Ripeness evaluation of kiwifruit by hyperspectral imaging

Ripeness evaluation of kiwifruit by hyperspectral imaging

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

Rapid, non-destructive fruit sorting techniques are increasingly being adopted to ensure that producers, industry, and consumers receive products that meet their quality requirements. Quality attributes typically used to assess fruit ripeness include soluble solids content (SSC) and flesh firmness (FF). In this study, hyperspectral imaging operating at 400–1000 nm (Vis/NIR) was adopted to evaluate the ripeness degree of 'Hayward' kiwifruit. Partial least square (PLS) regression models were developed to estimate SSC and FF, while two different types of PLS discriminant analysis (PLS-DA) were used to classify samples according to three repining classes (defined on the base of SCC and FF values). To reduce the computation complexity, and simplify the calibration models, two variable selection methods (genetic algorithm GA, and variable importance in projection VIP) were adopted. For SSC, the prediction R^2 values ranged from 0.85 (RMSE = 1.10 °Brix) to 0.94 (RMSE = 0.73 °Brix), and for FF from 0.82 (RMSE = 14.51 N) to 0.92 (RMSE = 9.87 N). Classification sensitivity reached values of 97% and 93%, for the model considering the SCC and FF classes, respectively. Prediction and classification performances remained substantially unchanged by reducing the number of wavelengths. Therefore, hyperspectral imaging appears to be suitable for prediction of kiwi quality attributes and their classification.

Keywords: Hyperspectral, kiwifruit, PLS, PLS-DA, ripening.

1. Introduction

Most fruits and vegetables are characterised by a short harvest season. Short or long-term storage is therefore necessary to ensure a continuous supply of quality raw material and to extend the processing and marketing periods of fresh produce (Mishra & Gamage, 2020a). Fruit is harvested when threshold values of some quality attributes, expressing their physiological maturity, are reached. With regard to kiwis, the soluble solids content (SSC) must reach 6.2–6.5 °Brix (from 7 to 9 °Brix for storage), while the flesh firmness (FF) must be equal to or greater than about 62 N (OECD, 2008) (Crisosto & Kader, 1999; Walsh et al., 2020). Regarding commercial ripening, to satisfy consumer taste, SSC should exceed 12.5 °Brix, while FF should fall below approximately 9–13 N (Crisosto & Kader, 1999).

The storage potential of kiwifruit in air at storage sites under semi-optimal conditions of temperature and relative humidity ranges from 4 to 8 weeks (Mishra & Gamage, 2020a). From postharvest to commercial maturity, several changes at the cellular level affect fruit texture, which becomes softer (Mishra & Gamage, 2020b). Composition, water content, turgor pressure and cell wall constituents undergo various transformations. In kiwi pulp cells, the structure of the middle lamella and the primary cell wall weakens mainly due to high pectin solubilisation; in addition, a moderate contribution to fruit softening is provided by the loss of pectic arabinan side chains. These phenomena result in cell separation and wall swelling (Watkins, 2017) (Mishra & Gamage, 2020b).

Early in ripening, there is often an increase in sugar concentration. In fruits and vegetables, sucrose comes from photosynthetic leaves and/or hydrolysis of starch reserves degraded by amylase, resulting in an increase in SSC. As ripening progresses, glucose and fructose increase due to the action of the enzyme invertase through glycolysis. The amount of sucrose may also increase due to gluconeogenesis of organic acids (Sánchez-Rodríguez et al., 2019). As fruits ripen, the chlorophyll content generally decreases, and other pigments such as carotenoids are usually biosynthesised and

53 accumulated. However, there are some exceptions to this general behaviour: some fruits retain
54 chlorophyll, such as the skin of 'Granny Smith' green apples and 'Conference' pears and the flesh of
55 'Hayward' green kiwi, in which the accumulation of carotenoids decreases with ripening. Kiwis are
56 also very rich in vitamin C (Alós et al., 2019) (Mellidou et al., 2019).
57 The use of the penetrometer and refractometer to measure FF and SSC, respectively, result in
58 destruction of the sample and are time consuming and expensive. To overcome this problem, several
59 techniques, many of which are non-destructive, have been introduced over the past few decades and
60 generally allow for faster analysis and lower costs.
61 Near-infrared (NIR) spectroscopy (McGlone & Kawano, 1998) (Schaare & Fraser, 2000) (Hu et al.,
62 2016a) and portable NIR instruments used directly in the field (Costa et al., 2015) (Cantin et al., 2011)
63 have been adopted with good results to predict dry matter , SSC and FF of kiwifruit.
64 Time-domain diffuse reflectance spectroscopy, a pulsed laser spectroscopic technique, was applied
65 to classify kiwifruits according to FF, sugar content and acidity, obtaining the best result with acidity
66 (Valero et al., 2004). A multifrequency magnetic induction spectroscopy system (156 kHz–2.5 MHz)
67 resulted in a good correlation of the bio-impedance conductivity measurement with SSC and FF
68 (Bauchot et al., 1999) (O'Toole et al., 2015). Ragni et al. (2012) adopted waveguide spectroscopy
69 (2–3 GHz and 15–16 GHz) to evaluate SSC and FF of kiwifruit during storage, obtaining the best
70 results by partial least squares (PLS). Using a device consisting of a xenon lamp whose emitted
71 radiation passes through kiwi samples and is captured by a NIR-sensitive camera, 8-bit greyscale
72 images (255 grey tones) have been used to predict the FF and SSC of kiwifruit (Berardinelli et al.,
73 2019).
74 FF penetrometric measurements correlated with measurements from a non-invasive system called
75 'intelligent firmness detector', which can detect the FF of a fruit as it rotates on a grading belt (Blanke,
76 2013). Li et al. (2016) correlated penetrometric measurements with those from four different non-
77 destructive devices. Hu et al. (2016b) observed a correlation between colour change and FF of

78 kiwifruit by adopting a palladium complex-based colourimetric sensor that can be used for real-time
79 monitoring of kiwifruit ripening.
80 Hyperspectral imaging (HSI) has been increasingly applied in the last decade for determination of
81 quality parameters in fruits and vegetables. This technique combines image analysis and
82 spectroscopy, obtaining absorbance/reflectance/transmittance spectra as single acquired pixels (in
83 this technique called voxels) that form the hyperspectral (HS) image. HSI has been applied for the
84 determination of quality parameters of fruits and vegetables in the laboratory and directly in the field
85 (Chandrasekaran et al., 2019). HSI was implemented by Zhu et al. (2017) for quality attributes
86 prediction of three kiwifruit varieties at commercial maturity. The best results were obtained in the
87 380–1023 nm range and applying multiple linear regression models, with a R_P = 0.981 and a residual
88 prediction deviation (RPD) of 5.17 N cm^{-2} for FF. R_P = 0.952 and RPD = 3.26 °Brix were obtained
89 for SSC. Serranti et al. (2017) proposed application of HSI (1000–1650 nm) to evaluate kiwifruit
90 ripeness. In particular, HS images of the inner pulp and outer surface of kiwifruit were acquired to
91 assess the degree of ripening. Fruits, divided into 4 batches, were analysed on the day of harvest and
92 after 7 and 14 days. Using principal component analysis (PCA) false-colour plots of the principal
93 components were obtained, which allowed to discriminate the ripening time of fruits.
94 In this study, HSI combined with variable selection methods and chemometric techniques was
95 implemented to evaluate quality parameters of kiwi in post-harvest conditions. PLS models were
96 developed to predict the SSC and FF, while PLS-discriminant analysis (PLS-DA) and soft PLS-DA
97 were used to classify samples into 3 ripening classes defined through observation of SSC and FF.
98
99 **2. Materials and methods**
100
101 **2.1 Samples**
102 130 unconditioned ‘Hayward’ kiwis harvested near Latina (Italy) and delivered on November 11th,
103 2019 to the storage centre, were adopted for the present research. During the analysis period, kiwis

were stored at 15 °C. The analyses were carried out on 5 different days starting from the day following delivery to the storage centre: day I (November 12th), II (November 19th), III (November 26th), IV (December 10th) and V (December 17th). The samples were weighed immediately before being placed in the storage chamber and before being analysed. As the samples were not homogeneous in terms of weight, they were divided into four weight classes (80–100 g, 100–120 g, 120–140 g and 140–160 g), random selected and equally distributed over the days of analysis.

