

On prefermentative maceration techniques: statistical analysis of sensory descriptors in Sangiovese wine

Alessandro Magrini¹, Ottorino L. Pantani², Alessandra Biondi Bartolini³,
Federico M. Stefanini¹

¹Department of Statistics, Computer Science, Applications, University of Florence, Italy,
e-mail: stefanini@disia.unifi.it

²Department of Soil Vegetation and Forest Environment, University of Florence, Italy

³R&D wine and sensory consultant

SUMMARY

The analysis of wine sensory descriptors is a fundamental step in the improvement of wine making, because the procedures are judged just before bottled wine is ready for consumption. Despite several contributions in the literature, traditional analysis of variance methods are not adequate to analyse sensory descriptors, because they are defined on ordinal scales. In this paper, we exploit cumulative link mixed models in a three-way full factorial design to assess the effect of prefermentative maceration, temperature and saignée on wine sensory descriptors. Using cumulative link mixed models, the bias introduced by assessors' judgement and the ordinal scale of sensory descriptors are taken into account. The results were the following: the application of prefermentative maceration techniques did not lead to an improvement in the sensory profile of wines after a year from bottling; wines treated with saignée showed greater intensity in olfactive descriptors; and higher fermentation temperatures resulted in wines that were generally more concentrated.

Key words: cumulative link mixed models, proportional odds, assessor bias, cold soak prefermentative maceration, cryomaceration.

1. Introduction

The final effect of different vinification procedures may be investigated by measuring changes in the chemical compounds related to colour and flavours over time, but the relevance of the hedonistic aspect of wine motivates sensory evaluations to characterise vinification procedures in terms of perceived quality. Descriptive sensory analysis is particularly suited for this purpose (Murray,

2001; Lesschaeve, 2007): a panel of assessors provides an evaluation of several sensory descriptors; a new vinification procedure is declared effective if statistically significant changes in one or more descriptors are detected with respect to the traditional procedure.

This study considers Sangiovese wine making using grapes from Tuscany (2008 vintage). A three-way full factorial design with three replications is investigated to assess the effect of three factors on wine quality:

- prefermentative maceration techniques: standard fermentation (CTRL), cold soak prefermentation (CSPF), cryomaceration (CRYO);
- temperature of fermentation: 20°C (t20), 30°C (t30);
- saignée: absent (noS), present (yeS).

The design encompasses 12 different vinification procedures (treatments), for a total of 36 vinifications. Eight olfactive and six gustative and tactile descriptors were evaluated on a four-level hedonistic scale, one year after bottling. In order to reduce both the costs and time needed to train assessors, sensory evaluation was performed by an ad-interim panel formed by people involved in the experiment, as suggested by Lesschaeve (2007).

Several studies in the literature use traditional analysis of variance (ANOVA) methods to analyse sensory descriptors (Runnebaum et al., 2011; Carlucci and Monteleone, 2008; Etaio et al., 2008; Parenti et al., 2004; Marais, 2003; Lawless et al., 1997). However, ANOVA is not adequate because, even if one can safely assume that the intervals between categories of the descriptors are equal, residuals are not normally distributed and homoscedasticity may not hold. Here, we exploit cumulative link mixed models in order to take into account the ordinal scale of sensory descriptors and the bias introduced by assessors' judgement.

In section 2 the experimental setup is described. In section 3 we provide details on panel formation, training and evaluation sessions. The statistical analysis is detailed in section 4. In section 5 the results are presented, and the findings are discussed in section 6.

2. Experimental setup

Sangiovese grapes from the area of Maremma, South Tuscany, were manually harvested at maturity (2008 vintage) and collected in 20 kg buckets. A randomisation procedure was used to allocate the buckets within bins (500 l each), which were transported to the experimental winery and randomly assigned to vinification tanks (1000 l each, stainless steel). A wine cellar was prepared to host the experiment (<http://www.ricercatuscania.it>). A programmable control unit (Parsec s.r.l.©2007) regulated the temperatures and pumping-overs of each individual tank.

The destemmed and crushed berries destined for standard fermentation were sent directly to the tanks and brought to the planned fermentation temperature. Cold soak prefermentation was applied by lowering the temperature to 5°C immediately after filling the tanks; this temperature was maintained for 48 hours. Cryomaceration was performed on the destemmed and drained berries using an experimental apparatus (Parsec s.r.l.©2007) consisting of a freezing tunnel containing a stainless steel conveyor belt, on top of which liquid nitrogen sprinklers were installed. The contact time of liquid nitrogen with the berries was approximately 5–10 seconds, such that the final temperature of the mass was below 0°C, as regulated by the speed of the conveyor belt. The temperature of the flowing mass was measured and found to be between +7°C and –5°C. The cryomacerated berries were then crushed and added to the previously drained must.

Saignée was applied by bleeding off 200 kg must from the bottom valve of each tank, immediately after filling, so that the tank volume was the same as for the other ferments. Each tank was then supplemented with potassium metabisulphite (50 mg/l) and the temperature was set at the appropriate level (20°C or 30°C). Finally, the tanks were inoculated with 250 mg/l of EC1118 Lalvin® yeast strain. CRYO tanks were inoculated once the temperature reached 20°C (18–24 h). During the maceration, which lasted for 14 days,

the tanks underwent pumping-over cycles, whose number and duration were adjusted daily on the basis of the must density.

A comprehensive description of the wine-making process is provided in Pantani et al. (2014).

3. Panel formation, training and evaluation sessions

The effect of the tested treatments on sensory descriptors was evaluated following a descriptive approach (Marais, 2003). For this purpose, a panel of tasters was formed and trained. The panel consisted of people involved in the experiment, including wine experts, winemakers and a non-expert wine taster.

The training activity consisted of six sessions in which assessors were trained to identify a descriptor in a genuine standard and in wine; they were then trained to identify and sort different concentrations of gustatory standards in water and in wine. Finally, they were taught to describe red wines' gustatory and tactile sensations, as provided by Delteil (2000) and by Granes et al. (2009) for the ASDQ method developed by ICV (Institute Cooperative du Vin, Lattes, France).

A descriptive evaluation card was specifically developed for this study, where visual, olfactive, gustatory and tactile descriptors were considered. The visual properties were synthesised in one descriptor: the perceived colour intensity. Olfactory descriptors were chosen according to the method of the Free Choice Profile: five wines randomly chosen among those included in the experiment were submitted to the assessors, who provided olfactory sensations using their own words. Those olfactory descriptions were later discussed in a round table, to determine an unambiguous association between common semantic labels and personal perceived sensations.

