Reprint of: In-field and non-destructive monitoring of grapes maturity by hyperspectral imaging

In-field and non-destructive monitoring of grapes maturity by hyperspectral imaging

Alessandro Benelli[a], Chiara Cevoli[a,b*], Luigi Ragni[a,b], Angelo Fabbri[a,b]

[a]*Department of Agricultural and Food Sciences, Alma Mater Studiorum, University of Bologna, Piazza Goidanich 60, Cesena, 47521, Italy*

[b]*Interdepartmental Centre for Agri-Food Industrial Research, Alma Mater Studiorum, University of Bologna, Via Quinto Bucci 336, Cesena, 47521, Italy*

[*]*Corresponding author: chiara.cevoli3@unibo.it*

Abstract

Monitoring the quality attributes of grapes is a practice that allows the state of ripeness to be checked and the optimal harvest time to be identified. A non-destructive method based on hyperspectral imaging (HSI) technology was developed. Analyses were carried out directly in the field on a 'Sangiovese' (*Vitis vinifera* L.) vineyard destined for wine production, by using a Vis/NIR (400–1000 nm) hyperspectral camera. One vineyard row was analysed on 13 different days during the pre-harvest and harvest time. The soluble solids content (SSC) expressed in terms of °Brix was measured by a portable digital refractometer. Afterwards, the grape samples were split in two classes: the first one composed by the samples characterised by a °Brix lower than 20 (not-ripe), while the second one by the samples with a °Brix higher than 20 (ripe). Grape mean spectra were extracted from each hyperspectral image and used to predict the SSC by partial least squares regression (PLS), and to classify the samples into the two classes by PLS discriminant analysis (PLS-DA). SSC was predicted with a R^2=0.77 (RMSECV=0.79 °Brix), and the samples were correctly classified with a percentage from 86 to 91%. Even if the number of wavelengths was limited, the percentages of correctly classified samples were again within the above-mentioned range. The present study shows the

potential of the use of HSI technology directly in the field by proximal measurements under natural light conditions for the prediction of the harvest time of the 'Sangiovese' red grape.

Keywords: hyperspectral, in-field, grape, wine, harvest, classification

1. Introduction

Italy is the largest wine producer in the world, with a production in 2018 of 5,480 million litres (18.8% of world production) (OIV International Organisation of Vine and Wine, 2019). In an industrialised wine growing system, monitoring the quality attributes, such as soluble solids content (SSC), acidity and anthocyanin content of grapes is extremely important: well-planned monitoring allows to check the growth and ripening of the grapes, and finally to decide when to proceed with the harvest (Delrot, Medrano, Or, Bavaresco, Grando, 2010). For instance, by managing irrigation through the use of techniques such as the regulated deficit irrigation, significant increases in SSC and anthocyanins can be achieved which together with a decrease in yield and berry size can lead to substantial improvements in grape quality (Acevedo-Opazo, Ortega-Farias, Fuentes, 2010; Pellegrino, Lebon, Simonneau, Wery, 2005).

Monitoring of grape quality attributes can be carried out directly in the field using traditional destructive techniques. Alternatively, quality attributes can be estimated by non-destructive techniques, such as near-infrared (NIR) spectroscopy. Portable NIR instruments were used to determine the following quality attributes of grapes: water content, SSC, reductant sugars, pH, titratable acidity, maturity index (sugar/acidity ratio), extractable anthocyanins, potential anthocyanins (Teixeira Dos Santos, Lopo, Páscoa, Lopes, 2013).

Over the last two decades the use of HSI technology in the quality assessment of fruits and vegetables has become of increasing interest (Chandrasekaran, Panigrahi, Ravikanth, Singh, 2019; Liu, Zeng, Sun, 2015). HSI was initially limited to controlled environments such as laboratories but gradually,

thanks to the miniaturisation and improved computing and data storage capabilities, it began to be used directly in the field (Benelli, Cevoli, Fabbri, 2020). Hyperspectral images can be captured either remotely by airborne vehicles and unmanned aerial vehicles (Ishida et al., 2018; Matese & Di Gennaro, 2015; Zarco-Tejada et al., 2013) or by ground vehicles (Deery, Jimenez-Berni, Jones, Sirault, Furbank, 2014; Gutiérrez, Fernández-Novales, Diago, Tardaguila, 2018a; Huang, Lee, Thomson, Reddy, 2016; Jay et al., 2017; Underwood, Wendel, Schofield, McMurray, Kimber, 2017; Wendel, Underwood, Walsh, 2018; Whetton, Waine, Mouazen, 2018), which produces proximal hyperspectral images with high spatial resolution. Proximal HSI could therefore allow non-destructive, contactless, and automated monitoring of grape quality attributes. Moreover, the acquisition of hyperspectral images can be performed continuously, enabling the rapid scanning of large areas. Hyperspectral data is characterised by two spatial and one spectral dimension, therefore only specific regions of interest (ROIs) can be selected, and the residual regions can be excluded. In this sense, hyperspectral analysis is useful to guide the choice of true multispectral less expensive solutions.

Concerning in-field grape studies, Gutiérrez et al. (2018a) adopted an on-the-go HSI system for the classification of 30 grapevine varietals directly acquired in the field. Through the development of classification models based on support vector machines (SVM) and multilayer perceptron (MLP), prediction performance (F1 score) up to 0.99 was achieved. The same on-the-go HSI system described above, combined with SVM, was adopted to estimate SSC and anthocyanin concentration of wine grapes (Gutiérrez et al., 2018b). Determination coefficients (R^2) of 0.92 (RMSE=1.274 °Brix) and 0.83 (RMSE=0.211 mg g^{-1}) were obtained for the prediction of SSC and anthocyanin concentration, respectively. These two studies highlight the potential of HSI to monitor the indices of grape ripening directly in the field and therefore to improve vineyard decisions and management. Non-linear statistical methods were used in both studies. Considering the results just mentioned, it would be interesting to investigate whether, even combining HSI with linear methods and reducing the number of wavelengths, it is possible to monitor the degree of grape maturity directly in the field.

Furthermore, considering a possible application of this technique, a simple binary model able to simply determine whether grapes are ripe for harvesting could be interesting.

Thus, the present study aims to identify the proper degree of ripeness, suitable for harvesting wine grapes, through the observation of the SSC evolution by means of HSI technology applied directly in the field. The mean spectra were extracted from each hyperspectral image and used to predict the SSC and to classify the samples into the two classes (not-ripe and ripe), by partial least squares regression (PLS) and discriminant analysis (PLS-DA), respectively.

2. Materials and methods

2.1 Samples

One side of a row of 'Sangiovese' (*Vitis vinifera* L.) grape vineyards, located near Cesena, Italy, was analysed on 13 different days in the period between August 20th and October 4th, 2019 (from pre-harvest to harvest time). The row was divided into 11 sections; from each section 3, grapes were taken for each day of analysis, for a total of 429 samples.

2.2 Hyperspectral acquisitions

The adopted push-broom hyperspectral camera (Nano-Hyperspec VNIR, Headwall Photonics, Inc., Fitchburg, MA, USA) scans single lines in a sequence, each one consisting of 640 voxels: the image is created by moving the camera along the scanning direction (Fig. 1a). Each voxel contains, in addition to the two spatial dimensions, a Vis/NIR spectrum (400–1000 nm) characterized by 272 spectral bands, with a nominal spectral resolution of 2.2 nm. The mounted lens has an effective focal length of 17 mm, with the optical axis perpendicular to the side of the vineyard row (scanned surface) analysed (Figure 1b). The camera was installed on a garden cart (Fig. 1b) 120 cm above the ground and it was powered by a 12 V, 45 Ah automotive battery through a DC to AC power inverter. The scans were performed at about 1.6 m from the side of the vineyard row.