2.2 Hyperspectral measurements

A push-broom linear array hyperspectral camera working in the 400–1000 nm spectral range (Nano-Hyperspec VNIR, Headwall Photonics, Inc., Fitchburg, MA, USA) was adopted to acquire HS images. The camera was equipped with a 17 mm effective focal length (EFL) lens, characterised by an angular field of view (FOW) of 15.3°.

The scan line has a spatial resolution of 640 points, each with 272 spectral bands and a nominal spectral resolution of 2.2 nm. The sensor exposure time and frame period were set at 28 ms, while the frame rate, which depends on the exposure time set, was 34.2 frames per second (FPS).

Figure 1 illustrates the HSI system adopted for the experimental trials. The camera was installed on a metal frame with the optical axis perpendicular to the underlying conveyor belt, which simulates an industrial fruit sorting line at a height of 54 cm. The speed of the conveyor belt, on which the sample runs, was approximately 8 mm s^{-1}. Two spotlights inclined by 15° and at a height of 32 cm from the conveyor belt plane were installed on the metal frame. The spotlights were equipped with 120 W halogen lamps. Ambient light was isolated using a specific box.

The white (R_W) and dark (R_D) reference reflectance spectra were obtained by acquiring a high-reflectance matte white panel and covering the camera lens with its cup, respectively. The raw diffuse reflectance spectrum (R_R) was obtained by scanning the sample. Each sample was scanned longitudinally at laboratory temperature (22 ± 1 °C). The calibrated diffuse reflectance spectrum (R_C) was calculated by applying the equation (1) (Guo et al., 2019):

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

2.3 Destructive measurements of quality parameters

The Magness-Taylor FF and SSC of the kiwi samples were measured after HS image acquisitions. FF was determined through a compression test performed on kiwi samples with a texture analyser (TA-HDi, Stable Micro System Ltd., Godalming, UK) equipped with a cylindrical steel probe 8 mm diameter with a hemispherical head (ASABE Standard, 2018). Before the analysis, a small portion of skin in the area subjected to the compression test was removed. Tests were performed with a penetration speed of 0.5 mm s^{-1} and a maximum deformation of 8 mm. Kiwi juice was analysed with a digital refractometer (PR-101 Digital Refractometer, ATAGO CO., LTD, Tokyo, Japan) to measure the SSC and expressed in °Brix.

Significant differences between the means of the quality indexes (FF and SSC) at different days of storage were evaluated through analysis of variance (one-way ANOVA with Tukey-HSD post-hoc test, p-level > 0.05). Homogeneity of variance was evaluated by the Levene test.

2.4 Data analysis

ROI determination

The background was removed from each HS image by k-means clustering, performed using 2 clusters (classification method: Euclidean distance) on the calibrated spectra (HyperCube v. 11.52, U.S. Army Engineer Research and Development Center (ERDC), USA). Cluster 1 (red in Figure 2b) represents the background to be discarded, while cluster 2 (green in Figures 2b and 2c) is the region of interest (ROI). It can be observed (Figure 2c) that the partially shaded edge of the kiwi sample was not included in cluster 2. For each sample, the mean spectrum was calculated considering spectra of the ROI.

155 *Multivariate analysis*

156 Due to a low signal-to-noise ratio produced by the sensor, spectral bands between 400 and 424 nm
157 were removed from the analysis. The spectra were subsequently smoothed to reduce noise from the
158 spectra and pre-processed by applying the standard normal variate (SNV) method, first derivative
159 (D1) and mean centred (MC). (Rinnan et al., 2009). The data were preliminary visualised according
160 to quality indices (SSC and FF) and days of analysis using PCA.

161

162 *Selection of wavelengths*

163 The selection of wavelengths from the original or pre-processed spectra can reduce the computation
164 complexity, improve the predictive ability of calibration models and simplify the calibration models.
165 Furthermore, a limited number of wavelengths can be directly used for developing on-line or portable
166 multispectral equipment. The wavelength selection methods used in this study are variable
167 importance in projection (VIP) and genetic algorithm (GA). VIP method is a filter method, which
168 estimates the importance of each variable in the projection used in a PLS model. The VIP score
169 calculates the contribution of each variable according to variance explained by each PLS component.
170 The 'greater than one' is conventionally used as the criterion for variable selection; accordingly, only
171 variables with VIP scores greater than one were selected (Wang et al., 2015). This method has been
172 extensively used in different fields and adopted for a variety of data types (Farrès et al., 2015).
173 GA is a wrapper method implemented by selecting the candidate of sensitive wavelengths and
174 optimising the number of evaluations in each run, while PLS is applied to perform and evaluate the
175 selected wavelengths (Zhu et al., 2017). The computation steps involved have been described by
176 Mehmood et al., (2012). In this study, 100 runs were set.

177

178 *Chemometric models*

PLS regression models were developed to estimate the quality indices (SCC and FF), adopting the Onion method to split the dataset (130 samples) in calibration (75% of samples, with venetian blinds as cross-validation method) and validation test set (25% of samples).

The onion method selects a ring of the most unique samples (based on Euclidian distance from the mean, like the outer-most layer of an onion.) These are used in the calibration set. Next, a ring of less unique samples, just inside the first set (the next onion layer), is put into the validation set. This is repeated two more times so that there are three outer rings of the most unique and less unique samples. Finally, all remaining samples are split randomly into calibration and validation (Gallagher and O'Sullivan, 2020; Gallagher et al., 2004).

To avoid over-fitting the model, the optimal number of latent variables (LV) were chosen by identifying the global minimum of root mean square error in cross validation (RMSECV).

PLS-DA and multi-class *soft* PLS-DA (Pomerantsev & Rodionova, 2018) (Zontov et al., 2020) were developed to classify samples into 3 ripening classes, according to Crisosto & Kader (1999). For the SSC, the following limits were selected: SSC<6.5 (unripe), 6.5≤SSC<12.5 (ripe for storage) and SSC≥12.5 (ripe for consumer), while for FF the following were used: FF≥62 (unripe), 13≤FF<62 (ripe for storage) and FF<13 (ripe for consumer). These ranges were chosen since they meet the requirements at both the harvesting stage (unripe to ripe for storage) and by consumers (ripe for storage to ripe for consumer). The dataset was spilt using the Onion method, as described for PLS.

PLS-DA assumes that each sample is assigned to a single class using a probabilistic approach based on Bayes rule. Multi-class *soft* PLS-DA, by adopting PCA of the matrix T of the predicted response ($\hat{Y}$), classifies samples according to the distance between rows of this matrix and rows of matrix Y (measured response). The procedure can be summarized as:

$$X, Y \xrightarrow{PLS} \hat{Y} \xrightarrow{PCA} T \quad (2)$$

where X is the matrix utilised as a predictor. Furthermore, with multi-class soft PLS-DA, the sample can be assigned to one or more classes, or it can not be assigned to any class (Zontov et al., 2020).

All the chemometric models were developed using PLS Toolbox for MATLAB R2020b, except for multi-class *soft* PLS-DA, where PLS-DA Toolbox for MATLAB R2019a was used.

Spatial distribution

To reduce calculation time, image resolution was decreased by averaging the spectra of adjacent pixels (3x3). PLS calibration models were developed considering the mean spectra of the ROIs and applied to groups of 3x3 pixels of each image. Consequently, predicted SCC and FF values were obtained for each pixel group. This allows to observe the difference attributes from sample to sample and even within the same sample.

3.1 Results

Means and standard deviations of quality attributes (SSC and FF) measured on kiwi samples during the 5 days of analysis are reported in Figure 3. Significant differences were observed for both attributes between the second and third day and between the third and fourth day, whereas no significant differences were found between the first and second dayand between the fourth and fifth day. The percentage variation between the first and last day of analysis was 124% and -82% for SSC and FF, respectively. The correlation between FF and SSC was negative, with a probable linear trend (R^2 = 0.84; FF = -10.22 SCC + 146.72).

