

# The use of hyperspectral imaging in the visible and near infrared region to discriminate between table grapes harvested at different times

Francesca Piazzolla, Maria Luisa Amodio, Giancarlo Colelli

Dipartimento di Scienze Agrarie, degli Alimenti e dell'Ambiente, Università di Foggia, Italy

## Abstract

Traditional analytical methods applied to the measurement of grape maturity and quality index in order to assess optimal harvest time have been proved to be slow and destructive. Therefore, non-destructive analytical techniques, including spectroscopy, can be a valid support for the choice of the best time to harvest. This study evaluated the feasibility of using a visible and near infrared spectral scanner (v. 1.4; DV Srl, Padova, Italy) with a detector in the region between 400-1000 nm to discriminate between grapes harvested at different times. Twelve clusters were harvested at 5 different times between October and December 2011. Spectra were acquired with a Spectral scanner on 3 intact berries from each bunch. These were randomly selected from top, medium and bottom zones, for a total of 180 spectra. Classification models were construed comparing 2 methods: soft independent modelling of class analogy (SIMCA) and partial least squares discriminant analysis (PLS-DA). The SIMCA model was developed building individual principal component analysis (PCA) models for the spectra of each harvest time. Different pre-treatment methods were tested in order to enhance the power of the model, thus enhancing the score differences among samples from different harvest times. The transformation that allowed the best statistical separation among scores of grapes from different harvest times was the second derivative of Norris. Therefore, the PCA model obtained from the spectra subjected to this pre-treatment was used for SIMCA classification. The PLS-DA model were developed applying the PLS2 algorithm. In order to construct discriminant models to classify bunch spectra according to the 5 harvest times, spectral variations were correlated with the 5 categories established. No pre-treatments were previously applied in this last case since they did not improve the final result. The SIMCA method was unable to correctly

classify grapes from harvest time 2 (59% of correct classification) and was less efficient compared to the PLS-DA model. Using the PLS-DA model, all the grapes were correctly classified (100%) with the exception of those from harvest time 5 (94%).

The overall results demonstrate that this method has excellent potential for discriminating grape quality.

## Introduction

Recorded consumption of table grapes places them as the most popular fruit in the world. According to the International Organization of Vine and Wine statistics, recent years have seen an increase in their production for fresh consumption (OIV, 2012).

Being a non-climacteric fruit, the choice of the harvest time is extremely critical to the final eating quality depending as it does on various characteristics. Visual attributes, such as colour, size and shape of the berry, are the main characteristics that consumers look for (Cliff *et al.*, 1996; Zeppa *et al.*, 1999) while sensory descriptors, such as skin friability, skin thickness and flesh firmness characterise commercial table grape cultivars (Cliff *et al.*, 1996; Vargas *et al.*, 2001). Organoleptic quality is related to the soluble solids' content (SSC), titratable acidity (TA) and SSC/TA ratio (Dokoozlian *et al.*, 1995; Jayasena and Cameron, 2008). These quality attributes change during maturation. Maturation involves (Coombe, 1992; Watson, 2003): i) growth of the berry; ii) modification of its mechanical consistency; iii) accumulation of simple sugars (glucose and fructose that at maturity may represent up to 20% of the fresh weight); iv) reduction in acidity and increase in the pH of the juice; v) degradation of chlorophyll; vi) accumulation of aromatic molecules and of their precursors.

As for most fruits and vegetables, in general, analytical methods to determine grape maturity, index and composition are destructive and involve a considerable amount of manual work that may require sophisticated equipment. In recent years, research has focused on the development of non-destructive techniques suitable to increase the number of pieces of fruit that can be analysed and which can be repeated on the same sample during its physiological evolution thus providing real-time information (Costa *et al.*, 2009).

Among these, visible and near infrared (VIS-NIR) (v. 1.4; DV Srl, Padova, Italy) spectroscopy is widely used for rapid, low-cost and non-destructive analysis of inner properties of fruits. The use of VIS-NIR technology has been extensively used to: i) predict SSC, pH, acidity, phenols and antioxidant activity of wine-grapes (Fernández-Novales *et al.*, 2009; González-Caballero *et al.*, 2012; Cao *et al.*, 2010); ii) to distinguish between different ripening stages with a view to optimising harvest times depending on the wine-grape variety and the type of the wine (González-Caballero *et al.*, 2012; Guidetti *et al.*, 2010); and iii) to discriminate between grape varieties (Kemps *et al.*, 2010). Fewer studies are available about the use of these technologies to assess quality and maturity of table grapes for which correct identification of the optimal harvest time is, from a sensorial point of view, very important.

