Manuscript Number:

Title: MODELLING PHENOLIC AND TECHNOLOGICAL MATURITIES OF GRAPES BY MEANS OF THE MULTIVARIATE RELATION BETWEEN ORGANOLEPTIC AND PHYSICOCHEMICAL PROPERTIES

Article Type: Special Issue : CAC 2012

Section/Category: CHEMOMETRICS

Keywords: grapes; phenolic maturity; technological maturity; sensory analysis; varimax rotation; PLS; D.O.C. Rioja

Corresponding Author: Dr. M.C. Ortiz,

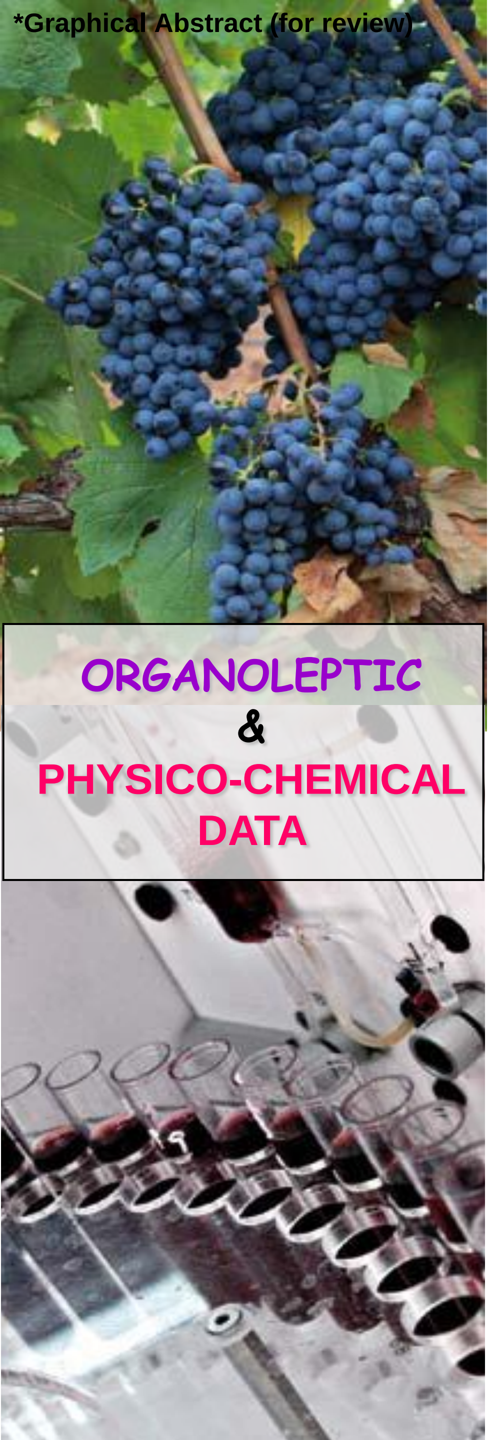
Corresponding Author's Institution: University of Burgos, Faculty of Sciences

First Author: E. Meléndez

Order of Authors: E. Meléndez ; M.C. Ortiz; L. A. Sarabia; M. Íñiguez ; P. Puras

Manuscript Region of Origin: SPAIN

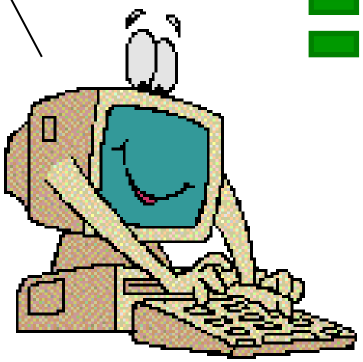
*Graphical Abstract (for review)



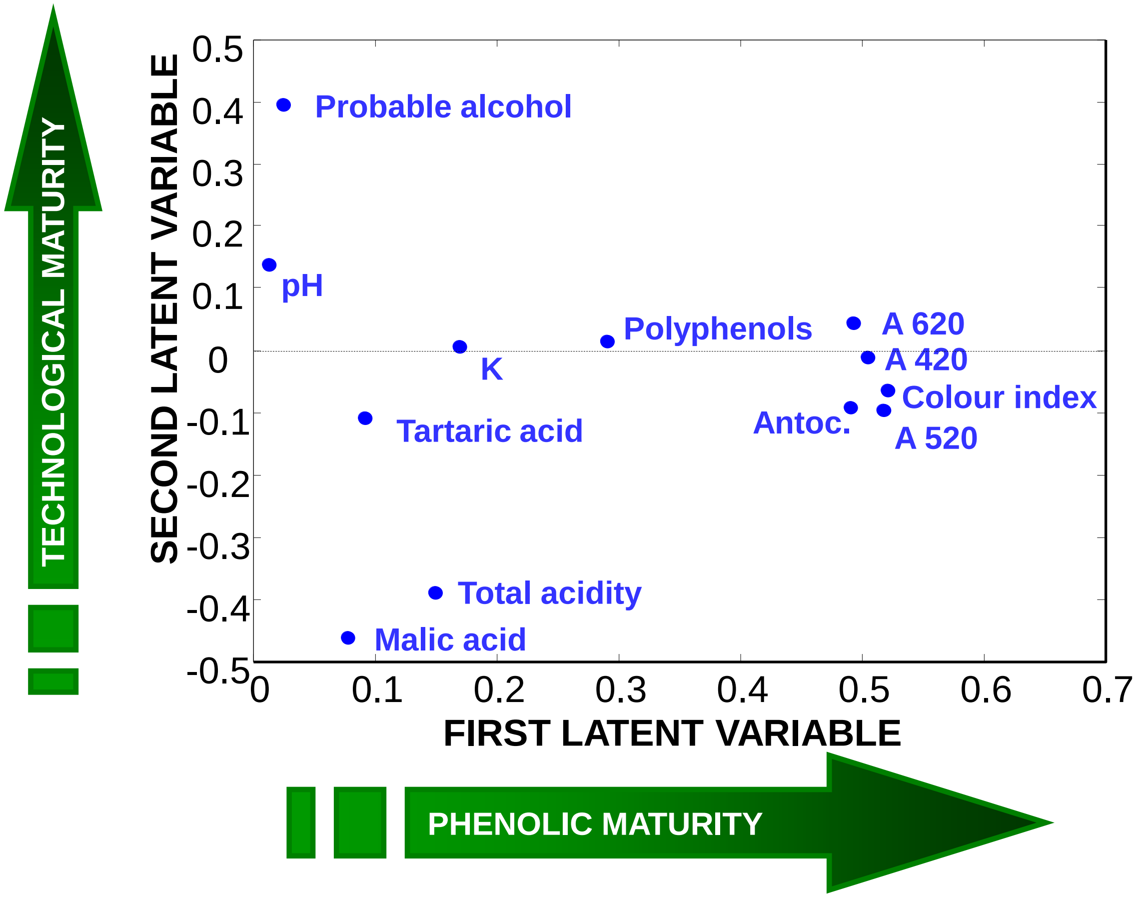
ORGANOLEPTIC

&

PHYSICO-CHEMICAL DATA



VARIMAX ROTATED LOADINGS OF PLS



***Highlights (for review)**

The ripeness of grapes at the harvest is interesting for obtaining high quality red wines

