Determination of technological maturity of grapes and total phenolic compounds of grape skins in red and white cultivars during ripening by near infrared hyperspectral image: A preliminary approach.

Julio Nogales-Bueno, José Miguel Hernández-Hierro, Francisco José Rodríguez-Pulido and Francisco José Heredia*

Food Colour & Quality Laboratory, Department of Nutrition & Food Science, Universidad de Sevilla, Facultad de Farmacia, 41012 Sevilla, Spain.

*** Corresponding author:** Francisco José Heredia

Phone: +34 9545 56495

Fax: +34 9545 56110

E-mail: heredia@us.es

ABSTRACT

Hyperspectral images of intact grapes during ripening were recorded using a near infrared hyperspectral imaging system (900 - 1700 nm). Spectral data have been correlated with grape skin total phenolic concentration, sugar concentration, titratable acidity and pH by modified partial least squares regression (MPLS) using a number of spectral pre-treatments and different sets of calibration. The obtained results (RSQ and SEP respectively) for the global model of red and white grape samples were: 0.89 and 1.23 mg g^{-1} of grape skin for total phenolic concentration, 0.99 and 1.37 ºBrix for sugar concentration, 0.98 and 3.88 g L^{-1} for titratable acidity and for pH 0.94 and 0.12. Moreover, separate calibration models for red and white grape samples were also developed. The obtained results present a good potential for a fast and reasonably inexpensive screening of these parameters in intact grapes and therefore, for a fast control of technological and phenolic maturity.

KEYWORDS: Technological maturity, phenolic maturity, grapes, near infrared hyperspectral imaging; chemometrics.

1. Introduction

Grape harvest time is one of the most fundamental aspects that have influence on the future of wine quality. A number of factors have influence on this decision, among them technological and phenolic maturity of grape, especially grape skins phenolic maturity. Technological maturity is mainly connected with sugar concentration, titratable acidity and pH. The sugar concentration determines the potential alcoholic strength. The titratable acidity and pH help to control the wine quality and colour. Phenolic maturity shows the ripeness degree for the skins, pulp and seeds taking into account its phenolic composition (Meléndez, Ortiz, Sarabia, Íñiguez, & Puras, 2013; Ferrer-Gallego, Hernández-Hierro, Rivas-Gonzalo, & Escribano-Bailón, 2012).

It is really important to the winemakers the determination of adequate technological and phenolic maturities. Nowadays, at wineries, the maturity of grapes is usually controlled using classic physical and chemical analyses. For determining the sugar concentration is common to carry out density studies in the grape must since the specific gravity is directly related to the contents of soluble solids (°Brix) or to the Baume scale. It is also possible to measure the sugar concentration in the must by means of its refractive index. The determination of the total acidity and pH is usually carried out by means of volumetric titrations using NaOH and selective electrodes respectively. The detailed phenolic profile is performed by High Pressure Liquid Chromatography (Xu et al., 2011) meanwhile spectrophotometric methods such as Folin-Ciocalteu are commonly selected for total phenolic determinations (Singleton, 1985).

All methods above mentioned are time consuming or destructive, or both. It would be interesting to replace these methods for new ones, not destructives and roughly reliable. Near infrared spectroscopy (NIRS) has been also used as analysis tool to replace traditional methods (Ferrer-Gallego, Hernández-Hierro, Rivas-Gonzalo, & Escribano-

62 Bailón, 2011). Moreover, hyperspectral techniques have been studied to replace
63 physicochemical analysis in several matrices (Baiano, Terracone, Peri, & Romaniello,
64 2012; Fernandes et al., 2011; Barbin, Elmasry, Sun, & Allen, 2013).
65 Characterizations of food quality, safety and composition have been accomplished using
66 the aforesaid analytical tool (Elmasry, Kamruzzaman, Sun, & Allen, 2012; Gowen,
67 O'Donnell, Cullen, Downey, & Frias, 2007; Lorente et al., 2012).
68 In the wine sector, it is really important to know critical parameters and attributes of
69 grapes, and is necessary to do it quickly and precisely. Near infrared hyperspectral
70 imaging could be an option to measure these parameters without sample destruction and
71 reagent consumption. In essence, hyperspectral imaging is a rapid, non-destructive,
72 rugged, multiparametric and flexible tool that potentially provides a suitable way to
73 analyse food (Gowen et al., 2007).
74 The aim of this study is to develop a useful and non-destructive hyperspectral method
75 for the determination of the principal parameters that compose phenolic and
76 technological maturity (i.e. pH, total acidity, sugar concentration and total phenols) in
77 white and red grapes. The samples used in this work have been collected in the
78 Condado de Huelva Designation of Origin D.O. (Andalusia, Spain) which is under the
79 typical climatic conditions of a warm area (Gordillo et al., 2012). To our knowledge,
80 this is the first time that near infrared hyperspectral imaging has been applied to grapes
81 to face the aforementioned goals.