Correspondence: Maria Luisa Amodio, Dipartimento di Scienze Agrarie, degli Alimenti e dell'Ambiente, Università di Foggia, via Napoli 25, 71122, Foggia, Italy. E-mail: marialuisa.amodio@unifg.it

Key words: soft independent modelling of class analogy, partial least squares discriminant analysis, table grapes, spectra, classification.

Received for publication: 16 May 2013.

Accepted for publication: 5 August 2013.

©Copyright F. Piazzolla *et al.*, 2013

Licencee PAGEPress, Italy

Journal of Agricultural Engineering 2013; XLIV:e7

doi:10.4081/jae.2013.e7

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License (by-nc 3.0) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Among them, Baiano *et al.* (2012) studied the prediction of physico-chemical and sensory attributes of white and red/black table grapes using the information provided by hyperspectral imaging.

Using the similarity between the spectral characteristics of a group of samples, classification methods can be applied to discriminate between different groups. Such a similarity can be expressed through different tools, such as spectral correlation, logical operations or calculating distances (Pérez-Marín, 2005). The most common methods of classification are: artificial neural networks, soft independent modelling of class analogy (SIMCA), and partial least squares-discriminant analysis (PLS-DA) (Naes and Indahl, 1998; Vandeginste *et al.*, 1998; Vigneau *et al.*, 2000; Naes *et al.*, 2002).

Classification models have been used to discriminate the harvest time of fruits and vegetables (Tavakolian *et al.*, 2013; González-Caballero *et al.*, 2012). PLS-DA was used to discriminate among classes of maturity of apples (based on the starch index; Menesatti *et al.*, 2009) or among orange varieties (Cen *et al.*, 2007). Both PLS-DA and SIMCA methods were compared to discriminate between non-bruised and bruised longan fruits (Polpho *et al.*, 2011), with PLS-DA giving the best performances.

The SIMCA model (Wold and Sjostrom, 1977) is based on a separate principal component analysis (PCA) model that defines the principal components for each class and calculates their optimal number. All the individual PCAs constitute the model. For the classification of a new unknown item, the *Euclidean distance* of the new item from each group is calculated in order to identify the nearest class.

The PLS-DA model uses a training group to develop qualitative prediction models that can be applied later to the classification of unknown samples. The model seeks to correlate spectral variations (X) with defined classes (Y), trying to maximise co-variance between the two types of variables. In this type of approach, the variables are artificial or fictitious categorical variables (*dummy*), created by assigning the value 0 to the sample of category A and the value 1 to the sample of category B (Heise and Winzen, 2002; Naes *et al.*, 2002; Kramer *et al.*, 2004). In the case of more than 2 classes, this means that each sample has a value of 1 in the class to which it belongs, and a value of 0 for the other classes (Vandeginste *et al.*, 1998; Heise and Winzen, 2002; Naes *et al.*, 2002).

The aim of this work was to study the feasibility of using data provided by spectral imaging extracted with an innovative hyperspectral scanner. This did not require any sample preparation and was able to discriminate between grapes from different harvest dates comparing PLS-DA and SIMCA classification methods. Hyperspectral imaging, that has a different acquisition system to the VIS-NIR sample preparation, spectral range (in this study 400-1000 nm) and resolution may be further developed for direct use in the field or on a sorting line, thus improving the entire process of table grape production and handling operations.

## Materials and methods

### Experimental design

Table grapes (*cv. Italia*) grown in the vicinity of Foggia, central Italy, cultivated with the Apulia canopy grape training system and covered with net and low density polyethylene plastic film (polyethylene low density, free of substances to thermal effect, 170  $\mu\text{m}$  thick) were harvested starting from 7<sup>th</sup> October 2011 (HT 1) and after 11 (HT 2), 27 (HT 3), 48 (HT 4) and 56 (HT 5) days. At each harvest time, 36 clusters (3 bunches from each of 12 plants) were harvested and transported to the Postharvest Laboratory of the University of Foggia.

### Quality determination

In order to monitor grape maturity stage and nutritional quality at each harvest, total soluble solids, acidity, pH, antioxidant activity and total phenol content were measured on grape juice as follows.

- Total soluble solids (TSS) were measured using a digital refractometer (Atago PR32-Palette; ATAGO CO., Ltd., Tokyo, Japan).

- Titratable acidity (TA) and pH were measured with an automatic titrator (TitroMatic CRISON 1S; Crison Instrument, Barcelona, Spain). Acidity was expressed as percentage of tartaric acid.