PLS models and a varimax rotation allows identifying technological and phenolic maturities in grapes

Organoleptic and physicochemical variables to study the degree of grapes maturity in D.O.C Rioja

Technological and phenolic maturities in grapes are not simultaneously reached

MODELLING PHENOLIC AND TECHNOLOGICAL MATURITIES OF GRAPES BY MEANS OF THE MULTIVARIATE RELATION BETWEEN ORGANOLEPTIC AND PHYSICOCHEMICAL PROPERTIES

E. Meléndez [a], M.C. Ortiz [b1], L.A. Sarabia [c], M. Íñiguez [a], P.Puras [a]
[a] Estación Enológica de Haro, Bretón de los Herreros,4 26200 Haro La Rioja (Spain)
[b] Departamento de Química, [c] Departamento de Matemáticas y Computación, Facultad de Ciencias, Universidad de Burgos, Pza. Misael Bañuelos s/n 09001 Burgos (Spain)

ABSTRACT

The ripeness of grapes at the harvest time is one of the most important parameters for obtaining high quality red wines. Traditionally the decision of harvesting is to be taken only after analyzing sugar concentration, titratable acidity and pH of the grape juice (technological maturity). However, these parameters only provide information about the pulp ripeness and overlook the real degree of skins and seeds maturities (phenolic maturity). Both maturities, technological and phenolic, are not simultaneously reached, on the contrary they tend to separate depending on several factors: grape variety, cultivar, adverse weather conditions, soil, water availability and cultural practices. Besides this divergence is increasing as effect of the climate change (larger quantities of CO_2, less rain, and higher temperatures).

247 samples collected in vineyards representative of the qualified designation of origin Rioja from 2007 to 2011 have been analyzed. Samples contain the four grape varieties usual in the elaboration of Rioja wines (‘tempranillo’, ’garnacha’, ‘mazuelo’ and ‘graciano’).

---

[1] Corresponding author. E-mail: mcortiz@ubu.es. Telephone number: +34 947 259571

The present study is the first systematic investigation of the maturity of grapes that includes the organoleptic evaluation of the degree of grapes maturity (sugars/acidity maturity, aromatic maturity of the pulp, aromatic maturity of the skins and tannins maturity) together with the values of the physicochemical parameters (probable alcohol degree, total acidity, pH, malic acid, K, total index polyphenolics, anthocians, absorbances at 420, 520 and 620 nm, colour index and tartaric acid) determined over the same samples. A varimax rotation of the latent variables of a PLS model between the physicochemical variables and the mean of four sensory variables allows identifying both maturities. Besides, the position of the samples in the first plane defines the effect that the different factors exert on both phenolic and technological maturities.

Keywords: Grapes; phenolic maturity; technological maturity; sensory analysis; varimax rotation; PLS; D.O.C. Rioja

INTRODUCTION

Relation between grape maturity and wine quality is evident, and thus knowing the optimum moment for harvesting has a great interest. There are two aspects on grape maturity that should be taken into account: (i) the technological maturity which is linked with the amount of sugar in the grape (alcohol degree), and (ii) the phenolic maturity related to grape colour.

Nowadays, the wine sector is very interested in defining the concept of phenolic maturity and its repercussions on the sensory parameters of the obtained wine (colour, astringency, bitterness, etc.) [1,2] as well as its interactions with the protein fraction of saliva [3]. Many works have been devoted to quantify the effect of maturity on several physicochemical parameters, for example, to distinguish between grape varieties [4], or the importance of the extraction methodology when monitoring seed maturity for prediction of the seed tannins in wine [5]. The increasing interest about the phenolic maturity has given rise to new techniques for its detection, such as the use of an electronic nose [6] or by computer vision [7].

However, the optimum of both maturities is not reached at the same time, and so if we wait to harvest until the phenolic maturity is reached, a wine with excessive alcohol degree will be obtained. Further, at present, an increasing difference between both processes of maturity is being confirmed, in part due to the climate change.

Usually the decision to harvest is based on the physicochemical analyses, but there is not doubt that the information supplied for the organoleptic analyses is useful for the decision. For this reason the Estación Enológica de Haro decided to incorporate the taste of the grape.

The present study is the first systematic investigation of the maturity of grapes that includes the organoleptic evaluation of the degree of grapes maturity and the values of the physicochemical parameters. A search in SCOPUS database with the keywords “grape” and “maturity” provides 290 documents published over the last 12 years, but none of them treats jointly the phenolic and technological maturities during the ripening.

For this study 12 physicochemical parameters were determined: probable alcohol degree, total acidity, pH, malic acid, K, total index polyphenolics, anthocians, absorbances at three wavelengths (420, 520 and 620 nm), colour index and tartaric acid. All these analyses were carried out at the laboratory of the Estación Enológica (Haro). The samples were collected during, at least, the last four consecutive weeks in the harvests corresponding to years 2007, 2008, 2009, 2010 and 2011. In the same samples, four levels of maturity were evaluated, namely sugars/acidity maturity, aromatic maturity of the pulp, aromatic maturity of the skins, and tannins maturity. They were scored by a panel of experts (tasters) in a scale from 1 to 4 in increasing order of maturity.

The need of separately evaluating the phenolic maturity of grape seeds and skins has been recognized also in a recent study about the effect of climatic conditions on the phenolic composition of grape [8], where clear differences between phenolic maturity pattern of skins and seeds were observed.

Grapes were collected in plots representative of the qualified designation of origin Rioja (D.O.C. Rioja), as well as from different varieties of grape: ‘tempranillo’,’garnacha’, ‘mazuelo’ and 'graciano'.

A description of the two types of maturity is obtained by means of a partial least squares regression model between physicochemical variables and the mean of the four values of the taste, followed by a varimax rotation of the first two latent variables.

2 THEORY

2.1 Partial least squares regression

Partial least squares (PLS) regression [9] is a biased multilinear regression based on latent variables that aims to obtain a linear model between a set of predictor variables, $\mathbf{X}_{\mathrm{NP}}$, and a set of response variables, $\mathbf{Y}_{\mathrm{NR}}$. The P values of the predictor variables for each of the N objects are the rows of matrix $\mathbf{X}$ and R is the number of responses for the same object (R can be equal to one). The PLS model with K factors or latent variables can be presented as:

$$\mathbf{X}_{\mathrm{NP}} = \mathbf{T}_{\mathrm{NK}}\mathbf{P}_{\mathrm{PK}}{}^{\mathrm{T}} + \mathbf{E}_{\mathrm{NP}} \tag{1}$$

$$\mathbf{Y}_{\mathrm{NR}} = \mathbf{U}_{\mathrm{NK}}\mathbf{C}_{\mathrm{RK}}{}^{\mathrm{T}} + \mathbf{F}_{\mathrm{NR}} \tag{2}$$

where, matrix **T** contains the *K* X-factor scores, **U** contains the *K* corresponding Y-factor scores, **P** stores the *K* X-factor loadings and **C** the *K* Y-factor loadings. **E** and **F** are the matrices of residuals not explained by the model in X and Y blocks, respectively.