Figure 4 shows the mean spectra of kiwi samples according to SSC and FF, and pre-processed by SNV and D1 filters, respectively. The portion of the spectrum from 400 to 700 nm (visible) is characterised by absorption bands of anthocyanins, carotenoids and chlorophyll-α (ElMasry et al., 2007), which have been adopted as quality attributes to assess fruit ripeness. The NIR part of the spectrum (700–1000 nm) shows absorption bands related to water (Nicolaï et al., 2007) and sugar (Pu et al., 2016). Manley et al. (2007) and Camps & Christen (2009) observed that in the 800–1000 nm range, SSC-related spectral information tends to be less covered by water, whose absorption peaks are broad and not very marked.

The score plots of principal components 1 and 2 resulting from PCA were obtained mean centring the calibrated spectra (Figure 5). Subsequently, some outliers were removed. Samples were distributed along PC1 from right to left for increasing SSC values (Figure5a) and decreasing FF values (Figure 5b).

The results of the PLS models developed to estimate SSC and FF were reported in Table 1. Two different pre-processing sequences were adopted: SNV+MC and D1+MC, as well as two variable selection methods, namely VIP and GA. For SSC, the prediction R^2 values ranged from 0.85 (RMSE = 1.10 °Brix) to 0.94 (RMSE = 0.73 °Brix) and the best results were achieved applying D1 pre-processing and selecting the variables by the GA. For FF, the best results, in terms of prediction R^2 (0.92) and RMSE (9.87 N), were obtained with SNV pre-processing and GA variable selection methods. In general, for both quality parameters, reduction of the variable number, especially by using the GA method, positively affected the results in terms of increasing of R^2 and reducing RMSE in prediction. Selected variables by GA method are shown in Figure 6 for SSC and FF prediction, respectively. Most are in the NIR region (700–1000 nm), especially for SSC. The spectral regions selected for the model prediction of SSC and FF were quite different. The region from 950 to 1000 nm, typically related to carbohydrate absorption, was selected only for SSC prediction, while the region around 750 nm (band assignments for the major water (O-H) vibration) was defined only for FF prediction. This may suggest that FF prediction is independent of SCC correlation. Considering the VIP method, the selected variables (VIP>1) were in similar spectral ranges (650–1000 nm) to those for the GA method, even if appreciable differences between the SSC and FF models were not observed (Figure 7). Overall, VIP scores obtained by PLS models developed after variable selection, were higher than one for all x-variables. This suggests that the contribution of the selected spectral variables on the model performance, is quite the same.

Measured vs. predicted values of SSC obtained by PLS regression applying SNV+MC pre-processing and GA variable section method are shown in Figure 8. These results confirm the feasibility of HSI in the Vis/NIR spectral region for the rapid prediction of kiwi quality parameters, such as SCC and

256 FF; furthermore, they are in good agreement with those reported in the literature (R of 0.9812 and
257 0.9523 for FF and SSC, respectively) by Zhu et al. (2017) in a very similar wavelength range (380–
258 1023 nm).
259 The spectral and spatial information of each pixel in HS images allowed evaluation of quality
260 parameters of each pixel with chemometric models. Using the PLS models developed after the GA
261 variable reduction, false colour images were obtained. The pixels having similar spectra showed
262 similar colours in the images, and consequently the hypercube allowed to visualise the biochemical
263 compounds of a sample in a pixel-wise manner (Zhu et al., 2017). The prediction maps of SSC and
264 FF of three representative kiwifruits (one for each ripening class) are shown in Figure 9. The colour
265 bars indicate the scales of the reference values. Measured (M) and mean predicted (P) values are also
266 reported. Spatial distribution of SSC (a) and FF (b) are in alignment with the measured values,
267 particularly passing from 5.9 °Brix/108.83 N to 13.6 °Brix/19.24 N, where the colour ranges from
268 blue/yellow to yellow/blue. As reported by Zhu et al. (2017), the noises of the HS image affected the
269 spectrum of each pixel, which may result in the predicted contents in maps exceeding the range of
270 the calibration set and prediction set.
271 Results of PLS-DA and *soft* PLS-DA in terms of class sensitivity (percentage of samples correctly
272 recognized as members of the class) are reported in Table 2. As for the PLS models, two different
273 pre-processing sequences were adopted, SNV+MC and D1+MC, as well as two variable selection
274 methods, VIP and GA. For both quality parameters, according to Crisosto & Kader (1999), three
275 classes were defined: SSC<6.5, 6.5≤SSC<12.5 and SSC≥12.5; FF≥62, 13≤FF<62 and FF<13.
276 Considering the PLS-DA models, regardless of pre-processing and variable selection method, lower
277 sensitivity values were achieved for the central classes, reporting values from 58 to 77% and from 55
278 to 86% for SSC and FF, respectively. For both quality parameters, considering the overall sensitivity
279 (SCC = 84% and FF = 93%), the best models were obtained applying the D1+MC pre-processing and
280 selecting the variables (78 and 75) by GA. PLS-DA prediction plots for FF are presented in figure
281 10. As shown in Figure 10a (Y predicted = FF≥62), there is a clear separation between samples

belonging to the FF≥62 class and the others samples. In Figures 10b (Y predicted = 13≤FF<62) and 10c (Y predicted = FF<13), some samples of classes 13≤FF<62 and FF<13 are overlapping, suggesting that the samples of the two groups, are quite similar. Nevertheless, it can be observed that almost all samples belonging to Y predicted class, were correctly classified.

Soft PLS-DA can simultaneously attribute a sample to multiple classes. In addition, a sample may not be attributed at all, if the distance between sample and class is greater than the set critical threshold. In this study, the outlier significance level (γ) was set to 0.01 (Pomerantsev & Rodionova, 2018). Regarding the *soft* PLS-DA models developed to classify the samples according to the SSC, the best results (overall sensitivity = 97%) was obtained applying SNV+MC pre-processing and selecting variables (47) by GA. In prediction, all samples of the classes SSC<6.5 and SSC≥12.5 were correctly classified, while 2 samples of the class 6.5≤SSC<12.5 were not (Figure 11a). For FF, the best result (overall sensitivity = 93%) was obtained by considering all the variables and applying D1 pre-processing. In Figure 11b, it can be observed that all samples of class FF≥62 were correctly classified. Furthermore, the separation of class FF≥62 from the other two classes is evident, as the acceptance area of the class bound by the ellipsoid is well separated from the others.

Comparing PLS-DA and *soft* PLS-DA results, it is possible to state that for SCC, regardless of pre-processing and variable selection method, notably higher sensitivity values were achieved by using *soft* PLS-DA. Instead, for FF, the sensitivity values obtained by the two discrimination methods are comparable. Overall sensitivities are higher than those reported in the literature for similar ripening classes defined on green-fleshed kiwi (82.0% for SSC and 74.0% for FF) and obtained using a detector working on nine wavelengths from 800 to 1100 nm (Yang et al., 2020).

The results show how the maturity class of kiwifruit can be predicted non-destructively even by limiting the number of wavelengths. Overall, the best models were obtained by applying the GA method, which allowed a decrease of wavelength number by more than 80%. This could be a first step in developing a multispectral equipment to sort kiwifruits directly on-line. In the light of what was observed during laboratory trials, it is necessary to highlight some choices that affected the

success of the experiment. The angle of inclination of the lamps affects the extension of the areas of light reflection and shadows on the surface of the fruit. In the case of the kiwi variety under analysis, light reflection is attenuated by the presence of hairs on the skin, so that it is possible to reduce the angle of inclination of lamps with respect to the plane. The black surface of the conveyor belt was useful as it was easy to exclude from HS images by 2-class k-means classification, which also allowed elimination of shaded or poorly illuminated parts of fruits.

Conclusions

The application of HSI combined with chemometric techniques (PLS and PLS-DA), allowed estimation of the FF and SSC of kiwifruit and to classify samples according to ripening classes. Prediction and classification performances remained substantially unchanged by reducing the number of wavelengths, and thus it is expected that a less expensive multispectral camera in the 400–1000 nm range would work just as well. Regarding SSC, the prediction R^2 values ranged from 0.85 (RMSE = 1.10 °Brix) to 0.94 (RMSE = 0.73 °Brix), and for FF from 0.82 (RMSE = 14.51 N) to 0.92 (RMSE = 9.87 N). Classification sensitivity reached of 97% and 93% for the model considering the SCC and FF classes, respectively. Considering the results obtained, HSI technology could be implemented in a sorting line of kiwifruit or portable multispectral equipment for the fresh market, but the segmentation techniques, dimensionality reduction, and automated real time analysis need to be improved upon.

