

Document downloaded from:

<http://hdl.handle.net/10251/155019>

This paper must be cited as:

Munera-Picazo, S.; Amigo, JM.; Aleixos Borrás, MN.; Talens Oliag, P.; Cubero-García, S.; Blasco Ivars, J. (2018). Potential of VIS-NIR hyperspectral imaging and chemometric methods to identify similar cultivars of nectarine. *Food Control*. 86:1-10.
<https://doi.org/10.1016/j.foodcont.2017.10.037>



The final publication is available at

<https://doi.org/10.1016/j.foodcont.2017.10.037>

Copyright Elsevier

Additional Information

Potential of VIS-NIR hyperspectral imaging and chemometric methods to identify similar cultivars of nectarine

Sandra Munera^a, Jose Manuel Amigo^{b,c}, Nuria Aleixos^d, Pau Talens^e, Sergio Cubero^a, José Blasco^{a*}

^{a)} Centro de Agroingeniería, Instituto Valenciano de Investigaciones Agrarias (IVIA). Carretera CV-315, Km 10.7, 46113 Moncada, Spain. *blasco_josiva@gva.es

^{b)} Department of Food Sciences, Faculty of Science, University of Copenhagen. Rolighedsvej 30, Frederikberg C DK-1958, Denmark.

^{c)} Department of Fundamental Chemistry. Federal University of Pernambuco, Av. Prof. Moraes Rego, 1235 – Cidade Universitária, Recife, Brazil.

^{d)} Departamento de Ingeniería Gráfica. Universitat Politècnica de València. Camino de Vera, s/n, 46022 Valencia, Spain.

^{e)} Departamento de Tecnología de Alimentos, Universitat Politècnica de València. Camino de Vera, s/n, 46022 Valencia, Spain.

ABSTRACT

Product inspection is essential to ensure good quality and to avoid fraud. New nectarine cultivars with similar external appearance but different physicochemical properties may be mixed in the market, causing confusion and rejection among consumers, and consequently affecting sales and prices. Hyperspectral reflectance imaging in the range of 450-1040 nm was studied as a non-destructive method to differentiate two cultivars of nectarines with a very similar appearance but different taste. Partial least squares discriminant analysis (PLS-DA) was used to develop a prediction model to distinguish intact fruits of the cultivars using pixel-wise and mean spectrum approaches, and then the model was projected onto the complete surface of fruits allowing visual inspection. The results indicated that mean spectrum of the fruit was the most accurate method, a correct discrimination rate of 94% being achieved. Wavelength selection reduced the dimensionality of the hyperspectral images using the regression coefficients of the PLS-DA model. An accuracy of 96% was obtained by using 14 optimal wavelengths, whereas colour imaging and a trained inspection panel achieved a rate of correct classification of only 57% of the fruits.

33 **Keywords:** *Stone fruit; quality control; cultivar discrimination; non-destructive; PLS-DA;*
34 *colour analysis; hyperspectral image.*

35

36 **Abbreviations**

37 ANOVA = analysis of variance

38 CCD = charge-coupled device

39 CV = cross validation

40 EU = European Union

41 F = firmness

42 LV = latent variables

43 NIR = near infrared

44 PC = principal component

45 PCA = principal component analysis

46 PLS-DA = partial least square discriminant analysis

47 RGB = red, green, blue

48 SC = skin colour

49 SNV = standard normal variate

50 TA = titratable acidity

51 TSS = total soluble solids

52 VIS = visible

53

54 **1. INTRODUCTION**

55 The surface area of the land devoted to the planting of peaches and nectarines (*Prunus*
56 *persica* L. Batsch) in the EU was around 232000 ha in 2015/16, with a production of nearly 3.7
57 million tons of fruit. Spain is the main producer with around 1.4 tons, which accounts for almost
58 40% of the total EU peach and nectarine production (USDA, 2016). Due to the importance of
59 nectarine (*Prunus persica* L. Batsch var. *nucipersica*) production, it is one of the fruits to which
60 most effort has been devoted by plant breeders in recent years in order to improve agronomic

61 performance, and enhanced fruit appearance and quality (Reig, Alegre, Gatius & Iglesias,
62 2013). This fact has resulted in a significant increase in the number of new cultivars available
63 to fruit growers. These cultivars are similar in appearance but present different sensory
64 properties and therefore different acceptance by the consumer (Iglesias & Echeverría, 2009). In
65 this context, one of the most widely accepted and cultivated nectarine cultivars in Europe is
66 ‘Big Top’ due to its presentation, size, sweet taste and low acidity (Echeverría, Cantín, Ortiz,
67 López & Graell, 2015). However, a stagnation of nectarine consumption is occurring owing to
68 early harvesting, which leads to flavourless fruits being offered with excessive F or irregular
69 quality (Iglesias & Echeverría, 2009). These authors also point out the lack of an adequate
70 identification of the product in the market. The mixture of sweet and acid cultivars on the shelf
71 could lead to consumer rejection, which in turn might affect sales and prices.

72 The internal quality assessment of stone fruits has traditionally been performed by
73 destructive methods, which are contaminating, time-consuming and only a few samples per
74 batch can be monitored (Pérez-Marín, Sánchez, Paz, González-Dugo & Soriano, 2011).
75 Moreover, there is an important lack of classification tools for differentiating cultivars that are
76 very similar to one another. There is therefore a strong need to develop non-destructive and
77 instantaneous methodologies that allow the correct identification of the cultivar in the
78 postharvest stage.

79 Hyperspectral imaging is a computer vision technique which combines conventional two-
80 dimensional digital imagery with spectroscopy to detect spectral features in regions of the
81 electromagnetic spectrum such as the ultraviolet, NIR or infrared regions (Lorente, Aleixos,
82 Gómez-Sanchis, Cubero, García-Navarrete, & Blasco, 2012). This technique is starting to be
83 used as a scientific tool for quality assurance of a wide range of food including bakery products
84 (Erkinbaev, Henderson, & Paliwal, 2017; Verdú, Vásquez, Grau, Ivorra, Sanchez & Barat,
85 2016), meat (Feng, Makino, Oshita, & García Martín, 2017; Iqbal, Sun, & Allen, 2014), or
86 vegetables (López-Maestresalas, Keresztes, Goodarzi, Arazuri, Jaren, & Saeys, 2016). Fruits are
87 of major interest for the use of this technology in the food industry (Keresztes, Goodarzi &
88 Saeys, 2016; Munera, Besada, Aleixos, Talens, Salvador, Sun, Cubero, Blasco, & 2017a).