Eight olfactive descriptors were finally agreed upon and included in the card: 'Fruity' (berries and tree fruit), 'Preserved Fruit', 'Spicy', 'Vegetal', 'Caramel' (candy), 'Chemical', 'Reductive Flavour' and 'Acetaldehyde' (bruised apple). Gustative and tactile profiles were assessed by means of the following six

descriptors: ‘Volume’, ‘Acidity’, ‘Tannic Intensity’, ‘Astringency’, ‘Dryness’ and ‘Bitterness’ (Delteil, 2000).

For each sensory descriptor, a four-level hedonistic scale was defined. The sensory evaluation was performed 12 months after wine bottling. All wines were tasted in three weekly sessions of panel testing. In each session, a sample from all treatments of one of the three replications (12 samples) was submitted to the panel of assessors in two series of six tastings, each separated from the other by a break of 20 minutes in order to reduce the potential bias due to fatigue (Bajec and Pickering, 2008). In the evaluation, descriptors were considered sequentially: the first three were evaluated when the wine was in the mouth, while the last three were evaluated after expelling the wine from the mouth. Bottles were taken from a stock kept at 18–20°C, a suitable temperature range for the serving of red wines. Samples were anonymised by assigning a three-digit numerical label and submitted for tasting in a random order.

4. Statistical analysis

A sensory descriptor is an ordinal variable that may take $J > 2$ different values, indicated by numerical values $j = 1, \dots, J$ called scores. Each treatment is coded by a configuration \mathbf{x} of 11 dummy variables:

$$\mathbf{x} = (\mathbf{x}_{\text{CSPF}}, \mathbf{x}_{\text{CRYO}}, \mathbf{x}_{\text{t20}}, \mathbf{x}_{\text{yeS}}, \mathbf{x}_{\text{CSPF:t30}}, \mathbf{x}_{\text{CSPF:yeS}}, \mathbf{x}_{\text{CRYO:t30}}, \mathbf{x}_{\text{CRYO:yeS}}, \mathbf{x}_{\text{t30:yeS}}, \mathbf{x}_{\text{CSPF:t30:yeS}}, \mathbf{x}_{\text{CRYO:t30:yeS}})$$

as shown in Table 1. Note that the treatment CTRL_t20_noS is taken as the reference due to its having the mildest vinification conditions; thus it is coded by a vector of zeros.

In this section, we provide technical details on cumulative link mixed models (subsection 4.1), discuss the interpretation of parameters representing the effect of the tested treatments (4.2), and explain the method used to quantify the statistical significance of each main effect and interaction term (4.3).

4.1. Cumulative link mixed models

Cumulative link mixed models were developed by Hedeker and Gibbons (1994) and by Tutz and Hennevogl (1996) by extending ordinal logit (proportional odds) models to clustered or longitudinal data. A cumulative link mixed model for a generic sensory descriptor Y has the following form:

$$G^{-1}(\Pr[Y \leq j | \mathbf{x}, \delta_k]) = \eta_j + \delta_k - \mathbf{x}'\boldsymbol{\beta} \quad (1)$$

$$\delta_k \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\delta^2), \quad j = 1, \dots, J, \quad k = 1, \dots, K$$

where $G^{-1}(\pi) = \log(\pi/(1 - \pi))$ is the quantile function of the standard logistic distribution, often called logit link; η_j is an intercept depending on the score under consideration (the j -th); δ_k is a random effect depending on the assessor under consideration (the k -th); and $\boldsymbol{\beta}$ is a vector of parameters, one for each of the 11 dummy variables encoding the tested treatments. This model is a linear regression where the response is the logit transformation of the cumulative probability for each score. The minus sign before $\mathbf{x}'\boldsymbol{\beta}$ is introduced so that a positive value of any parameter implies a greater probability for higher scores. Note that the parameters in $\boldsymbol{\beta}$ do not depend either on the score, or on the assessor. See Agresti (2002, Chapter 7) for further details.

Table 1. Codes of the tested treatments

Prefermentative maceration	Temperature	Saignée	Code of the treatment
CTRL			(0,0,0,0,0,0,0,0,0,0,0)
CSPF	t20		(1,0,0,0,0,0,0,0,0,0,0)
CRYO			(0,1,0,0,0,0,0,0,0,0,0)
CTRL		noS	(0,0,1,0,0,0,0,0,0,0,0)
CSPF	t30		(1,0,1,0,1,0,0,0,0,0,0)
CRYO			(0,1,1,0,0,1,0,0,0,0,0)
CTRL			(0,0,0,1,0,0,0,0,0,0,0)
CSPF	t20		(1,0,0,1,0,0,1,0,0,0,0)
CRYO			(0,1,0,1,0,0,0,1,0,0,0)
CTRL		yeS	(0,0,1,1,0,0,0,0,1,0,0)
CSPF	t30		(1,0,1,1,1,0,1,0,1,1,0)
CRYO			(0,1,1,1,0,1,0,1,1,0,1)

4.2. Interpretation of parameters

Consider the odds of the event $Y > j$ given treatment \mathbf{x} and assessor k defined as:

$$\text{odds}[Y > j | \mathbf{x}, \delta_k] = \frac{\Pr[Y > j | \mathbf{x}, \delta_k]}{\Pr[Y \leq j | \mathbf{x}, \delta_k]}$$

For example, if the event $Y > j$ has probability 1/6, then its odds are 1/5. An odds ratio is the ratio of the odds for the same score and assessor under two different treatments, say \mathbf{x}_1 and \mathbf{x}_2 :

$$\text{OR}[Y > j | \mathbf{x}_1, \mathbf{x}_2, \delta_k] = \frac{\Pr[Y > j | \mathbf{x}_2, \delta_k]/\Pr[Y \leq j | \mathbf{x}_2, \delta_k]}{\Pr[Y > j | \mathbf{x}_1, \delta_k]/\Pr[Y \leq j | \mathbf{x}_1, \delta_k]}$$

For example, an odds ratio equal to 2 implies that, by applying the treatment \mathbf{x}_2 , the odds of the event $Y > j$ are expected to be twice the odds obtained applying the treatment \mathbf{x}_1 .

