texture properties during ripening. *J. Agric. Food Chem.* 59, 8796–8805.
- Rolle, L., Torchio, F., Lorrain, B., Giacosa, S., Río Segade, S., Cagnasso, E., Gerbi, V., Teissedre, P.L., 2012. Rapid methods for the evaluation of total phenol content and extractability in intact grape seeds of Cabernet-Sauvignon: Instrumental mechanical properties and FT-NIR spectrum. *J. Int. Sci. Vigne Vin* 46, 29-40.
- Rolle L., Torchio F., Zeppa G., Gerbi V., 2008. Anthocyanin extractability assessment of grape skins by texture analysis. *J. Int. Sci. Vigne Vin* 42, 3, 157-162.
- Sun, A.Y., Simonyi, A., Sun, G.Y., 2002. The “french paradox” and beyond: neuroprotective effects of polyphenols. *Free Radical Bio Med*, 32 (4), 314-318.
- Torchio F., Giacosa S., Río Segade S., Mattivi F., Gerbi V., Rolle L., 2012. Optimization of a method based on the simultaneous measurement of acoustic and mechanical properties of winegrape seeds for the determination of the ripening stage. *J. Agric. Food. Chem.*, 60, 9006-9016.
- Torchio F., Rio Segade, S., Giacosa, S., Gerbi, V., Rolle, L., 2013. Effect of growing zone and vintage on the prediction of the extractable flavanols in winegrape seeds by a FT-NIR

method. J. Agric. Food Chem., 2013, 61 (38), 9076–9088.

Zouid, I., Siret, R., Mehinagic, E., Maury, C., Chevalier, M., Jourjon, F., 2010. Evolution of grape berries during ripening: investigations into the links between their mechanical properties and the extractability of their skin anthocyanins. J. Int. Sci.Vigne Vin. 44, 87–99.

Zsófi, Zs, Sz. Villangó, Z. Pálfi, E. Tóth, B. Bálo, 2014. Texture characteristics of the grape berry skin and seed (*Vitis vinifera* L. cv. Kékfrankos) under postveraison water deficit. Sci. Horticult. 172, 176–182.

Table 1. Variable importance for all predictors in the five MARS models, based on the reduction of the GCV statistics during the backward selection routine. Predictors that are not selected at all in the model get a value of 0, while predictors that most contribute to the error in prediction get a values of 100; intermediate predictors are scaled respect to the one with the highest influence.

	Brancellao	Cabernet Franc	Mencia	Merenzaio	Whole Dataset
BW	100	100	100	100	100
Co	0	0	59.12	44.31	22.26
Ch		0	52.95		
Esk	29.26	0	52.95	0	12.16
R	0	0	0	100	58.86
Spsk	0	0	64.18	0	58.86
Wsk	16.45	0	0	0	43.51

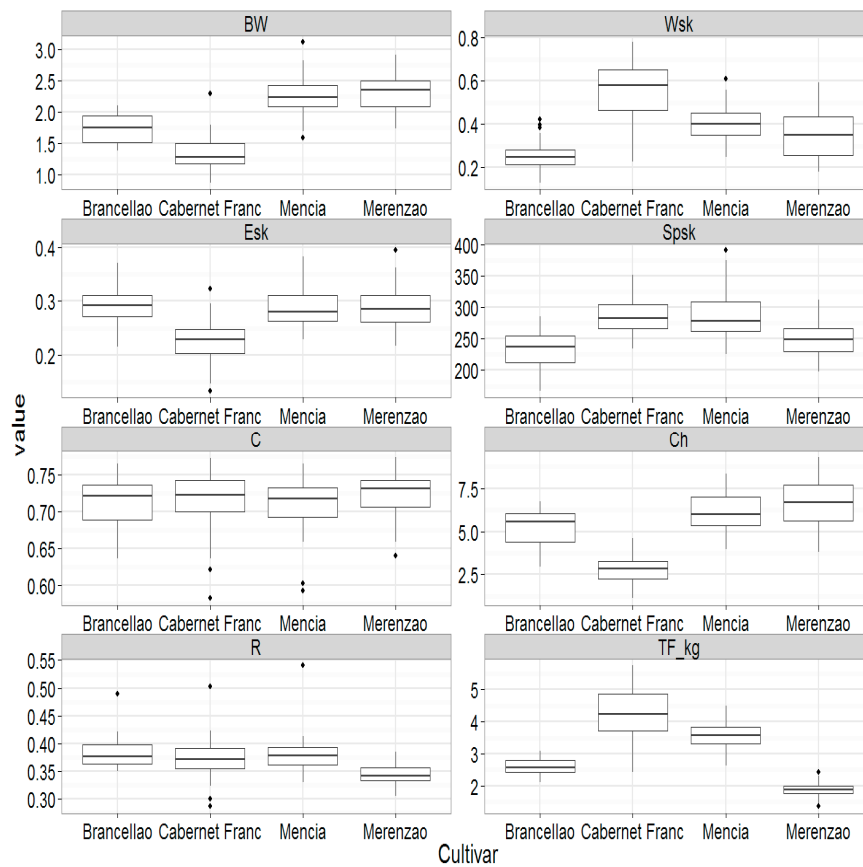


Figure 1. Boxplots showing descriptive statistics for the whole dataset. BW is berry weight (g), W_{sk} is skin break energy (mJ), E_{sk} is skin Young's modulus ($N\ mm^{-1}$), Sp_{sk} is skin thickness (μm), C is berry cohesiveness (adimensional), Ch is berry chewiness (mJ), R is berry resilience (adimensional), TF is the amount of total flavonoids ($mg\ g^{-1}$).

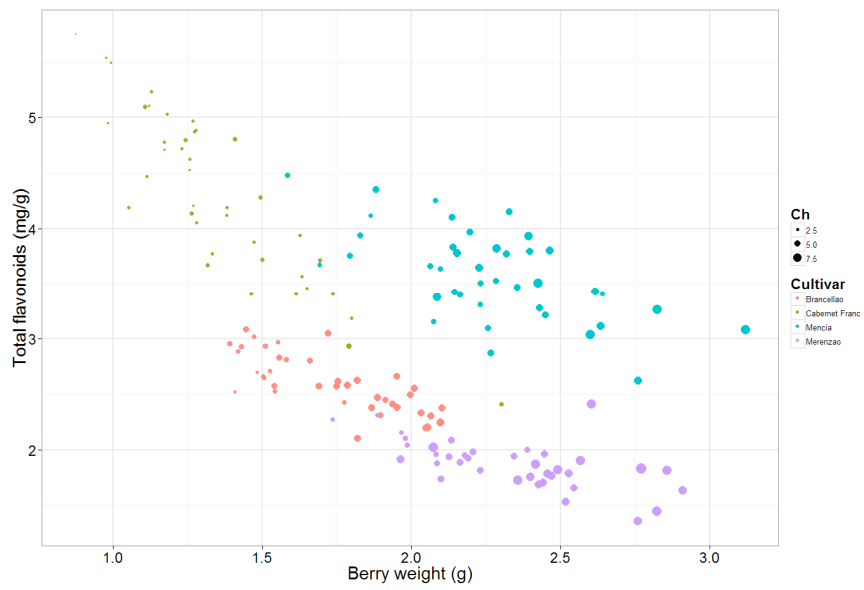


Figure 2. Raw relationships between berry weight (x axis) and total flavonoid amount (y axis) where color is mapped to the type of cultivar and size of point to berry chewiness. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

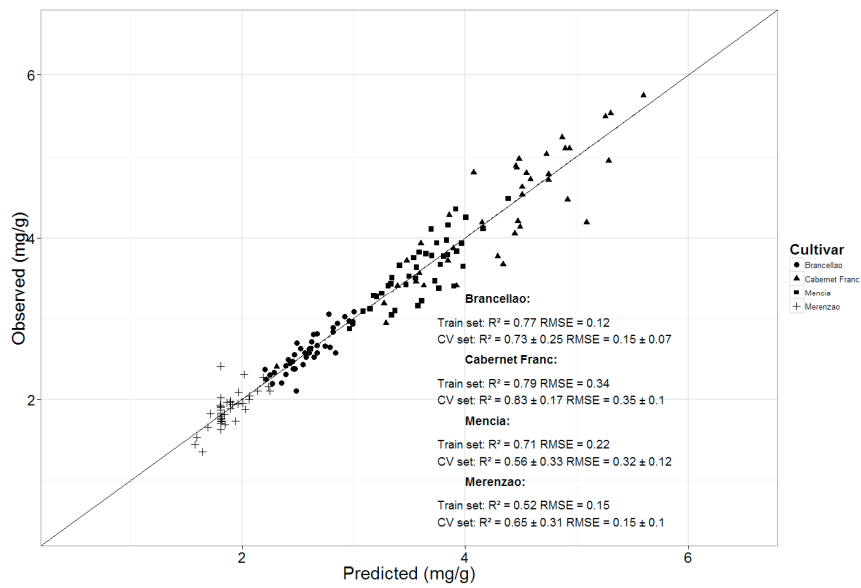


Figure 3. Scatterplot of the predicted versus observed total flavonoid amount for all cultivars in the dataset (train set). Data are on a straight line with intercept zero and slope one. Legend in figure reports model performances for all cultivars both on the train set and in the cross-validation assessment procedure (25 repetitions of 10-folds cross-validation).