89 However, due to the high importance of other fruits such as citrus or apples, few scientific
90 studies have been done for quality control of stone fruit quality assessment using hyperspectral
91 imaging. Herrero-Langreo, Lunadei, Lleó, Diezma and Ruiz-Altisent (2011) assessed the
92 ripeness of peaches by using multispectral indexes. Lu and Peng (2006) assessed the F of
93 peaches and Zhu, Lin, Nie, Wu and Chen (2016) obtained F distribution maps inside the peach
94 pulp, while Zhang et al. (2015), Li et al. (2016), Pan et al. (2016) and Sun et al. (2017) detected
95 different types of defects and injuries, including decay. Regarding nectarine, Huang et al. (2015)
96 used the same technique to detect defective features and Munera et al. (2017b) to monitor its
97 ripeness.

98 Hyperspectral imaging generates a huge amount of redundant and frequently highly
99 correlated data that need to be processed (Vélez-Rivera et al., 2014; Sun, Zhang, Liu & Wang,
100 2017). To handle such an amount of data and extract the useful information, it must be assisted
101 by chemometric methods. These methods connect chemical measurements with the essential
102 spectral information in order to classify and/or quantify important characteristics. PCA is one of
103 the most popular methods commonly used both to reduce the dimensionality of data and to
104 obtain an overview of all the relevant information in the dataset. It is an unsupervised projection
105 method which summarises data by forming new independent linear combinations of the original
106 variables (Jolliffe, 2002).

107 PLS-DA is a variant of PLS regression in which the independent variable is categorical,
108 expressing the class membership of the samples. It is performed in order to sharpen the
109 separation between groups of observations by maximising the covariance between the spectra
110 and the independent variable such that a maximum separation among classes is obtained.
111 Furthermore, it is commonly used to understand which variables contain the discriminating
112 information (Lorente et al., 2012). Some examples of the use of this method include the
113 detection of decay lesions in citrus fruits (Folch-Fortuny, Prats-Montalbán, Cubero, Blasco, &
114 Ferrer, 2016), classification of oat kernels (Serranti, Cesare, Marini, & Bonifazi, 2013), the
115 classification of edible fennel heads based on the harvest time (Amodio, Capotorto, Chaudhry,

116 & Colelli, 2017), and the examination of aflatoxin on corn kernels (Kandpal, Lee, Kim, Bae, &
117 Cho, 2015).

118 In this paper, we put forward a novel approach based on VIS-NIR hyperspectral imaging
119 and chemometric methods to develop statistical predictive models capable of distinguishing
120 cultivars of nectarines with a very similar appearance but different taste. Previous studies have
121 been conducted to differentiate among nectarine cultivars using colour images (Font et al.,
122 2014). However, they use fruits with clearly different appearance. In this work, ‘Diamond Ray’
123 and ‘Big Top’ cultivars have been used due to their similar skin and flesh appearance.
124 Furthermore, these cultivars are grown and marketed at the same time and become a problem
125 for producers when they are mixed, either accidentally or intentionally, in the market.

126 In addition, using the spectral and spatial information provided by the hyperspectral
127 images, two approaches are further investigated: the first based on the analysis of the individual
128 spectrum of each pixel and the second based on the mean spectrum of each fruit. Finally,
129 visualisation of the result of the classification model over the images of nectarines is proposed
130 to establish a practical tool for nectarine classification in the packing houses.

131

132 **2. MATERIAL AND METHODS**

133 **2.1. Fruit samples**

134 Nectarines cv. ‘Diamond Ray’ and ‘Big Top’ were selected as reference cultivars of sweet
135 and acid cultivars, respectively (Reig, Iglesias & Echeverría, 2009), due to their similar skin and
136 flesh appearance. These two cultivars are difficult to distinguish by the naked eye, which is
137 problematic for producers when they are mixed in the market.

138 Fruits were harvested in a commercial orchard in Lerida (Spain) at the commercial maturity
139 stage in the summer season of 2016. A total of 125 fruits of each cultivar without defects or
140 bruises were selected and stored under controlled conditions (1 °C; 90% relative humidity) in
141 order to avoid the further ripening of either cultivar during the experiment.

142

143 **2.1. Hyperspectral image acquisition and processing**

144 The hyperspectral imaging system consisted of an industrial camera (CoolSNAP ES,
 145 Photometrics, AZ, USA), coupled to two liquid-crystal tuneable filters (Varispec VIS-07 and
 146 NIR-07, Cambridge Research & Instrumentation, Inc., MA, USA). The camera was configured
 147 to acquire images with a size of 1392 x 1040 pixels and a spatial resolution of 0.14 mm/pixel at
 148 60 different wavelengths every 10 nm, in the working spectral range of 450 nm – 1040 nm. In
 149 order to avoid problems of unfocused images due to the refraction of light across this wide
 150 spectral range, the focus was adjusted on the central band of the acquisition interval (740 nm)
 151 and the images were captured using lenses capable of covering the whole spectral range without
 152 going out of focus (Xenoplan 1.4/23, Schneider Optics, Hauppauge, NY, USA). To optimise the
 153 dynamic range of the camera, prevent saturated images and correct the spectral sensitivity of the
 154 different elements of the system, a calibration of the integration time of each band was
 155 performed by capturing the averaged grey level of a white reference target (Spectralon 99%,
 156 Labsphere, Inc, NH, USA) corresponding to 90% of the dynamic range of the camera.

157 The scene was illuminated by indirect light from twelve halogen spotlights (37 W) (Eurostar
 158 IR Halogen MR16. Ushio America, Inc., CA, USA) powered by direct current (12 V) and
 159 arranged equidistant from each other inside a hemispherical aluminium diffuser. The inner
 160 surface of the aluminium diffuser was painted white with a rough texture to maximise its
 161 reflectivity, the rough texture being applied in order to minimise directional reflections, which
 162 could cause bright spots, thus resulting in highly homogeneous light.

163 The fruits were introduced manually into a fruit holder, with the stem-calyx axis lying
 164 horizontal. Two images of each fruit were acquired using customised software developed at
 165 IVIA. A total of 250 images of each cultivar were imported into MATLAB R2015a (The
 166 MathWorks, Inc. MA, USA) to be pre-processed using the customised toolbox HYPER-Tools
 167 (Amigo, Babamoradia & Elcoroaristizabal, 2015).