The effect of the tested treatments is assessed by computing the odds ratio between each treatment and the reference. If the odds ratio is 1, there is no difference between the treatment considered and the reference treatment. If the odds ratio is greater (lower) than 1, the treatment considered is expected to provide higher (lower) response scores than the reference. In a cumulative link mixed model, the odds ratio between a generic treatment \mathbf{x} and the reference treatment \mathbf{x}_0 for score j and assessor k is:

$$\begin{aligned} \text{OR}[Y > j | \mathbf{x}, \mathbf{x}_0, \delta_k] &= \frac{\Pr[Y > j | \mathbf{x}, \delta_k]/\Pr[Y \leq j | \mathbf{x}, \delta_k]}{\Pr[Y > j | \mathbf{x}_0, \delta_k]/\Pr[Y \leq j | \mathbf{x}_0, \delta_k]} = \\ &= \frac{\exp(\eta_j + \delta_k - \mathbf{x}'\boldsymbol{\beta})}{\exp(\eta_j + \delta_k - \mathbf{x}'_0\boldsymbol{\beta})} = \exp(\mathbf{x}'\boldsymbol{\beta}) \end{aligned}$$

that is, it is the same for each score and for each assessor.

4.3. Analysis of deviance

The statistical significance of each main effect and interaction term is quantified by comparing the deviance of the reduced models obtained by constraining the

parameters in $\boldsymbol{\beta}$ of the model in Equation 1 to the value 0. Any reduced model obtained in this way is coded by a set of dummy variables similarly to the treatments: the value 0 indicates that a parameter is constrained to 0, and the value 1 indicates that a parameter is freely estimated. The deviance explained by a main effect or an interaction term is given by the difference in the deviance between the reduced model with the minimum number of parameters including that main effect or interaction, say \mathcal{M}_1 , and the reduced model obtained from \mathcal{M}_1 by constraining the regression coefficients associated with that main effect or interaction to the value 0, say \mathcal{M}_0 . Table 2 shows the codes of the reduced models that are to be compared in order to compute the deviance explained by each main effect and interaction term. The deviance between \mathcal{M}_1 and \mathcal{M}_0 is tested by considering that the log-likelihood ratio statistic:

$$\mathcal{D}_{\mathcal{M}_1|\mathcal{M}_0} = -2 (l_{\mathcal{M}_1} - l_{\mathcal{M}_0})$$

is asymptotically distributed as a χ^2 random variable with a number of degrees of freedom equal to the difference in the number of parameters between \mathcal{M}_1 and \mathcal{M}_0 , where $l_{\mathcal{M}_1}$ and $l_{\mathcal{M}_0}$ are, respectively, the log-likelihood of models \mathcal{M}_1 and \mathcal{M}_0 computed at the maximum likelihood estimates of parameters.

Further deviance terms may be considered: the deviance explained by the heterogeneity in the assessors' judgement, which is computed by comparing the reduced model where $\sigma_\delta^2 = 0$ with the model in Equation 1; and the residual deviance, which is obtained by comparing the model in Equation 1 with the model having one parameter per observation (saturated model).

An estimation of the odds ratio for the generic treatment \boldsymbol{x} with respect to the reference treatment is obtained as $\exp(\boldsymbol{x}'\hat{\boldsymbol{\beta}})$, where $\hat{\boldsymbol{\beta}}$ is the maximum likelihood estimate of $\boldsymbol{\beta}$. An asymptotic 95% confidence interval for the odds ratio of \boldsymbol{x} is:

$$\left(\exp(\boldsymbol{x}'\hat{\boldsymbol{\beta}}) - 1.96\sqrt{\boldsymbol{x}'\hat{\boldsymbol{S}}\boldsymbol{x}}, \exp(\boldsymbol{x}'\hat{\boldsymbol{\beta}}) + 1.96\sqrt{\boldsymbol{x}'\hat{\boldsymbol{S}}\boldsymbol{x}} \right)$$

where $\hat{\boldsymbol{S}}$ is the estimated variance and covariance matrix of $\hat{\boldsymbol{\beta}}$.

Table 2. Codes of reduced models to be compared in order to compute the deviance explained by each main effect and interaction term. ‘macer’: prefermentative maceration; ‘temp’: temperature of fermentation; ‘df’: degrees of freedom of the log-likelihood ratio statistic. Colons indicate interactions between experimental factors

	\mathcal{M}_1	\mathcal{M}_0	df
macer	(1,1,0,0,0,0,0,0,0,0)	(0,0,0,0,0,0,0,0,0,0)	2
temp	(0,0,1,0,0,0,0,0,0,0)	(0,0,0,0,0,0,0,0,0,0)	1
saignée	(0,0,0,1,0,0,0,0,0,0)	(0,0,0,0,0,0,0,0,0,0)	1
macer:temp	(1,1,1,0,1,1,0,0,0,0)	(1,1,1,0,0,0,0,0,0,0)	2
macer:saignée	(1,1,0,1,0,0,1,1,0,0)	(1,1,0,1,0,0,0,0,0,0)	2
temp:saignée	(0,0,1,1,0,0,0,1,0,0)	(0,0,1,1,0,0,0,0,0,0)	1
macer:temp:saignée	(1,1,1,1,1,1,1,1,1,1)	(1,1,1,1,1,1,1,1,0,0)	2

5. Results

The statistical analysis was performed using the function `clmm` contained in the R package **ordinal** (Christensen, 2015). Wine characteristics found to be significantly affected by at least one experimental factor include the visual descriptor ‘Colour’, the olfactive descriptors ‘Fruity’, ‘Spicy’, ‘Reductive Flavour’, and the gustative and tactile descriptors ‘Volume’, ‘Tannic Intensity’, ‘Astringency’, ‘Dryness’. The descriptors ‘Preserved fruit’, ‘Vegetable’, ‘Candy’, ‘Chemical’, ‘Bruised Apple’, ‘Acidity’ and ‘Bitterness’ were found not to be significantly affected by any experimental factors. The analysis of deviance for sensory descriptors which were significantly affected by at least one experimental factor is shown in Table 3. A summary of the parameter estimation for each sensory descriptor is given in the Appendix.

5.1. Evaluation of treatments

Table 4 shows estimated odds ratios for each treatment concerning sensory descriptors which were found to be significantly affected by at least one experimental factor (reference treatment: CTRL_t20_noS).