Direct sunlight with clear sky conditions, was used as a light source. To reduce the fluctuation of the sample temperature, all the acquisitions were carried during the same period of the day, from 10:30 a.m. to 12:00 p.m.

The frame rate was set to approximately 100 frames s^{-1}. The exposure time was set from 6 to 8 ms, depending on the light intensity, and was achieved through calibration, by framing a white high-reflectance matter panel, placed at the same distance as the vineyard row, to cover the entire angle of view of the camera. Given the clear sky conditions and the short time required, about 10 min, acquisition of hyperspectral images of the 11 vineyard row sections and calibration was carried out only once per day.

The raw diffuse reflectance spectrum (R_R) was extracted from the HS images. The calibrated diffuse reflection spectrum (R_C) was calculated by applying the following equation:(Guo et al., 2019):

$$R_C = \frac{R_R - R_D}{R_W - R_D} \quad (1)$$

where R_D is the reflectance spectrum of dark reference, obtained by applying the cap on the lens; R_W is the reflectance spectrum of white reference, obtained by means calibration with the white high-reflectance matte panel reported above.

A hyperspectral image from each of the 11 sections obtained from the vineyard row, was acquired per day of analysis (Fig. 1c): therefore, during the 13 d of analysis, a total of 143 vineyard row sections were scanned.

2.3 Soluble solids content measurement

After the acquisition of the images, the SSC, expressed in °Brix, was measured on 3 grape berries (randomly selected) from each of the 11 sections by using a portable digital refractometer (PR-101 Digital Refractometer, ATAGO CO., LTD, Tokyo, Japan). Subsequently, the mean for each section was calculated.

One-way analysis of variance (ANOVA) with Tukey-HSD post-hoc test (p-level < 0.05) was applied to evaluate significant differences between SSC means over the different days of analysis.

130 *2.4 Hyperspectral images elaboration*

131 ROI selection was made using the software HyperCube, v. 11.52 (U.S. Army Engineer Research and
132 Development Center (ERDC), USA). For each image, 5 points and a maximum of a further 120
133 adjacent points (11 x 11 voxels matrix, with the selected point in the centre of the matrix) were
134 selected on 5 different berries directly illuminated by the sun, not in the shade. The (reference) points
135 included in the classification were characterized by a metric distance from the mean (signature)
136 spectrum of the manually selected points within the range (tolerance threshold) [0,0.04] (Fig. 2). The
137 metric distance was calculated by applying the absolute difference (Manhattan) function (Eq. (2))
138 (Deborah, Richard, Hardeberg, 2015) and normalised in the range [0,1].

139 $$d(R_1, R_2) = \sum_{\lambda} \left| R_{1,\lambda} - R \right| \quad (2)$$

140 where R_1, R_2 are two reflectance spectra.

141 From the spectra of the classified points, the mean spectrum for each hyperspectral image was
142 calculated.

143 The spectral bands between 400–424 nm were omitted as a result of the low signal-to-noise ratio
144 produced by the sensor, as reported in Wendel et al. (2018). The spectra were smoothed (Savitzky-
145 Golay method; polynomial order: 2; smoothing points: 15) to reduce noise from the spectra and
146 following pre-treated by the standard normal variate (SNV) method, first derivative (D1) and finally
147 mean centred (MC). The SNV is one of the most common pre-processing method used to correct
148 spectra for changes in optical path length and light scattering, while the derivatives have the capability
149 to remove both additive and multiplicative effects in the spectra (Rinnan, van den Berg, Engelsen,
150 2009). After SNV, each spectrum will have a mean of 0 and a standard deviation of 1.

151 Principal component analysis (PCA) was applied to the mean spectra as exploratory technique to
152 visualize the data according to °Brix and time evolution. Subsequently, a preliminary PLS regression
153 model was built to estimate the SSC. The validation was carried out by the venetian blind cross-
154 validation method (segments: 10).

PLS-DA models were built to classify the samples according to °Brix. In particular, classification models with 2 categories (not-ripe and ripe) were developed: according to Bucelli, Costantini and Storchi (2010), the first class was composed by the samples characterized by a °Brix lower than 20 (0), while the second one by the samples with a °Brix equal or higher than 20 (1).

The sample dataset (n = 143) was split in calibration (venetian blinds cross-validation, including 75% of the samples) and external validation set (25% of the samples) by using the Onion method (Gallagher et al., 2004). The threshold value, able to identify the belonging category of each sample into one of the groups, was defined by using a probabilistic approach based on Bayes's rule. To avoid the model over-fitting, the optimal number of latent variables were chosen by plotting the root mean square error of cross-validation (RMSECV) as a function of the number of components and by identifying where the curve reaches a local minimum. The receiver operating characteristic (ROC) curves in prediction were evaluated to assess the goodness of the models.

All the chemometrics models were developed by using PLS Toolbox for Matlab2018a.

3. Results and discussion

Means and standard deviation of the °Brix measured during the 13 d of analysis are reported in Table 1. The increase between the first and last day was of 27.5%, from 17.8 °Brix (day I) to 22.7 °Brix (day XIII). Several significant differences between the °Brix mean values were achieved over the different days of analysis. The whole data set was characterized by a mean value of 20.6 ± 1.7 °Brix, which make suitable to split it in two subsets with a threshold of 20 °Brix.

Raw and pre-treated (smoothing and SNV) mean spectra of all the samples by day of analysis are presented in Fig. 3. In the Vis/NIR region (400–1000 nm), the visible spectrum (400–700 nm) presents the absorption bands of some substances used as ripening indexes of fruit: anthocyanins at around 500 nm, carotenoids at 570–590 nm, and chlorophyll *a* at 680–710 nm (ElMasry, Wang, ElSayed, Ngadi, 2007; Munera et al., 2017). In the NIR region (700–1000 nm), absorption bands of water at 760 nm and 960–970 nm are characterised by the overtone of O-H bonds (McGlone &

Kawano, 1998; Nicolaï et al., 2007): since the water content of a ripe wine grape is 70–80% (FAO Food and Agriculture Organization, 2009), it can be expected that the water related absorption band will prevail. Absorption band around 840 nm was associated with sugar (Pu, Liu, Wang, Sun, 2016); moreover, peaks observed in the 950–1000 nm region were related to both water and carbohydrates, as the second overtone of O-H and N-H, a combination band of O-H bonds and the third overtone of C-H, were found in the region (Camps & Christen, 2009). As observed, the water absorption peaks in the NIR (700–1000 nm) spectral region are not very marked and wide. Therefore, spectral information from SSC in the 800–1000 nm range will tend to be less covered by water (Camps & Christen, 2009; Manley, Joubert, Myburgh, Kidd, 2007).

The score plot of the first two principal components (PC1: 59%; PC2: 28%) resulting from PCA shows the samples distribution according to the day of analysis. A tendency to place the samples from left to right can be observed on PC1, starting from day I (quadrant III) to day XIII (quadrant IV) (Figure 4a).

A similar distribution can be observed according to the SSC: the samples are distributed along PC1, mainly on the left those with SSC < 20 °Brix (quadrants II and III), on the right those with SSC ⩾ 22 °Brix (quadrants I and IV) (Fig. 4b).

The best PLS results were obtained pre-treating the spectra by SNV + MC. Particularly, $R^2 = 0.768$ and RMSECV = 0.79 °Brix were achieved in cross-validation with 7 latent variables. Figure 5 shows the predicted versus measured °Brix values.