- Total phenol content and antioxidant activity were measured using the same extraction procedure. Five grams of berries were homogenised in an Ultraturrax (IKA T18 basic; IKA®-Werke GmbH & Co. KG, Staufen, Germany) after the addition of 3 mg/g of methanol plus 3% formic acid. The extracts were then centrifuged at 5°C and 9000 rpm for 10 min. Total phenols were determined according to the method of Singleton and Rossi (1965). Each extract (100  $\mu\text{L}$ ) was mixed with 1.58 mL water, 100  $\mu\text{L}$  of Folin-Ciocalteu reagent and 300  $\mu\text{L}$  of sodium carbonate solution (200 g/L). After the solution had been standing for 2 h in the dark, the absorbance was read at 725 nm against a blank using a spectrophotometer (UV-1700; Shimadzu Corp., Jiangsu, China). The content of total phenols was calculated on the basis of the calibration curve of gallic acid and was expressed as grams of gallic acid per kilogram of fresh weight (g GA/kg). Antioxidant assay was performed following the procedure described by Brand-Williams *et al.* (1995) with minor modifications. The diluted sample (50  $\mu\text{L}$ ) was pipetted into 0.95 mL of diphenylpicrylhydrazyl solution to initiate the reaction. The absorbance was read after 24 h at 515 nm. Trolox (6-Hydroxy-2, 5, 7, 8-tetramethylchromane-2-carboxylic acid) was used as a standard and the antioxidant activity was reported in milligrams of Trolox equivalents per kilogram of fresh weight (g TE/kg).

### Visible and near infrared image and spectra acquisition

Hyperspectral images were taken on 3 berries per bunch, selected from top, medium and bottom zones, in order to control variability due to the berry position in the cluster with a lab-scale hyperspectral imaging system (v. 1.4.5; DV Srl). The analysis was performed on fruits at room temperature (approx. 20°C). One scan per sample (3 berries) was acquired at an acquisition speed of 3 mm/s.

Hyperspectral imaging system consisted of a charge-coupled device (CCD), a 12-bit camera connected to a V10 type spectrograph (400-1000 nm, 25  $\mu\text{m}$  slit, resolution 5 nm; ImSpector V10, Specim Ltd., Haarlem, The Netherlands) coupled with a standard C-mount f16 mm lens. The optics of this imaging system allowed us to study the fruit properties associated to the spectral range 400-1000 nm of reflectance with 5 nm of resolution. The target was placed at a distance of 360 mm from the camera. The light source consisted of a 150 W halogen lamp (EKE 21V150 W, Tokyo, Japan) mounted at an angle of 45° to the horizontal plane, and of an optic fibre that transfers the radiation to a linear light diffuser. The camera spectrograph assembly was supplied with a stepper motor to move the unit through the field of view of the camera and carry out a line-by-line scan of the berry. The spectral images were collected in a dark room where the halogen light was the only light source. The hyperspectral images were first corrected with a white and a dark reference. The dark reference was used to remove the effect of dark current of the CCD detectors that are thermally sensitive.

All the spectra were extracted using the Spectral Scanner software (v. 1.4.5; DV Srl). A total of 180 spectra were obtained (3 berries from each of 12 bunches harvested at 5 different times). A region of interest corresponding to the maximum inscribed rectangle was manually selected on each berry.

## Spectra pre-treatment

Spectral data were analysed using the Unscrambler packing software version 9.1 (CAMO ASA, Oslo, Norway). All the reflectance measurements were first transformed to absorbance values using  $\log(I/R)$  according to the law of Lambert-Beer.

The spectra were analysed with a PCA central model to identify and eliminate defective spectral outliers (Massart *et al.*, 1998; Naes *et al.*, 2002).

## Development of classification models

Classification models were constructed to classify grapes according to harvest time, comparing 2 classification methods: SIMCA using the Unscrambler 9.1 software, and PLS-DA using the WinISI II software package version 1.50 (Infrasoft Int., Port Matilda, PA, USA).

## Soft independent modelling of class analogy

The SIMCA classification is a method based on disjoint PCA modelling realised for each class in the calibration set. Unknown samples are then compared to the class models and assigned to classes according to their analogy with the calibration samples. Each class is modelled using separate PCA models. A number  $K$  of principal components is used to build the model. These  $K$  components define the inner space and the space of the structure, and the other principal components are the outer space and the space of the noise. The SIMCA model is a hyper volume in the space of the significant components, delimited by the range of the scores (normal range). The sensitivity of a class model is the fraction of the objects belonging to that class accepted by the model. The specificity of a model is the fraction of objects belonging to other classes rejected by the studied class model. An unknown sample is compared with each group in turn by computing two distances: the Euclidean distance from the spectrum to its projection into the model for that group (distance from the model) and the Mahalanobis distance from the projected spectrum to the group mean (distance within the model), and comparing these distances with thresholds derived from the training data. If both distances are less than the threshold, the unknown is a possible member of that group. After all the comparisons have been made, the unknown sample may be identified as a possible member of none, one, or more than one of the groups.