Table 3. Analysis of deviance tables for sensory descriptors found to be significantly affected by at least one experimental factor. ‘macer’: prefermentative maceration; ‘temp’: temperature of fermentation; ‘df’: degrees of freedom of the log-likelihood ratio statistic. Colons indicate interactions between experimental factors. ‘*’: statistically significant at the level 0.05; ‘**’: statistically significant at 0.01; ‘***’: statistically significant at 0.001

Source	df	Colour	Fruity	Spicy	Reductive Flavour
assessors	1	118.59 ***	92.37 ***	198.74 ***	47.81 ***
macer	2	4.30	2.52	0.42	3.22
Temp	1	51.62 ***	2.49	2.91 **	7.32 **
saignée	1	21.38 ***	5.76 *	1.18	5.89 *
macer:temp	2	4.20	2.84	1.67	2.23
macer:saignée	2	1.34	2.51	1.50	1.89
temp:saignée	1	1.94	0.03	5.17 ***	0.41
macer:temp:saignée	2	17.35 ***	3.42	0.78	8.32 *
residuals	24	96.91	26.81	21.36	25.96

Source	df	Volume	Tannic Intensity	Astringency	Dryness
assessors	1	149.90 ***	173.79 ***	79.69 ***	88.83 ***
macer	2	5.29	0.19	1.64	1.69
temp	1	10.49 **	11.95 ***	5.05 *	7.87 **
saignée	1	0.13	1.63	0.25	0.46
macer:temp	2	0.20	0.39	2.53	5.05
macer:saignée	2	0.02	0.63	2.82	1.40
temp:saignée	1	2.87	0.23	0.23	1.76
macer:temp:saignée	2	3.18	2.90	2.12	1.46
residuals	24	28.31	28.40	18.26	15.93

The application of saignée entails a significantly positive effect on the descriptor ‘Fruity’, as shown by the estimated odds ratio of 3.16. This means that the odds for ‘Fruity’ to be rated with higher scores in wines treated only with saignée (CTRL_t20_yeS) are more than one-and-a-half times the odds for wines not so treated (CTRL_t20_noS). Similar results were found for the descriptors ‘Colour’ (odds ratio 1.39) and ‘Reductive Flavour’ (odds ratio 7.03).

Conversely, wines fermented at 30°C were rated significantly higher in the perception of ‘Colour’ (odds ratio 15.46), ‘Reductive Flavour’ (odds ratio 1.61), ‘Volume’ (odds ratio 1.88), ‘Tannic Intensity’ (odds ratio 2.71), ‘Dryness’ (odds ratio 1.43) and ‘Astringency’ (odds ratio 1.9) than wines fermented at 20°C. On the other hand, the higher temperature of fermentation resulted in lower

Table 4. Estimated odds ratios for each treatment (reference: CTRL_t20_noS) for sensory descriptors found to be significantly affected by at least one experimental factor. 95% confidence intervals are shown within brackets

Treatment	Colour	Fruity	Soiciv	Reductive Flavour
CTRL_t20_yes	7.0377 (3.0819, 16.0253)	3.1641 (1.4363, 6.9700)	0.7195 (0.2974, 1.7407)	1.3948 (0.5483, 3.5480)
CTRL_t30_noS	15.4617 (6.8339, 34.9817)	0.8585 (0.3980, 1.8519)	0.5855 (0.2379, 1.4412)	1.6065 (0.6467, 3.9907)
CTRL_t30_yes	29.4592 (12.7824, 67.8938)	1.2019 (0.5571, 2.5931)	1.2445 (0.5133, 3.0173)	6.4005 (2.8179, 14.5376)
CSPF_t20_noS	9.2671 (4.0678, 21.1119)	1.0512 (0.4905, 2.2530)	0.8310 (0.3316, 2.0824)	1.6483 (0.6644, 4.0888)
CSPF_t20_yes	5.2804 (2.3180, 12.0289)	0.6997 (0.3210, 1.5250)	1.1801 (0.4876, 2.8561)	4.0560 (1.7363, 9.4749)
CSPF_t30_noS	8.2581 (3.6092, 18.8950)	0.7769 (0.3524, 1.7128)	0.3874 (0.1555, 0.9655)	5.6304 (2.3673, 13.3915)
CSPF_t30_yes	72.1453 (30.5969, 170.1135)	1.2712 (0.5943, 2.7189)	0.8462 (0.3439, 2.0823)	3.3906 (1.4542, 7.9053)
CRYO_t20_noS	4.0316 (1.7820, 9.1212)	1.1745 (0.5367, 2.5701)	1.6449 (0.6925, 3.9073)	2.3387 (0.9886, 5.5326)
CRYO_t20_yes	7.7616 (3.3698, 17.8772)	2.0020 (0.9088, 4.4100)	0.8237 (0.3339, 2.0323)	1.9881 (0.8124, 4.8655)
CRYO_t30_noS	11.4578 (5.1097, 25.6925)	0.6737 (0.3148, 1.4416)	0.4954 (0.1947, 1.2605)	1.4880 (0.5850, 3.7851)
CRYO_t30_yes	28.4649 (12.4065, 65.3084)	1.2855 (0.6066, 2.7244)	0.9053 (0.3669, 2.2339)	5.0031 (2.1481, 11.6528)

Treatment	Volume	Tannic Intensity	Astringency	Drvness
CTRL_t20_yes	0.8589 (0.3640, 2.0270)	1.7077 (0.7563, 3.8562)	0.7875 (0.3607, 1.7190)	0.8133 (0.3670, 1.8021)
CTRL_t30_noS	1.8844 (0.8002, 4.4379)	2.7151 (1.2161, 6.0617)	1.9009 (0.8908, 4.0560)	1.4296 (0.6600, 3.0966)
CTRL_t30_yes	2.0495 (0.8697, 4.8298)	3.0000 (1.3229, 6.8031)	1.3369 (0.6233, 2.8676)	1.3607 (0.6207, 2.9833)
CSPF_t20_noS	1.9642 (0.8276, 4.6619)	1.5903 (0.7071, 3.5765)	0.5485 (0.2513, 1.1972)	0.6113 (0.2804, 1.3329)
CSPF_t20_yes	0.7485 (0.3171, 1.7667)	1.7242 (0.7821, 3.8011)	1.0299 (0.4841, 2.1910)	0.6109 (0.2783, 1.3410)
CSPF_t30_noS	1.4228 (0.6010, 3.3685)	2.7951 (1.2021, 6.4993)	1.2670 (0.5926, 2.7089)	1.6679 (0.7637, 3.6428)
CSPF_t30_yes	2.9753 (1.2504, 7.0793)	2.6782 (1.1606, 6.1803)	1.7517 (0.8034, 3.8196)	2.0012 (0.9255, 4.3272)
CRYO_t20_noS	0.6819 (0.2823, 1.6472)	1.5094 (0.6789, 3.3557)	1.6931 (0.7963, 3.6000)	1.4488 (0.6717, 3.1246)
CRYO_t20_yes	0.5970 (0.2530, 1.4083)	1.3257 (0.6085, 2.8882)	1.1851 (0.5555, 2.5284)	1.2509 (0.5862, 2.6694)
CRYO_t30_noS	1.2787 (0.5508, 2.9685)	1.7678 (0.7883, 3.9644)	1.0761 (0.5039, 2.2981)	0.8851 (0.4055, 1.9319)
CRYO_t30_yes	1.2587 (0.5251, 3.0170)	4.4669 (1.9919, 10.0173)	1.9004 (0.9023, 4.0023)	2.3546 (1.0903, 5.0852)