82 ## 2. Material and methods

83 ### 2.1. Samples

84 *Vitis vinifera* L. cv. Zalema, Tempranillo and Syrah were collected from four vineyards
85 located in the Condado de Huelva Designation of Origin D.O. (Andalusia, Spain).
86 Zalema is a white cultivar autochthonous to the South of Spain where it represents over

90% of the overall production (Hernanz et al., 2009). Zalema grapes were collected from two vineyards which present different types of soil, sand and clay. Tempranillo is the most often grown red grape cultivar in Spain for producing quality red wines and Syrah is a resistant cultivar to warm climatic conditions (Gordillo et al., 2012).

Both, white and red grapes were collected at different dates from mid-July to early September in the 2012 vintage. In this way, grapes were collected at different stages of maturity. There were different numbers of samples for each variety due to the earlier ripening of the red cultivars. Sixteen dates were taken into account for Tempranillo, seventeen for Syrah and eighteen for each Zalema type of soil. Three groups of berries were collected for each date and vineyard. With the aim of achieving representative samples, these were collected from the top, middle and bottom of the cluster and in the sunlight and shade side of this. After that, the samples were immediately frozen and stored at -20 ºC until analyses were performed. Two subsamples were taken from each sample, one to determine the reference parameters and the other one for the hyperspectral analysis.

2.2. Determination of reference parameters

Reference parameters were total phenolic concentration in grape skins, sugar concentration, titratable acidity and pH. Total phenolic concentration was determined using the Folin-Ciocalteu method (Singleton, 1985). In order to perform this determination, grape skins were separated manually from the whole grapes. Afterwards, one gram of grape skin was macerated in 10 mL of methanol containing 0.1% of 12M HCl. Methanolic phases were centrifuged (3000 rpm, 10 min) and successively pooled, approxymately 2 mL millilitres of water were added and the extract was concentrated under vacuum at 30 ºC until methanol was removed and finally made up to 10 mL with ultrapure water. The total phenolic concentration was determined using the Folin-

112 Cicalteu method in this aqueous extract and it was expressed as gallic acid equivalents
113 per gram of grape skin.
114 Technological maturity parameters were determined using the analytical methods
115 recommended by the O.I.V. (1990) using the must obtained after crushing the grapes.

116 **2.3. Hyperspectral imaging analysis**

117 **Fig. 1** shows the main components of the hyperspectral imaging device (Infaimon S.L.,
118 Barcelona, Spain) which were the illumination source, optics (mirror scanner and lens),
119 spectrograph, camera and computer. The system comprised a Xenics® XEVA-USB
120 InGaAs camera (320 × 256 pixels; Xenics Infrared Solutions, Inc., Leuven, Belgium), a
121 spectrograph (Specim ImSpector N17E Enhanced; Spectral Imaging Ltd., Oulu,
122 Finland) covering the spectral range between 900 and 1700 nm (spectral resolution of
123 3.25 nm), two 70 W tungsten iodine halogen lamps (Prilux ®, Barcelona, Spain)
124 mounted as source light, a mirror scanner (Spectral Imaging Ltd., Oulu, Finland) and a
125 computer system. Hyperspectral images were recorded using a 50 Hz frame rate and an
126 exposure time of 9 ms using the instrument acquisition software SpectralDAQ v. 3.62
127 (Spectral Imaging Ltd., Oulu, Finland).
128 A two point reflectance calibration was used. A Spectralon® ceramic tile (Labsphere
129 Inc., North Sutton, USA) was used as a white reference while dark current was recorded
130 by taking a measurement after covering the spectrograph lens with a cup and closing the
131 shutter. Corrected reflectance values (R) were calculated taking into account the
132 relationship between sample (S), white standard (W) and dark current (D) absolute
133 signal intensities using the following formula:
134 $R=[(S-D)/(W-D)]$ (1)
135 Thereafter, the samples were thawed and tempered at room temperature and the
136 hyperspectral images of the intact grapes on a polyethylene plastic were recorded. The