In our case, the SIMCA model was developed by building PCA models for each harvest time and these were subsequently used to classify external spectra.

Spectra were pre-treated by different mathematical methods. In order to find the mathematical transformation which best differentiated between spectra from different harvest times, different pre-treatment methods were tested. Pre-treatment methods (smoothing, normalisation, multiple scatter correction, noise, Norris derivative, Savitzky-Golay derivative and baseline) were tested individually. Transformed data were used to build PCA models and the obtained scores of the PC1 were then subjected to a one-way analysis of variance

(ANOVA) for the harvest time. Means were compared using Tukey's test ( $P \leq 0.05$ ).

The PCA scores represent the weighted sums of the original variables without significant loss of useful information, whereas loadings (weighting coefficients) identified major variables responsible for specific features appearing in the scores.

Based on Tukey test results, the transformation that provided the best differentiation between scores of grapes from different harvest times was chosen to build the individual PCA models for each harvest time. The test was performed by applying also the random method of crossvalidation with 10 segments and 17 samples per segment in order to prevent model over-fitting.

## Partial least squares-discriminant analysis model

The PLS-DA model was developed applying the PLS2 algorithm and using the *discriminant equations* option in the WinISI II (version 1.50 software package, 2000; Infrasoft Int.). While PLS was initially built for quantitative analysis, now it is also used for qualitative classification. This supervised analysis is based on the relation between spectral intensity and sample characteristics, in our case, using spectral variations ( $X$ ) and 5 established categories, corresponding to the harvest dates ( $Y$ ), and trying to maximise covariance between the two types of variables. PLS-DA is, therefore, performed using an exclusive binary coding from PLS regression (PLS2) that uses  $M$  spectral variables as predictors and  $q$  variables (0 or 1) as variables response.

During the calibration process, the PLS-DA method is trained to compute the *membership values*, one for each class. The sample is then assigned to one class when the value is above a specific prediction threshold (Musumarra *et al.*, 2005; Liu *et al.*, 2007).

All models were constructed using full crossvalidation (*leave-one-out*), suitable for small sample sets (Naes *et al.*, 2002).

## Statistical analysis

One-way ANOVA was performed on quality attributes at harvest and on PC1 scores after mathematical pre-treatment. Mean values were separated with Tukey's test ( $P < 0.05$ ) after testing the hypothesis of the normal distribution of the data and of the homogeneity of the variance (Levene test with  $P = 0.05$ ). Data were analysed with the StatGraphics Centurion software (v. 16.1.11; StatPoint Technologies, Inc., Warrenton, VA, USA).

## Results and discussion

### Grape composition

Table 1 shows the evolution of quality parameters as harvesting proceeds over time. Harvest time influenced TSS and TA at the harvest time point, while for pH, phenols, and antioxidant activity, any change

**Table 1. Chemical composition of *Italia* table grapes at different harvest times.**

Harvest times (HT)	TSS (°Brix)	pH	Titrateable acidity (% tartaric acid)	Phenols (gallic acid mg/100 g)	Antioxidant activity (Trolox mg/100 g)
1	21.57 <sup>a</sup>	4.21	0.37 <sup>a</sup>	51.76	296.38
2	20.86 <sup>a</sup>	4.30	0.34 <sup>ab</sup>	44.16	268.01
3	21.89 <sup>a</sup>	4.27	0.30 <sup>b</sup>	58.76	273.39
4	17.97 <sup>b</sup>	4.28	0.30 <sup>b</sup>	43.80	264.07
5	18.21 <sup>b</sup>	4.37	0.29 <sup>b</sup>	46.59	264.36

TSS, total soluble solids. <sup>ab</sup>Within each column different letters indicate significantly different mean values ( $P < 0.05$ ).

was observed after the first harvesting. Total soluble solids decreased from HT 1, 2 and 3 (approx. 21° Brix) to HT 4 and 5 (approx. 18° Brix) while no significant difference was found according to the position in the cluster (*data not shown*). Titratable acidity decreased during ripening, with grapes of HT 1 showing a higher titratable acidity than grapes from the other harvest times, except for grapes of HT 2 that showed intermediate values. The decrease in acidity can be related to the progressive maturation. However, since this was not accompanied by an increase in soluble solids (Watson, 2003), it can be hypothesised that grapes up to HT 3 had already reached optimal maturity and that only changes related to overripening and senescence were observed as harvesting proceeded over time. Furthermore, it can be concluded that there was no statistical difference in maturity stage for the entire harvest time, and that possibly only 2 classes of maturity could be differentiated.

### Spectral analysis

Principal component analysis showed the presence of 6 outliers; therefore, the number of initial samples was reduced to 174 and these were used for classification. These outliers were found according to the *Mahalanobis distance* (H), using a threshold of 3.0 (Williams and Norris, 2001).