Table 5. Estimated random effects for sensory descriptors found to be significantly affected by at least one experimental factor. 95% confidence intervals are shown within brackets

Assessor ID	Colour	Fruity	Spicy	Reductive Flavour
1	-0.0380 (-0.7717, 0.6957)	-1.2918 (-2.0063, -0.5773)	-0.0605 (-0.8155, 0.6945)	0.1498 (-0.6020, 0.9016)
2	0.0256 (-0.7340, 0.7852)	-0.0192 (-0.7358, 0.6973)	-0.3455 (-1.0901, 0.3990)	-0.0461 (-0.8858, 0.7936)
3	-0.6109 (-1.3423, 0.1205)	-0.8671 (-1.6667, -0.0676)	0.1077 (-0.6222, 0.8376)	-0.4999 (-1.4016, 0.4018)
4	-0.7956 (-1.4041, -0.1870)	0.6565 (0.0958, 1.2173)	1.0451 (0.4233, 1.6668)	0.7206 (0.1334, 1.3078)
5	2.4485 (1.7739, 3.1232)	1.8381 (1.2167, 2.4594)	-0.6913 (-1.3432, -0.0393)	-2.3691 (-3.6150, -1.1233)
6	-2.7616 (-3.4035, -2.1198)	0.3975 (-0.1651, 0.9600)	-0.3320 (-0.9680, 0.3041)	0.0267 (-0.6484, 0.7018)
7	1.7153 (0.6476, 2.7831)	0.6990 (-0.2851, 1.6831)	1.0467 (-0.9334, 1.2268)	-0.4121 (-1.5231, 0.6989)
8	-1.7029 (-2.7128, -0.6930)	1.7484 (0.7000, 2.7967)	3.2010 (2.1443, 4.2577)	1.5951 (0.5575, 2.6328)
9	-0.1234 (-0.7497, 0.5029)	-0.9231 (-1.5338, -0.3124)	0.4823 (-0.1618, 1.1263)	-1.7197 (-2.7376, -0.7017)
10	-1.1222 (-1.9144, -0.3300)	-0.9498 (-1.6969, -0.2028)	-2.9586 (-4.3840, -1.5333)	1.1482 (0.4570, 1.8393)
11	0.6625 (-0.0022, 1.3272)	1.6193 (1.0209, 2.2177)	4.0046 (3.3788, 4.6304)	0.7024 (0.1090, 1.2958)
12	0.1944 (-0.4844, 0.8731)	-0.6191 (-1.2291, -0.0091)	-2.9200 (-4.0930, -1.7471)	0.5539 (-0.0984, 1.2063)
13	-0.3188 (-0.9135, 0.2758)	-1.5501 (-2.1996, -0.9005)	1.1245 (0.4941, 1.7548)	-0.0979 (-0.7879, 0.5921)
14	0.8823 (0.1125, 1.6521)	-0.0033 (-0.7022, 0.6957)	-0.1533 (-0.8963, 0.5897)	1.1406 (0.4213, 1.8599)
15	1.4967 (0.7439, 2.2496)	-0.7053 (-1.4328, 0.0222)	-2.1775 (-3.2628, -1.0923)	-0.4080 (-1.2615, 0.4456)
Assessor ID	Volume	Tannic Intensity	Astringency	Dryness
1	0.8038 (0.0492, 1.5584)	0.0221 (-0.6718, 0.7161)	0.2596 (-0.4119, 0.9310)	-0.2015 (-0.8978, 0.4949)
2	0.7203 (-0.0450, 1.4856)	0.5699 (-0.1524, 1.2921)	0.3910 (-0.2746, 1.0566)	-0.2642 (-0.9522, 0.4238)
3	1.5355 (0.7438, 2.3271)	0.5382 (-0.2073, 1.2838)	0.7474 (0.0212, 1.4737)	1.4932 (0.7325, 2.2538)
4	-1.1896 (-1.8404, -0.5389)	-0.6115 (-1.2364, 0.0134)	-0.3515 (-0.9221, 0.2192)	0.2337 (-0.3230, 0.7945)
5	1.3575 (0.6979, 2.0170)	3.3393 (2.6674, 4.0113)	0.8082 (0.1426, 1.4738)	1.3363 (0.6901, 1.9825)
6	-3.6032 (-4.3157, -2.8907)	-2.6307 (-3.2462, -2.0153)	-2.5507 (-3.1499, -1.9514)	-2.1186 (-2.7455, -1.4916)
7	0.8443 (-0.2305, 1.9192)	0.9340 (-0.0196, 1.8877)	-0.8180 (-1.7434, 0.1073)	-1.5803 (-2.6131, -0.5475)
8	2.1529 (0.9618, 3.3440)	0.2319 (-0.7763, 1.2400)	2.1931 (1.1507, 3.2354)	1.1104 (0.1436, 2.0773)
9	0.6849 (0.0536, 1.3141)	0.5863 (-0.0258, 1.1983)	-0.5484 (-1.1427, 0.0458)	0.2612 (-0.3086, 0.8309)
10	-0.1501 (-0.9650, 0.6648)	0.2135 (-0.4970, 0.9240)	-0.6625 (-1.3880, 0.0630)	-0.4158 (-1.1786, 0.3470)
11	0.4656 (-0.1928, 1.1239)	0.5174 (-0.1176, 1.1525)	0.3162 (-0.2678, 0.9001)	0.1198 (-0.4792, 0.7188)
12	-2.4478 (-3.1324, -1.7632)	-0.0564 (-0.7325, 0.6196)	1.3122 (0.6530, 1.9715)	1.0338 (0.3954, 1.6722)
13	-0.4256 (-1.0590, 0.2078)	-2.4038 (-3.0182, -1.7893)	-0.7891 (-1.3741, -0.2042)	-0.0018 (-0.5796, 0.5759)
14	0.5232 (-0.2287, 1.2750)	-1.8128 (-2.5447, -1.0810)	-0.5504 (-1.2417, 0.1408)	-2.3080 (-3.0895, -1.5265)
15	-1.2964 (-2.1114, -0.4814)	0.5223 (-0.2103, 1.2549)	0.2363 (-0.4998, 0.9724)	1.3068 (0.5887, 2.0250)