Declarations of interest: none.

References

Alós, E., Rodrigo, M. J., & Zacarias, L. (2019). Ripening and senescence. In E. M. Yahia (Ed.), *Postharvest Physiology and Biochemistry of Fruits and Vegetables* (pp. 131–155). Elsevier. doi:10.1016/B978-0-12-813278-4.00007-5

ASABE Standard. (2018). ASAE S368.4 DEC2000 (R2018) - Compression Test of Food Materials

of Convex Shape. American Society of Agricultural and Biological Engineers, 2000.

Bauchot, A. D., Harker, F. R., & Arnold, W. M. (1999). The use of electrical impedance spectroscopy to assess the physiological condition of kiwifruit. Postharvest Biology and Technology, 18(1), 9–18. doi:10.1016/S0925-5214(99)00056-3

Berardinelli, A., Benelli, A., Tartagni, M., & Ragni, L. (2019). Kiwifruit flesh firmness determination by a NIR sensitive device and image multivariate data analyses. Sensors and Actuators, A: Physical, 296, 265–271. doi:10.1016/j.sna.2019.07.027

Blanke, M. M. (2013). Non-invasive Assessment of Firmness and NIR Sugar (TSS) Measurement in Apple, Pear and Kiwi Fruit. Erwerbs-Obstbau, 55(1), 19–24. doi:10.1007/s10341-013-0181-3

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.

Cantin, C. M., Soto, A., Crisosto, G. M., & Crisosto, C. H. (2011). Evaluation of a non-destructive dry matter sensor for kiwifruit. Acta Horticulturae, 913, 627–632. doi:10.17660/ActaHortic.2011.913.86

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

Costa, G., Vidoni, S., Rocchi, L., Cellini, A., Buriani, G., Donati, I., & Spinelli, F. (2015). Innovative non-destructive device for fruit quality assessment and early disease diagnosis. Acta Horticulturae, 1096, 69–78. doi:10.17660/ActaHortic.2015.1096.4

Crisosto, C. H., & Kader, A. A. (1999). Kiwifruit Postharvest Quality Maintenance Guidelines. Department of Pomology, University of California, Davis, CA 95616, US. Retrieved from https://ucanr.edu/sites/kac/files/123823.pdf

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

Farrés, M., Platikanov, S., Tsakovski, S., & Tauler, R. (2015). Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation. *Journal of Chemometrics*, *29*(10), 528–536. https://doi.org/10.1002/cem.2736

Gallagher, N. B., and O'Sullivan, D. (2020). Selection of Representative Learning and Test Sets Using the Onion Method. Eigenvector Research.

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

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

Hu, X. F., Lin, M., Fu, J. H., Jiao, L., & Liu, H. J. (2016a). Nondestructive hardness assessment of kiwi-fruit using near-infrared spectroscopy. In Proceedings - 2016 6th International Conference on Instrumentation and Measurement, Computer, Communication and Control, IMCCC 2016 (pp. 69–72). IEEE. doi:10.1109/IMCCC.2016.109

Hu, X. G., Li, X. L., Park, S. H., Kim, Y. H., & Yang, S. I. (2016b). Nondestructive monitoring of kiwi ripening process using colorimetric ethylene sensor. Bulletin of the Korean Chemical Society, 37(5), 759–762. doi:10.1002/bkcs.10745

Li, J., Xue, L., Liu, M., Li, Z., & Yang, Y. (2008). Application of laser to nondestructive detection of fruit quality. Photonics and Optoelectronics Meetings (POEM) 2008: Laser Technology and Applications, 7276, 72760R. doi:10.1117/12.823295

Manley, M., Joubert, E., Myburgh, L., Lotz, E., & 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

386 Mellidou, I., Georgiadou, E. C., Kaloudas, D., Kalaitzis, P., Fotopoulos, V., & Kanellis, A. K. (2019).
387 Vitamins. In E. M. Yahia (Ed.), *Postharvest Physiology and Biochemistry of Fruits and*
388 *Vegetables* (pp. 359–383). Elsevier. doi:10.1016/B978-0-12-813278-4.00017-8
389 McGlone, V. A., & Kawano, S. (1998). Firmness, dry-matter and soluble-solids assessment of
390 postharvest kiwifruit by NIR spectroscopy. Postharvest Biology and Technology, 13(2), 131–
391 141. doi:10.1016/S0925-5214(98)00007-6
392 Mehmood, T., Liland, K. H., Snipen, L., & Sæbø, S. (2012). A review of variable selection methods
393 in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems*, *118*, 62–
394 69. https://doi.org/10.1016/j.chemolab.2012.07.010
395 Mishra, V. K., & Gamage, T. V. (2020a). Postharvest Handling and Treatments of Fruits and
396 Vegetables. In M. S. Rahman (Ed.), Handbook of Food Preservation (3rd Ed., pp. 45–60). CRC
397 Press. doi:10.1201/9780429091483-6
398 Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., & Lammertyn, J. (2007).
399 Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A
400 review. Postharvest Biology and Technology, 46(2), 99–118.
401 https://doi.org/10.1016/j.postharvbio.2007.06.024
402 O'Toole, M. D., Marsh, L. A., Davidson, J. L., Tan, Y. M., Armitage, D. W., & Peyton, A. J. (2015).
403 Non-contact multi-frequency magnetic induction spectroscopy system for industrial-scale bio-
404 impedance measurement. Measurement Science and Technology, 26(3). doi:10.1088/0957-
405 0233/26/3/035102
406 OECD. (2008). Kiwifruit - International Standardisation of Fruit and Vegetables. OECD. OECD.
407 doi:10.1787/9789264044272-en-fr
408 Pomerantsev, A. L., & Rodionova, O. Y. (2018). Multiclass partial least squares discriminant
409 analysis: Taking the right way—A critical tutorial. Journal of Chemometrics, 32(8), 1–16.
410 https://doi.org/10.1002/cem.3030
411 Pu, H., Liu, D., Wang, L., & Sun, D. W. (2016). Soluble Solids Content and pH Prediction and

412 Maturity Discrimination of Lychee Fruits Using Visible and Near Infrared Hyperspectral
413 Imaging. Food Analytical Methods, 9(1), 235–244. https://doi.org/10.1007/s12161-015-0186-7
414 Ragni, L., Cevoli, C., Berardinelli, A., & Silaghi, F. A. (2012). Non-destructive internal quality
415 assessment of "hayward" kiwifruit by waveguide spectroscopy. Journal of Food Engineering,
416 109(1), 32–37. doi:10.1016/j.jfoodeng.2011.10.002
417 Rinnan, Å., van den Berg, F., & Engelsen, S. B. (2009). Review of the most common pre-processing
418 techniques for near-infrared spectra. TrAC Trends in Analytical Chemistry, 28(10), 1201–1222.
419 https://doi.org/10.1016/J.TRAC.2009.07.007
420 Sánchez-Rodríguez, L., Syd Ali, N., Cano-Lamadrid, M., Noguera-Artiaga, L., Lipan, L., Carbonell-
421 Barrachina, Á. A., & Sendra, E. (2019). Flavors and aromas. In E. M. Yahia (Ed.), *Postharvest*
422 *Physiology and Biochemistry of Fruits and Vegetables* (pp. 385–404). Elsevier.
423 doi:10.1016/B978-0-12-813278-4.00019-1
424 Schaare, P. N., & Fraser, D. G. (2000). Comparison of reflectance, interactance and transmission
425 modes of visible-near infrared spectroscopy for measuring internal properties of kiwifruit
426 (Actinidia chinensis). Postharvest Biology and Technology, 20(2), 175–184.
427 doi:10.1016/S0925-5214(00)00130-7
428 Serranti, S., Bonifazi, G., & Luciani, V. (2017). Non-destructive quality control of kiwi fruits by
429 hyperspectral imaging. In Sensing for Agriculture and Food Quality and Safety IX (Vol. 10217,
430 p. 102170O). doi:10.1117/12.2255055
431 Valero, C., Ruiz-Altisent, M., Cubeddu, R., Pifferi, A., Taroni, P., Torricelli, A., Dover, C. J. (2004).
432 Detection of Internal Quality in Kiwi with Time-Domain Diffuse Reflectance Spectroscopy.
433 *Applied Engineering in Agriculture*, *20*(2), 223–230. doi:10.13031/2013.15879
434 Wang, Z. X., He, Q. P., & Wang, J. (2015). Comparison of variable selection methods for PLS-based
435 soft sensor modeling. *Journal of Process Control*, *26*(2015), 56–72.
436 https://doi.org/10.1016/j.jprocont.2015.01.003
437 Walsh, K.B., Mcglone, V.A., Han, D.H. (2020). The uses of near infra-red spectroscopy in

postharvest decision support: Areview. Postharvest Biol. Technol, 163, 111139

Zhu, H., Chu, B., Fan, Y., Tao, X., Yin, W., & He, Y. (2017). Hyperspectral Imaging for Predicting the Internal Quality of Kiwifruits Based on Variable Selection Algorithms and Chemometric Models. Scientific Reports, 7(1), 1–13. doi:10.1038/s41598-017-08509-6