### Development of classification models

Before constructing the individual PCA for each harvest time, and carrying out the classification process with the SIMCA model, different mathematical transformations were tested in order to maximise differences among spectra of grapes from different harvest times. After each

transformation, a PCA with all samples (from all harvests) was performed and mean scores of the first principal components (PC1) grouped for harvest time were separated using Tukey's test.

From the ANOVA results on PC1 score obtained after different pre-treatment methods, it was observed that the second derivative of Savitzky-Golay showed the poorest performance of separation of the PC1. The noise, moving average smoothing and first derivative of Norris separated the 5 harvest times into 3 groups, while the normalisation, MSC, baseline and first derivative of Savitzky-Golay put the PC1 score into 4 groups (*data not shown*). Finally, the second derivative of Norris completed the separation of the PC1 score among harvest dates (Table 2) fulfilling also the ANOVA hypothesis. In Figure 1, the spectrum of the samples is shown before and after the pre-treatments with second derivative of Norris. In Figure 2, the results of the PCA are shown. Therefore, after transformation by the second derivative of Norris, individual PCA for samples of each harvest were saved and used to classify internal data in full crossvalidation (Figure 2). Despite the difference in PC1 score, the PCA plot shows that it was not possible to discriminate completely between grapes from HT 1, 2 and 5.

Table 3 shows the performance of SIMCA in classifying grapes from different harvest times. Results are shown in the form of a matrix in which numbers of correctly classified harvest times are shown on the diagonal.

It was found that SIMCA could not effectively classify all samples according to harvest time. The percentage of correctly classified samples was over 75% at all harvest times except for HT 2 (59%).

In order to try to improve the classification performance, 2 different ranges of the spectrum were individually tested: 400-775 and 780-980. This left out only the very first and the last part of the spectra. This step

**Table 2. Analysis of variance of first principal component mean scores for spectra of *Italia* table grapes transformed with the second derivative of Norris.**

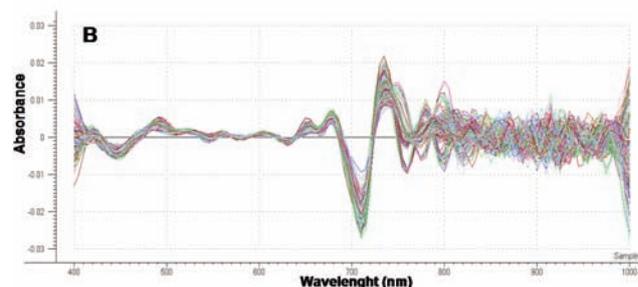
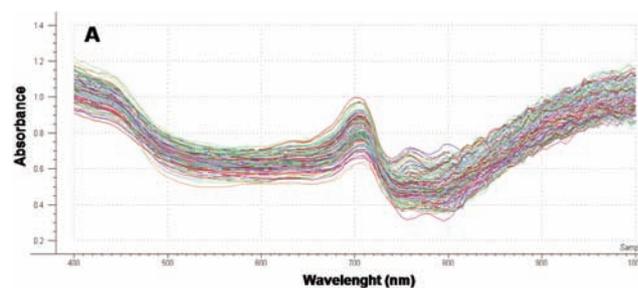
Factor	PC1
HT 1	0.009 <sup>b</sup>
HT 2	-0.004 <sup>c</sup>
HT 3	0.021 <sup>a</sup>
HT 4	-0.017 <sup>e</sup>
HT 5	-0.010 <sup>d</sup>

PC1, first principal component; HT, harvest times. <sup>a,b,c,d,e</sup>Mean values with different letters show significant statistical difference according to Tukey's test ( $P < 0.05$ ).

**Table 3. Classification results of soft independent modelling of class analogy analysis applied to discriminate between samples of *Italia* grapes from different harvest times using the whole spectra interval (400-1000 nm).**

Qualitative groups	HT 1	HT 2	HT 3	HT 4	HT 5
HT 1	33	-	-	-	-
HT 2	-	20	-	-	-
HT 3	-	-	36	-	-
HT 4	-	-	-	26	-
HT 5	-	-	-	-	27
% correctly classified	94	59	100	79	75

HT, harvest times.



**Figure 1. Spectra plot of *Italia* grapes before (A) and after (B) second derivative of Norris on the entire 400-1000 nm interval.**

could, in fact, clarify whether the part of the spectrum containing the peaks at 710 and 735 contained more information than the second region of the spectrum, and if a part of the spectra not affected by the harvest time was inducing the low performance of the classification.