perceptions of the descriptor ‘Spicy’ (odds ratio 0.58). In particular, the (negative) effect was larger if saignée was also applied (odds ratio 0.72).

The full-order interaction proved to be significant for the descriptors ‘Colour’ and ‘Reductive Flavour’. The effect of joint application of 30°C fermentation and saignée on the descriptor ‘Colour’ was greater if cold soak prefermentation was applied (odds ratio 72.14 against 29.46). On the other hand, the effect of joint application of 30°C fermentation and saignée on the descriptor ‘Reductive Flavour’ was lower if cold soak prefermentation or cryomaceration was applied (odds ratios 3.39 and 5.00 against 6.40).

5.2. Evaluation of heterogeneity in assessors’ judgement

For each sensory descriptor, random effects quantify the deviation of the assessors’ judgement from the ‘average’ opinion ($\delta_k = 0$). Since random effects do not depend either on the response score or on the treatment, the cumulative link mixed model makes it possible to separate the heterogeneity in the assessors’ judgement from the effect of the tested treatments.

Table 5 shows estimated random effects for sensory descriptors which were found to be significantly affected by at least one experimental factor. The heterogeneity in the assessors’ judgements was confirmed by the significance of the deviance explained by random effects (see Table 3).

The bias introduced by each assessor may be quantified in terms of the deviation of the respective random effect from value 0, which corresponds to the random effect of the ‘average’ assessor. Although random effects were ranked differently across sensory descriptors, it is apparent that some assessors tended to award higher scores than the average (numbers 6, 10 and 13), while others had the opposite tendency (numbers 5, 8 and 11). Assessors 1, 2 and 3 were close to the ‘average’ value for each wine descriptor.

Two natural interpretations of the heterogeneity in the assessors’ judgement are either that each score means different things to different assessors, or that assessors have a different perception of the same sensory descriptor.

6. Discussion

A descriptive sensory approach was adopted to evaluate the effect of different fermentation techniques on Sangiovese quality after one year from bottling, including cold soak prefermentation, cryomaceration, 30°C temperature and saignée. For this purpose, an ad-interim panel was formed by people involved in the experiment (15 assessors). The analysis was performed using cumulative link mixed models, in order to take into account the ordinal scale of sensory descriptors and bias in assessors' judgement: on one hand, a linear regression model is applied to the logit of the cumulative probability of each score, instead of to the score itself; on the other hand, a random effect is estimated for each assessor.

Wines treated with saignée resulted in a greater intensity in the descriptors 'Colour' and 'Fruity'. However, the descriptor 'Reductive Flavour' was also found to be significantly affected by saignée, suggesting that the treatment should be applied carefully. It is noteworthy that the increase in the score for 'Reductive Flavour' can be prevented by applying cold soak prefermentation or cryomaceration.

Higher fermentation temperatures resulted in greater expression of all of the gustatory descriptors except for 'Fruity', that is in wines with a highly concentrated profile. However, we must consider that some assessors might perceive high scores for the descriptor 'Dryness' as a depreciation of wine quality.

The application of prefermentative maceration techniques did not lead to an improvement in the sensory profile of wines after a year from bottling, but it amplified the effect of the joint application of 30°C fermentation and saignée on the visual descriptor 'Colour', while also preventing the increase in the score for 'Reductive Flavour'. The lack of a main effect due to prefermentative maceration techniques is in contrast with the results of related studies performed on wines drawn earlier. For instance, Marais (2003) reported a marked effect of cold soak prefermentation on the overall quality of Pinotage after six months

from bottling (on 'Fruity', in particular), while Parenti et al. (2004) found an increase in 'Colour', 'Tannic Intensity' and 'Astringency' for cryomacerated Sangiovese after two months from bottling. In our study, the wines were tasted far later (one year from bottling); we therefore believe that the differences in the results are due to the rapid loss of these quality characteristics in the first months after bottling.

Random effects made it possible not only to characterise each assessor with respect to the bias introduced in his or her judgements, but also to individuate assessors with a definite tendency in awarding scores lower than, higher than or close to the average. Furthermore, estimated random effects could be useful in developing ad-hoc training for each assessor before future sessions.

Acknowledgements

This study was part of a research project funded by the Italian Ministry of Economic Development and Consorzio Toscana, Florence, Italy (<http://www.ricercatuscانيا.it>).

REFERENCES

- Agresti A. (2002): *Categorical data analysis*. Second edition. Wiley.
- Bajec M.R., Pickering G.J. (2008): Astringency: mechanisms and perception critical reviews. *Food Science and Nutrition* 48(9): 858-875.
- Carlucci A., Monteleone E. (2008): A procedure of sensory evaluation for describing the aroma profile of single grape variety vines. *Journal of Sensory Studies* 23(6): 817-834.
- Christensen R.H.B. (2015): *ordinal: regression models for ordinal data*. R package version 2015-06-28. <http://www.cran.r-project.org/package=ordinal>
- Delteil D. (2000): Evaluation sensorielle du profil gustatif des vins. *Revue des Oenologues* 94: 21-23.
- Etaio I., Pérez Elortondo F.J., Albisu M., Gaston E., Ojeda M., Schlich P. (2008): Development of a quantitative sensory method for the description of young red wines from Rioja Avalesa. *Journal of Sensory Studies* 23(5): 631-655.
- Hedeker D., Gibbons R.D. (1994): A random-effects ordinal regression model for multilevel analysis. *Biometrics* 50(4): 933-944.
- Lawless H., Liu Y., Goldwyn C. (1997): Evaluation of wine quality using a small-panel hedonic scaling method. *Journal of Sensory Studies* 12(4): 317-322.