168 The image processing started with the correction of the relative reflectance by using
 169 equation (1) (Gat, 2000):

$$170 \quad \rho_{xy}(x, y, \lambda) = \frac{R_{white}^{abs}}{R_{white}^{abs}} = \rho^{Ref}(\lambda) \frac{R(x,y,\lambda) - R_{black}(x,y,\lambda)}{R_{white}(x,y,\lambda) - R_{black}(x,y,\lambda)} \quad (1)$$

171 where $\rho^{Ref}(\lambda)$ is the standard reflectance of the white reference target (99% in this work),
172 $R(x,y,\lambda)$ is the reflectance of the fruit captured by the CCD sensor of the camera, $R_{white}(x,y,\lambda)$ is
173 the reflectance captured by the CCD of the white reference target, and $R_{black}(x,y,\lambda)$ is the
174 reflectance captured by the CCD while avoiding any light source in order to quantify the
175 electronic noise of the CCD. The images were then clipped and spatially compressed to reduce
176 the computation time, and a proper removal of the background was performed using K-means
177 clustering. Thus, the relative reflectance spectrum of all the pixels in each fruit image was
178 extracted.

179

180 **2.2. Colour image acquisition and processing**

181 Before image acquisition, the SC was analysed to obtain the L^* , a^* and b^* colour
182 coordinates (CIELAB colour space) of each fruit, also using a colorimeter (MINOLTACM-
183 700d, Minolta Co. Tokyo, Japan) configured with the standard illuminant D65 and the observer
184 10° . The SC was obtained as the average of the values of two measurements, one in the blush
185 zone (reddish colour) and another in the ground zone (yellowish colour).

186 The colour imaging system consisted of a digital camera (EOS 550D, Canon Inc, Japan)
187 arranged inside a square inspection chamber that included a calibrated and uniform illumination
188 system composed of four lamps, each containing two fluorescent tubes BIOLUX 18W/965
189 (Osram GmbH, Germany) with a colour temperature of 6500 K. The angle between the axis of
190 the lens and the sources of illumination was approximately 45° , and polarising filters were
191 placed in front of the lamps and in the camera lenses to eliminate specular bright spots that
192 could alter the true colour.

193 The fruits were introduced manually upon a fruit holder, with the stem-calyx axis lying
194 horizontal. Two images were acquired for each fruit, corresponding to each of the two sides
195 delimited by the suture of the fruit. Then, a total of 250 images of each cultivar were imported
196 into customised software developed at IVIA (FoodImage-Inspector v4.0, freely available at
197 <http://www.cofilab.com>, Spain) to analyse the SC and to obtain the percentage of the reddish and
198 yellowish zones on the fruit. This segmentation was based on the Bayes theorem to assign all

199 the pixels in the image to the two classes used in a previous training. The RGB colour
200 coordinates of the acquired images were converted to the L^* a^* b^* coordinates and then
201 corrected using a colour reference target (ColorChecker Digital SG, X-Rite, MI, USA).

202

203 **2.3 Visual analysis with trained panel**

204 The panel was composed of five panellists, ages 29 to 50 years (three male and two female),
205 with expertise in fruit quality and marketing. The panellists were trained using 20 colour images
206 of nectarines of the calibration set (10 from each cultivar chosen at random). A total of 40
207 colour images of fruits of the validation set (20 from each cultivar chosen at random) were
208 presented with randomised order to each panellist to be classified as belonging to the ‘Diamond
209 Ray’ or ‘Big Top’ cultivar.

210

211 **2.4 Reference analysis**

212 The characterisation of the physicochemical properties of the samples using reference
213 methods was performed immediately after the acquisition of the images. F was registered on
214 opposite sides of the fruits using an XT2 Stable texturometer (MicroSystems Haslemere, UK)
215 equipped with a 6 mm flat plunger. The crosshead speed during the puncture test was 1 mm/s.
216 The maximum force was expressed in Newton (N). Immediately after SC and F measurements,
217 a juice sample was taken from each fruit for TSS and TA measurements. TSS were determined
218 using a digital refractometer RFM330+VWR (Internacional Eurolab S.L., Barcelona, Spain) at
219 20 °C and results were expressed as percentage of TSS. TA was determined using a Crison pH-
220 Burette 24 automatic titrator (Crison, Barcelona, Spain) and NaOH 0.5 N, according to standard
221 UNE34211:1981 (AENOR, 1981). The results were expressed as the percentage of malic acid.

222 The ANOVA was conducted using the software Statgraphics (Manugistics Corp., Rockville,
223 USA) in order to determine significant differences in the physicochemical properties (F, TSS,
224 TA and L^* , a^* and b^* colour coordinates) between cultivars.

225

226 **2.4 Chemometric methods**

227 To identify both nectarine cultivars with high precision, two approaches were studied for
228 setting up the classification models: i) including in the model the individual spectrum of each
229 pixel in the nectarine image, and ii) using only the mean spectrum of all the pixels
230 corresponding to each fruit. Thus, 512828 pixel spectra were used in the first approach, and the
231 mean spectra of 500 fruits were used for the second. The data of all the fruits of both cultivars
232 were collected and randomly partitioned into two sets: two thirds of the samples were used to
233 calibrate the models (calibration set) and for cross-validation, while the remaining third was
234 used for independent test prediction (validation set).

235 Both the directly acquired spectrum of each pixel and that obtained as an average for each
236 fruit were pre-processed using SNV in order to reduce the physical variability between samples
237 due to light scatter (Rinnan et al., 2009). This correction was performed using the equation (2):

$$238 \quad x_{corr} = \frac{x_{org} - a_0}{a_1} \quad (2)$$

239 Where x_{corr} and x_{org} are the corrected and raw spectra, respectively, a_0 is the average value of the
240 sample spectrum to be corrected and a_1 is the standard deviation of the sample spectrum.

241 Later, mean centring was applied to normalise the full spectrum. Multivariate analyses were
242 then performed using the PLS_Toolbox (Eigenvector Research Inc., USA) and the HYPER-
243 Tools toolbox (Amigo et al., 2015) both working under MATLAB R2015a.

244 PCA was used to explore the differences between the two cultivars using the pixel and mean
245 spectra of the calibration set previously pre-processed by means of SNV and mean centring.
246 Later, PLS-DA models were built to sort the fruits into one of the two studied cultivars. The
247 models were also calibrated using the pre-processed pixel and mean spectra of the calibration
248 set and tested using only samples of the validation or prediction set.

249 In order to compare the performance of the hyperspectral imaging with the colour imaging
250 system, a PLS model was also built using the mean value of the $L^*a^*b^*$ colour coordinates.