Table 4 shows the percentage of correctly classified samples obtained using the 2 individual ranges. In particular, using the spectral range from 400 to 775 the percentage of correctly classified samples increased only for HT 4 (from 79% to 82%), and decreased from 59% to 38% in HT 2 and from 75% to 0% in HT 5, while there was no change for HT 1 and 3. When using the spectral range from 780 to 790, the percentage of correctly classified samples increased for all classes except for HT 2 and 3. However, the model performance was generally improved. Classification increased from 94% to 100% in the HT 1 group, from 79% to 88% in HT 4 and from 75% to 83% in HT 5. In HT 2, 53% of the samples were correctly classified compared to 59% of the previous

model, while HT 3 correctly classified 92% of the samples compared to 100% of the previous model. Finally, this model was tested using only 117 samples in the calibration set and 57 samples for external validation. In this case, there was a slight improvement in classification performance. Table 5 shows the percentage of correctly classified samples obtained from spectral range between 780-980 nm using external data. In this case, the ability of the model to correctly classify samples of HT 2 increased from 53% to 73%, decreasing from 97% to 92% for HT 3, and from 88% and 83% to 82% for samples of HT 4 and 5, respectively. These results highlight the fact that the region between 780-980 nm of the spectra contained the highest variability related to the time of harvest. This was probably because it contained the band of sugars (840 nm), as reported by Baiano *et al.* (2012), and the band of water (980 nm), as reported by Cynkar *et al.* (2009) in white grapes.

Table 6 shows the classification performance of the PLS-DA method.

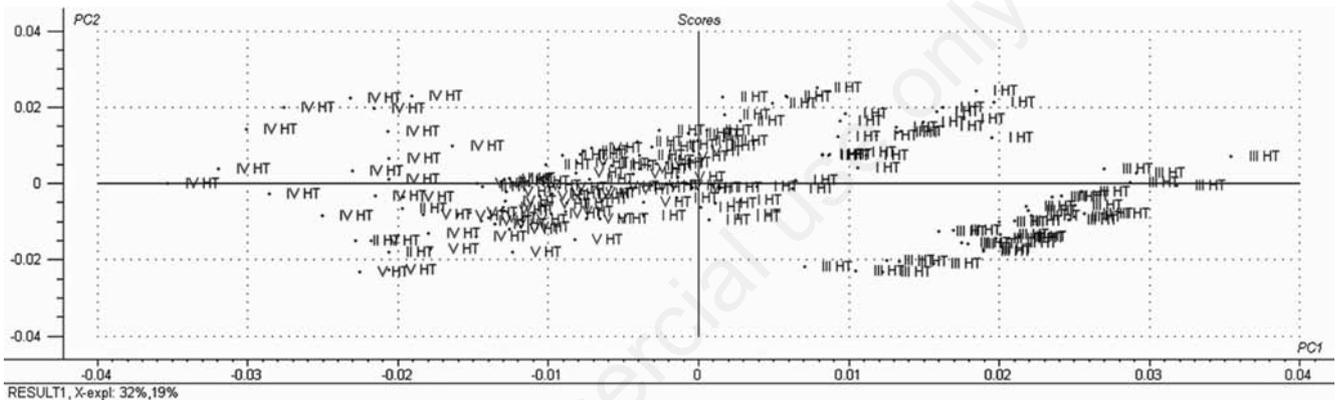


Figure 2. Principal component analysis result for spectra of *Italia* grapes transformed with the second derivative of Norris.

Table 4. Classification results by soft independent modelling of class analogy analysis applied to discriminate between samples of *Italia* grapes from different harvest times using the 2 different ranges of the spectrum (400-775 and 780-980 nm).

Qualitative groups	Spectral range 400-775 nm				
	HT 1	HT 2	HT 3	HT 4	HT 5
HT 1	32	-	-	-	-
HT 2	-	10	-	-	-
HT 3	-	-	36	-	-
HT 4	-	-	-	18	-
HT 5	-	-	-	-	0
% correctly classified	94	38	100	82	0
Qualitative groups	Spectral range 780-980 nm				
	HT 1	HT 2	HT 3	HT 4	HT 5
HT 1	35	-	-	-	-
HT 2	-	18	-	-	-
HT 3	-	-	35	-	-
HT 4	-	-	-	28	-
HT 5	-	-	-	-	30
% correctly classified	100	53	97	88	83

HT, harvest times.

**Table 5. Results of classification by soft independent modelling of class analogy analysis applied to discriminate between external samples of *Italia* grapes from different harvest times in the 780-980 nm spectral range.**

Qualitative groups	External samples				
	HT 1	HT 2	HT 3	HT 4	HT 5
HT 1	12	-	-	-	-
HT 2	-	8	-	-	-
HT 3	-	-	11	-	-
HT 4	-	-	-	9	-
HT 5	-	-	-	-	9
% correctly classified	100	73	92	82	82

HT, harvest times.

**Table 6. Classification results by partial least squares discriminant analysis at different harvest times for *Italia* grapes.**

Qualitative groups	External samples				
	HT 1	HT 2	HT 3	HT 4	HT 5
HT 1	35	-	-	-	-
HT 2	-	34	-	-	-
HT 3	-	-	36	-	-
HT 4	-	-	-	33	-
HT 5	-	-	-	-	34
% correctly classified	100	100	100	100	94

HT, harvest times.

Results achieved with this model may be considered excellent, achieving 100% correct classification on 4 classes out of 5, and 94% of corrected classified samples in HT 5.

Spectral data provided information on differences between grapes from different harvest dates which were not strictly dependent on the maturity index and that could not be revealed by composition at harvest. This indicates that the spectra of the fruits significantly changed with harvest time. Higher performance of PLS-DA *versus* SIMCA was reported by Pholpho *et al.* (2011), discriminating between bruised and non-bruised longan fruits, and by González-Caballero *et al.* (2012) who evaluated the ability of NIR spectroscopy to distinguish between different ripening stages of wine-grapes.

## Conclusions

The overall results demonstrate that spectral information obtained by hyperspectral imaging of the whole berries has excellent potential to evaluate grapes as a function of time of harvest. In particular, the classification of harvest time using the PLS-DA model provided extremely satisfying results, being able to correctly classify almost all samples. It may be interesting in further studies to apply this algorithm for: i) real-time classification of the table grapes in the vineyard or on a sorting line; ii) to further develop a method to identify the optimal harvest time; or iii) to discriminate between table grapes to define length of storage times.

This approach represents a fast, clean method to allow producers to improve table grape quality in the whole chain. However, the cost-benefit ratio will need to be evaluated.

## 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. *Comput. Electron. Agric.* 87:142-51.
- Brand-Williams W., Cuvelier M.E., Berset C. 1995. Use of a free radical method to evaluate antioxidant activity. *Food Sci. Technol. Lab.* 28:25-30.
- Cao F., Wu D., He Y. 2010. Soluble solids content and pH prediction and varieties discrimination of grapes based on visible-near infrared spectroscopy. *Comput. Electron. Agric.* 71S:S15-8.
- Cen H., Yong H., Min H. 2007. Combination and comparison of multivariate analysis for the identification of orange varieties using visible and near infrared reflectance spectroscopy. *Eur. Food Res. Technol.* 6:699-705.
- Cliff M.A., Dever M.C., Reynolds A.G. 1996. Descriptive profiling of new and commercial British Columbia table grape cultivars. *Am. J. Enol. Vitic.* 47:301-8.
- Coombe B.G. 1992. Research on development and ripening of the grape berry. *Am. J. Enol. Viticult.* 43:101-10.
- Costa G., Noferini M., Fiori G., Torrigiani P. 2009. Use of Vis/NIR spectroscopy to assess fruit ripening stage and improve management in post-harvest chain. *GBS J. Fresh Produce.* 3:35-41.
- Cynkar W., Cozzolino D., Damberg R.G. 2009. The effect of sample storage and homogenisation techniques on the chemical composition and near infrared spectra of white grapes. *Food Res. Int.* 42:653-8.
- Dokoozlian N., Luvisi D., Moriyama M., Schrader P. 1995. Cultural practices improve color, size of 'Crimson Seedless'. *Calif. Agr.* 49:36-40.
- Fernández-Navales J., López M.I., Sánchez M.T., Morales J., González-Caballero, V. 2009. Shortwave-near infrared spectroscopy for determination of reducing sugar content during grape ripening, wine-

- making, and aging of white and red wines. *Food Res. Int.* 42:285-291.
- González-Caballero V., Sánchez M.T., Fernández-Navales J., López M.I., Pérez-Marín D. 2012. On-vine monitoring of grape ripening using near-infrared spectroscopy. *Food Anal. Methods.* 5:1377-85.
- Guidetti R., Beghi R., Bodria L. 2010. Evaluation of grape quality parameters by a simple Vis/NIR system. *Trans. ASABE.* 53:477-84.
- Heise H.M., Winzen R. 2002. Chemometrics in near-infrared spectroscopy. In: H.W. Siesler, Y. Ozaki, S. Kawata, H.M. Heise (eds.), *Near-infrared spectroscopy. Principles, instruments, applications.* Wiley-VCH, Weinheim, Germany, pp 125-62.
- Jayasena V., Cameron I. 2008. Brix/acid ratio as a predictor of consumer acceptability of Crimson seedless table grapes. *J. Food Qual.* 31:736-50.
- Kemps B., Leon L., Best S., De Baerdemaeker J., De Ketelaere B. 2010. Assessment of the quality parameters in grapes using VIS/NIR spectroscopy. *Biosyst. Engine.* 105:507-13.
- Kramer R., Workman J., Reeves J.B. III. 2004. Qualitative analysis. In: C.A. Roberts, J. Workman, J.B. Reeves III (eds.), *Near-infrared spectroscopy in agriculture.* American Society of Agronomy, Crop Science Society of America and Soil Science Society of America, Madison, WI, USA, pp 175-206.
- Liu L., Cozzolino D., Cynkar W.U., Damberg R.G., Janik L., O'Neill B.K., Colby C.B., Gishen M. 2007. Preliminary study on the application of visible-near infrared spectroscopy and chemometrics to classify Riesling wines from different countries. *Food Chem.* 106:781-6.
- Massart D.L., Vandeginste B.G.M., Deming S.M., Michotte Y., Kaufman L. 1998. *Chemometrics: a textbook. Data handling in science and technology 2.* Elsevier Science, Amsterdam, The Netherlands.
- Menesatti P., Zanella A.S., D'Andrea S., Costa C., Paglia G., Pallottino F. 2009. Supervised multivariate analysis of hyper-spectral NIR images to evaluate the starch index of apples. *Food Bioproc. Technol.* 2:308-14.
- Musumarra G., Barresi V., Condorelli D.F., Fortuna C.G., Scire S. 2005. Genome-based identification of diagnostic molecular markers for human lung carcinomas by PLS-DA. *Comput. Biol. Chem.* 29:183-95.
- Naes T., Indahl U. 1998. A unified description of classical classification methods for multicollinear data. *J. Chemometr.* 12:200-5.
- Naes T., Isaksson T., Fearn T., Davies A. 2002. *A user-friendly guide to multivariate calibration and classification.* NIR Publications, Chichester, UK.
- OIV (International Organisation of Vine and Wine). 2012. *Statistical Report on world vitiviniculture.* Available from: <http://www.oiv.int/oiv/cms/index?rubricId=44538631-7ad2-49cb-9710-ad5b957296c7>
- Pérez-Marín D. 2005. *Tecnología NIRS para la Certificación y Trazabilidad de Piensos Compuestos.* Tesis Doctoral, Universidad de Córdoba, España.
- Polpho T., Pathaveerat S., Sirisomboon P. 2011. Classification of longan fruit bruising using visible spectroscopy. *J. Food Engine.* 104:169-72.
- Singleton S.L., Rossi J.A. 1965. Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. *Am. J. Enol. Vitic.* 16:144-58.
- Tavakolian M.S.S., Silaghi F.A., Fabbri A., Molari G., Giunchi A., Guarnieri A. 2013. Differentiation of post harvest date fruit varieties non-destructively using FT-NIR spectroscopy. *Int. J. Food Sci. Technol.* 48:1282-8.
- Vargas A., Pérez J., Zoffoli J.P., Pérez A. 2001. Comparación de variables de textura en la medición de firmeza de bayas de uva Thompson seedless. *Cienc. Investig. Agrar.* 28:37-42.
- Vandeginste B.G.M., Massart D.L., Buydens L.M.C., De Jong S., Lewi P.J., Smeyers-Verbeke J. 1998. *Handbook of chemometrics and qualimetrics: part B.* Elsevier Science B.V., Amsterdam, The Netherlands, pp 649-99.
- Vigneau E., Qannari E.M., Devaux M.F. 2000. Méthodes prédictives. In: D. Bertrand, E. Dufour (eds.), *La spectroscopie infrarouge et ses applications analytiques.* Technique & Documentation, Paris, France, pp 295-332.
- Watson B. 2003. Evaluation of wine grape maturity. In: E.W. Hellman (ed.), *Oregon viticulture.* Oregon State University Press, Corvallis, OR, USA, pp 235-45.
- Williams P.C., Norris K. 2001. Variables affecting near-infrared spectroscopic analysis. In: Williams P. and Norris K. (eds.), *Near-infrared technology in the agricultural and food industries,* 2nd edition. American Association of Cereal Chemists, Inc., St. Paul, MN, USA, pp 171-85.
- Wold S., Sjostrom M. 1977. SIMCA: A method for analyzing chemical data in terms of similarity and analogy. In: B.R. Kowalski (ed.), *Chemometrics theory and application,* American Chemical Society Symposium - Series 52, Washington, DC, USA, pp 243-82.
- Zeppa G., Rolle L., Gerbi V. 1999. Valutazione mediante consumer test dell'attitudine al consumo diretto di un'uva a bacca rossa. *Industrie Alimentari* 38:818-24.