137 characteristic spectral profile of this surface was useful in segmentation process for
138 recognising the region of interest. Noisy wavebands at both extremes of the spectra
139 range were removed and only spectral data in the resulting effective wavelength 950 -
140 1650 nm regions were used in data analysis due to reduced efficiency outside this range
141 in the used device.

142 **2.4. Image processing and data analysis**

143 *Image processing.* Image treatment was carried out using Matlab (R2010b; The Math
144 Works, Inc. USA). Prior to the quantitative analysis, a discriminant method was applied
145 to the grape images to isolate the grapes from other parts of image. Firstly, three regions
146 of interest (ROIs) were selected (background, grape and pedicel) to develop a stepwise
147 lineal discriminant model. The aforementioned discriminant model classified each pixel
148 into two classes (grape or no grape pixel) using the reflectance values from six
149 wavelengths (979, 1034, 1073, 1314, 1386 and 1550 nm). After that, the average
150 spectrum of the grape region was extracted and then transformed into Log (1/R) units.
151 The procedure was repeated for each sample and the obtained spectra were combined
152 into the spectral matrix.

153 *Data analysis.* Prior to quantitative analysis, an unsupervised pattern recognition
154 technique, principal component analysis (PCA), was used in order to provide
155 information about the latent structure of spectral matrix. The spectral matrix was
156 constructed from the red grape spectra, white grape spectra or both. This method
157 provides not only information related to spectral outliers, the distribution of samples in
158 the newly-created space and their possible separations in different spectral groups but is
159 also an important source of knowledge with which to create cross-validation groups
160 used in the calibration process (Shenk & Westerhaus, 1995; Brereton, 2003).

161 Calibrations were performed using modified partial least squares regression (MPLS).
162 For achieving this task, the corresponding total phenolic concentration, sugar
163 concentration, titratable acidity and pH values were allocated to the raw spectrum of
164 each sample, and then different spectral pre-treatments were tested. They were also used
165 different calibration sets (i.e red grapes, white grapes or both). In this method, the group
166 of calibration samples is divided into a series of subsets in order to perform cross-
167 validation to set the number of PLS factors, reduce the possibility of overfitting (Shenk
168 & Westerhaus, 1995) and remove chemical outliers. Using the T≥2.5 criterion, samples
169 that presented a high residual value when they were predicted were eliminated from the
170 set. Finally, validation errors are combined into a single figure, the standard error of
171 cross-validation (SECV).
172 Spectral pre-treatments are usually applied to NIR raw data; scattering effects were
173 removed using multiplicative scatter correction (MSC), standard normal variate (SNV),
174 and detrending (Geladi, MacDougall, & Martens, 1985; Dhanoa, Lister, & Barnes,
175 1995). Moreover, the effect of differentiation and variations in spectral ranges were
176 tested in the development of the NIRS calibrations.
177 The software used was Win ISI® (v1.50) (Infrasoft International, LLC, Port. Matilda,
178 PA, USA). This software allowed the data pre-treatment and development of
179 quantitative and qualitative models. From the three samples of each date, one (33%)
180 was randomly allocated to the validation set and the other two (66%) to the calibration
181 set. Consequently, from the 213 spectral samples (99 red and 114 white grape spectral
182 samples), 142 were allocated in the calibration set and the remaining 71 were allocated
183 in the validation set (two and one thirds of white and red grape spectral samples
184 respectively).

185 **3. Results and discussion**

3.1. Chemical analysis

Total phenolic concentration ranged from 2.2 to 15.8 mg g^{-1} of grape skin with a standard deviation value of 2.9 mg g^{-1} of grape skin.

Sugar concentration, titratable acidity and pH ranged from 4.1 to 25.4 ºBrix, 2.7 to 52.9 g L^{-1} and 2.5 to 3.8 respectively.