251 A single 10-fold venetian blind CV was used to choose the optimal number of LV as well as
252 to obtain an estimation of the error rate of the models. All models were statistically validated by
253 using the sensitivity, specificity, class error and accuracy (Eq. 3, 4, 5 and 6):

$$254 \quad \text{Sensitivity} = \frac{TP}{TP+FN} \quad (3)$$

$$255 \quad \text{Specificity} = \frac{TN}{TN+FP} \quad (4)$$

$$256 \quad \text{Class error} = 1 - \left(\frac{\text{Sensitivity} + \text{Specificity}}{2} \right) \quad (5)$$

$$257 \quad \text{Accuracy (\%)} = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (6)$$

258 where TP and TN stand for true positive and true negative, respectively, accounting for the
259 samples that have been correctly assigned as belonging (TP), or not belonging (TN), to a
260 specific class. FP and FN stand for false positive and false negative, respectively, accounting for
261 the samples that have been wrongly assigned as belonging (FP), or not belonging (FN), to a
262 specific class.

263 The ANOVA, using the software Statgraphics, was also conducted in order to determine
264 significant differences in the accuracy of the models.

265

266 **3. RESULTS AND DISCUSSION**

267 **3.1 Cultivar characterisation**

268 **3.1.1 Physicochemical properties**

269 Table 1 shows the results obtained from the reference analysis of the physicochemical
270 properties. F is one of the physicochemical properties commonly used to assess ripeness. In this
271 work, the measures of F obtained for both cultivars showed no statistical differences, which
272 means that they were in a similar stage of ripeness. According to the mean value of F measured
273 for each cultivar, these fruits were considered as being within the group that Valero, Crisosto
274 and Slaughter (2007) described as 'ready to buy'.

275 As noted above, the principal difference between these two cultivars is the flavour; i.e. the
276 typical TSS values for 'Big Top' being higher than in 'Diamond Ray' and vice versa for TA.

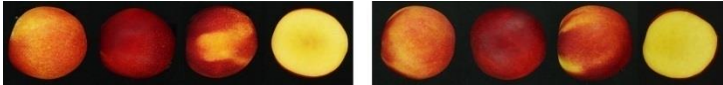
277 The measured values (Table 1) agreed with Crisosto et al. (2006), who found that ‘Diamond
 278 Ray’ had 0.8% TA and 10.3% TSS. The difference in TSS content between these cultivars may
 279 be attributable to the stage of maturity, the season or the production area (Crisosto, 1994).
 280 Regarding the ‘Big Top’ cultivar, Giné-Bordonaba et al. (2014) reported results similar to those
 281 in the present study, i.e. 0.3% TA and TSS between 12.2% and 13.5%.

282

283 Table 1. Results of analysis of physicochemical properties of both cultivars of nectarine

Property	‘Diamond Ray’		‘Big Top’		
	Mean	SD	Mean	SD	
Firmness (N)	33.8 ^a	9.5	34.8 ^a	7.1	
Total soluble solids (%)	11.9 ^b	1.6	12.7 ^a	2.3	
Titratable acidity (%)	0.7 ^a	0.1	0.4 ^b	0.1	
Skin colour by colorimeter	<i>L</i> *	36.9 ^a	6.6	36.5 ^a	6.0
	<i>a</i> *	27.0 ^a	4.2	26.2 ^a	3.9
	<i>b</i> *	13.3 ^a	5.1	13.4 ^a	4.9
Skin colour by imaging	<i>L</i> *	28.0 ^a	8.6	27.0 ^a	8.2
	<i>a</i> *	44.9 ^a	5.4	41.0 ^b	5.5
	<i>b</i> *	27.2 ^a	8.8	24.8 ^b	8.6
	Blush zone (%)	67.0 ^a	21.4	66.3 ^a	18.4
	Ground zone (%)	33.0 ^a	21.4	33.7 ^a	18.4

External and internal appearance



284 *Different superscript letters in the same row indicate significant differences between cultivars (p-*
 285 *value<0.05). SD = standard deviation;*
 286

287 The mean *L**, *a**, and *b** colour coordinates of the SC using the colorimeter were not
 288 statistically different between cultivars (Table 1). However, colorimeters measure small regions
 289 only, which can be a major limitation in applications where distinguishing the colours all over
 290 the sample is of interest. This means that they are not well suited to measuring objects with a
 291 heterogeneous colour (Gardner, 2007), such as nectarines of these cultivars. However, a colour
 292 camera provides images in which the colours of the pixels are determined individually (Cubero,
 293 Aleixos, Moltó, Gómez-Sanchis & Blasco, 2011), along with their spatial distribution. The
 294 analysis of the colour of the nectarines using imaging enable the evaluation of the SC of the
 295 different colour zones separately and calculation of the relative distribution (percentage) of
 296 reddish or yellowish colour in the whole fruit.

297 Using this percentage, a mean value of the L^* , a^* , b^* coordinates was calculated from the
298 images. On average, a reddish colour was present on 67% of the fruit surface and a yellowish
299 colour on 33% in both cultivars (Table 1). Even so, the mean colour using imaging indicated
300 that the a^* and b^* scores were statistically different in the two cultivars, i.e. both were higher in
301 ‘Diamond Ray’. However, the differences were too small to be detected visually by the human
302 eye, especially during a rapid fruit-sorting process.

303

304 **3.1.2 Spectral analysis**

305 Differences between cultivars were observed in their hyperspectral spectra (Fig. 1). The pre-
306 processed (SNV) mean spectra of the two cultivars followed a similar spectral pattern but had
307 clear differences at specific wavelengths.

308 In the VIS region, no apparent differences could be visualised in the range between 400-
309 600 nm where carotenoids are present. In contrast, the ‘Big Top’ cultivar had lower reflectance
310 (higher absorbance) than ‘Diamond Ray’ near 680 nm, which is associated with chlorophylls
311 (Lleó et al., 2011; Rajkumar, Wang, Elmasry, Raghavan & Gariepy, 2012), suggesting a higher
312 content of this molecule. This agrees with the differences in the values of a^* and b^* found in the
313 colour analysis (Table 1).

314 In the NIR region, the absorption bands for acids and sugars are usually found around 800
315 nm and 840 nm respectively, attributable to the hydroxyl groups of these compounds (Malegori
316 et al., 2017; Yang, Sun, Pu, Wang & Zhu, 2015). However, only small differences are usually
317 observable due to the water absorption bands which dominate the spectrum (Nicolai et al.,
318 2007). In this region, the main differences observed in the spectra were at wavelengths above
319 850 nm and, in particular, around 970 nm, where Lu & Peng (2006) described a peak associated
320 with water absorption, which in this case was more pronounced in ‘Big Top’ nectarines (Fig. 1).

321

322 **3.2. Overview of the spectral data**

323 A PCA was performed in order to obtain an overview of the distribution of the spectral data
324 information from the samples of both cultivars. The PCA results from the individual pixel
325 spectra and the mean spectra of each fruit are shown in Figures 2 and 3, respectively.

326 Forty samples of each cultivar were randomly selected to provide individual pixel spectra
327 and this data was used to generate a score image plot. The first two PCs explained 87.8% of the
328 total variance (76.5% and 11.3%, respectively). The variations in the colour within each fruit
329 showed the distribution or content of the biochemical constituents. A possible trend was
330 discerned in PC2, where pixels with low values (dark blue) were found mostly in ‘Big Top’
331 samples; however, there was little difference in individual fruit spectra of the ‘Diamond Ray’
332 and ‘Big Top’ cultivars.

333 In the PCA of the mean spectra of the calibration set, the first two PCs (Fig. 3A) explained
334 93.3% of the variance (81.4% and 11.9%, respectively). The ellipses for the two cultivars
335 appeared distinct, but discrimination between them was not possible because of overlap (Fig.
336 3A).

337 Although the loadings obtained for PC1 and PC2 (Fig. 3B) might have offered information
338 on the most important wavelengths to distinguish the cultivars, this was not useful because
339 separation was not evident in the preceding plot (Fig. 3A). PCA maximises the variance in the
340 first components, which may or may not be related to the segregation of the classes; this does
341 not guarantee the class separability of data due to its unsupervised nature (Jolliffe, 2002).

342

343

344 **3.3 Cultivar classification using individual pixel spectrum**

345 A PLS-DA model was performed using the spectral range of 450–1040 nm and the
346 spectrum of the individual pixels of each fruit of the calibration set. The values obtained for
347 sensitivity and specificity (Table 2) indicated that the number of samples correctly identified as
348 belonging to a specific cultivar, or not, was above 0.80 in the CV set, using five LV. Sensitivity

349 of 0.83 and 0.86 was determined for ‘Diamond Ray’ and ‘Big Top’ respectively being the
 350 accuracy of classification 84.8% and error 0.15.

351 Using the spatial data collected by the imaging system the combined results were applied to
 352 the calibration set. The predicted class of each pixel was obtained by introducing the spectrum
 353 measured for those pixels into the previously built model, and visualising the result. Each pixel
 354 was coloured blue if it was assigned to ‘Diamond Ray’ or red if it was assigned to ‘Big Top’, as
 355 shown in Figure 4A. The accuracy of this classification was 83.8% and error 0.16.

356 To classify each fruit using this approach, the whole fruit was assigned to the class found in
 357 the majority of its pixels (Fig. 4B). In this case, the accuracy and the classification error were
 358 84.4% and 0.16. In both cases, ‘Big Top’ was also the best discriminated, with a sensitivity of
 359 about 0.90.

360

361 Table 2. Cultivar discrimination using the pixel spectrum approach

#V	#LV	Set	Class	Sensitivity	Specificity	Error	Accuracy (%)
60	5	Calibration	‘DR’	0.83	0.86	0.15	84.8
			‘BT’	0.86	0.83		
		Cross Validation	‘DR’	0.83	0.86	0.15	84.8
			‘BT’	0.86	0.83		
		Validation pixel	‘DR’	0.79	0.89	0.16	83.8
			‘BT’	0.89	0.79		
Validation object	‘DR’	0.78	0.91	0.16	84.4		
	‘BT’	0.91	0.78				

362 *V=Variables; LV=Latent variables; ‘DR’=Diamond Ray; ‘BT’=‘Big Top’*

363

364

365

366 3.4 Cultivar classification using mean fruit spectrum

367 The sensitivity and specificity in the results of calibration using the mean fruit spectrum
 368 approach giving values above 0.90 using six LV (Table 3). In this case, both cultivars were
 369 discriminated similarly and the accuracy of classification of the CV was 93.2% and error 0.07.

370 In order to get a graphical view of the veracity of the classification obtained using the
 371 validation set, the class for each fruit was predicted by introducing the mean spectrum measured
 372 into the previously built model. The result was visualised showing the fruit coloured blue if the

373 mean value was assigned by the model to ‘Diamond Ray’ or red if it was assigned to ‘Big Top’
 374 (Fig. 5A). The results for the validation set were similar to those obtained in the calibration,
 375 showing an accuracy of 94.4% with a classification error of 0.06. The ANOVA results indicated
 376 that the mean spectra model was significantly better than the pixel model ($p < 0.05$) to classify
 377 the fruits

378 As Williams and Kucheryavskiy (2016) pointed out, using properly computed object
 379 features as the mean spectrum decreases the amount of data, leading to more stable
 380 classification models. Furthermore, this approach avoids classifying by pixels when objects
 381 from different classes contain many similar pixels and are easily miss-assigned to the opposite
 382 class, such as for the cultivars studied in this work. On the other hand, it is important to include
 383 the negative influence of the sphericity of the fruits on the reflectance of the light. As seen in
 384 Figure 4A, most errors occur at the borders of the fruit, since the centres are usually well
 385 illuminated. The pixels near the borders are therefore more likely to be wrongly classified, thus
 386 affecting the overall result. In contrast when using the mean fruit spectrum, the averaging
 387 minimises these errors.

388

389

Table 3. Cultivar discrimination using the mean spectrum approach

#V	#LV	Set	Class	Sensitivity	Specificity	Error	Accuracy (%)
60	6	Calibration	‘DR’	0.94	0.94	0.06	93.8
			‘BT’	0.94	0.94		
		Cross Validation	‘DR’	0.93	0.94	0.07	93.2
			‘BT’	0.94	0.93		
		Validation	‘DR’	0.94	0.94	0.06	94.4
			‘BT’	0.94	0.94		

390

V=Variables; LV=Latent variables; ‘DR’=Diamond Ray; ‘BT’=‘Big Top’

391

392 3.5 Selection of the optimal wavelengths

393 In order to optimise the algorithms for an automatic in-line sorting system working at high
 394 speed, it is important to reduce the computational complexity generated by the huge amount of
 395 data obtained by hyperspectral imaging systems. This problem is commonly alleviated by
 396 techniques that retain the information in the few bands that reveal the most variability and
 397 therefore most significant information in the hyperspectral image (Du & Sun, 2006). The

398 method used in this study was the vector of the regression coefficients. This measures the
 399 association between each variable and the response and selects variables in two steps: (i) the
 400 PLS-DA model is fitted to the data, and (ii) the variable selection is based on a threshold
 401 (Mehmood, Liland, Snipen & Sæbø, 2012). Variables with a high absolute value can be selected
 402 because they make the highest contribution to cultivar classification and those with a small
 403 absolute value can be ignored. In this study, the regression coefficients were obtained from the
 404 PLS-DA model using the mean fruit spectrum approach, due to its higher accuracy in the
 405 classification of both cultivars.

406 Figure 6 shows the vector of regression coefficients. Those peaks where the absolute value
 407 was highest were selected as important wavelengths. In the VIS region the selected wavelengths
 408 were at 630, 650, 680 and 720 nm while in the NIR region they were 750-770, 790, 810-840,
 409 860 and 900 nm.

410 The optimised PLS-DA model was performed using the 14 selected wavelengths as input.
 411 The sensitivities and specificities in the CV were similar to the full model using six LV (Table
 412 4). In the prediction set, using only the 14 wavelengths, the sensitivity for the two cultivars
 413 increased from 0.94 for both to 0.95 and 0.98, in ‘Diamond Ray’ and ‘Big Top’ respectively.
 414 Figure 5 shows the results of both classifications, using the full spectrum (Fig. 5A) and the
 415 optimal wavelengths (Fig. 5B) in which more fruits were coloured as they should be when the
 416 wavelengths selected as the most important. However, the accuracy obtained, 96.3%, was not
 417 statistically different ($p > 0.05$) from the accuracy of the full model (96.3 and 94.4%,
 418 respectively)

419

420 Table 4. Cultivar discrimination using the mean spectrum and the optimal wavelengths methods

#V	#LV	Set	Class	Sensitivity	Specificity	Error	Accuracy (%)
14	6	Calibration	‘DR’	0.94	0.94	0.06	93.8
			‘BT’	0.94	0.94		
		Cross Validation	‘DR’	0.93	0.94	0.07	93.2
			‘BT’	0.94	0.93		
		Validation	‘DR’	0.95	0.98	0.04	96.3
			‘BT’	0.98	0.95		

421 *V=Variables; LV=Latent variables; ‘DR’=Diamond Ray; ‘BT’=‘Big Top’*

422

423 **3.6 Hyperspectral imaging vs. colour and visual analysis**

424 When the validation set was classified visually by the trained panel, the same fraction of
425 each cultivar was identified correctly (Table 5). However, the accuracy was very low, i.e. 54.5%
426 with a classification error of 0.46. This demonstrates difficulty of the human eye to distinguish
427 between the similar external appearances of these cultivars.

428 Classification by the colour data had similar accuracy (p-value > 0.05) to that achieved by
429 the trained panel (Table 5), i.e. 56.9% accuracy and error of 0.43. This is especially poor in
430 comparison with the results of the hyperspectral imaging using 14 wavelengths, i.e. 96.3%,
431 error 0.04 (Table 4).

432 These results are in agreement with the work carried out by Nogales-Bueno, Rodríguez-
433 Pulido, Heredia and Hernández-Hierro (2015) that used NIR hyperspectral and colour imaging
434 to discriminate between four red grape cultivars. Only 52% of the samples were correctly
435 classified using colour imaging but this figure increased to 86% using hyperspectral imaging.
436 Furthermore, Font et al. (2014) described an in-line system for verification of nectarine cultivars
437 with close harvest times using different colour space layers of the SC histogram. The success of
438 their technique was 100% in comparing fruits of three cultivars with a single cultivar for
439 reference. In the same experiments, human classification achieved 86% accuracy, likely
440 attributable to the large differences in the SC of the cultivars tested.

441 The high rate of accuracy in classification of these cultivars using hyperspectral imaging
442 was important because of the external similarity of the cultivars studied. This makes it difficult
443 to accurately identify the cultivars by colour features, although they appear very different to
444 consumers at the table. This is a genuine problem for the industry. Although colour imaging is a
445 rapid and inexpensive tool, it has lower discrimination power for cultivars with very similar
446 appearance, which necessitates the use of more VIS wavelengths and optimal wavelengths in
447 the NIR region.

448

449

Table 5. Cultivar discrimination using colour imaging and by a trained panel

	#V	#LV	Set	Class	Sensitivity	Specificity	Error	Accuracy (%)
Colour imaging (PLS-DA)	3	2	Calibration	'DR'	0.75	0.61	0.32	68.0
				'BT'	0.61	0.75		
			Cross Validation	'DR'	0.75	0.62	0.32	68.3
				'BT'	0.62	0.75		
Validation	'DR'	0.65	0.49	0.43	56.9			
	'BT'	0.49	0.65					
Trained panel	-	-	Validation	'DR'	0.54	0.55	0.46	54.5
				'BT'	0.55	0.54		

450

V=Variables; LV=Latent variables; 'DR'=Diamond Ray; 'BT'='Big Top'

451

452

453

4. CONCLUSIONS

454

The capability of VIS-NIR hyperspectral imaging to discriminate very similar cultivars of nectarine has been demonstrated in this work.

455

456

The classification of these two cultivars by colour imaging or by a trained panel was very poor, achieving an accuracy of only 56.9% and 54.5% respectively. However, hyperspectral imaging supported by chemometric methods and optimised by reduction of the spectral and spatial information enabled classification more accurately than by traditional manual or colour-based systems, and it is also faster than destructive methods.

457

458

459

460

461

The use of the mean spectrum of the fruit as input of the predictive models provided classification accuracy of 94.4%. To cope with the huge amount of data captured by the hyperspectral systems, the vector of the regression coefficients of a PLS-DA model identified 14 wavelengths which were selected as optimal, producing the best classification model with a classification accuracy of 96.3%.

462

463

464

465

466

This technique may have potential as a tool for rapid and non-destructive cultivar discrimination, allowing the selection of fruit that is better suited to the consumer's preferences. Nevertheless, the results of this study should be confirmed on a larger sample set of fruits grown in different areas and harvested at different stages of ripeness before they can be implemented in an in-line system.

467

468

469

470

471

472 **ACKNOWLEDGEMENTS**

473 This work was partially funded by INIA and FEDER funds through project RTA2015-
474 00078-00-00. Sandra Munera thanks INIA for the FPI-INIA grant num. 43 (CPR2014-0082),
475 partially supported by European Union FSE funds. The authors wish to thank Fruits de Ponent
476 (Lleida) for providing the fruit.

477

478 **REFERENCES**

- 479 AENOR. (1981). Productos derivados de frutas y verduras, determinación de la acidez
480 valorable. UNE 34211: 1981
- 481 Amigo, J.M., Babamoradia, H. & Elcoroaristizabal, S. (2015). Hyperspectral image analysis. A
482 tutorial. *Analytica Chimica Acta* 896, 34–51.
- 483 Amodio, M.L., Capotorto, I., Chaudhry, M.M.A. & Colelli, G. (2017). The use of hyperspectral
484 imaging to predict the distribution of internal constituents and to classify edible fennel heads
485 based on the harvest time. *Computers and Electronics in Agriculture* 134, 1-10
- 486 Crisosto, C.H. (1994). Stone fruit maturity indices: a descriptive. *Postharvest News and*
487 *Information* 6, 65-68.
- 488 Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J. & Blasco, J. (2011). Advances in
489 Machine Vision Applications for Automatic Inspection and Quality Evaluation of Fruits and
490 Vegetables. *Food Bioprocess Technology* 4, 487–504
- 491 Du, C.J. & Sun, D.W. (2006). Learning techniques used in computer vision for food quality
492 evaluation: a review. *Journal of Food Engineering*, 72, 39–55.
- 493 Echeverría, G., Cantín, C.M., Ortiz, A., López, M.L. & Graell, J. (2015). The impact of maturity,
494 storage temperature and storage duration on sensory quality and consumer satisfaction of
495 ‘Big Top’ nectarines. *Scientia Horticulturae* 190, 179-186.
- 496 Erkinbaev, C., Henderson, K., & Paliwal, J. (2017). Discrimination of gluten-free oats from
497 contaminants using near infrared hyperspectral imaging technique, *Food Control* 80, 197-
498 203.

499 Feng, C.H., Makino, Y., Oshita, S., & García-Martín J.F. (2017). Hyperspectral Imaging and
500 Multispectral Imaging as the Novel Techniques for Detecting Defects in Raw and Processed
501 Meat Products: Current State-of-the-Art Research Advances. *Food Control*, In press, DOI:
502 /10.1016/j.foodcont.2017.07.013.

503 Folch-Fortuny, A., Prats-Montalbán, J.M., Cubero, S., Blasco, J. & Ferrer, A. (2016). VIS/NIR
504 hyperspectral imaging and N-way PLS-DA models for detection of decay lesions in citrus
505 fruits. *Chemometrics and Intelligent Laboratory Systems* 156, 241-248

506 Font, D., Tresanchez, M., Pallejà, T., Teixidó, M., Martínez, D., Moreno, J. & Palacín, J. 2014.
507 An image processing method for in-line nectarine variety verification based on the
508 comparison of skin feature histogram vectors. *Computers and Electronics in Agriculture*
509 102, 112–119.

510 Gardner, J.L. (2007). Comparison of calibration methods for tristimulus colorimeters. *Journal*
511 *of Research of the National Institute of Standards and Technology* 112, 129–138.

512 Gat, N. (2000). Imaging spectroscopy using tunable filters: A review. Technical report, Opto-
513 Knowledge Systems Inc. OKSI.

514 Giné-Bordonaba, J., Cantina, C.M., Larrigaudière, C., López, L., López, R. & Echeverría, G.
515 (2014). Suitability of nectarine cultivars for minimal processing: The role of genotype,
516 harvest season and maturity at harvest on quality and sensory attributes. *Postharvest Biology*
517 *and Technology* 93, 49-60.

518 Herrero-Langreo, A., Lunadei, L., Lleó L., Diezma, B. & Ruiz-Altisent, M. (2011).
519 Multispectral Vision for Monitoring Peach Ripeness. *Journal of Food Science* 2, 178-187.

520 Huang, F., Zhang, S., Yang, Y., Man, Z., Zhang, X. & Wu, Y. (2015). Application of
521 hyperspectral imaging for detection of defective features in nectarine fruit. *Transactions of*
522 *the Chinese Society for Agricultural Machinery* 11, 252-259.

523 Iglesias, I. & Echeverría, G. (2009). Differential effect of cultivar and harvest date on nectarine
524 colour, quality and consumer acceptance. *Scientia Horticulturae* 120, 41–50.

525 Iqbal A., Sun, D.-W., & Allen, P. (2014). An overview on principle, techniques and application
526 of hyperspectral imaging with special reference to ham quality evaluation and control. *Food*
527 *Control* 46, 242-254.

528 Jolliffe, I.T. (2002). Principal component analysis (2nd ed.). New York: Springer.

529 Kandpal, L. M., Lee, S. , Kim, M.S., Bae, H. & Cho, B.K. (2015). Short wave infrared (SWIR)
530 hyperspectral imaging technique for examination of aflatoxin B1 (AFB1) on corn kernel.
531 *Food Control* 51, 171-176.

532 Keresztes, J.C., Goodarzi, M., & Saeys, W. (2016). Real-time pixel based early apple bruise
533 detection using short wave infrared hyperspectral imaging in combination with calibration
534 and glare correction techniques. *Food Control* 66, 215-226.

535 Li, J., Chen, L., Huang, W., Wang, O., Zhang, B., Tian, X., Fan, S. & Li, B. (2016).
536 Multispectral detection of skin defects of bi-colored peaches based on VIS–NIR
537 hyperspectral imaging. *Postharvest Biology and Technology* 112, 121–133.

538 López-Maestresalas, A., Keresztes, J.C., Goodarzi, M., Arazuri, S., Jaren, C., & Saeys W.
539 (2016). Non-destructive detection of blackspot in potatoes by Vis-NIR and SWIR
540 hyperspectral imaging. *Food Control* 70, 229-241.

541 Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O. L. & Blasco, J.
542 (2012). Recent advances and applications of hyperspectral imaging for fruit and vegetable
543 quality assessment. *Food Bioprocess Technology* 5, 1121–1142.

544 Lu, R. & Peng, Y. (2006). Hyperspectral scattering for assessing peach fruit firmness.
545 *Biosystems Engineering* 93, 161–171.

546 Malegori, C., Nascimento, E. J., Tonetto de Freitas, S., Pimentel, M. F. & Casiraghi, E. (2017).
547 Comparing the analytical performances of Micro-NIR and FT-NIR spectrometers in the
548 evaluation of acerola fruit quality, using PLS and SVM regression algorithms. *Talanta*, 165,
549 112-116.

550 Mehmood, T., Liland, K. H., Snipen, L. & Sæbø, S. (2012). A review of variable selection
551 methods in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory*
552 *Systems* 118, 62–69.