These results agree with those present in literature and developed by using spectra in the same wavelength range and acquired directly in-field. Furthermore, regardless of the R^2 values, the RMSE is substantially lower. Diezma-Iglesias, Barreiro, Blanco and García-Ramos (2008) predicted SSC on 480 samples of 'Cabernet Sauvignon' by using a portable hand-held spectrometer (590–1090 nm), reporting a $R^2 = 0.72$, while Guidetti, Beghi and Bodria (2010) working with a simple Vis/NIR system in the range from 400 to 1000 nm, obtained in prediction r=0.82 (R^2=0.67) and RMSEP=1.48°Brix. Gutiérrez et al. (2018b) used the HSI technique (400–1000 nm) to measure SSC in wine grapes in

real time. R^2 of 0.91 (RMSE = 1.358 °Brix) and 0.92 (RMSE = 1.274 °Brix) were achieved in cross validation and prediction, respectively, by using the SVM techniques.

PLS-DA results, in terms of percentage of correctly classified samples, are reported in Table 2. The percentages ranged from 86 to 91%. Considering the prediction set, the best result was obtained applying as pre-treatment the SNV + MC. The results obtained with the hyperspectral technique improved those achieved by Guidetti et al. (2010) by using of a portable contact system. The authors combined Vis/NIR spectroscopy in the wavelength range 450 – 980 nm with PLS-DA to classify grape samples in two groups based on SSC (threshold = 21 °Brix), obtaining a percentage of samples correctly classified (in prediction) of 77.1% (Guidetti et al., 2010).

The ROC curves in prediction (Fig. 6) summarise the trade-off between specificity (number of samples predicted to not be in the class divided by the actual number not in the class) and sensitivity (number of samples predicted to be in the class divided by number actually in the class) for the PLS-DA classification models. The area under the curves (AUC = 0.9855 for SNV + MC and AUC = 0.9578 for SNV + D1 + MC) suggest that the models were characterised by a high degree of discrimination, confirming that the best model was those developed considering only the SNV + MC as pre-treatment.

Results, in terms of probability (Bayes's rule) of belonging to the class °Brix < 20, are shown in Fig. 7. The higher a sample is placed, the higher the probability that it will be classified as a member of the °Brix < 20 class. Consequently, samples classified as members of the other class (°Brix $\geqslant$ 20) are placed at the bottom of the graph. The threshold value (dotted red line) was set at 0.5 (probability of 50%); samples with a probability lower than this value are considered improperly classified.

Considering the SNV + MC and SNV + D1 + MC pre-treatments, 89% and 77% of the samples (in prediction) were classified with a probability higher than 70%.

Figure 8 shows the VIP (variable importance in projection) scores obtained by the PLS-DA models. These scores estimate the importance of each variable in the projection used in the PLS-DA model. A variable with a VIP score close to or higher than 1 can be considered important in a given model.

For both the pre-treatments, similar regions with VIP score higher than 1 were obtained, suggesting that the wavelengths with the highest contribution are in the NIR region of the spectrum (from 700 to 1000 nm). Consequently, the variable selection method based on the VIP scores higher than 1 was used to reduce the original data set.

PLS-DA results (percentage of correctly classified samples) obtained by using the reduced number of variables (wavelengths), are reported in Table 3. In particular, 93 (SNV + MC) and 88 (SNV + D1 + MC) x-variables were used to develop the new PLS-DA. The percentages of correctly classified samples are slightly lower (from 86 to 91%) than those obtained considering the whole spectrum. However, the results are still completely acceptable. This confirms also by the high AUC values (0.939 and 0.942).

This work presents a solution for in-field and non-destructively determination of grape maturity degree by using HSI combined with linear chemometric techniques. Particularly, the results confirm the suitability for estimating the soluble solid content of red grapes.

The positioning of the vineyard row did not create any obstacles regarding the direct solar lighting of the grape bunches. The images along the vineyard row were acquired in the NNE-SSO direction, with the row on the right-hand side of the scanning direction. Consequently, the suitable period for analysis was during the late morning and not later than midday. To scan the other side of the row, it would have been necessary to proceed after midday. Shadows also did not present any problems, as the sun was behind the camera at the time of the acquisitions. To correctly acquire hyperspectral images, the vines should be stripped to ensure that the upper leaves do not overshadow the grapes below. Grapes often had reflection and shaded areas, which need to be excluded during the selection of the ROIs, along with fully or partially shaded grape bunches. In the presence of clouds or even just a slight cloud cover, a significant variation in light intensity was observed: this means that the camera would need to be recalibrated every time a section of the row was scanned. In addition, it is possible that the brightness conditions change soon after calibration, so it would be necessary to recalibrate and immediately proceed with the scan of the row section.

The variability grape of a vineyards, in terms of SCC, often is quite high. This depends on many factors, such us the vineyards orientation, unevenness of the land and meteorological phenomena. For this reason, the grapes are harvested at different times even along the same row. Consequently, to optimise the harvest, a technique that allows to have a mapping of the SCC for all grapes, would certainly be an advantage.

4. Conclusions

Hyperspectral imaging technology, usually adopted in laboratories with auxiliary artificial lighting, was used in-field under natural lighting conditions to monitor the maturity degree of 'Sangiovese' (*Vitis vinifera* L.) grapes. The results achieved confirm that it is possible to predict the soluble solid content and to classify grape samples into two classes (not-ripe and ripe) using a linear technique to elaborate the spectral data. Furthermore, the classification performance remained substantially unchanged by reducing the number of wavelengths, so it is expected that a less expensive multispectral camera in the 400–1000 nm range can work just as well. The implementation of a hyperspectral imaging system on an agricultural vehicle coupled to a gimbal stabilisation system, together with the development of hyperspectral image segmentation techniques, would allow on-the-go analysis of large vineyard extensions. Attention should be paid to the presence of water on the surface of the sample under analysis, to the presence of variable cloudiness and, if the leaves are analysed, to the presence of wind.

Declarations of interest: none.

References

Acevedo-Opazo, C., Ortega-Farias, S., & Fuentes, S. (2010). Effects of grapevine (Vitis vinifera L.) water status on water consumption, vegetative growth and grape quality: An irrigation scheduling

application to achieve regulated deficit irrigation. *Agricultural Water Management, 97*(7), 956–964. https://doi.org/10.1016/j.agwat.2010.01.025

Benelli, A., Cevoli, C., & Fabbri, A. (2020). In-field hyperspectral imaging: An overview on the ground-based applications in agriculture. *Journal of Agricultural Engineering, LI*(1030), 129–139. https://doi.org/10.4081/jae.2020.1030

Bucelli, P., Costantini, E. A. C., & Storchi, P. (2010). It is possible to predict sangiovese wine quality through a limited number of variables measured on the vines. *Journal International Des Sciences de La Vigne et Du Vin, 44*(4), 207–218. https://doi.org/10.20870/oeno-one.2010.44.4.1473

Camps, C., & Christen, D. (2009). On-tree follow-up of apricot fruit development using a hand-held NIR instrument. *Journal of Food, Agriculture and Environment, 7*(2), 394–400.