Zontov, Y. V., Rodionova, O. Y., Kucheryavskiy, S. V., & Pomerantsev, A. L. (2020). PLS-DA – A MATLAB GUI tool for hard and soft approaches to partial least squares discriminant analysis. Chemometrics and Intelligent Laboratory Systems, 203(March), 104064. https://doi.org/10.1016/j.chemolab.2020.104064

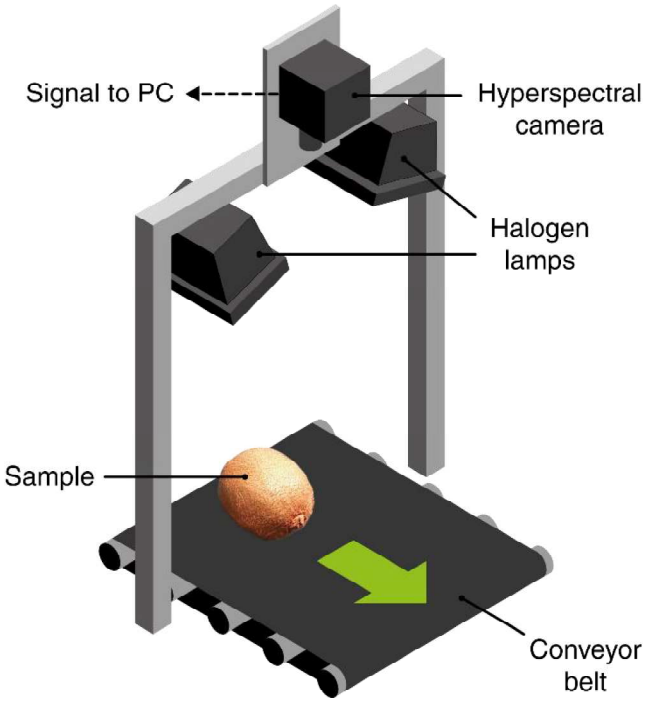


Fig.1

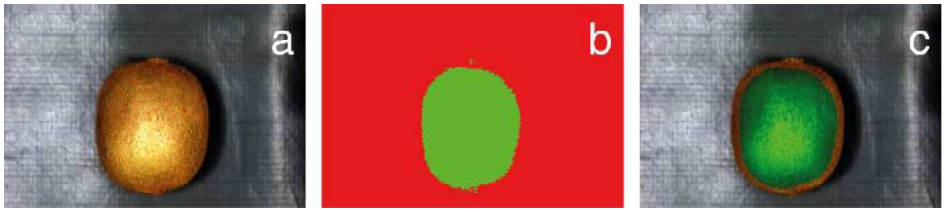


Fig.2

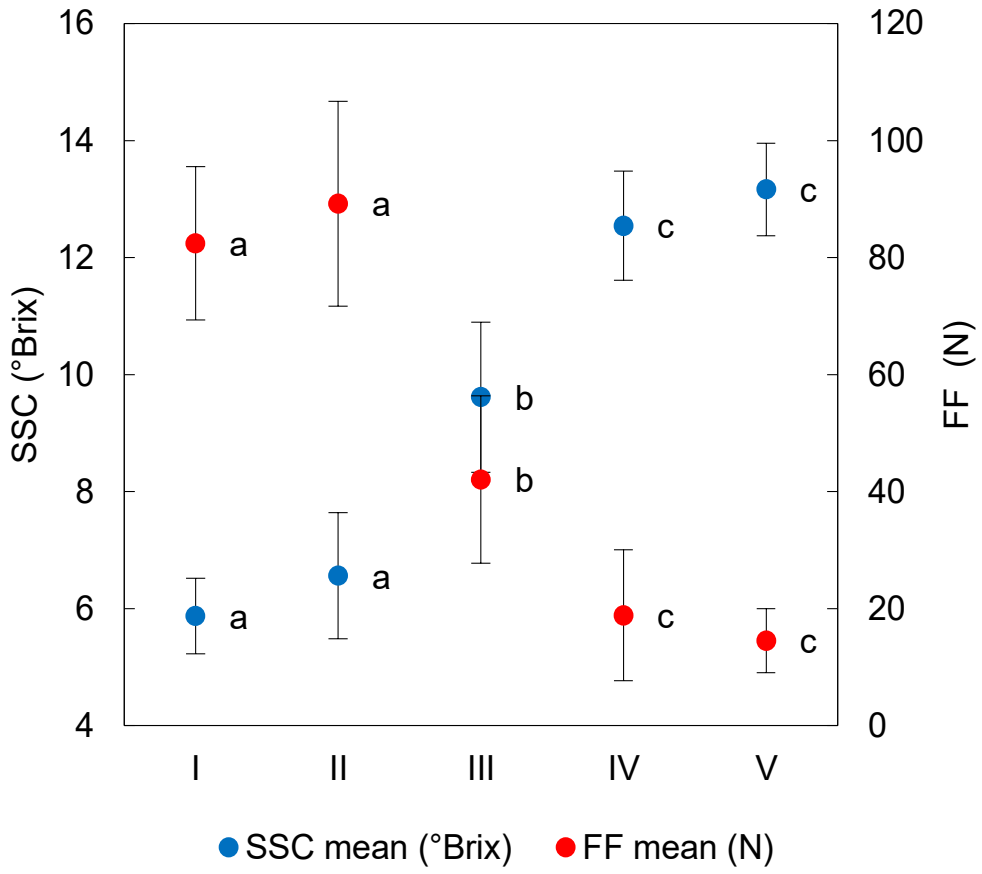


Fig.3

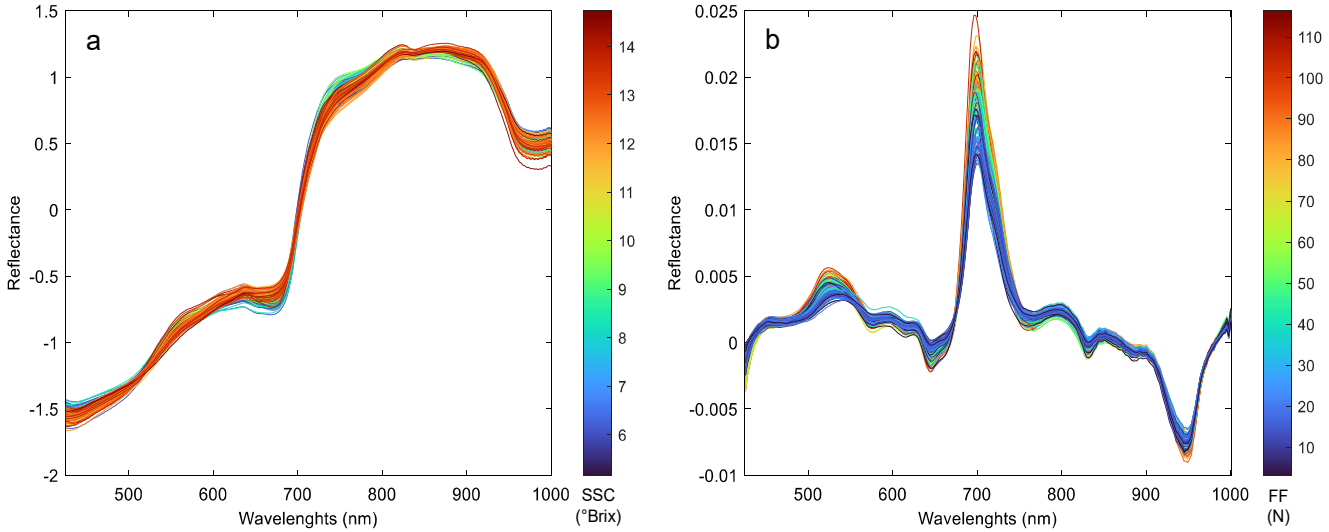


Fig.4

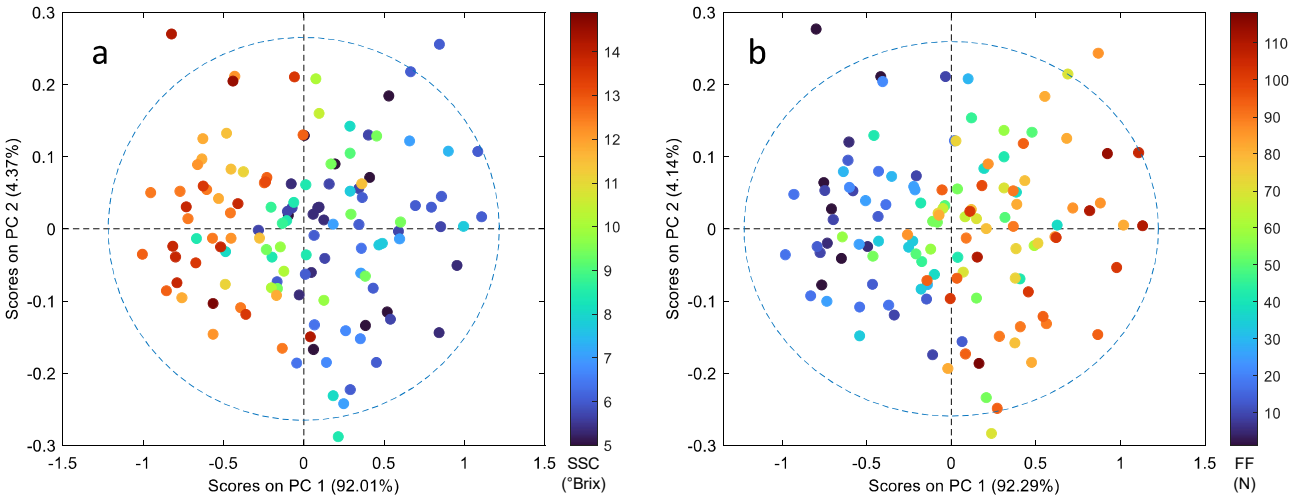


Fig.5

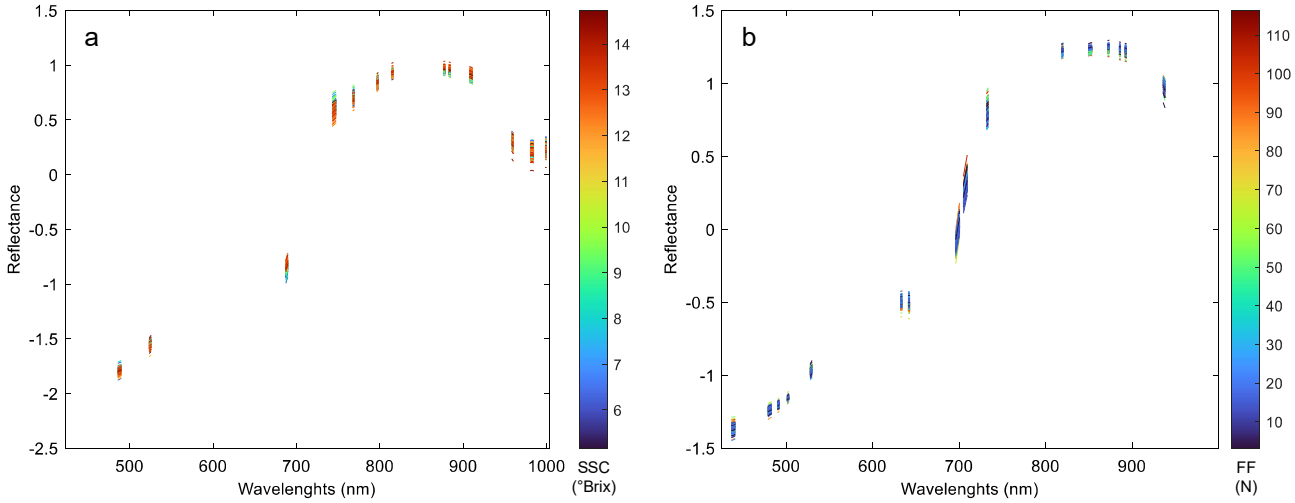
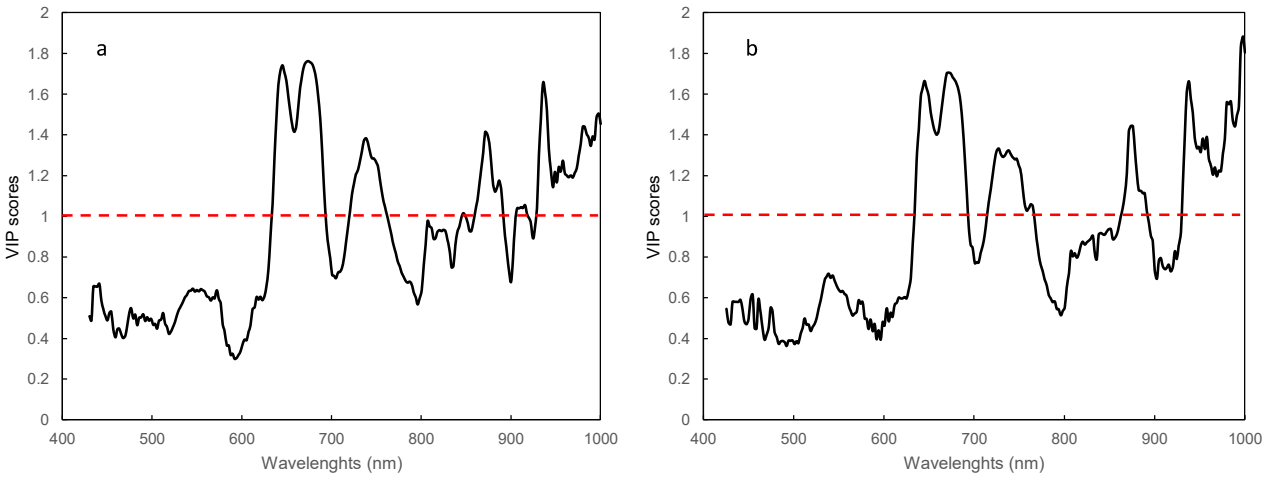


Figure 6.