553 Munera, S., Amigo, J.M., Blasco, J., Cubero, S., Talens, P. & Aleixos, N. (2017a). Ripeness
554 monitoring of two cultivars of nectarine using VIS-NIR hyperspectral reflectance imaging.
555 *Journal of Food Engineering* 214, 29-39.

556 Munera, S., Besada, C., Aleixos, N., Talens, P., Salvador, A., Sun, D.-W., Cubero, C. & Blasco,
557 J. (2017b). Non-destructive assessment of the internal quality of intact persimmon using
558 colour and VIS/NIR hyperspectral imaging. *LWT - Food Science and Technology* 77C, 241-
559 248.

560 Nicolai, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I., Lammertyn, J.,
561 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR
562 spectroscopy: a review. *Postharvest Biology and Technology* 46, 99–118.

563 Nogales-Bueno, J., Rodríguez-Pulido, F. J., Heredia, F. J. & Hernández-Hierro J. M. (2015).
564 Comparative study on the use of anthocyanin profile, color image analysis and near-infrared
565 hyperspectral imaging as tools to discriminate between four autochthonous red grape
566 cultivars from La Rioja (Spain). *Talanta* 131, 412–416.

567 Pan, L., Zhang, Q., Zhang, W., Sun, Y., Hua, P. & Tu, K. (2016). Detection of cold injury in
568 peaches by hyperspectral reflectance imaging and artificial neural network. *Food Chemistry*
569 192, 134–141.

570 Pérez-Marín, D., Sánchez, M. T., Paz, P., González-Dugo, V. & Soriano, M. A. (2011).
571 Postharvest shelf-life discrimination of nectarines produced under different irrigation
572 strategies using NIR-spectroscopy. *LWT - Food Science and Technology* 44, 1405-1414.

573 Rajkumar, P., Wang, N., Elmasry, G., Raghavan, G.S.V. & Garipey, Y. (2012). Studies on
574 banana fruit quality and maturity stages using hyperspectral imaging. *Journal of Food*
575 *Engineering* 108, 194–200.

576 Reig, G., Alegre, S., Gatiús, F., & Iglesias, I. (2013). Agronomical performance under
577 Mediterranean climatic conditions among peach [*Prunus persica* (L.) Batsch] cultivars
578 originated from different breeding programs. *Scientia Horticulturae* 150, 267–277.

579 Reig, G., Iglesias, I., & Echeverría, G. (2009). Agronomical performance, fruit quality and
580 sensory attributes of several flat peach and flat nectarine cultivars. VII International Peach
581 Symposium 962, 563-569.

582 Rinnan, Å., van den Berg, F., Engelsen, S.B. (2009). Review of the most common pre-
583 processing techniques for near-infrared spectra. *Trends in Analytical Chemistry* 28, 1201-
584 1222.

585 Serranti, S., Cesare, D., Marini, F. & Bonifazi, G. (2013). Classification of oat and groat
586 kernels using NIR hyperspectral imaging. *Talanta* 103, 276–284.

587 Sun, Y., Wang, Y., Xiao, Y., Gu, X., Pan, L. & Tu, K. (2017). Hyperspectral imaging detection
588 of decayed honey peaches based on their chlorophyll content. *Food Chemistry* 235, 194–202

589 USDA. (2016). EU-28 Stone Fruit Annual. <https://gain.fas.usda.gov/> Accessed 15.04.17.

590 Valero, C., Crisosto, C. H. & Slaughter, D. (2007). Relationship between nondestructive
591 firmness measurements and commercially important ripening fruit stages for peaches,
592 nectarines and plums. *Postharvest Biology and Technology* 44, 248–253.

593 Vélez-Rivera, N., Gómez-Sanchis, J., Chanona-Pérez, J.J., Carrasco, J.J., Millán-Giraldo, M.,
594 Lorente, D., Cubero, S. & Blasco, J. (2014). Early detection of mechanical damage in mango
595 using NIR hyperspectral images and machine learning. *Biosystems Engineering* 122, 91-98.

596 Verdú, S., Vásquez, F., Grau, R., Ivorra, E., Sanchez, A.J., & Barat J.M. (2016). Detection of
597 adulterations with different grains in wheat products based on the hyperspectral image
598 technique: The specific cases of flour and bread. *Food Control* 62, 373-380.

599 Yang, C.H., Sun, D.W., Pu, H., Wang, N.N. & Zhu, Z. (2015). Rapid detection of anthocyanin
600 content in lychee pericarp during storage using hyperspectral imaging coupled with model
601 fusion. *Postharvest Biology and Technology* 103, 55–65.

602 Zhang, B., Li, J., Fan, S., Huang, W., Zhao, C., Liu, C. & Huang, D. (2015). Hyperspectral
603 imaging combined with multivariate analysis and band math for detection of common
604 defects on peaches (*Prunus persica*). *Computers and Electronics in Agriculture* 114, 14–24.

605 Zhu, N., Lin, M., Nie, Y., Wu, D. & Chen, K. (2016). Study on the quantitative measurement of
606 firmness distribution maps at the pixel level inside peach pulp. *Computers and Electronics in*
607 *Agriculture 130*, 48–56.
608

609 **FIGURES**

610

611 Figure 1. Mean hyperspectral image spectra of ‘Diamond Ray’ and ‘Big Top’ cultivars

612

613 Figure 2. Score image of the two first PC of the PCA model using pixel spectra of 40 fruits of
614 each cultivar from the calibration set

615 Key for Figure 2: The percentages indicate the explained variance (87.8% of the total variance).

616 The variations in the colour in both score plots show features linked to the distribution or

617 content of the biochemical constituents in each fruit and cultivar. ‘DR’=‘*Diamond Ray*’;

618 ‘BT’=‘*Big Top*’

619

620 Figure 3. Score (A) and loadings plot (B) of the PCA of the mean spectra of the calibration set

621

622 Figure 4. Visualisation of cultivar classification using individual pixel spectrum: A) Pixel
623 classification method; B) Object classification method

624 Key for Figure 4: *Blue* = ‘*Diamond Ray*’; *Red* = ‘*Big Top*’

625

626 Figure 5. Visualisation of cultivar classification using mean spectrum: A) Classification using
627 the full range; B) Classification using 14 optimal wavelengths

628 Key for Figure 5: *Blue* = ‘*Diamond Ray*’; *Red* = ‘*Big Top*’

629

630 Figure 6. Vector of regression coefficients of the PLS-DA model using mean spectra and with
631 the optimal wavelengths selected

632