3.2 Hyperspectral imaging analysis

It was expected a clear difference between red and white grape spectra, however the average spectra of red and white grapes are really similar **(Fig. 2)**. For this reason it was initially decided to work with both groups together. This figure shows the average spectra of red and white grapes over the 950-1650 nm range. Standard deviation spectra for each group are also represented and for display reasons they have been multiplied by a factor of 10. Spectral intensities were low and well within the linear response range of the instrument detector range. A strong feature of the sample spectra was the absorbance pattern around 1250 and 1450 nm wavelengths.

A SNV (2,5,5,1) spectral pre-treatment was applied to the spectra of both red and white grapes in the 950-1650 nm regions, where the hyperspectral system has revealed greater efficiency. Mathematical treatment is denoted as a,b,c,d, where the first digit is the number of the derivative; the second is the gap over which the derivative is calculated; the third is the number of data points in a running average or smoothing, and the fourth is the second smoothing (Shenk & Westerhaus, 1995). This spectral pre-treatment was applied only in the calibration set and after that, principal component analysis was carried out in order to look for spectral outliers and create cross-validation groups. Overall, the spectral variability explained was 99% using 15 principal components and Mahalanobis distances for each sample were calculated. Samples were ranked in order

of their H (Mahalanobis) distance from the mean spectrum of the entire sample set and the H > 3 criterion was applied. Only one H-outlier was found, a Zalema sample which spectrum did not meet this criterion and it was eliminated from the calibration set. **Fig. 3** shows the scores of the grape samples in the space defined by the first and second principal components which described 57.38% (PC1) and 22.06% (PC2) of the variability in the data. In this plot (**Fig. 3a**) it is not possible to separate completely red and white grape samples, however, it is possible to see some differences between both groups. Furthermore, it is also possible to find some semi-separation between the early days of ripening and the rest of samples (**Fig. 3b**). It is not shown a cultivar comparison (i.e. Zalema, Tempranillo and Syrah) because the different varieties were overlapped in this plane.

Finally, quantitative calibrations were developed by modified partial least squares (MPLS) regression. As described above, to perform this calibration all grape spectra, red and white, were used as the independent (X) variables. Total phenolic concentration, sugar concentration, titratable acidity and pH were used as dependent (Y) variables. The statistical parameters of the final calibration equations are shown in **Table 1** where N is the number of samples used to obtain the calibration equation after eliminating samples for chemical reasons (T criterion). The best of the different mathematical treatments, the range of application, and standard deviations are also shown.

The robustness of the selected models was tested using a set of 71 samples, which did not belong to the calibration set, as external validation. In the case of titratable acidity, two samples presented reference values outside the applicability of the obtained models and then should not be used in this procedure. As result of this external validation they were obtained the standard errors of prediction (SEP) for each reference variable, these values were also included in **Table 1**. A relevant aspect of this method was observed,

235 the prediction of all the reference parameters was worse for the white grape samples
236 than for the red ones. This can be observed in **Fig. 4a**, which shows the SEP (expressed
237 as percentages) when these global models were used to predict the reference parameters
238 in all, white or red grape samples. Considering these results it was decided to develop
239 separate calibration models for red and white grape samples. These models were
240 developed following the same methodology described above. The spectral samples were
241 randomly allocated to the validation and calibration sets. In the calibration set a SNV
242 (2,5,5,1) spectral pre-treatment and a principal component analysis were carried out in
243 order to look for spectral outliers and create cross-validation groups. It was not found
244 any outlier for red grape model, however, a spectral sample was eliminated from the
245 white calibration set. The same sample had been eliminated in global model following
246 the Mahalanobis distance criterion. For each calibration set (i.e. red and withe samples),
247 a quantitative calibration was developed by modified partial least squares (MPLS)
248 regression. Finally, the robustness of the selected models was tested using the external
249 validation sets (33 samples for the red grape model and 38 samples for the white one).
250 In the case of the red grape model all sample presented reference values inside the
251 applicability of the obtained model, however, one and three samples were not used in
252 the white grape model for total phenols and titratable acidity validations respectively.
253 As result of these external validations they were obtained the standard errors of
254 prediction (SEP) for each reference variable. The statistical parameters of these
255 calibrations and the standard errors of prediction are described in **Table 1**. Better results
256 were achieved in the prediction of red grape parameters using these new models,
257 however, the prediction of white ones was worse for total phenolic concentration, sugar
258 concentration and pH. **Fig. 4b** shows this behaviour, here it is possible to observe the
259 SEP (%) when global, red grape and white grape methods are used to predict their

respective validation sets. Since there were not differences neither in the applied reference methods nor in the hyperspectral analysis used for white and red grape samples it was not possible to indicate the source of the larger error for the white samples than for the red ones.

If differences in determination of analytical and spectral data are discarded, these findings may be due to a tinier spectral variability in red grape samples than in white ones. Samples set with tiny spectral variability can generate calibration models that predict sample parameters with small errors at the expense of less applicability (Shenk & Westerhaus, 1995). In **Fig. 3a** it is possible to see that the red grape samples are less dispersed than white ones. This theory is also supported by the fact that global model is better than white grape model and worse than red one. This might be because spectral samples used in the global calibration are as dispersed as the samples in white set but the global calibration set is greater than the white one. However, the results obtained in the external validation, the SEP values, are comparatively similar to the errors previously reported for these parameters using classic near infrared spectroscopy taking into the account the applicability range (Cozzolino, 2009; Cozzolino, Dambergs, Janik, Cynkar, & Gishen, 2006; Ferrer-Gallego et al., 2011; González-Caballero, Sánchez, López, & Pérez-Marín, 2010; Kemps, Leon, Best, De Baerdemaeker, & De Ketelaere, 2010).

4. Conclusion

The procedure reported here using near infrared hyperspectral imaging presents a good potential for a fast and reasonably inexpensive screening of total phenolic concentration, sugar concentration, titratable acidity and pH in intact grapes, and therefore, for a fast control of technological and phenolic maturity. Nonetheless, a comprehensive study

should be made in order to evaluate factors, such as different production areas and grape varieties, in the complete development of these models.

Abbreviations

H, Mahalanobis distance; MPLS, modified partial least squares; MSC, multiplicative scatter correction; NIRS, near infrared spectroscopy; PC, principal component; PCA, principal component analysis; PLS, partial least squares; RSQ, coefficient of determination; SEC, standard error of calibration; SECV, standard error of cross-validation; SEP, standard error of prediction; SNV, standard normal variate.

Acknowledgments

The Spanish MICINN is thanked for J. Nogales-Bueno, F.J. Rodríguez-Pulido FPI grants (BES-2012-060192 and BES-2009-025429 respectively), J.M. Hernández-Hierro Juan de la Cierva contract (JCI-2011-09201) and project AGL2011-30254-C02. Junta de Andalucía is also thanked for financial support (project P10-AGR6331).

References

Baiano, A., Terracone, C., Peri, G., & Romaniello, R. (2012). Application of hyperspectral imaging for prediction of physico-chemical and sensory characteristics of table grapes. *Computers and Electronics in Agriculture, 87,* 142-151.

Barbin, D. F., Elmasry, G., Sun, D. W., & Allen, P. (2013). Non-destructive determination of chemical composition in intact and minced pork using near-infrared hyperspectral imaging. *Food Chemistry, 138,* 1162-1171.

Brereton, R. G. (2003). *Chemometrics : data analysis for the laboratory and chemical plant*. Chichester, West Sussex, England: J. Wiley.

Cozzolino, D. (2009). Near Infrared Spectroscopy in Natural Products Analysis. *Planta Med., 75*.

Cozzolino, D., Dambergs, R. G., Janik, L., Cynkar, W. U., & Gishen, M. (2006). Review: Analysis of grapes and wine by near infrared spectroscopy. *Journal of near Infrared Spectroscopy, 14*.

Dhanoa, M. S., Lister, S. J., & Barnes, R. J. (1995). On the Scales Associated with Near-Infrared Reflectance Difference Spectra. *Applied Spectroscopy, 49*.

Elmasry, G., Kamruzzaman, M., Sun, D. W., & Allen, P. (2012). Principles and Applications of Hyperspectral Imaging in Quality Evaluation of Agro-Food Products: A Review. *Critical Reviews in Food Science and Nutrition, 52*, 999-1023.