1



2

3

4 **Fig 7.**

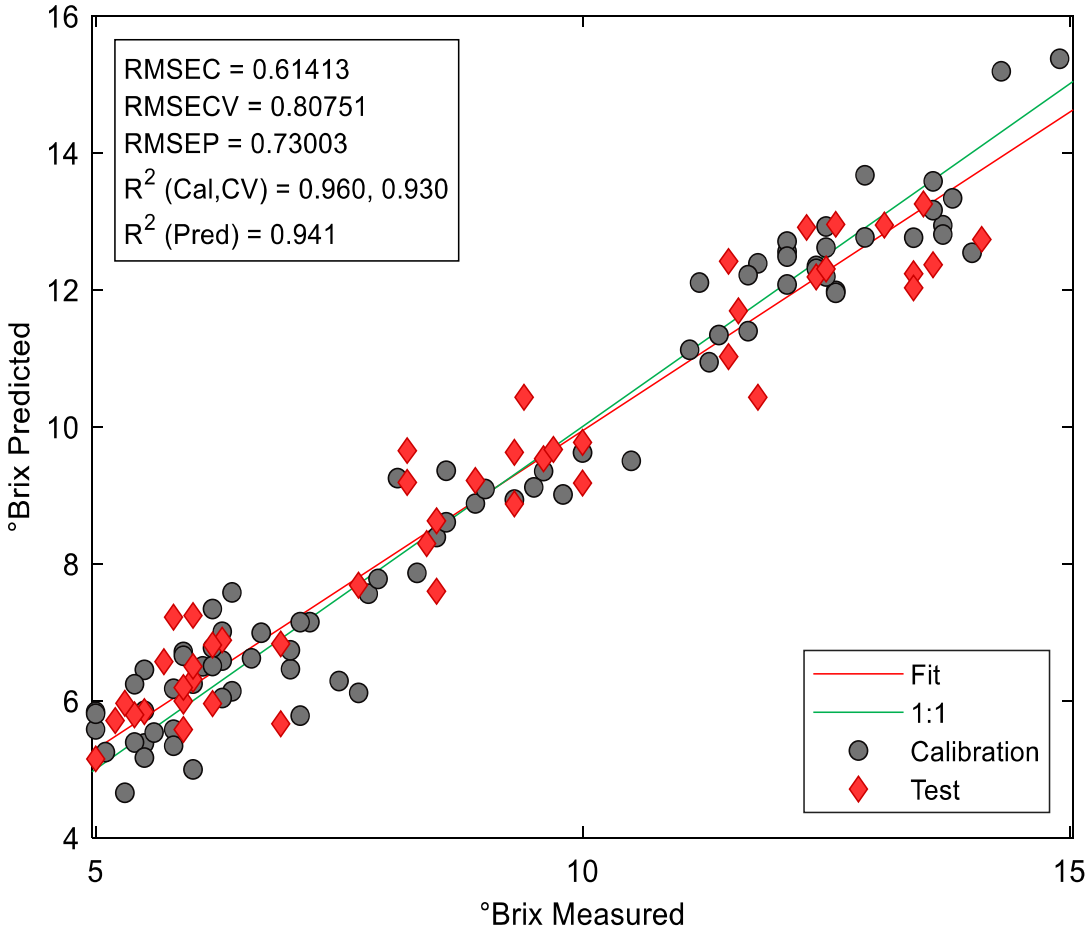


Fig.8

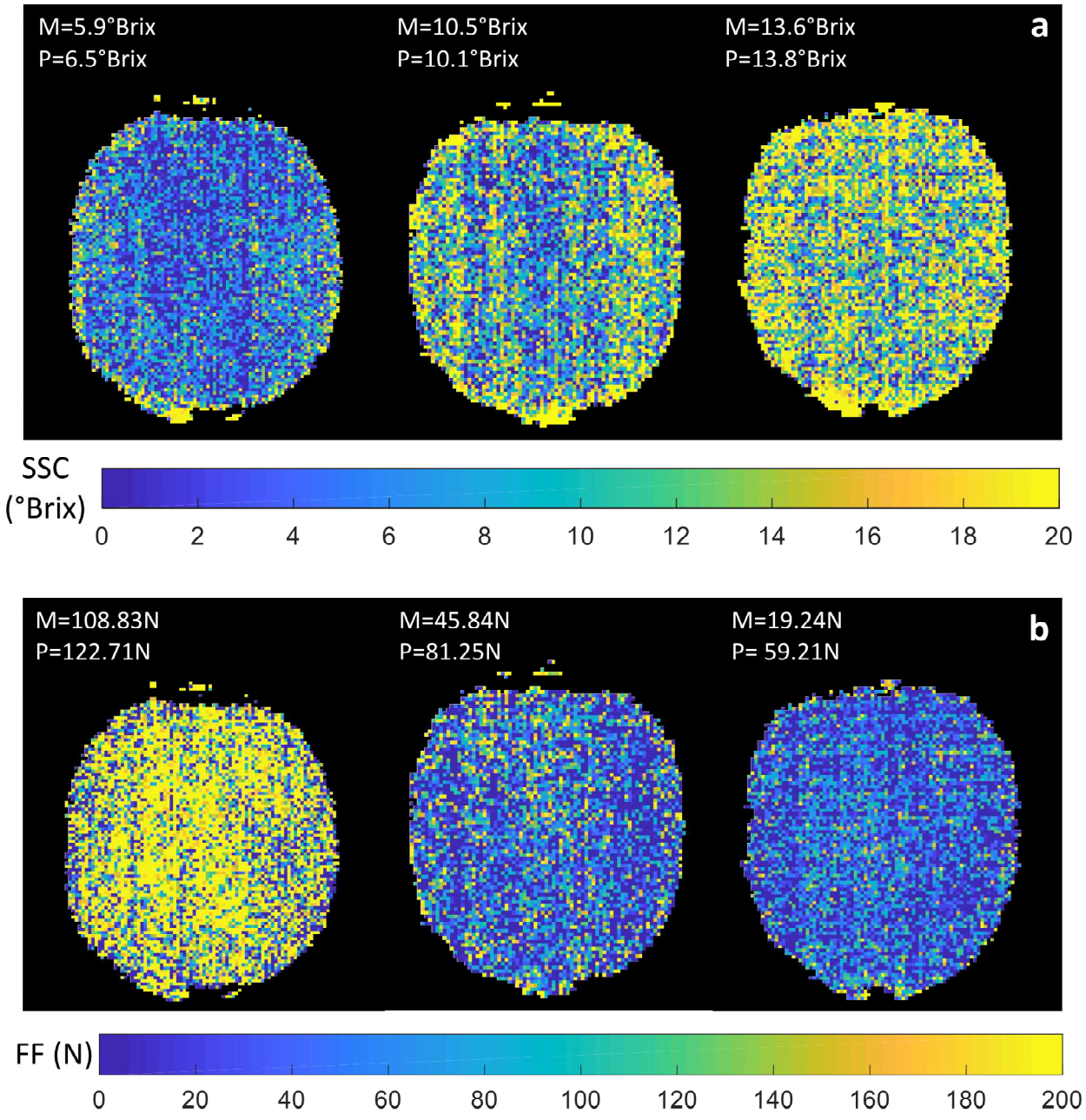


Fig.9

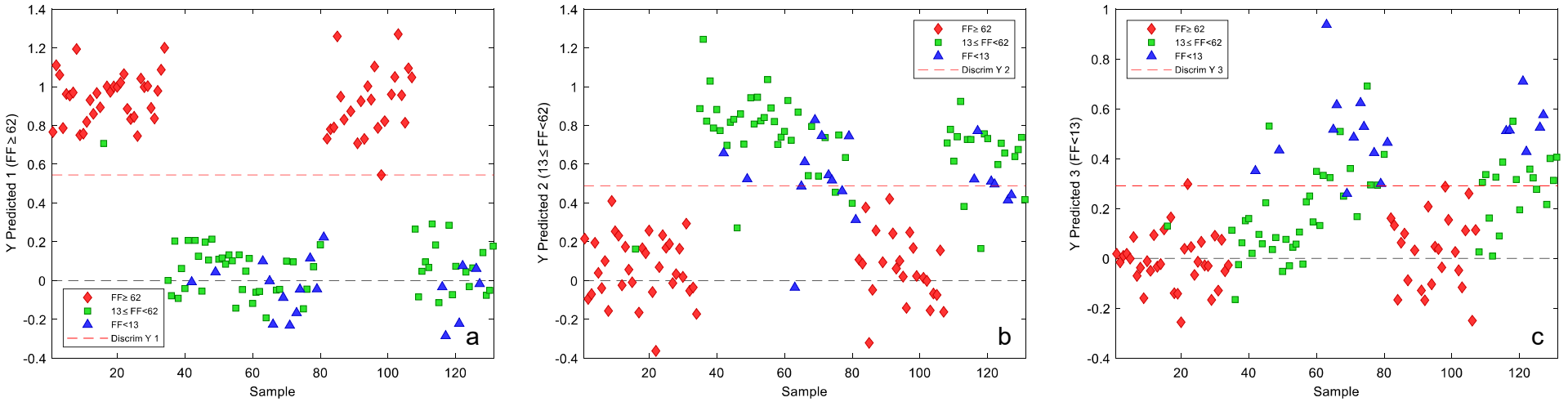


Fig.10

Figure 11

Click here to access/download;Figure;fig.11_rev.docx

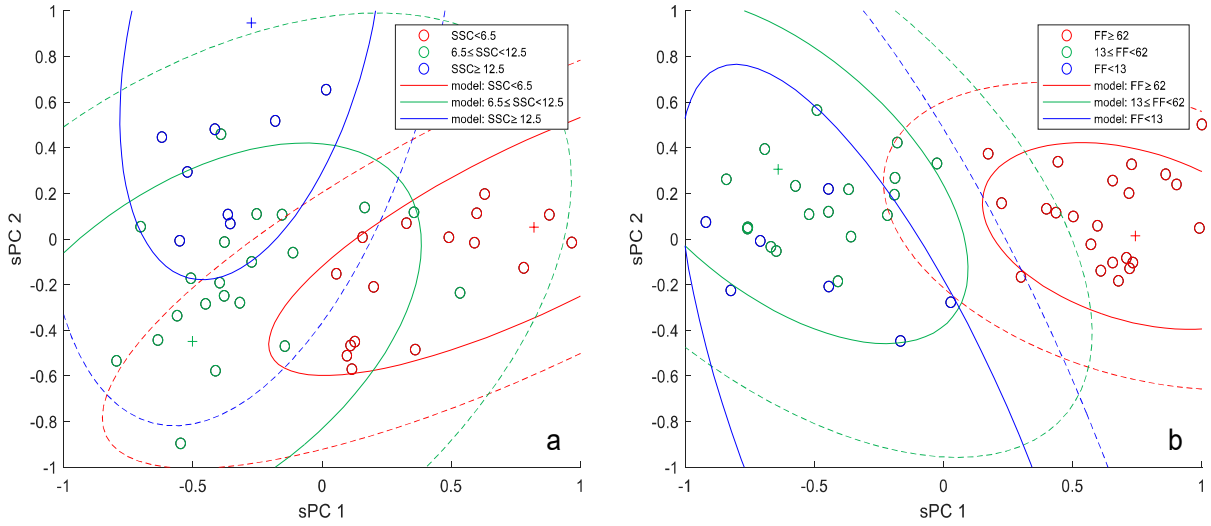


Fig.11

Figure captions

Figure 1. Representation of the lab-scale hyperspectral imaging system used to acquire hyperspectral images of kiwi samples.

Figure 2. (a) RGB representation obtained from a HS image of a kiwi sample; (b) ROI (green) and background (red) obtained by k-means clustering; (c) ROI overlay (semi-transparent green) on the RGB representation in (a).

Figure 3. Mean and standard deviation of quality attributes (SSC and FF) of the sample of kiwifruits during the five days of analysis.

Figure 4. Mean spectra of kiwi samples pre-processed by: (a) SNV (colour scale related to SSC); (b) D1 (colour scale related to FF).

Figure 5. Score plots obtained by PCA according to: (a) SSC; (b) FF. The colour scales are reported on the right of the relative score plots.

Figure 6. GA selections of mean spectra of kiwi samples pre-processed by SNV related to: (a) SSC and (b) FF.

Figure 7. VIP scores of the PLS models developed to predict SSC (a) and FF (b) (spectra pre-processed by SNV).

Figure 8. Measured vs. predicted values of SSC obtained by PLS regression (10 LVs) applying SNV+MC pre-processing, and then selecting the variables with GA.

Figure 9. Prediction maps of SSC (a) and FF (b) of three representative kiwifruits. (M: measured value; P: mean predicted value).

Figure 10. PLS-DA prediction plots for FF applying D1+MC pre-processing and GA variables selection method.

Figure 11. *Soft* PLS-DA prediction plots for :(a) SSC applying SNV+MC pre-processing and GA variable selection method; (b) FF applying D1+MC pre-processing and selecting all variables.

Table 1. Results of the PLS models developed to estimate SSC e FF.

Qualitative index	Spectra pre-processing	Variable selection (nv)	LVs	R^2 RMSE		
				Calibration	Cross-validation	Prediction
SSC (°Brix)	SNV+MC	All (260)	12	0.96 0.59 °Brix	0.92 0.83 °Brix	0.85 1.10 °Brix
		VIP (104)	10	0.92 0.83 °Brix	0.88 1.03 °Brix	0.86 1.14 °Brix
		GA (47)	11	0.96 0.61 °Brix	0.93 0.81 °Brix	0.94 0.73 °Brix
	D1+MC	All (260)	12	0.95 0.66 °Brix	0.90 0.97 °Brix	0.90 0.94 °Brix
		VIP (58)	6	0.90 0.94 °Brix	0.85 1.12 °Brix	0.88 1.08 °Brix
		GA (78)	7	0.92 0.87 °Brix	0.86 1.13 °Brix	0.90 1.00 °Brix
FF (N)	SNV+MC	All (260)	14	0.93 8.56 N	0.86 12.43 N	0.85 13.10 N
		VIP (97)	13	0.91 9.76 N	0.82 14.17 N	0.82 14.51 N
		GA (72)	14	0.94 7.66 N	0.87 11.50 N	0.92 9.87 N