Chandrasekaran, I., Panigrahi, S. S., Ravikanth, L., & Singh, C. B. (2019). Potential of Near-Infrared (NIR) Spectroscopy and Hyperspectral Imaging for Quality and Safety Assessment of Fruits: an Overview. *Food Analytical Methods, 12*(11), 2438–2458. https://doi.org/10.1007/s12161-019-01609-1

Deborah, H., Richard, N., & Hardeberg, J. Y. (2015). A Comprehensive Evaluation of Spectral Distance Functions and Metrics for Hyperspectral Image Processing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8*(6), 3224–3234. https://doi.org/10.1109/JSTARS.2015.2403257

Deery, D., Jimenez-Berni, J., Jones, H., Sirault, X., & Furbank, R. (2014). Proximal remote sensing buggies and potential applications for field-based phenotyping. *Agronomy, 4*(3), 349–379. https://doi.org/10.3390/agronomy4030349

Delrot, S., Medrano, H., Or, E., Bavaresco, L., & Grando, S. (2010). Preface. In *Methodologies and Results in Grapevine Research* (pp. 1–448). Springer Netherlands. https://doi.org/10.1007/978-90-481-9283-0

Diezma-Iglesias, B., Barreiro, P., Blanco, R., & García-Ramos, F. J. (2008). Comparison of robust modeling techniques on NIR spectra used to estimate grape quality. *Acta Horticulturae, 802*, 367–372. https://doi.org/10.17660/ActaHortic.2008.802.48

ElMasry, G., Wang, N., ElSayed, A., & Ngadi, M. (2007). Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry. *Journal of Food Engineering, 81*(1), 98–107. https://doi.org/10.1016/j.jfoodeng.2006.10.016

FAO (Food and Agriculture Organization). (2009). Grapes Wine. In FAO *Agribusiness handbook* (p. 49). http://www.fao.org/3/al176e/al176e.pdf

Gallagher, N. B., Shaver, J. M., Martin, E. B., Morris, J., Wise, B. M., & Windig, W. (2004). Curve resolution for multivariate images with applications to TOF-SIMS and Raman. *Chemometrics and Intelligent Laboratory Systems, 73*(1 SPEC. ISS.), 105–117. https://doi.org/10.1016/j.chemolab.2004.04.003

Guidetti, R., Beghi, R., & Bodria, L. (2010). Evaluation of grape quality parameters by a simple VIS/NIR system. *Transactions of the ASABE, 53*(2), 477–484.

Guo, W., Li, W., Yang, B., Zhu, Z. Z., Liu, D., & Zhu, X. (2019). A novel noninvasive and cost-effective handheld detector on soluble solids content of fruits. *Journal of Food Engineering, 257*(March), 1–9. https://doi.org/10.1016/j.jfoodeng.2019.03.022

Gutiérrez, S., Fernández-Novales, J., Diago, M. P., & Tardaguila, J. (2018a). On-the-go hyperspectral imaging under field conditions and machine learning for the classification of grapevine varieties. *Frontiers in Plant Science, 9*(July), 1–11. https://doi.org/10.3389/fpls.2018.01102

Gutiérrez, S., Tardaguila, J., Fernández-Novales, J., & Diago, M. P. (2018b). On-the-go hyperspectral imaging for the in-field estimation of grape berry soluble solids and anthocyanin concentration. *Australian Journal of Grape and Wine Research, i*, 127–133. https://doi.org/10.1111/ajgw.12376

Huang, Y., Lee, M. A., Thomson, S. J., & Reddy, K. N. (2016). Ground-based hyperspectral remote sensing for weed management in crop production. *International Journal of Agricultural and Biological Engineering, 9*(2), 98–109. https://doi.org/10.3965/j.ijabe.20160902.2137

Ishida, T., Kurihara, J., Viray, F. A., Namuco, S. B., Paringit, E. C., Perez, G. J., Takahashi, Y., & Marciano, J. J. (2018). A novel approach for vegetation classification using UAV-based hyperspectral imaging. *Computers and Electronics in Agriculture, 144*(December 2017), 80–85. https://doi.org/10.1016/j.compag.2017.11.027

Jay, S., Gorretta, N., Morel, J., Maupas, F., Bendoula, R., Rabatel, G., Dutartre, D., Comar, A., & Baret, F. (2017). Estimating leaf chlorophyll content in sugar beet canopies using millimeter- to centimeter-scale reflectance imagery. *Remote Sensing of Environment, 198*, 173–186. https://doi.org/10.1016/j.rse.2017.06.008

Liu, D., Zeng, X. A., & Sun, D. W. (2015). Recent Developments and Applications of Hyperspectral Imaging for Quality Evaluation of Agricultural Products: A Review. *Critical Reviews in Food Science and Nutrition, 55*(12), 1744–1757. https://doi.org/10.1080/10408398.2013.777020

Manley, M., Joubert, E., Myburgh, L., & Kidd, M. (2007). Prediction of soluble solids content and post-storage internal quality of Bulida apricots using near infrared spectroscopy. *Journal of Near Infrared Spectroscopy, 15*(3), 179–188. https://doi.org/10.1255/jnirs.725

Matese, A., & Di Gennaro, S. F. (2015). Technology in precision viticulture: A state of the art review. *International Journal of Wine Research, 7*(1), 69–81. https://doi.org/10.2147/IJWR.S69405

McGlone, V. A., & Kawano, S. (1998). Firmness, dry-matter and soluble-solids assessment of postharvest kiwifruit by NIR spectroscopy. *Postharvest Biology and Technology, 13*(2), 131–141. https://doi.org/10.1016/S0925-5214(98)00007-6

Munera, S., Besada, C., Blasco, J., Cubero, S., Salvador, A., Talens, P., & Aleixos, N. (2017). Astringency assessment of persimmon by hyperspectral imaging. *Postharvest Biology and Technology, 125*, 35–41. https://doi.org/10.1016/j.postharvbio.2016.11.006

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology, 46*(2), 99–118. https://doi.org/10.1016/j.postharvbio.2007.06.024

OIV (International Organisation of Vine and Wine). (2019). *2019 Statistical Report on World Vitiviniculture* (p. 23). https://doi.org/64/19/6835 [pii]\n10.1158/0008-5472.CAN-04-1678

Pellegrino, A., Lebon, E., Simonneau, T., & Wery, J. (2005). Towards a simple indicator of water stress in grapevine (Vitis vinifera L.) based on the differential sensitivities of vegetative growth components. *Australian Journal of Grape and Wine Research, 11*(3), 306–315. https://doi.org/10.1111/j.1755-0238.2005.tb00030.x

Pu, H., Liu, D., Wang, L., & Sun, D. W. (2016). Soluble Solids Content and pH Prediction and Maturity Discrimination of Lychee Fruits Using Visible and Near Infrared Hyperspectral Imaging. *Food Analytical Methods, 9*(1), 235–244. https://doi.org/10.1007/s12161-015-0186-7

Rinnan, Å., van den Berg, F., & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry, 28*(10), 1201–1222. https://doi.org/10.1016/J.TRAC.2009.07.007

Teixeira Dos Santos, C. A., Lopo, M., Páscoa, R. N. M. J., & Lopes, J. A. (2013). A review on the applications of portable near-infrared spectrometers in the agro-food industry. *Applied Spectroscopy, 67*(11), 1215–1233. https://doi.org/10.1366/13-07228

Underwood, J., Wendel, A., Schofield, B., McMurray, L., & Kimber, R. (2017). Efficient in-field plant phenomics for row-crops with an autonomous ground vehicle. *Journal of Field Robotics, 34*(6), 1061–1083. https://doi.org/10.1002/rob.21728

Wendel, A., Underwood, J., & Walsh, K. (2018). Maturity estimation of mangoes using hyperspectral imaging from a ground based mobile platform. *Computers and Electronics in Agriculture, 155*(June), 298–313. https://doi.org/10.1016/j.compag.2018.10.021

Whetton, R. L., Waine, T. W., & Mouazen, A. M. (2018). Hyperspectral measurements of yellow rust and fusarium head blight in cereal crops: Part 2: On-line field measurement. *Biosystems Engineering, 167*, 144–158. https://doi.org/10.1016/j.biosystemseng.2018.01.004