- Granes D., Pic-Blateyron L., Negrel J., Bonnefond C. (2009): L'analyse sensorielle descriptive quantifiée (ASDQ): une méthode pour un langage commun. *Revue Française d'Oenologie* 238: 16-21.
- Lesschaeve I. (2007): Sensory evaluation of wine and commercial realities: review of current practices and perspectives. *American Journal of Enology and Viticulture* 58: 252-258.
- Marais J. (2003): Effect of different wine-making techniques on the composition and quality of Pinotage wine. I. Low temperature skin contact prior to fermentation quadrature. *South African Journal for Enology and Viticulture* 24: 70-75.
- Murray J.M., Delahunty C.M., Baxter I.A. (2001): Descriptive sensory analysis: past, present and future. *Food Research International* 34: 461-471.
- Pantani O.L., Stefanini F.M., Lozzi I., Calamai L., Biondi Bartolini A., Di Blasi S. (2014): Pre-maceration, saignée and temperature affect daily evolution of pigment extraction during vinification. COBRA Preprint Series, Working Paper 107. <http://biostats.bepress.com/cobra/art107>
- Parenti A., Spugnoli P., Calamai L., Ferrari S., Gori C. (2004): Effects of cold maceration on red wine quality from Tuscan Sangiovese grape. *European Food Research and Technology* 218: 360-366.
- Runnebaum R.C., Boulton R.B., Powell R.L., Heymann H. (2011): Key constituents affecting wine body – an exploratory study. *Journal of Sensory Studies* 26: 62-70.
- Tutz G., Hennevogl W. (1996): Random effects in ordinal regression models. *Computational Statistics and Data Analysis* 2: 537-557.

Appendix

<i>Response: Colour</i>	<i>Estimate</i>	<i>Std. Error</i>	<i>z value</i>	<i>p value</i>
(Intercept 1 2)	-1.5487	0.5232	-2.9599	0.0031
(Intercept 2 3)	1.2612	0.5181	2.4342	0.0149
(Intercept 3 4)	4.3539	0.5570	7.8167	0.0000
macer_mpf	2.2265	0.5069	4.3923	0.0000
macer_crio	1.3942	0.4967	2.8067	0.0050
temp_t30	2.7384	0.5099	5.3708	0.0000
saignée_yeS	1.9499	0.5051	3.8607	0.0001
macer_mpf:temp_t30	-2.8536	0.6995	-4.0796	0.0000
macer_crio:temp_t30	-1.6939	0.6815	-2.4854	0.0129
macer_mpf:saignée_yeS	-2.5123	0.6979	-3.6001	0.0003
macer_crio:saignée_yeS	-1.2948	0.6912	-1.8732	0.0610
temp_t30:saignée_yeS	-1.3052	0.6928	-1.8840	0.0596
macer_mpf:temp_t30:saignée_yeS	4.0352	0.9894	4.0784	0.0000
macer_crio:temp_t30:saignée_yeS	1.5602	0.9639	1.6186	0.1055

Variance of random effects: 1.845667

Log-likelihood: -406.2734 AIC: 842.5467

<i>Response: Fruity</i>	<i>Estimate</i>	<i>Std. Error</i>	<i>z value</i>	<i>p value</i>
(Intercept 1 2)	-1.1386	0.4471	-2.5468	0.0109
(Intercept 2 3)	1.2060	0.4479	2.6926	0.0071
(Intercept 3 4)	4.3321	0.5363	8.0782	0.0000
macer_mpf	0.0499	0.4557	0.1096	0.9127
macer_crio	0.1608	0.4645	0.3462	0.7292
temp_t30	-0.1525	0.4585	-0.3326	0.7394
saignée_yeS	1.1519	0.4704	2.4487	0.0143
macer_mpf:temp_t30	-0.1498	0.6539	-0.2292	0.8187
macer_crio:temp_t30	-0.4033	0.6513	-0.6192	0.5358
macer_mpf:saignée_yeS	-1.5589	0.6599	-2.3623	0.0182
macer_crio:saignée_yeS	-0.6185	0.6662	-0.9285	0.3532
temp_t30:saignée_yeS	-0.8154	0.6564	-1.2424	0.2141
macer_mpf:temp_t30:saignée_yeS	1.7148	0.9296	1.8446	0.0651
macer_crio:temp_t30:saignée_yeS	0.9283	0.9257	1.0028	0.3160
Variance of random effects: 1.318172				
Log-likelihood: -440.3359 AIC: 910.6717				

<i>Response: Spicy</i>	<i>Estimate</i>	<i>Std. Error</i>	<i>z value</i>	<i>p value</i>
(Intercept 1 2)	-0.5564	0.6204	-0.8969	0.3698
(Intercept 2 3)	2.0614	0.6295	3.2745	0.0011
(Intercept 3 4)	4.9503	0.7176	6.8980	0.0000
macer_mpf	-0.1852	0.5078	-0.3646	0.7154
macer_crio	0.4977	0.4833	1.0297	0.3031
temp_t30	-0.5353	0.4998	-1.0710	0.2842
saignée_yeS	-0.3292	0.4915	-0.6698	0.5030
macer_mpf:temp_t30	-0.2278	0.7219	-0.3155	0.7523
macer_crio:temp_t30	-0.6648	0.7122	-0.9335	0.3506
macer_mpf:saignée_yeS	0.6800	0.7080	0.9603	0.3369
macer_crio:saignée_yeS	-0.3624	0.6954	-0.5211	0.6023
temp_t30:saignée_yeS	1.0832	0.7037	1.5394	0.1237
macer_mpf:temp_t30:saignée_yeS	-0.6528	1.0082	-0.6475	0.5173
macer_crio:temp_t30:saignée_yeS	0.2113	1.0046	0.2103	0.8334
Variance of random effects: 3.820375				
Log-likelihood: -370.3745 AIC: 770.749				

<i>Response: Reductive Flavour</i>	<i>Estimate</i>	<i>Std. Error</i>	<i>z value</i>	<i>p value</i>
(Intercept 1 2)	1.6393	0.5385	3.0441	0.0023
(Intercept 2 3)	3.5753	0.5663	6.3135	0.0000

(Intercept 3 4)	5.6891	0.6864	8.2885	0.0000
macer_mpf	0.4997	0.5912	0.8452	0.3980
macer_crio	0.8496	0.5730	1.4826	0.1382
temp_t30	0.4741	0.5916	0.8013	0.4230
saignée_yeS	0.3327	0.6012	0.5535	0.5799
macer_mpf:temp_t30	0.7544	0.7995	0.9435	0.3454
macer_crio:temp_t30	-0.9262	0.8043	-1.1516	0.2495
macer_mpf:saignée_yeS	0.5677	0.8007	0.7090	0.4783
macer_crio:saignée_yeS	-0.4952	0.7995	-0.6194	0.5357
temp_t30:saignée_yeS	1.0496	0.7944	1.3212	0.1864
macer_mpf:temp_t30:saignée_yeS	-2.4572	1.0868	-2.2610	0.0238
macer_crio:temp_t30:saignée_yeS	0.3255	1.0936	0.2976	0.7660

Variance of random effects: 1.297362

Log-likelihood: -341.1959 AIC: 712.3919

<i>Response: Volume</i>	<i>Estimate</i>	<i>Std. Error</i>	<i>z value</i>	<i>p value</i>
(Intercept 1 2)	-2.4312	0.5641	-4.3100	0.0000
(Intercept 2 3)	1.0280	0.5440	1.8898	0.0588
(Intercept 3 4)	5.3420	0.7100	7.5243	0.0000
macer_mpf	0.6751	0.4994	1.3518	0.1764
macer_crio	-0.3829	0.5063	-0.7561	0.4496
temp_t30	0.6336	0.4956	1.2785	0.2011
saignée_yeS	-0.1520	0.4958	-0.3067	0.7591
macer_mpf:temp_t30	-0.9561	0.7015	-1.3628	0.1729
macer_crio:temp_t30	-0.0049	0.6996	-0.0070	0.9944
macer_mpf:saignée_yeS	-0.8128	0.7008	-1.1599	0.2461
macer_crio:saignée_yeS	0.0190	0.7051	0.0269	0.9785
temp_t30:saignée_yeS	0.2360	0.6968	0.3387	0.7348
macer_mpf:temp_t30:saignée_yeS	1.4665	0.9897	1.4818	0.1384
macer_crio:temp_t30:saignée_yeS	-0.1187	0.9893	-0.1200	0.9045

Variance of random effects: 2.452081

Log-likelihood: -363.6736 AIC: 757.3472

<i>Response: Tannic Intensity</i>	<i>Estimate</i>	<i>Std. Error</i>	<i>z value</i>	<i>p value</i>
(Intercept 1 2)	-1.8768	0.5193	-3.6138	0.0003
(Intercept 2 3)	0.5701	0.5076	1.1232	0.2613
(Intercept 3 4)	3.3761	0.5381	6.2747	0.0000
macer_mpf	0.4639	0.4666	0.9943	0.3201
macer_crio	0.4117	0.4610	0.8932	0.3718
temp_t30	0.9988	0.4646	2.1498	0.0316

saignée_yeS	0.5352	0.4684	1.1426	0.2532
macer_mpf:temp_t30	-0.4349	0.6690	-0.6501	0.5157
macer_crio:temp_t30	-0.8408	0.6538	-1.2861	0.1984
macer_mpf:saignée_yeS	-0.4543	0.6556	-0.6930	0.4883
macer_crio:saignée_yeS	-0.6650	0.6481	-1.0261	0.3048
temp_t30:saignée_yeS	-0.4354	0.6616	-0.6581	0.5104
macer_mpf:temp_t30:saignée_yeS	0.3118	0.9442	0.3303	0.7412
macer_crio:temp_t30:saignée_yeS	1.4921	0.9263	1.6109	0.1072
Variance of random effects: 2.151023				
Log-likelihood: -444.2782 AIC: 918.5565				

Response: Astringency	Estimate	Std. Error	z value	p value
(Intercept 1 2)	-3.2159	0.4859	-6.6183	0.0000
(Intercept 2 3)	-0.2615	0.4380	-0.5970	0.5505
(Intercept 3 4)	2.0096	0.4489	4.4768	0.0000
macer_mpf	-0.6006	0.4635	-1.2960	0.1950
macer_crio	0.5266	0.4516	1.1660	0.2436
temp_t30	0.6423	0.4533	1.4170	0.1565
saignée_yeS	-0.2390	0.4625	-0.5166	0.6054
macer_mpf:temp_t30	0.1950	0.6481	0.3009	0.7635
macer_crio:temp_t30	-1.0955	0.6415	-1.7078	0.0877
macer_mpf:saignée_yeS	0.8691	0.6547	1.3275	0.1844
macer_crio:saignée_yeS	-0.1178	0.6462	-0.1823	0.8554
temp_t30:saignée_yeS	-0.1130	0.6487	-0.1742	0.8617
macer_mpf:temp_t30:saignée_yeS	-0.1932	0.9217	-0.2096	0.8340
macer_crio:temp_t30:saignée_yeS	1.0384	0.9094	1.1419	0.2535
Variance of random effects: 1.278738				
Log-likelihood: -462.8125 AIC: 955.6249				

Response: Dryness	Estimate	Std. Error	z value	p value
(Intercept 1 2)	0.4996	-5.9272	-0.0843	0.9328
(Intercept 2 3)	0.4603	0.1848	2.4908	0.0127
(Intercept 3 4)	0.4748	4.9201	0.0965	0.9231
macer_mpf	-0.4921	0.4608	-1.0681	0.2855
macer_crio	0.3707	0.4558	0.8133	0.4161
temp_t30	0.3574	0.4575	0.7812	0.4347
saignée_yeS	-0.2067	0.4675	-0.4422	0.6584
macer_mpf:temp_t30	0.6463	0.6500	0.9943	0.3201
macer_crio:temp_t30	-0.8502	0.6475	-1.3131	0.1891
macer_mpf:saignée_yeS	0.2061	0.6582	0.3130	0.7542

macer_crio:saignée_yeS	0.0598	0.6464	0.0926	0.9262
temp_t30:saignée_yeS	0.1573	0.6561	0.2398	0.8105
macer_mpf:temp_t30:saignée_yeS	0.0255	0.9244	0.0276	0.9780
macer_crio:temp_t30:saignée_yeS	0.9680	0.9169	1.0557	0.2911

Variance of random effects: 1.518573

Log-likelihood: -459.4391 AIC: 948.8782