Fernandes, A. M., Oliveira, P., Moura, J. P., Oliveira, A. A., Falco, V., Correia, M. J. et al. (2011). Determination of anthocyanin concentration in whole grape skins using hyperspectral imaging and adaptive boosting neural networks. *Journal of Food Engineering, 105*, 216-226.

Ferrer-Gallego, R., Hernández-Hierro, J. M., Rivas-Gonzalo, J. C., & Escribano-Bailón, M. T. (2011). Determination of phenolic compounds of grape skins during ripening by NIR spectroscopy. *Lwt-Food Science and Technology, 44*, 847-853.

Ferrer-Gallego, R., Hernández-Hierro, J. M., Rivas-Gonzalo, J. C., & Escribano-Bailón, M. T. (2012). Influence of climatic conditions on the phenolic composition of *Vitis vinifera* L. cv. Graciano. *Analytica Chimica Acta, 732*, 73-77.

329 Geladi, P., MacDougall, D., & Martens, H. (1985). Linearization and Scatter-Correction
330 for Near-Infrared Reflectance Spectra of Meat. *Applied Spectroscopy, 39*.

331 González-Caballero, V., Sánchez, M. T., López, M. I., & Pérez-Marín, D. (2010). First
332 steps towards the development of a non-destructive technique for the quality
333 control of wine grapes during on-vine ripening and on arrival at the winery.
334 *Journal of Food Engineering, 101,* 158-165.

335 Gordillo, B., Rodríguez-Pulido, F. J., Mateus, N., Escudero-Gilete, M. L., González-
336 Miret, M. L., Heredia, F. J. et al. (2012). Application of LC-MS and tristimulus
337 colorimetry to assess the ageing aptitude of Syrah wine in the Condado de
338 Huelva D.O. (Spain), a typical warm climate region. *Analytica Chimica Acta,*
339 *732,* 162-171.

340 Gowen, A. A., O'Donnell, C. P., Cullen, P. J., Downey, G., & Frias, J. M. (2007).
341 Hyperspectral imaging - an emerging process analytical tool for food quality and
342 safety control. *Trends in Food Science & Technology, 18,* 590-598.

343 Hernanz, D., Gallo, V., Recamales, Á. F., Meléndez-Martínez, A. J., González-Miret,
344 M. L., & Heredia, F. J. (2009). Effect of storage on the phenolic content, volatile
345 composition and colour of white wines from the varieties Zalema and
346 Colombard. *Food Chemistry, 113,* 530-537.

347 Kemps, B., Leon, L., Best, S., De Baerdemaeker, J., & De Ketelaere, B. (2010).
348 Assessment of the quality parameters in grapes using VIS/NIR spectroscopy.
349 *Biosystems Engineering, 105*.

350 Lorente, D., Aleixos, N., Gomez-Sanchis, J., Cubero, S., Garcia-Navarrete, O. L., &
351 Blasco, J. (2012). Recent Advances and Applications of Hyperspectral Imaging

for Fruit and Vegetable Quality Assessment. *Food and Bioprocess Technology, 5,* 1121-1142.

Meléndez, E., Ortiz, M. C., Sarabia, L. A., Íñiguez, M., & Puras, P. (2013). Modelling phenolic and technological maturities of grapes by means of the multivariate relation between organoleptic and physicochemical properties. *Analytica Chimica Acta, 761,* 53-61.

Organisation internationale de la vigne et du vin (O.I.V.) Le recueil des méthodes internationales d'analyse des vibs et des moûts (1990). O.I.V.

Shenk, J. S. & Westerhaus, M. O. (1995). *Routine Operation,Calibration, Development and Network System Management Manual, NIRSystems*. MD, USA: Silver Spring.

Singleton, V. L. (1985). Citation Classic - Colorimetry of Total Phenolics with Phosphomolybdic-Phosphotungstic Acid Reagents. *Current Contents/Agriculture Biology & Environmental Sciences,* 18.

Xu, Y. P., Simon, J. E., Welch, C., Wightman, J. D., Ferruzzi, M. G., Ho, L. et al. (2011). Survey of Polyphenol Constituents in Grapes and Grape-Derived Products. *Journal of Agricultural and Food Chemistry, 59,* 10586-10593.