	D1+MC	All (260)	8	0.90 10.02 N	0.85 12.41 N	0.84 13.62
		VIP (41)	10	0.90 10.03 N	0.86 12.30 N	0.86 12.26 N
		GA (75)	10	0.92 9.64 N	0.87 12.26 N	0.89 11.31 N

Note: SSC = soluble solids content; FF = flesh firmness; SNV = standard normal variate; MC = mean centred; D1 = first derivative; nv = number of variables selected; All = all variables; VIP = variable importance in projection; GA = genetic algorithms; LVs = latent variables.

Table 2. Results of PLS-DA and *soft* PLS-DA, in terms of sensitivity, developed to classify the samples according to SSC (class 1 = SSC<6.5; class 2 = 6.5≤SSC<12.5; class 3 = SSC≥12.5).

Spectra pre-processing	Classes of qualitative index	Variable selection (nv)	Sensitivity (%)			
			PLS-DA		Soft PLS-DA	
			Cal.	Pred.	Cal.	Pred.
SNV+MC	class 1	All (260)	97	83	97	67
	class 2		79	62	97	92
	class 3		94	75	100	100
	class 1	VIP (104)	97	62	97	77
	class 2		84	77	92	77
	class 3		100	100	100	100
	class 1	GA (47)	93	77	93	100
	class 2		78	58	92	92
	class 3		78	58	92	92
D1+MC	class 1	All (260)	91	100	94	91
	class 2		77	62	100	92
	class 3		93	70	100	100
	class 1	VIP (58)	86	93	97	100
	class 2		78	68	98	100
	class 3		87	89	100	89
	class 1	GA (78)	94	100	94	100
	class 2		72	71	97	88
	class 3		89	82	100	83

Note: SNV = standard normal variate; MC = mean centred; D1 = first derivative; SSC = soluble solids content; nv = number of variables selected; All = all variables; VIP = variable importance in projection; GA = genetic algorithms; PLS-DA = partial least squares discriminant analysis; Cal. = calibration; Pred. = prediction.

Table 3. Results of PLS-DA and *soft* PLS-DA, in terms of sensitivity, developed to classify the samples according to FF (class 1 = FF≥62; class 2 = 13≤FF<62; class 3 = FF<13).

Spectra pre-processing	Classes of qualitative index	Variable selection (nv)	Sensitivity (%)			
			PLS-DA		Soft PLS-DA	
			Cal.	Pred.	Cal.	Pred.
SNV+MC	class 1	All (260)	100	100	97	100
	class 2		94	86	97	95
	class 3		92	50	100	67
	class 1	VIP (97)	100	95	94	85
	class 2		91	61	94	89
	class 3		100	75	100	100
	class 1	GA (72)	94	100	97	78
	class 2		89	77	97	88
	class 3		83	83	100	100
D1+MC	class 1	All (260)	91	84	100	86
	class 2		72	55	97	94
	class 3		91	93	100	100
	class 1	VIP (41)	97	95	95	89
	class 2		81	77	97	91
	class 3		83	100	92	83
	class 1	GA (75)	100	96	97	78
	class 2		89	83	94	94
	class 3		92	100	100	100

Note: SNV = standard normal variate; MC = mean centred; D1 = first derivative; FF = flesh firmness; nv = number of variables selected; All = all variables; VIP = variable importance in projection; GA = genetic algorithms; PLS-DA = partial least squares discriminant analysis; Cal. = calibration; Pred. = prediction.