PLS was devised to find a few linear combinations (*K* latent variables or factors) of the predictor variables in order to explain the values of the response variables. In the case of only one response, $R = 1$, the m-th latent variable (m-th column of **P**), $\mathbf{p}_{\mathrm{m}}$, is the result of maximizing the product of $\mathrm{corr}^2(\mathbf{y},\mathbf{X}\boldsymbol{\alpha})$ by $\mathrm{Var}(\mathbf{X}\boldsymbol{\alpha})$, contrained to $\|\boldsymbol{\alpha}\| = 1$ and $\mathbf{p}_{\mathrm{i}}^{\mathrm{T}}\ \mathbf{S}\boldsymbol{\alpha} = 0\ (\mathrm{i} = 1,...,\mathrm{m}\text{-}1)$, to ensure that $\mathbf{X}\boldsymbol{\alpha}$ is uncorrelated with all the previous linear combination $\mathbf{X}\,\mathbf{p}_{\mathrm{i}}$. Therefore, PLS searches the directions in the predictor space with the maximum variance but avoiding these that are not correlated with the responses to achieve the highest prediction capacity. PLS has gained importance for the prediction of wine characteristics. According to the SCOPUS database, since 1988 [10] 139 papers have been published (110 in the last seven years) which contain 'wine' and 'PLS' among the keywords and that cover such diverse aspects as the determination of copper content [11], the modeling of their colour [12], ageing [13], the kind of outliers and how to detect them [14] or the characterization of 'compliant' wines according to quality characteristics [15].

2.2 Varimax rotation.

The task of interpreting latent variables is not always straightforward. The technique of Varimax rotation [16] has been developed to make the loadings in factorial (or principal components) analysis more interpretable. But it is also possible to use it for the X-block loadings of a PLS model. The Varimax rotation perturbs the X-block loadings so as to maximize the variance within each latent variable. As a result, in each latent variable the number of variables with intermediate loadings is decreased, and the number with either very large (absolute magnitude) or very small loadings is increased.

The varimax criterion produces an orthogonal rotation in the space of the *K* X-block latent variables. The matrix $\mathbf{P}_{\mathrm{PK}}{}^{\mathrm{T}}$ is rotated by means of an orthogonal matrix $\mathbf{R}_{\mathrm{KK}}$ to obtain a matrix

$$\mathbf{A}_{\mathrm{PK}}{}^{\mathrm{T}} = \mathbf{R}_{\mathrm{KK}}\mathbf{P}_{\mathrm{PK}}{}^{\mathrm{T}} \qquad (3)$$

with the maximum row simplicity. The simplicity of the *k*-th latent variable (*k*-th column of $\mathbf{A}_{\mathrm{PK}}$) is defined as the variance of the squares of the elements a_{pk}, that is:

$$\mathrm{sim}_{\mathrm{k}} = \frac{1}{\mathrm{P}}\left[\sum_{\mathrm{p}=1}^{\mathrm{P}} \mathrm{a}_{\mathrm{pk}}^{4} - \frac{\sum_{\mathrm{p}=1}^{\mathrm{P}} \mathrm{a}_{\mathrm{pk}}^{2}}{\mathrm{P}}\right] \qquad (4)$$

Therefore the varimax rotation is the one that maximizes $\sum_{\mathrm{k}=1}^{\mathrm{K}} \mathrm{sim}_{\mathrm{k}}$

The new *K* X-block loadings $\mathbf{A}_{\mathrm{PK}}{}^{\mathrm{T}}$ are multiplied by the new *N* scores $\mathbf{V}_{\mathrm{NK}}$ so that:

$$\mathbf{V}_{\mathrm{NK}} = \mathbf{X}_{\mathrm{NK}} \mathbf{A}_{\mathrm{PK}}{}^{\mathrm{T}} \tag{5}$$

and finally equation (1) is transformed in

$$\mathbf{X}_{\mathrm{NP}} = \mathbf{V}_{\mathrm{NK}} \mathbf{A}_{\mathrm{PK}}{}^{\mathrm{T}} + \mathbf{E}_{\mathrm{NP}} \tag{6}$$

with the same variance explained in X-block.

3 EXPERIMENTAL

3.1 Samples

After selecting the representative plots, the samples used for the present study are picked up from among those that have a minimum of maturity (at least with 10% in probable alcohol degree), and then the physicochemical and organoleptic analyses were carried out. Table 1 shows the temporary sampling distribution of the selected samples.

Table 2 shows the geographical distribution of the 247 samples, as well as their (grape) variety, and the basic characteristics of the plot (year of plantation and altitude) from which the samples come. Figure 1 is a map of La Rioja map where the situation of every zone is shown. It is interesting to indicate that the grape maturity in La Rioja is progressive from East to West and, within each zone, the grape variety matters, being the tempranillo the first, and then garnacha, mazuelo and graciano, in this order.

3.2 Variables

In every plot, 100 berries were collected. Subsequently, 100 grams of grapes were chopped during 10 seconds, and then centrifuged for 5 minutes at 10,000 rpm. The clean grape juice was collected to perform the analyses.

In every sample, 12 variables were determined: 1) probable alcohol degree, 2) total acidity, 3) pH, 4) malic acid, 5) K, 6) total index polyphenolics, 7) anthocians, 8) absorbances at 420 nm, 9) absorbance at 420 nm, 10) absorbance at 620 nm), 11) colour index and 12) tartaric acid. The analyses were carried out at the laboratory of the Estación Enológica of Haro following the methods in [17].

The Institut Coopératif du Vin, ICV, has developed a method that describes and quantifies the sensory analysis of grapes according to sensory metrology and ISO 11035. Each part of the grape (pulp, skin and seeds) is characterized by 20 descriptors quantified on a graduated scale from 1 to 4 [18,19]. In the scale from 1 to 4, the following maturities have been scored in the samples: (i) Sugars/acidity maturity, (ii) Aromatic maturity of the pulp, (iii) Aromatic maturity of the skins, and (iv) Tannins maturity. The tasting of grapes was carried out by a panel of experts in the 'Estación Enológica de Haro'.

3.3 Software

PLS models and correlation analysis have been done with the PLS Toolbox [20] and the varimax rotation with the Statistics Toolbox, version 7.1 (2009) for Matlab 7.0 [21].

4 RESULTS AND DISCUSION

4.1 Models for every year.

A global PLS model has been built, using the 12 physicochemical variables as predictor variables, and the average value of the four values of tasting as response, separately for each year. Several additional models, which are not showed here, have been built for each value of taste individually. These models have provided structures similar to those observed with the average of tasting.

In all cases the 12 physicochemical variables were autoscaled, and the responses were centred. The purpose of the analysis is to study the structure of the physicochemical variables, when they are related to the sensory analysis, it is to say the loadings structure corresponding to the X block. In all the models for the different years, it was observed that the first two latent variables were related to the two processes of maturity (technological and phenolic). Table 3 shows the percentage of variance explained by these two latent variables, observing high stability among years, varying from 44.4% to 61.4% in the X block and from 50% to 73.5% in the Y-block.

Figure 2 depicts the new loadings obtained after the varimax rotation of the loadings of the PLS models. The sub-graphs in columns depict the first and two rotated latent variable, RLV, while in rows, the models for the different years are shown. The numbers identify the variables, as numbered in section 3.2.

The loadings on the first RLV have a similar structure in the years 2007, 2009 and 2010, fig.2a), e) and g) respectively, showing that the physicochemical variables with larger weight in the three cases are: total index polyphenolics, anthocians, the three absorbances and colour index. That is to say, this first RLV is linked to the phenolic maturity. As their loadings are positive, these raw variables increase or decrease simultaneously. The second RLV for the same years, fig. 2b), f) and h), has the same similar structure. The most relevant aspect about theses loadings is the opposition between the probable alcohol degree (1) and the total acidity (2) and the amount of malic acid (4). The presence of malic acid indicates lack of technological maturity (inversely correlated to the amount of probable alcohol degree). This second RLV clearly represents the technological maturity in the grape.

The structure corresponding, for these two latent variables, to years 2008 and 2011 is clearly different from the other three. In 2008 the differences can be attributable to the atypical weather conditions in the months previous to harvesting. It was a cool summer which provokes a lack of maturity of grapes in some zones, being the zones to the east where the maturity was better. In the East zones the vegetative cycle lasts 10 or 15 days more compared to the high Rioja zone (in the west). Besides, in the west zone, harvesting was made in adverse conditions due to the rain threat at the end of the maturity, and the grape had to be quickly picked up. As a consequence, a problem due to herbaceous flavours in many wines elaborated in year 2008 was found. The reason was a big amount of chlorophyll in the grape, as well as the malic acid (lack of maturity). The climate anomaly in 2008 has also been described in ref. [8] by using a principal component analysis of (monthly) temperature and rain as well as the humidity and sun radiation during 2008 and 2009.

Finally in the latent rotated structure in the data of 2011, it was observed that the probable alcohol degree and malic acid variables had large weight in the first RLV but not in the second. The possible interpretation is that the sampling was not adequate to represent the whole maturity cycle. In fact the small sample size is due to the lack of grapes with a low level of maturity, as can be seen in table 4 that shows that the frequencies of high scores (3 or 4) is much greater than for low scores. Also, taking into account that phenolic immaturity is more linked to the third and fourth tasting variables, in almost all the samples the phenolic maturity has not been reached (note that 10 samples achieved a score of 4 in the skin maturity and only 4 in the tannins maturity). On the contrary, when looking at the technological maturity, at least 16 samples have a score of 4 in the first two sensory attributes.

4.2 Correlation analysis

The origin of the anomaly in data of year 2008 can be corroborated by analyzing the Pearson's correlation between the physicochemical variables because the climate effect in the maturity will be directly noticeable on them. It is important to take into consideration that the correlations between the variables describe the relative evolution when the process of grape maturity goes through. In this way they describe the maturity process, so that the meaning had been different if the sampling of the same physicochemical parameters had been done in a short period of time (instead of the four or five weeks considered).

Figure 3 shows the correlations in a colour scale, fig. 3a) corresponds to the samples of 2007, 2009, 2010 and 2011 and fig. 3b) to the samples of 2008. It is evident the different correlation between the variables directly linked to the phenolic maturity: total index polyphenolics, anthocians, absorbances at three wavenlegths (420, 520 and 620 nm) and colour index which is positive and high for years 2007, 2009, 2010 and 2011 (figure 3a) and very different for year 2008 (figure 3b). Note also the huge differences in the correlation between the malic and tartaric acids with colour variables.

The Box's M test [22,23], to check the equality of the two covariance matrices, has a p-level less than 10^{-10} (much smaller than 0.05), therefore both covariance matrices are significantly different which confirms the visual impression of figure 3.

4.3 Structure of grapes maturity

Excluding the data of 2008, so with only 190 samples, a PLS regression model with the physicochemical variables and the average of the four tasting values was fitted. Like in the previous cases the predictor variables have been autoescaled and the response was centred. For the crossvalidation 10 random sets were used.

To detect outlier data the following iterative process was used: To eliminate all those objects that have values of Hotelling's T^2 and Q statistics greater than the threshold values at 99% confidence level. Afterwards, objects with standardized residuals greater than 2.5 in absolute value are eliminated.

Twelve samples were removed with this procedure. The model fitted with the remaining 172 has a minimum of the Root Mean Square Error in Cross Validation (RMSECV) of 0.4019 which is reached with 2 latent variables. This is the model chosen with has a RMSEC (Root Mean Square Error in Calibration) equal to 0.3855, explains 55.45 % of the variance in the X-block (33.13% with the first latent variables and 22.32 % with the second), and 41.87% of the variance in the Y-block (30.80% with the first and 11.07 % with the second latent variables). Compared to the annual models (table 3) the percentage of variance explained in the predictor variables block is similar, but with a little less percentage in the response. This can be due to the tasting variability among years.

The varimax rotation of these two factors keeps the percentage of explained variance corresponding to X block, but the variance is redistributed in such a way that the first RLV explains the 48.31% and the second one only the 7.41%. The rotated plane is shown in figure 4, where it is seen that all variables have positive loadings in the first RLV. The variables linked to phenolic evolution (total index polyphenolics, anthocians, absorbances and colour index) have large loadings in the first RLV and very small loadings in the second RLV. Looking at the second RLV, it is observed the opposition between the probable alcohol degree and pH (both with positive loadings) versus the malic acid and the total acidity (with negative loadings). Besides these four variables have small loadings in the first RLV, so that this second RLV defines the technological maturity. Consequently, the interpretation of this plane allows describing both maturities in an orthogonal way.

The position of the samples in this rotated plane is shown in figure 5. Each sample has been labelled with the four tasting values. The first and the second are related to the technological maturity while the third and the fourth are more linked to the phenolic maturity. The samples that have reached the two maturities (labelled as ‘4444’ and marked with red and bigger characters) have a rotated score high in the first RLV, except for three of them that have a score very high in the second RLV. These samples are mostly placed in the 'external' (top and right) zone of the cloud of points. The samples with very small score in the 3th and

4th tasting variables are placed in the opposite zone in the graph, with negative scores in both RLV, so these are grapes that have not reached the phenolic maturity.

The majority of samples, marked in magenta in figure 5, with tasting scores equal to 4 and 4 in the first two tasting variables, i.e., samples with adequate maturity in relation to sugar/acidity in the pulp (technological maturity), have large positive scores in the second RLV, and very few of them have also a large score in the first RLV. In any case, all of them are samples that have reached enough technological maturity but not enough phenolic one.

CONCLUSIONS

For the first time an orthogonal varimax rotation has been used for the descriptive analysis of the latent variables of a PLS regression model. This PLS regression model is computed between 12 physicochemical variables and the average of the tasting of grapes of the D.O. C. Rioja.

The deformation of the latent structure allows seeing the impact due to the adverse weather conditions in the year 2008. This fact considerably changes some correlations between the physicochemical variables.

The structure obtained through the rotated PLS latent variables, can not be noticed when the physicochemical variables are analyzed by using principal components (this analysis is not shown in this paper). The key to define the two aspects (phenolic and technological) of the grape maturity is the correlation that the PLS regression imposes between the physicochemical and organoleptic variables.

ACKNOWLEDGMENTS

The authors thank the financial support provided by Ministerio de Economía y Competitividad (CTQ2011-26022) and Junta de Castilla y León (BU108A11-2). Also, acknowledge the work of the experts (tasters) of the Oenological Station of Haro: A. Bastida, D. Hernando, B. Pamplona and of L.Dulau of Laffort España.

REFERENCES

[1] N. Kontoudakis, M. Esteruelas, F. Fort, J. M. Canals, F. Zamora, Anal. Chim. Acta, 660 (2010) 127.

[2] Y. Cadot, S. Caillé, A. Samson, G. Barbeau, V. Cheynier, Analytica Chimica Acta, 732, (2012) 91.

[3] E. Obreque-Slier, R. López-Solís, A. Peña-Neira International Journal of Food Science and Technology, (2012) Article in Press DOI: 10.1111/j.1365-2621.2012.03051.x

[4] E. Obreque-Slier, R.López-Solís, L. Castro-Ulloa, C. Romero-Díaz, A. Peña-Neira, Food Science and Technology, 48 (2012) 134

[5] A.B. Bautista-Ortín, P. Rodríguez-Rodríguez, R. Gil-Muñoz, E. Jiménez-Pascual, N. Busse-Valverde, A. Martínez-Cutillas, J.M. López-Roca, E. Gómez-Plaza, Australian Journal of Grape and Wine Research, 18 (2012) 123.

[6] S. Vallone, N.W. Lloyd, S.E. Ebeler, F. Zakharov, Journal of visualized experiments, 61 (2012), DOI: 10.3791/3821.

[7] F.J. Rodríguez-Pulido, R. Ferrer-Gallego, M. L. González-Miret, J.C. Rivas-Gonzalo, M.T. Escribano-Bailón, F.J. Heredia, Analytica Chimica Acta, 732 (2012) 78.

[8] R. Ferrer-Gallego, J. M. Hernández-Hierro, J. C. Rivas-Gonzalo, M.T. Escribano-Bailón, Anal. Chim. Acta, 732 (2012) 73.

[9] H. Martens, T. Naes, Multivariate calibration, John Wiley & Sons, New York, 1989.

[10] B.R. Buchanan, D.E. Honigs, J. Cynthia, W. Roth, Applied Spectroscopy, 42 (1988) 1106.

[11] A. Herrero, M.C. Ortiz, J. Arcos, J. Lopez-Palacios, L. Sarabia, Analytica Chimica Acta, 293 (1994) 277.

[12] M.C. Ortiz, A. Herrero, M.S. Sánchez, L. Sarabia, M. Íñiguez Chemometrics and Intelligent Laboratory Systems, 28 (1995) 273.

[13] M.C. Ortiz, L. Sarabia, C. Symington, F. Santamaría, M. Iñiguez, Analyst, 121 (1996) 1009.

[14] R. Lletí, E. Meléndez, M.C. Ortiz, L.Sarabia, , M.S. Sánchez Analytica Chimica Acta, 544 (2005) 60.

[15] M.S. Sanchez, M.C.Ortiz, L. Sarabia, V.Busto, Chemometrics and Intelligent Laboratory Systems, 103 (2010) 25.

[16] H.H. Harmann, Modern Factor Analysis 3rd ed., Chicago, University of Chicago Press, 1976.

[17] E. Meléndez, M. Íñiguez, P. Puras, M. C. Ortiz, L. A. Sarabia ,L. Dulau, Cuadernos de Campo, ISSN 1137-2095, 46 (2011) 34.

[18] J. Rousseau, Bulletin O.I.V., 74 (2001) 719.

[19] E. Meléndez, M. Íñiguez, P. Puras, M. C. Ortiz, L. A. Sarabia ,L. Dulau, Cuadernos de Campo, ISSN 1137-2095, 45 (2010) 34.

[20] B.M. Wise, N.B. Gallagher, R. Bro, J.M. Shaver, W. Windig, R.S. Koch, (2008) PLS Toolbox 5. Eigenvector Research Inc., Wenatchee, WA, USA.

[21] Matlab version 7.8.0.347 (R 2009a). The MathWorks, Inc. 3 Apple Hill Drive Natick, MA 01760-2098

[22] G.E.P. Box, Biometrika 36 (1949) 317.

[23] G.A.F. Seber, Multivariate observations. New York: John Wiley & Sons, Inc. 1984.

FIGURE CAPTIONS

Figure 1. Map of La Rioja in Spain. The coding of zones is the same as in table 2.

Figure 2. Loadings of varimax rotated latent variables. a) first, and b) second RLV for year 2007; c) first, and d) second RLV for year 2008; e) first, and f) second RLV for year 2009; g) first, and h) second RLV for year 2010; i) first, and j) second RLV for year 2011.

Figure 3. Map of correlations between physicochemical variables. a) Years 2007, 2009, 2010 and 2011, b) Year 2008.

Figure 4. Varimax rotated loadings on the first two latent variables of the PLS regression models for the years 2007, 2009, 2010 and 2011.

Figure 5. Varimax rotated scores on the first two latent variables of the PLS regression model for the years 2007, 2009, 2010 and 2011.

Table 1. Dates of sampling distribution

Year	Days	Number of samples
2007	September (4,11,18,29); October (2)	31
2008	September (17, 25, 30); October (10)	57
2009	September (1, 8, 15, 24, 30); October (7)	74
2010	September (7, 14, 24, 28); October (13)	56
2011	August (30); September (6, 12, 20, 27)	29

Table 2: Geographic distribution of samples

	Zone	Location	Variety	Year	Altitude (m)
HIGH RIOJA	I - Obarenes	Haro	Tempranillo	1993	438
	III - Sonsierra	San Vicente	Tempranillo	1987	440
	IV - Valpierre.	San Asensio	Tempranillo	1985	457
	V - Bajo Najerilla	Cenicero	Tempranillo	1998	434
	VI - Centro	Fuenmayor	Tempranillo	2000	428
LOW RIOJA	X-Iregua-Leza	Murillo	Mazuelo	1986	460
		Murillo	Tempranillo	1997	460
		Murillo	Garnacha	1997	460
	XI-Valle de Ocón	Alcanadre	Tempranillo	2000	400
		Alcanadre	Garnacha	1997	400
		Ausejo	Tempranillo	1984	565
		Ausejo	Garnacha	1984	565
		Ausejo	Graciano	1987	565
	XIV-Alhama-Aldeanueva	Aldeanueva	Tempranillo	1999	397
		Aldeanueva	Garnacha	1992	397
		Aldeanueva	Graciano	1996	350

Table 3 RMSEC and percentage of the variance explained on the first and second latent variables of the yearly PLS models computed between the physicochemical variables and the organoleptic tasting of grapes.

	Year				
	2007	2008	2009	2010	2011
RMSEC	0.344	0.506	0.804	0.284	0.306
X-Block					
1th latent variable	36.96	30.95	40.34	43.25	30.93
2th latent variable	24.45	25.11	20.49	14.76	13.50
Total	61.42	55.06	60.83	58.01	44.43
Y-Block					
1th latent variable	55.31	40.04	53.92	37.39	45.11
2th latent variable	18.19	9.97	12.07	13.70	20.71
Total	73.50	50.01	65.98	51.09	65.82

Table 4 Frequency of the scores of the sensory tasting in the 29 samples of year 2011.

Scores	Sensorial attribute			
	Sugars/acidity maturity	Aromatic maturity of the pulp	Aromatic maturity of the skins	Tannins maturity
2	1	3	4	6
3	8	10	15	19
4	20	16	10	4

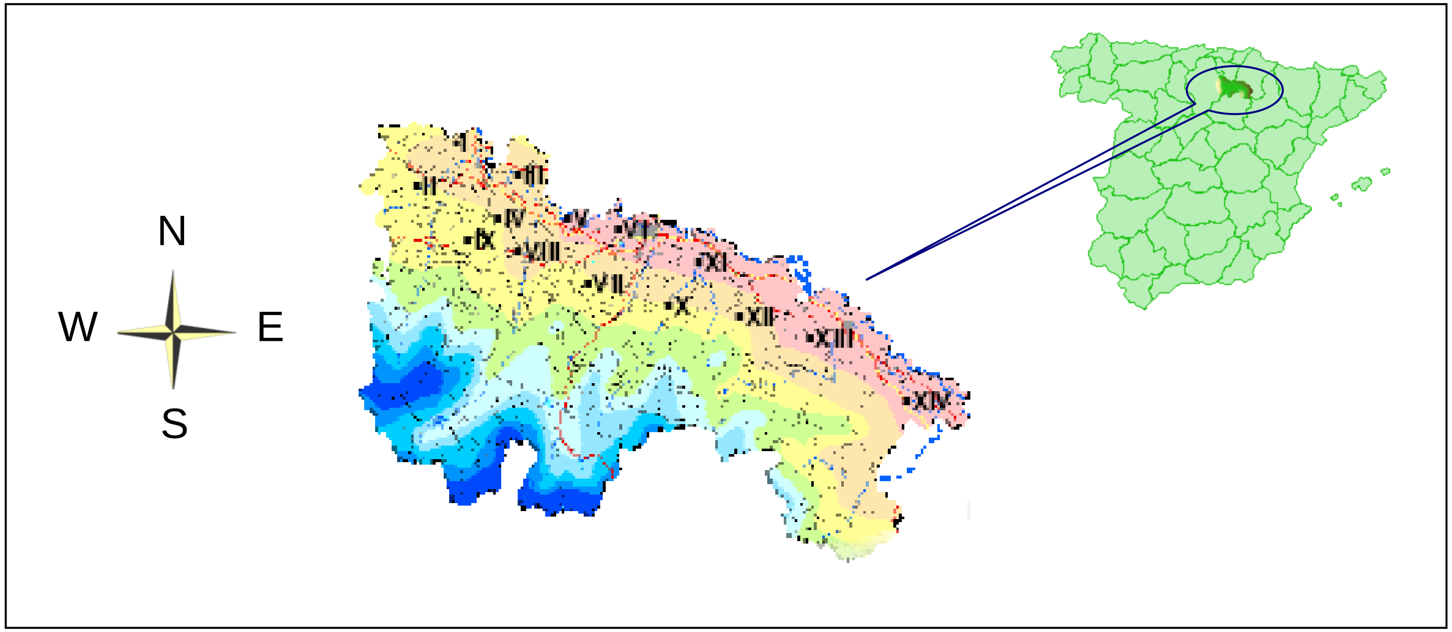


Figure 1

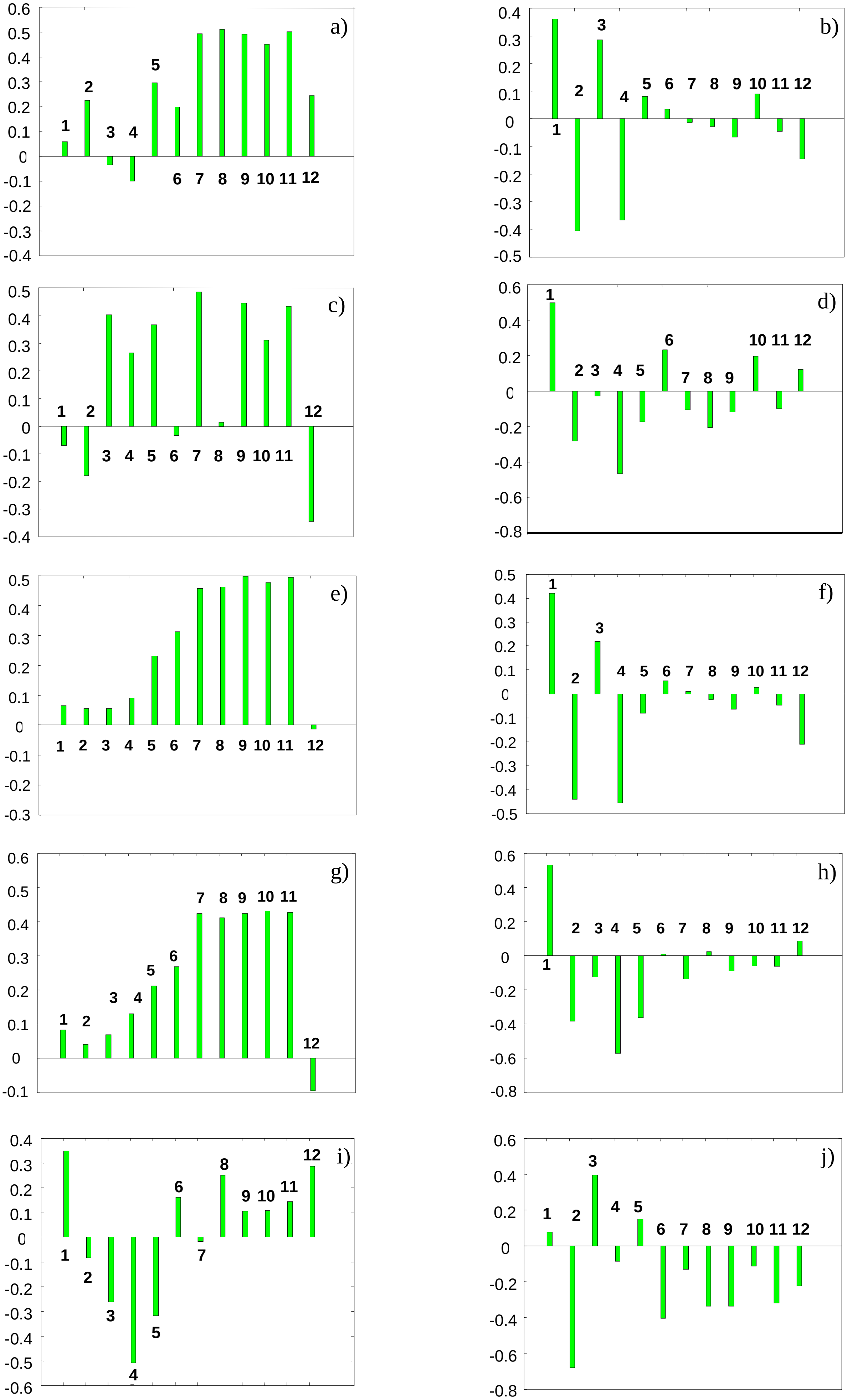

a)
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
1 2 3 4 5 6 7 8 9 10 11 12
b)
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
1 2 3 4 5 6 7 8 9 10 11 12
c)
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
1 2 3 4 5 6 7 8 9 10 11 12
d)
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
1 2 3 4 5 6 7 8 9 10 11 12
e)
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
1 2 3 4 5 6 7 8 9 10 11 12
f)
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
1 2 3 4 5 6 7 8 9 10 11 12
g)
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
1 2 3 4 5 6 7 8 9 10 11 12
h)
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
1 2 3 4 5 6 7 8 9 10 11 12
i)
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
1 2 3 4 5 6 7 8 9 10 11 12
j)
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
1 2 3 4 5 6 7 8 9 10 11 12

Figure 2

Figure3

Click here to download Figure: FIG3.doc

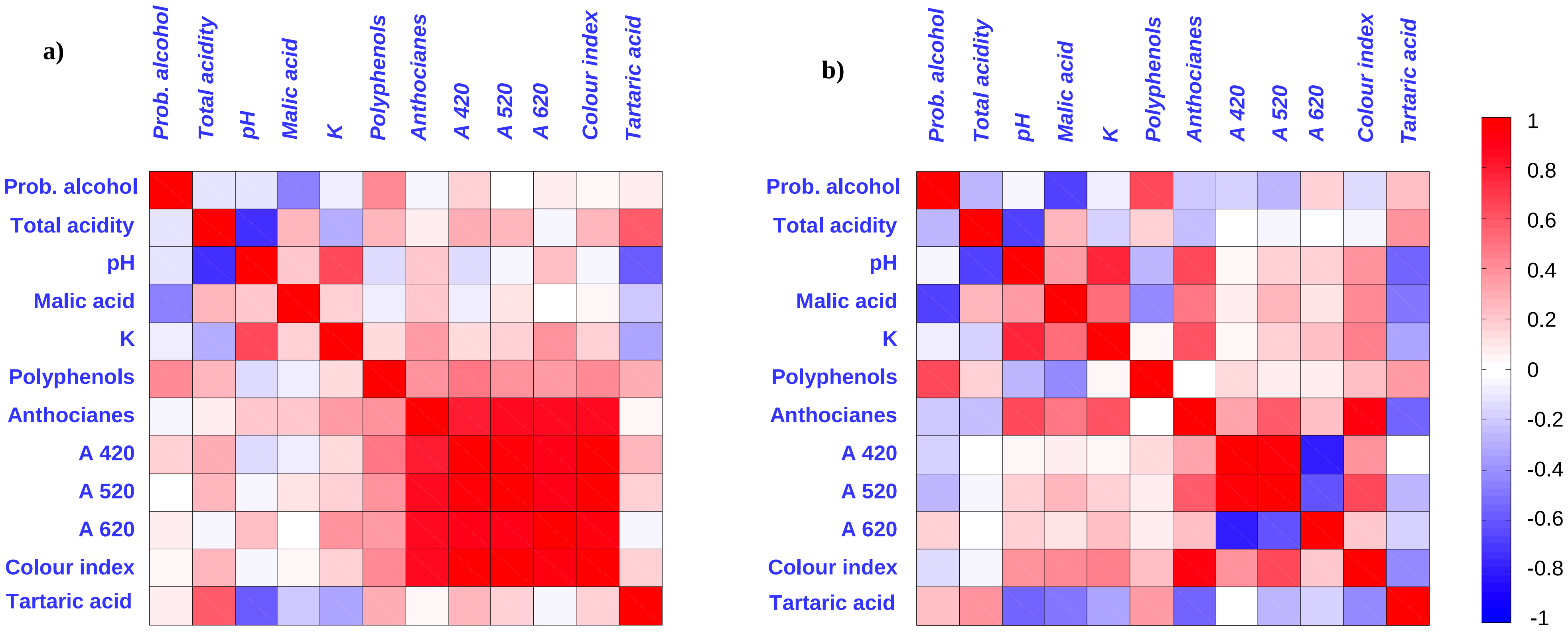


Figure 3

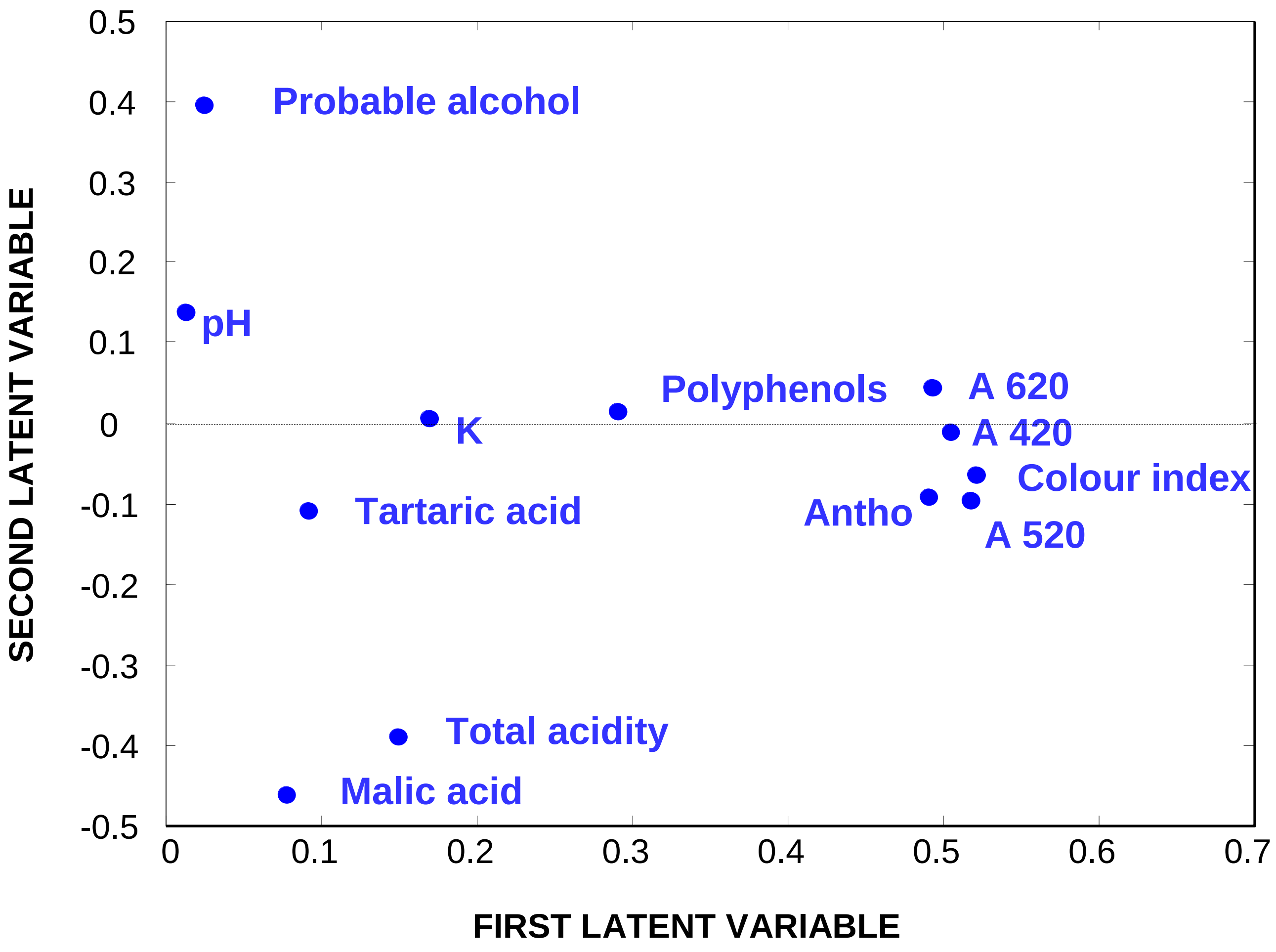

SECOND LATENT VARIABLE
FIRST LATENT VARIABLE
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Probable alcohol
pH
K
Polyphenols
A 620
A 420
Colour index
Tartaric acid
Antho
A 520
Total acidity
Malic acid

Figure 4

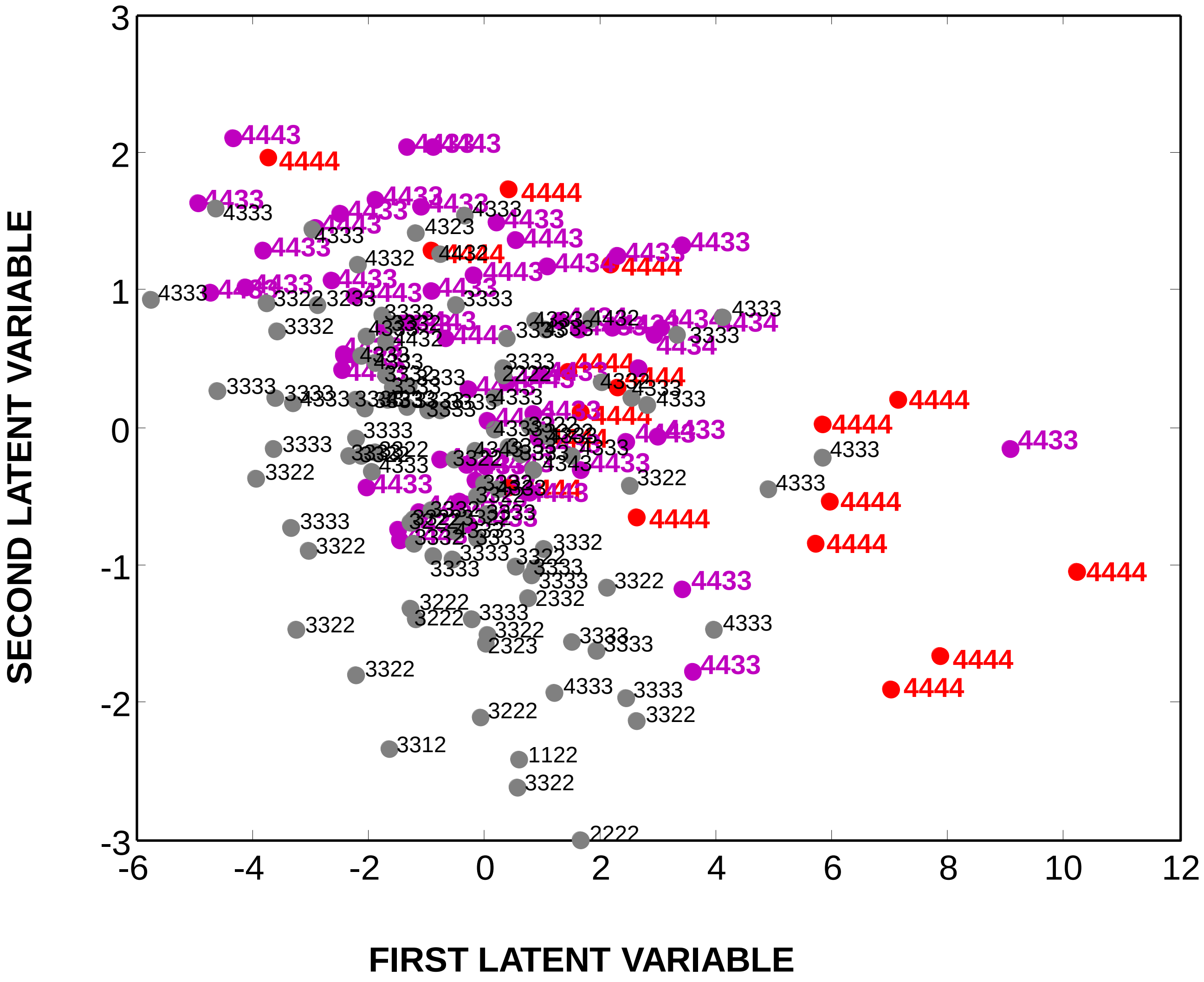

SECOND LATENT VARIABLE
FIRST LATENT VARIABLE
3
2
1
0
-1
-2
-3
-6
-4
-2
0
2
4
6
8
10
12

Figure 5