Zarco-Tejada, P. J., Guillén-Climent, M. L., Hernández-Clemente, R., Catalina, A., González, M. R., & Martín, P. (2013). Estimating leaf carotenoid content in vineyards using high resolution hyperspectral imagery acquired from an unmanned aerial vehicle (UAV). *Agricultural and Forest Meteorology, 171–172*, 281–294. https://doi.org/10.1016/j.agrformet.2012.12.013

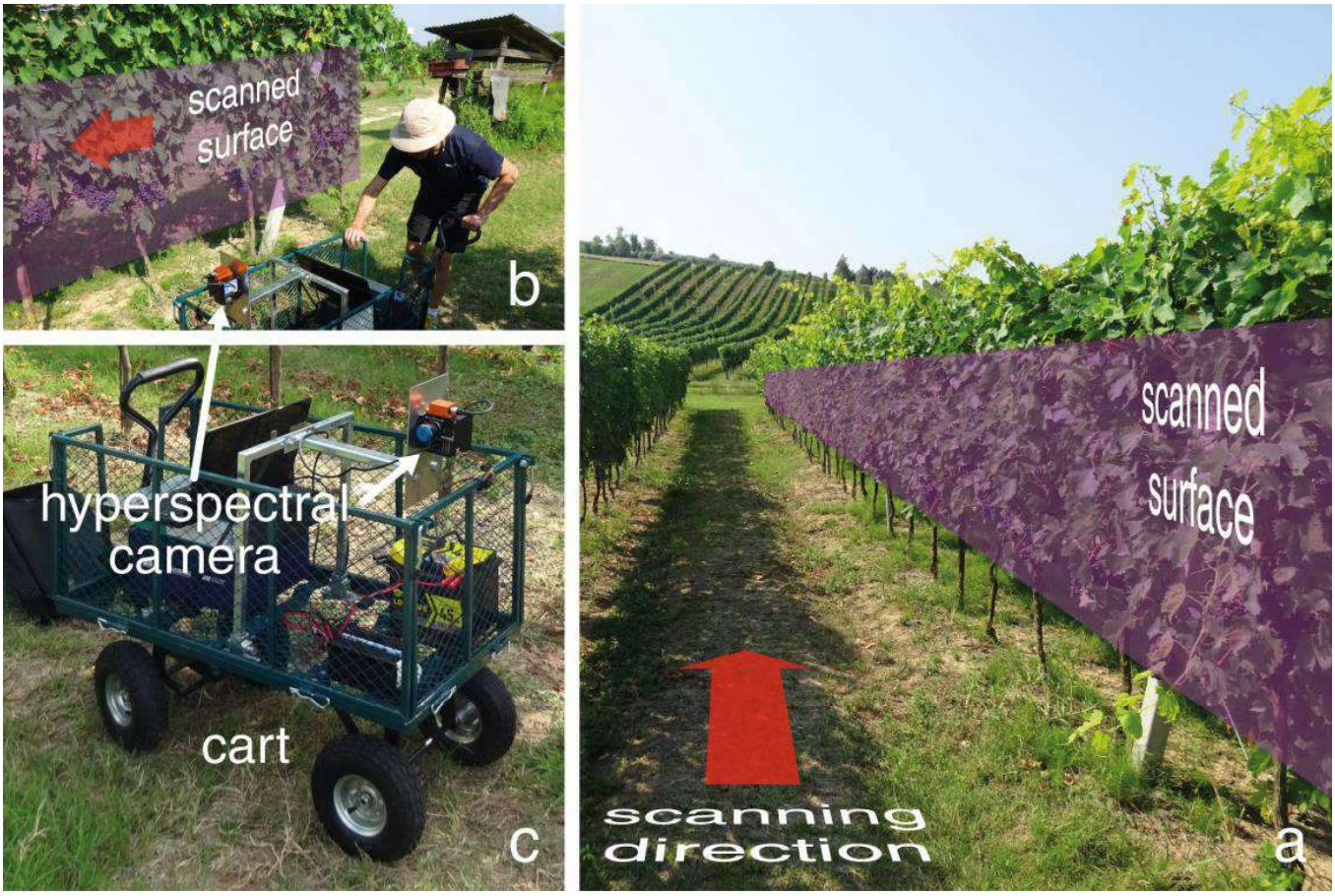

Fig.1

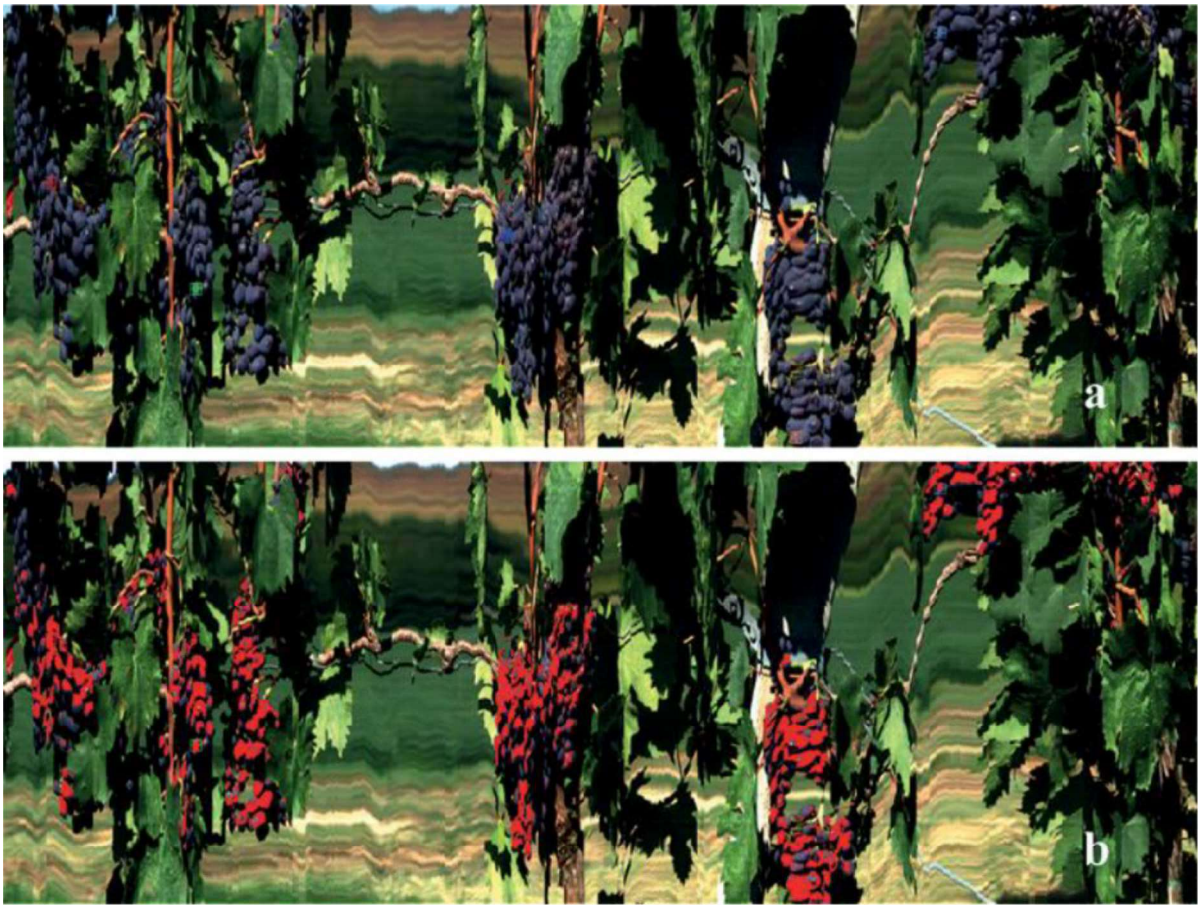

Fig. 2

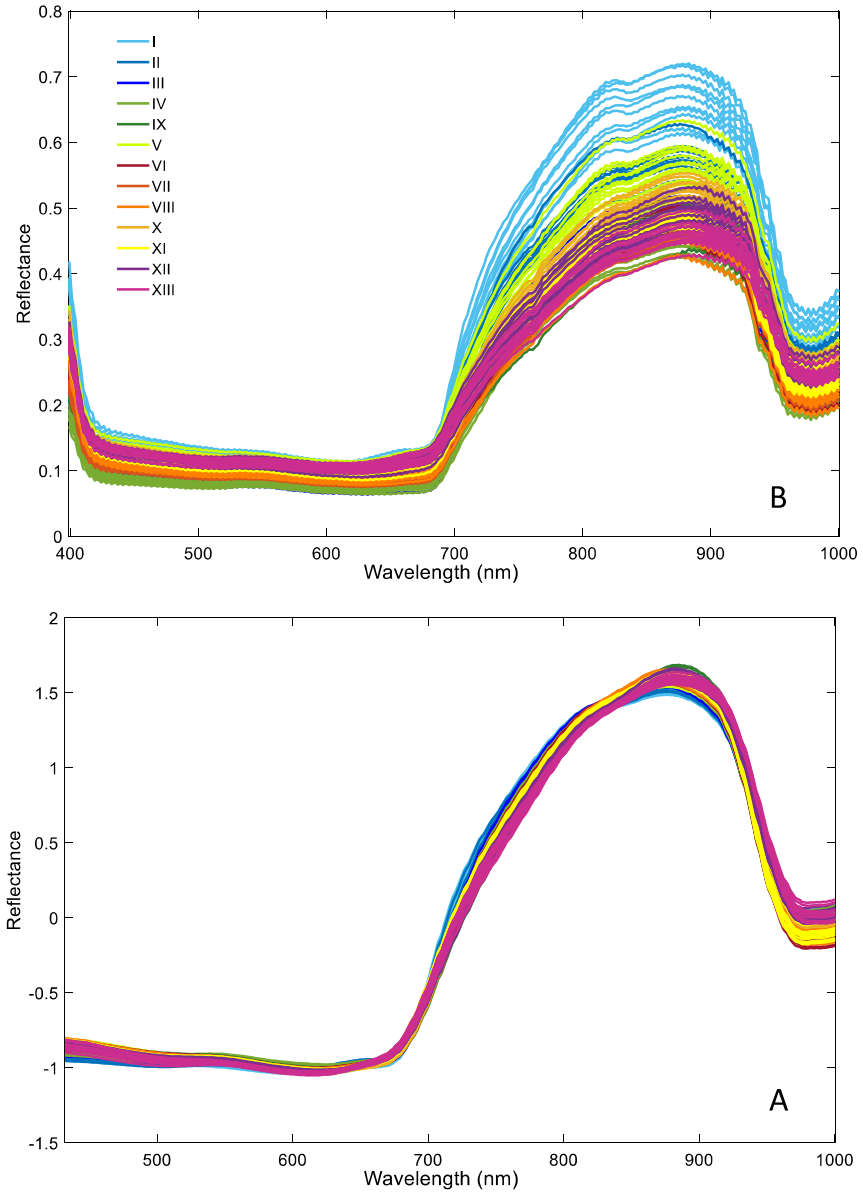

Fig. 3

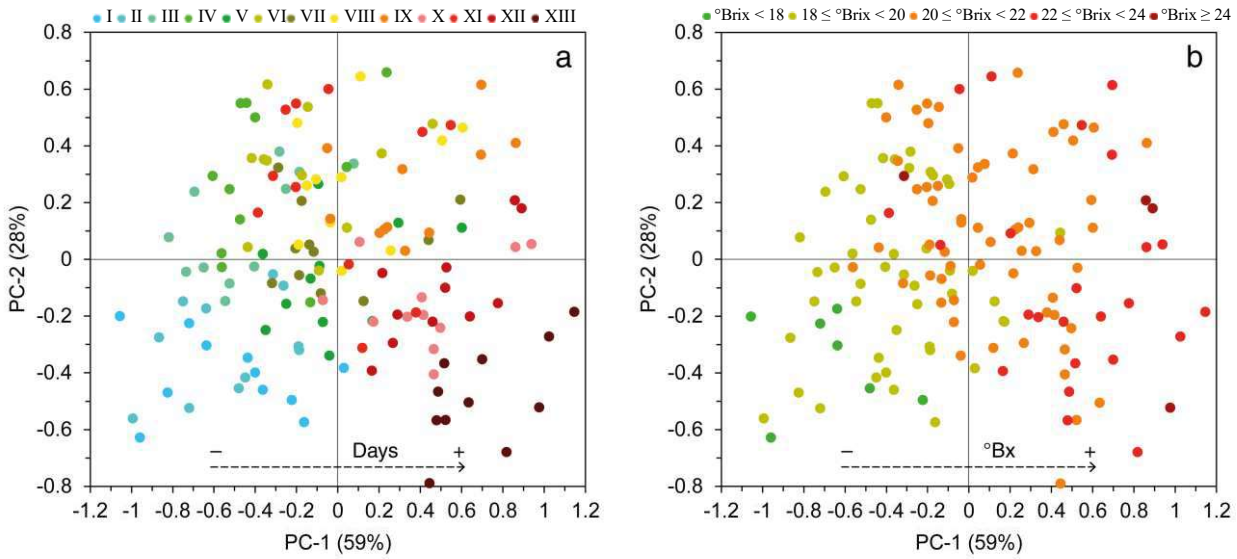

Fig. 4

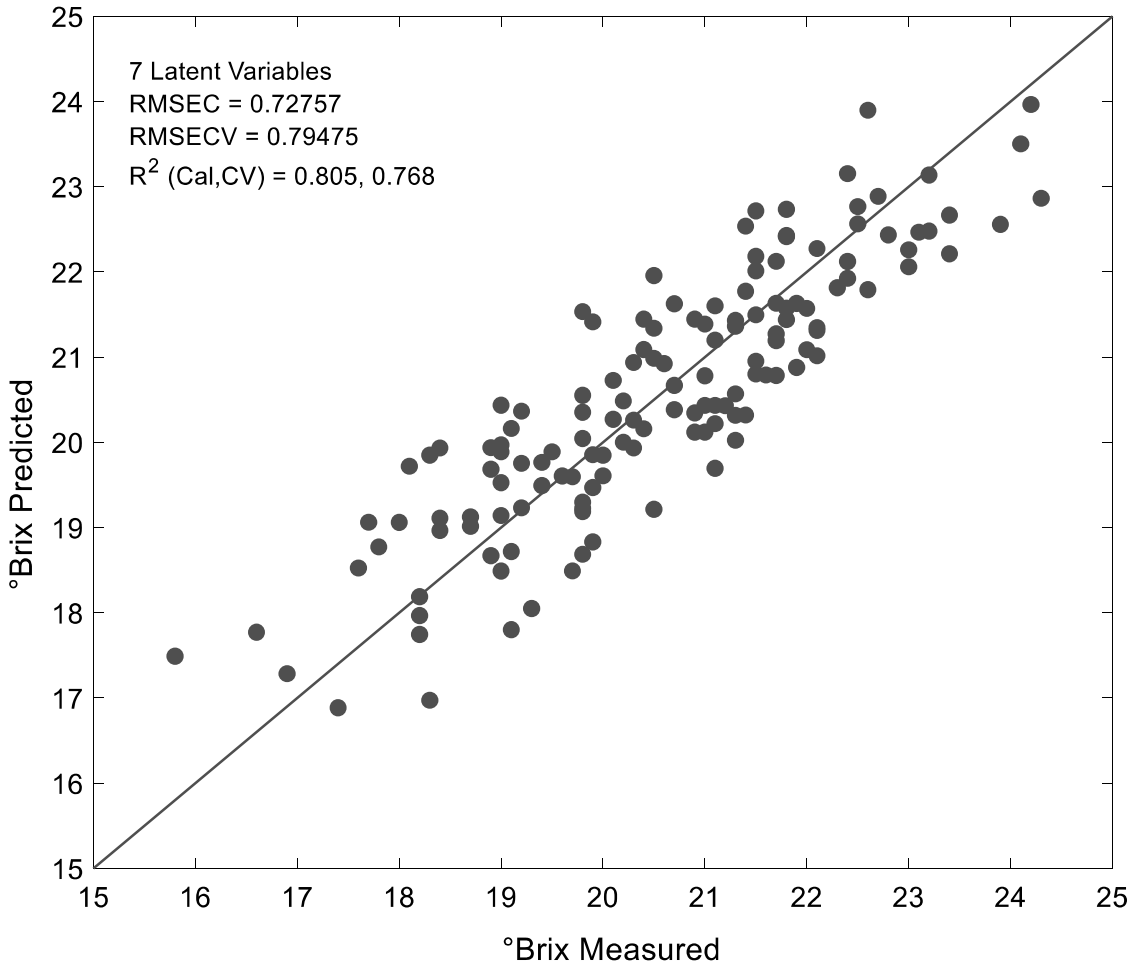

Fig. 5

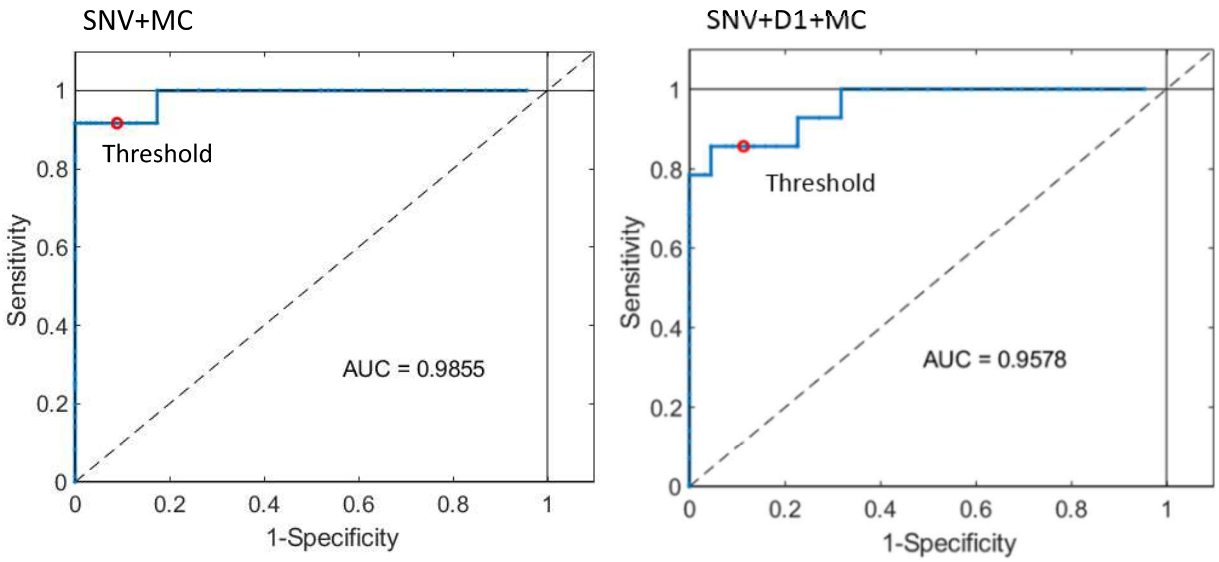

Fig. 6

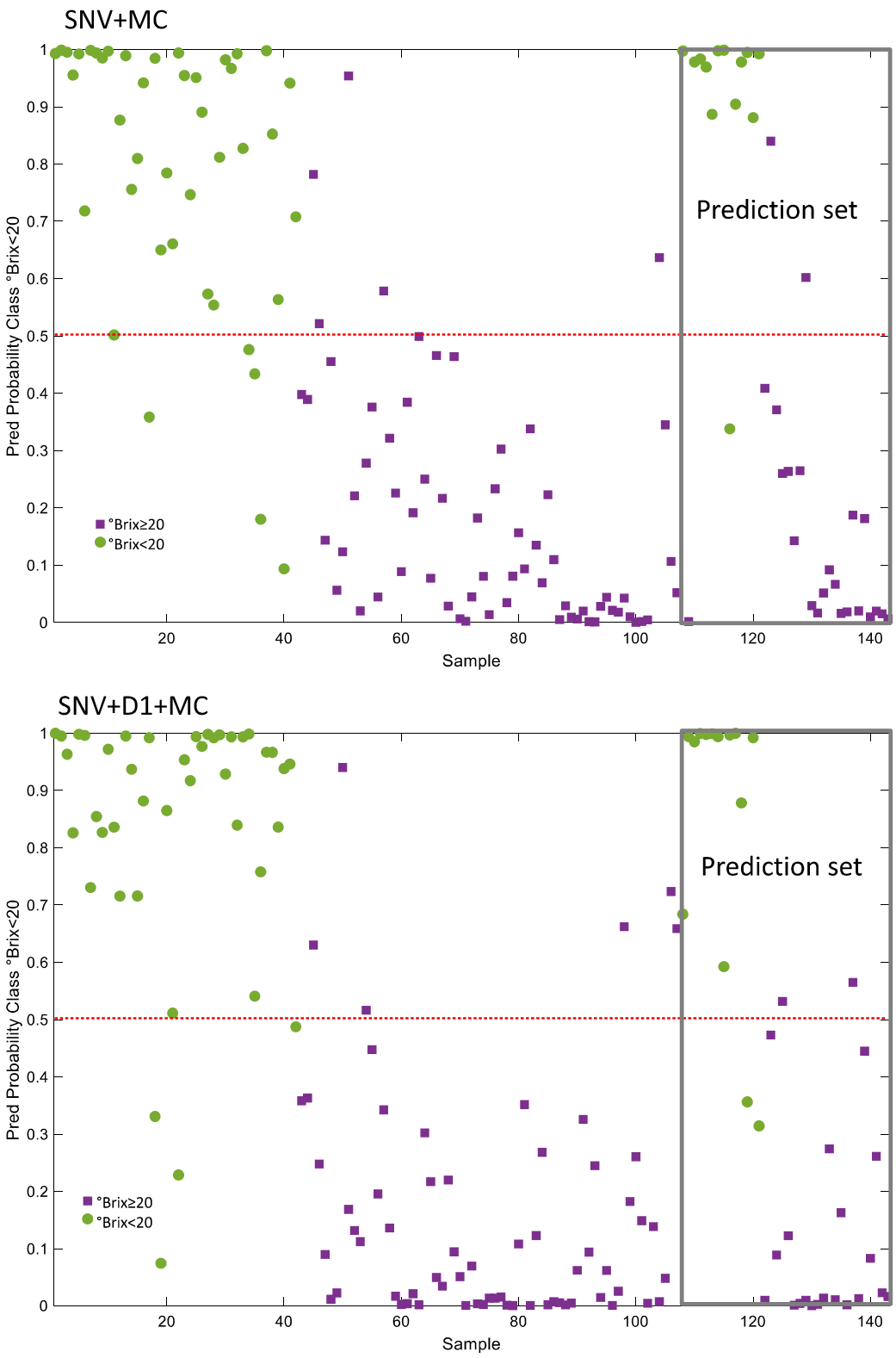

Fig. 7

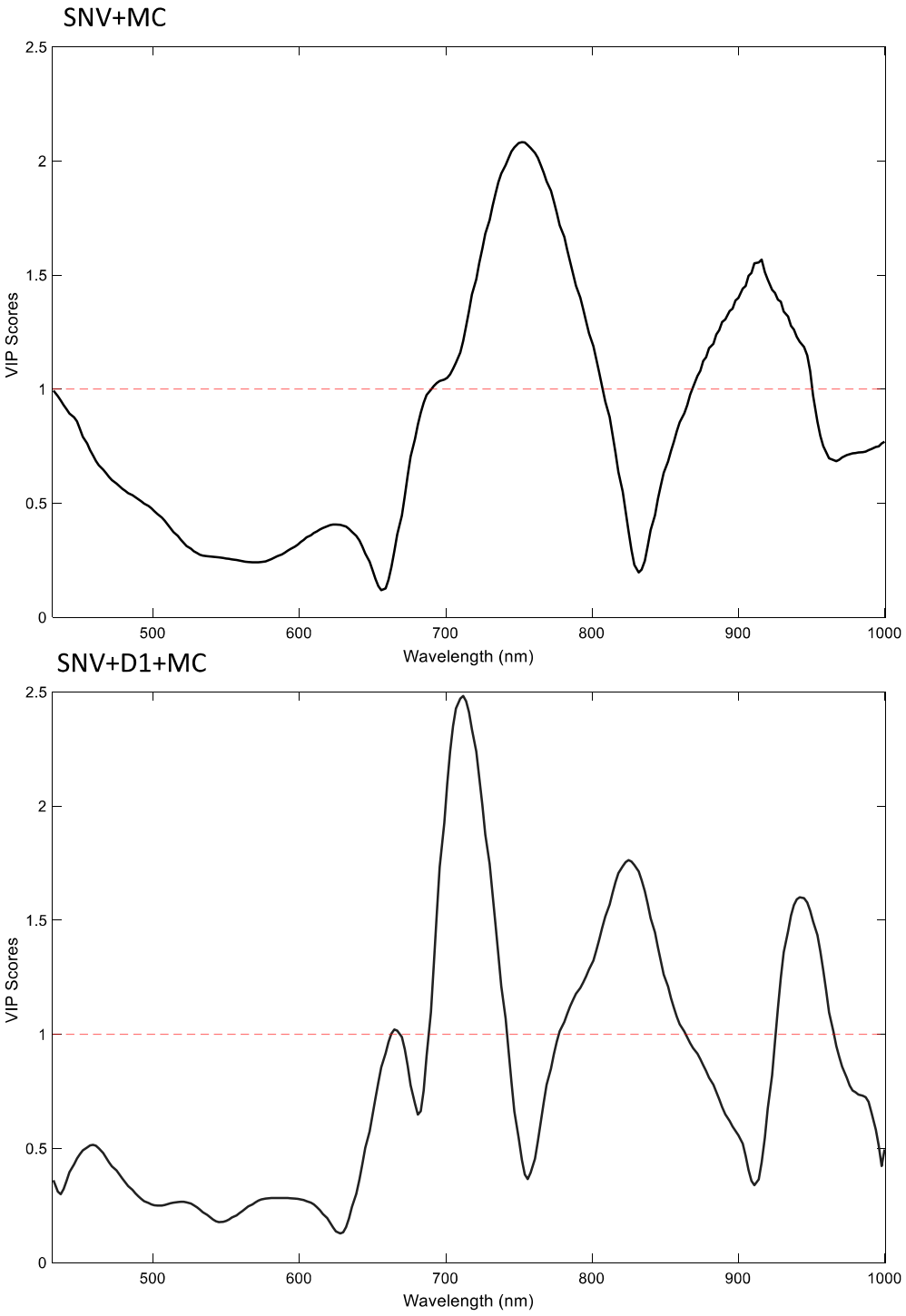

Fig. 8

Figure captions

Fig. 1. (a) Vineyard row, highlighted by the purple surface, scanned with the Vis/NIR hyperspectral camera; red arrow indicates the direction of scanning; (b) Garden cart mounted with hyperspectral camera. (c) In-field hyperspectral imaging to measure soluble solids content of wine grape berries during ripening.

Fig. 2. (a) RGB image from hyperspectral image of a scanned vineyard row section. (b) Representation of the ROI (in red) resulting from the classification obtained by the Manhattan function.

Fig. 3. Raw (A) and pre-treated by smoothing and SNV (B) spectra of all samples on different days of analysis (from I to XIII).

Fig. 4. Score plot obtained by the PCA according to: (a) days of analysis (from I to XIII); (b) soluble solids content (°Brix).

Fig. 5. Measured vs predicted values of solid soluble content (°Brix) obtained by PLS regression (cross validation).

Fig. 6. ROC curve of the PLS-DA models in prediction.

Fig. 7. Probability values of belonging to the class °Brix<20.

Fig. 8. VIP scores of the PLS-DA.

Table 1. Mean and standard deviation values of soluble solids content (°Brix) as a function of days of analysis.

Day of analisys	Mean (°Brix)	Standard deviation (°Brix)
I	17.8[a]	1.05
II	19.0[a,b]	0.81
III	19.2[b,c]	0.75
IV	19.9[b,c,d]	1.1
V	20.1[b,c,d]	1.16
VI	20.2[b,c,d,e]	0.95
VII	20.5[c,d,e,f]	1.11
VIII	20.7[d,e,f]	0.82
IX	21.5[e,f,g]	0.76
X	21.6[e,f,g]	0.89
XI	21.7[f,g]	1.17
XII	22.7[g]	0.94
XIII	22.7[g]	0.88

Note: means with the same letter are not significant different at p-level < 0.05.

Table 2. PLS-DA results in terms of percentages of correctly classified samples (whole spectral range).

Spectra pretreatment	Class	Calibration (n=107)	Cross-validation (n=107, 10 segments)	Prediction (n=36)	LV
SNV+MC	°Brix<20	89%	89%	91%	4
	°Brix≥20	91%	91%	91%	
SNV+D1+ MC	°Brix<20	91%	89%	86%	4
	°Brix≥20	91%	91%	91%	

Note: SNV=Standard Normal Variate; MC=Mean Cantered; D1=first derivative, LV=Latent Variable.

Table 3. PLS-DA results in terms of percentages of correctly classified samples (reduced spectral range).

Spectra pretreatment	Class	Calibration (n=107)	Cross-validation (n=107, 10 segments)	Prediction (n=36)	LV
SNV+MC	°Brix<20	84%	83%	86%	4
	°Brix≥20	93%	92%	91%	
SNV+D1+ MC	°Brix<20	89%	89%	88%	3
	°Brix≥20	90%	87%	87%	

Note: SNV=Standard Normal Variate; MC=Mean Cantered; D1=first derivative, LV=Latent Variable.

Nomenclature

AUC	area under the curve
CV	cross validation
D1	first derivative
HSI	hyperspectral imaging
MC	mean centring
NIR	near-infrared
PCA	principal component analysis
PLS	partial least squares
PLS-DA	partial least squares discriminant analysis
R	reflectance spectrum
RMSE	root mean square error
ROC	receiver operating characteristic
ROI	region of interest
SNV	standard normal variate
SSC	soluble solids content
SVM	support vector machines
VIP	variable importance in projection

Conflict of Interest Statement

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: