Scientia Horticulturae 304 (2022) 111328

Contents lists available at ScienceDirect

Scientia Horticulturae

journal homepage: www.elsevier.com/locate/scihorti

Assessment of watermelon maturity using portable new generation NIR spectrophotometers

Miguel Vega-Castellote [a], María-Teresa Sánchez [a,*], Irina Torres [b], María-José de la Haba [a], Dolores Pérez-Marín [b,*]

[a] Department of Bromatology and Food Technology, University of Cordoba, Rabanales Campus, Córdoba 14071, Spain
[b] Department of Animal Production, University of Cordoba, Rabanales Campus, Córdoba 14071, Spain

ARTICLE INFO

Keywords:
In situ analysis
Spectral sensors
Optimum harvesting time
Prediction sugar content
Watermelon classification

ABSTRACT

The non-destructive evaluation of internal maturity in watermelons, which are large fruits with a thick rind, during their development on the vine can be considered as a challenge for the growing sector. Near infrared spectroscopy (NIRS) was used to assess *in situ* soluble solid content (SSC), the main parameter to establish full maturity, in 249 intact watermelons, of which, 152 had a striped light green rind and 97 a solid dark green rind. Two handheld new generation spectrophotometers were compared for this purpose. Different pre-processing methods and the partial least squares (PLS) regression algorithm were used to build global calibration models and specific calibration models for each one of the two types of watermelon analysed. The results obtained for the global models showed that NIRS is a suitable technology for screening the fruit for maturity, and that the linear variable filters (LVF) sensor is the best equipment for this purpose (SECV = 1.02%; RPD_p = 1.36). Moreover, the best results were obtained when different models were used depending on the type of watermelon. Additionally, near infrared (NIR) classification models were developed to discriminate the samples by stage of maturity for each type of watermelon available, using partial least squares discriminant analysis (PLS-DA). The optimum threshold values for the striped light green and solid dark green rind watermelons (0.82 and 0.65, respectively) were displaced from the mean value of the discriminant variable due to the differences in terms of number of samples per class used. A total of 66.4% and 82.2% of the striped light green and solid dark green rind watermelons, respectively, were correctly classified. The results of this study demonstrated the viability of using NIRS technology as a decision-making support tool to measure the maturity of watermelons and to establish the optimum harvest time of watermelons and therefore meet the consumers' demand for sweet-tasting fruits. Further studies will be needed to improve calibration robustness, and to further interpret outdoor applications in fruits with thick rind.

1. Introduction

The watermelon [*Citrullus lanatus* (Thunb.)] is a fruit that must be harvested when fully mature, since once separated from the vine, neither the sugar content nor the internal colouration increase (Sun et al., 2010). In watermelons, the soluble solids content (SSC) can vary widely at the time of harvesting, and depends mainly on the cultivar harvested (Suslow, 1997; Yativ et al., 2010). However, watermelons are considered to be ready to be harvested when the SSC of the pulp is over 10% (Suslow, 1997).

Measuring SSC while the fruits are developing is, therefore, key to establishing the optimal moment for harvest, since, as mentioned above, the physiological maturation process finishes at harvest. It should be also considered the fact that, unlike other fruit, the internal maturity of watermelons cannot be established by observing their external appearance and that it is difficult to recognize until the fruit is cut open (Flores et al., 2008; Jie et al., 2019).

The traditional way of establishing if watermelons are mature is to sense sound or vibration by slapping or rapping them. However, this method is tedious, time-consuming, and highly subjective, so therefore subject to error, and must also be carried out by specially trained teams (Abbaszadeh et al., 2011). Moreover, since consumer acceptance of watermelons is based, among other factors, on their sweetness, measuring objectively this quality parameter of the fruits on the vine

* Corresponding authors.
E-mail addresses: teresa.sanchez@uco.es (M.-T. Sánchez), dcperez@uco.es (D. Pérez-Marín).

https://doi.org/10.1016/j.scienta.2022.111328
Received 5 April 2022; Received in revised form 21 June 2022; Accepted 3 July 2022
Available online 9 July 2022

would allow them to be harvested selectively and then sorted according to their quality.

Near infrared spectroscopy (NIRS) is a non-invasive, multiparameter analytical technology that can be used to facilitate the rapid individual internal quality control of intact fruits (Arendse et al., 2018; Cattaneoand Stellari, 2019; Torres et al., 2019; Walsh et al., 2020), thus meeting the needs of the industry and the producers. The recent development of portable and light-weight instruments has led to this technology being considered ideally suited for use in the field (Pasquini, 2018; Yan and Siesler, 2018; Bec et al., 2020). However, the fact that watermelons have a thick rind makes it difficult for light to penetrate the fruit to evaluate the inside, which could imply certain limitations in the use of this technology for field analyses, together with their high water content (90% w/w) (Bianchi et al., 2018) and the fact that in this fruit, the sweetness varies from the rind to the center of the pulp (Lammertyn et al., 2000; Magwaza and Opara, 2015).

In fact, NIRS sensor applications in watermelons –mainly concerned with measuring SSC– have been developed using different benchtop NIRS instruments working in transmittance mode (Abebe, 2006; Tian et al., 2007; Jie et al., 2013, 2014; Qi et al., 2014), since this method of analysis has been traditionally recommended for hard-rind fruits (Saranwong and Kawano, 2007). However, these instruments use intense illumination and involve long integration times, which potentially could lead to heat damage in the fruit (Long and Walsh, 2006).

Only Flores et al. (2008) and Tamburini et al. (2017) have measured SSC in products already harvested using benchtop instruments working in reflectance mode. The results obtained by these authors allow to confirm that light penetrated the fruit enough to measure the parameters tested accurately, thanks to improvements in photodetectors and measuring devices. However, it is worth noting that no research articles exist in the scientific literature which demonstrate the feasibility of using new generation portable NIRS instruments working in reflectance mode to monitor the pre-harvest internal quality of intact watermelons in the field.

The aim of this research was to investigate the potential of NIRS technology to be used at a field level, *i.e., in situ* analysis, using two new generation handheld NIRS sensors of different optical designs and technical specifications, working in reflectance mode to gage the optimum moment for the watermelons to be harvested, based on the soluble solid content, and thus meet the consumers' demands.

2. Materials and methods

2.1. Watermelon sampling

A total of 249 watermelons (*Citrullus lanatus* Thunb.) – of which 152 had a striped light green rind (cv. 'Bazman', 'Bengala', 'Boston', 'Premium', 'Premium-Frilly' and 'Red Jasper') and 97 a solid dark green rind (cv. 'Fashion', 'Fenway' and 'Jamaica') (Table 1) were grown in greenhouses and in open fields in different regions of Spain (Andalusia, Canarias, Castilla La Mancha, and Murcia). The fruits were harvested at different stages of maturity, ranging from early to full maturity, between May and October 2021. The fruits were taken to the laboratory at the Faculty of Agricultural and Forestry Engineering at the University of Cordoba, Spain, and kept at 10 °C and 90% relative humidity until the following day, when laboratory testing was performed.

2.2. Instrumentation and NIR spectra acquisition

The acquisition of the near infrared (NIR) spectra of intact watermelons was carried out using two portable handheld NIRS instruments of different optical designs (Table 2). These instruments were suitable for the *in situ* analysis of the product in the field.

Spectra of the intact watermelons were initially taken using a Linear Variable Filters (LVF) spectrophotometer (MicroNIR™ Pro 1700, VIAVI Solutions, Inc., San Jose, California, USA). Spectra acquisition was carried out using the VIAVI MicroNIR Pro software version 2.2 (VIAVI Solutions, Inc., San Jose, California, USA). The instrument's performance was checked every 10 min. Four spectra were taken per fruit on the equator region, and the fruit was rotated through 90° after each measurement. The four spectra were averaged to provide a mean spectrum for each fruit.

Additionally, following the same procedure, four spectra were taken per fruit and later averaged, using a compact, handheld instrument based on diode-array (DA) technology (Aurora spectrophotometer, GraiNit S.r.l., Padova, Italia). This instrument has an internal reference which is taken automatically prior to the measurement of each sample. The UCal 4™ software (Unity Scientific LLC, Milford, MA, USA) was used for spectra acquisition.

Table 1
Rind type, number of samples (N) and mean (standard deviation) values of the rind thickness (mm) and soluble solid content (%) parameters of the different watermelon cultivars analysed.

Rind type	Cultivar	N	Rind thickness (mm)	SSC (%)
Striped light green	Bazman	5	14.74 (0.97) d	6.58 (0.31) a
	Bengala	26	10.35 (1.42) ab	9.37 (1.05) d
	Boston	33	12.51 (5.26) cd	8.06 (1.40) b
	Premium	68	9.34 (2.95) a	8.19 (1.68) bc
	Premium-Frilly	17	8.71 (1.87) a	7.74 (1.06) ab
	Red Jasper	3	13.03 (0.48) bcd	9.43 (0.39) bcd
Solid dark green	Fashion	54	11.19 (1.67) bc	8.88 (1.25) d
	Fenway	38	11.14 (3.56) bc	8.73 (1.49) cd
	Jamaica	5	8.76 (2.23) ab	8.70 (0.90) bcd

Standard deviation in bracket.
Means in the same column followed by the same letter showed no significant differences between them ($P > 0.05$).

Table 2
Technical features of the linear variable filters and the diode-array spectrophotometers.

Property	Instrument MicroNIR™ Pro 1700	Aurora
Analysis mode	Reflectance	Reflectance
Detector type	128-pixel InGaAs photodiode array	256-pixel InGaAs detector
Dispersion element	Linear variable filters	Diode-array
Wavelength range (nm)	908–1676	950–1650
Resolution (nm)	6.2	2
Sampling integration time (ms)	11	6.57
Scanning time per measurement (s)	2–3	2–3
Weight (kg)	$64{\cdot}10^{-3}$	2
Optical window size (mm^2)	227	1256

2.3. Reference analysis

The watermelons were then cut in half along the equatorial axis and the rind thickness of the half containing the stem was measured at four equidistant points using a digital precision calliper (0–300 ± 0.01 mm; Comecta, Barcelona, Spain). These 4 values were averaged to obtain a mean rind thickness per fruit. After that, the same half was again cut into two identical pieces, one of which was used to measure the SSC (%). To achieve this, the rind was removed, the flesh was squeezed, and the SSC of the extracted juice was measured by refractometry using a temperature-compensated digital Abbé-type refractometer (model B, Zeiss, Oberkochen, Würt, Germany). The samples were analysed in duplicate, the data per fruit were averaged and the standard error of

laboratory (SEL) was calculated from these duplicates.

2.4. Data processing

The WinISI II software package version 1.50 (Infrasoft International LLC, Port Matilda, PA, USA) (ISI, 2000) and the Matlab software version R2019a were used to carry out the chemometric treatment of the data. The CENTER algorithm (ISI, 2000) was applied to the sets of intact watermelons analysed with both instruments to study the structure and variability of the population. This algorithm performs a principal component analysis (PCA) and calculates the Mahalanobis distance (GH) between each sample and the center of the population (Shenk and Westerhaus, 1995). Prior to the application of this algorithm, a combination of mathematical pre-treatments was applied — standard normal variate (SNV) and de-trending (DT) for scatter correction (Barnes et al., 1989), together with the first derivative treatment. Any samples showing GH values over 4 were studied as potential spectral outliers and were excluded from their sets if the removal was justified.

2.4.1. Quantitative models for the prediction of the SSC in intact watermelons

Once the spectral outliers were removed and the samples ordered by spectral distances to the centre of the population, a structured sample selection was carried out according to Shenk and Westerhaus (1991) to build the calibration and validation sets using the sets of watermelons analysed with both instruments. To achieve this, one out of every four samples were selected from the set of intact watermelons to be part of the validation set, and the remaining samples were used to build the calibration set.

The quantitative models developed to predict the SSC in intact watermelons analysed with both portable instruments were carried out using the Matlab software and the partial least squares (PLS) regression method. We tested the first and the second derivative treatments in combination with SNV and DT for scatter correction. The best models were selected by assessing their performance using the coefficient of determination for cross validation (r^2_{cv}), the standard error of cross validation (SECV), and the residual predictive deviation for cross validation (RPD_{cv}), calculated as the ratio of the standard deviation (SD) of the reference data for calibration to the SECV. The RPD_{cv} values obtained for both instruments were statistically compared using Fisher's F test (Mark and Workman, 2003).

The samples included in the validation set were then used to evaluate the best models developed and the validation results were assessed following the protocol outlined by Windham et al. (1989).

In this way, the most suitable spectrophotometer was identified to predict the internal quality in intact watermelons, and new quantitative models were made, following the same procedure explained above, to predict the SSC by separating the set of samples into striped light green and solid dark green rind samples. In this case, given the number of watermelons available of each type, the external validation procedure was skipped, and in this case, the cross-validation approach was followed to evaluate the models obtained.

2.4.2. Classification models according to internal maturity of the fruit

The viability of using NIRS technology for the *in situ* classification of intact watermelons in the field based on their stage of internal maturity was assessed by developing classification models for the striped light green and solid dark green rind watermelons, respectively, once the spectral outliers were removed. These two sets of striped light green and solid dark green rind watermelons were in turn divided into two groups, as determined by Suslow (1997), who reported that an SSC of 10% in the watermelon flesh was an indicator of full maturity in the fruit. Consequently, the category of mature watermelons was made up of those samples showing an SSC equal to or over 10%, while the category of immature watermelons included those samples with an SSC lower than 10%.

The discriminant models were constructed using Matlab software and PLS discriminant analysis (PLS-DA) for supervised classification (Naes et al., 2002), in particular the PLS2 algorithm, which generates as many discriminant variables as there are classes in the learning group. The same signal pre-treatments described for quantitative analysis were used for qualitative model development. As the models were unbalanced in terms of the number of samples included in each class, an optimum threshold value was calculated using the Receiver Operating Characteristic (ROC) curves (Brereton, 2009). The performance of these models was assessed using the number of correctly classified samples, both for the global model and for each class.

2.4.3. Statistical analysis

In order to study potential differences in terms of rind thickness (mm) and SSC (%) between the cultivars, a one-factor analysis of variance (ANOVA) was carried out per parameter using Matlab software version R2019a. Next, the differences between the means were compared using the Fisher's Least Significant Difference (LSD) test, with differences at $P < 0.05$ considered as significant.

3. Results and discussion

3.1. Spectral analysis of the watermelon population

The spectra of the samples were plotted, and three of them were removed from the set of samples analysed using the DA instrument due to problems in the spectral acquisition process. Fig. 1 shows the average spectra of the $N = 249$ and $N = 246$ intact watermelons available for the LVF and DA instruments, respectively. Both instruments yielded similar spectral patterns (absorption peaks aligned on the horizontal axis). The most relevant common absorption bands observed in the LVF and DA spectra were found at 970 nm, 1160–1200 nm, 1440–1450 nm. The spectra displayed water-related absorption peaks at 970 nm and 1450 nm (Osborne et al., 1993; Shenk et al., 2008), as is usual in the case of fruit, and particularly watermelons, which are 90% water (Bianchi et al., 2018). Other peaks, at around 1200 nm and 1440 nm, were characteristic of the sugar-related absorption band (Osborne et al., 1993; Shenk et al., 2008), since sugar is the second largest component after water (Flores et al., 2008).

After applying the CENTER algorithm to the sets of intact watermelons analysed using the LVF and DA instruments, one outlier sample was identified and eliminated for each of the sets. The sample removed from the LVF set showed a GH value over the limit (GH = 7.27) and presented a yellow stain from the middle of the fruit to the area near the pistil scar. Likewise, the sample removed from the DA set showed a GH

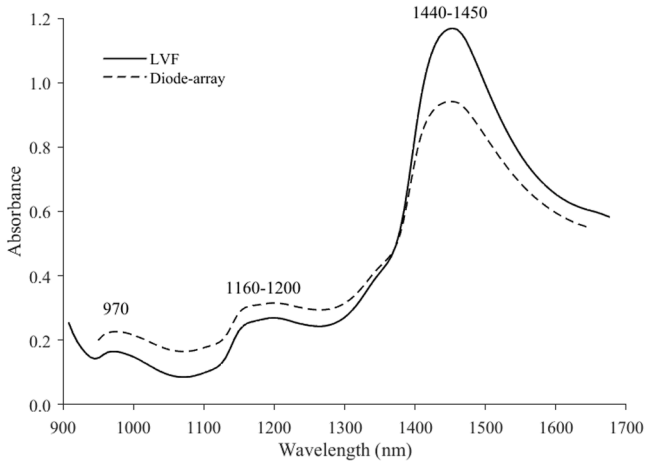

Fig. 1.. Mean NIR spectra for watermelons analysed using the linear variable filters (LVF) and diode-array instruments.

= 11.14 and external abrasions made during the postharvest management of the fruit. Both defects on the watermelon rind could have interfered with the spectra acquisition and could have led to GH values over 4 being obtained for these samples. Consequently, the number of watermelons available in the LVF and DA instruments sets after removing the spectral outliers was 248 and 245, respectively.

3.2. Quantitative prediction of the SSC in intact watermelon samples

Table 3 shows the calibration and validation sets of samples analysed using the LVF and DA instruments. These sets showed, for both instruments, similar values for SSC in terms of mean, standard deviation (SD), and coefficient of variation (CV). In addition, the values of the validation set range for this parameter laid within those of the calibration set. The CV parameter showed values ranging from 15.40 to 17.87% and 15.86–17.95% for the LVF and DA instruments, respectively. These values reflect the great variability in terms of the SSC parameter of the watermelons used in this study, which included samples harvested from early to full maturity which, in turn, is of great importance when the objective is to carry out a follow-up of this internal quality parameter while the fruit is developing on the vine and establish the optimal moment for harvest.

Table 4 shows the cross-validation results for the best models developed to predict SSC in intact watermelon samples using the LVF and DA instruments. For both instruments, the best models were obtained using the SNV + DT + 2nd derivative treatment. These results show that the models would enable to distinguish between low and high SCC values in intact watermelons using the LVF and DA instruments (Shenk and Westerhaus, 1996; Williams, 2001), which can be of great interest for preharvest decision-making in the fields, facilitating the *in situ* discrimination of watermelons at an optimal stage of maturity from those ones at an early stage. Significant differences ($P < 0.05$) were found when comparing the RPD_{cv} values obtained for both instruments, which could be due, among other factors, to the difference in terms of the window size of both spectrophotometers and to the optical design (Table 2). The greater window size of the DA instrument could hinder the spectra acquisition process, due to the difficulty in covering the whole window with the fruit analysed due to its round shape.

The regression coefficients for the model developed using the LVF instrument for the prediction of the SSC are illustrated in Fig. 2. The region between 1193 and 1360 nm showed several significant peaks which corresponds to the second overtone of the C–H stretching bonds (Shenk et al., 2008). In addition, the peaks at around 1385 nm and 1441 nm could be related to sugars absorption (Osborne et al., 1993).

Only two studies exist in the literature which developed models to predict SSC in intact watermelons using NIRS instruments working in reflectance mode. Flores et al. (2008) measured SSC in $N = 203$ intact watermelons belonging to the 'Fashion' cultivar using a Perten DA-700 spectrometer (Perten Instruments North America, Inc., Springfield IL, USA) covering the 400–1700 nm range. These authors reported better results in terms of RPD_{cv} (1.53) than those obtained in this study. Nevertheless, it must be stressed that the model used by these authors involved a single cultivar and a benchtop instrument covering a wider spectral range. In addition, Tamburini et al. (2017) predicted the SSC in

Table 3
Number of samples (N), range, mean, standard deviation (SD) and coefficient of variation (CV) for the soluble solid content (%) in the calibration and validation sets of watermelons analysed using the linear variable filters and diode-array instruments.

Instrument	Set	N	Range	Mean	SD	CV (%)
Linear variable filters	Calibration	187	4.40–12.05	8.34	1.49	17.87
	Validation	61	5.30–11.45	8.96	1.38	15.40
Diode-array	Calibration	184	4.40–12.05	8.41	1.51	17.95
	Validation	61	5.35–11.45	8.70	1.38	15.86

Table 4
Calibration statistics for the best equations obtained to predict the soluble solid content (%) in intact watermelons.

Instrument	[a] N	Range	[b] Mean	[c] SD	[d] r^2_{cv}	[e] SECV	[f] RPD_{cv}
Linear variable filters	180	4.40–11.45	8.39	1.39	0.47	1.02	1.36
Diode-array	181	4.65–11.30	8.40	1.47	0.33	1.22	1.20

[a] Number of samples.
[b] Mean of the calibration set.
[c] Standard deviation of the calibration set.
[d] Coefficient of determination of cross validation.
[e] Standard error of cross validation.
[f] Residual predictive deviation for cross validation.

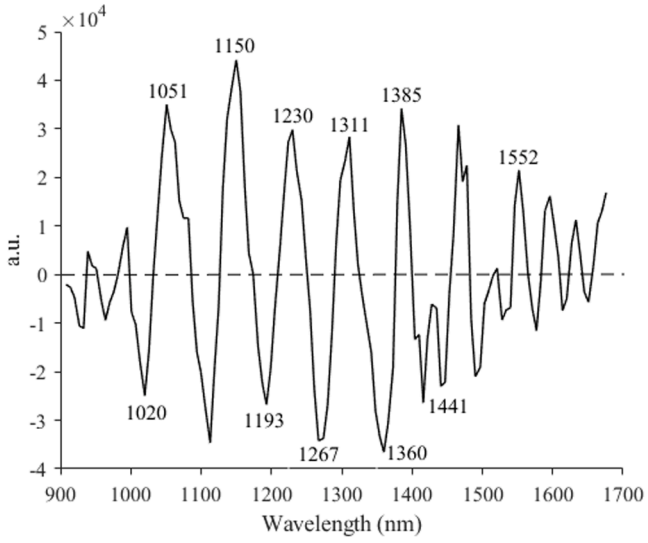

Fig. 2. Regression coefficients for the best model developed for the prediction of the SSC in intact watermelons using the LVF instrument. a.u.: arbitrary units.

$N = 135$ intact watermelons belonging to the 'Minirossa' cultivar using a NIR On-Line X-One instrument (Buchi, Flawil, Switzerland) working in the 900–1700 nm range and reported an RPD_{cv} value (3.04) greater than that here obtained. Again, these authors worked with a benchtop device and a single cultivar, which was particularly small in size (diameter, 100–150 mm) with a very thin (< 0.5 mm), striped, green rind. It is therefore important to stress the novelty of our research, since there are no works available in the literature dealing with the non-destructive measurement of the SSC in intact watermelons in the field using portable NIRS instruments.

The calibration models developed for both instruments were subjected to an external validation procedure. For the LVF instrument, good prediction results were obtained, with the samples at the ends of the SSC (%) range showing the highest residual predicted values, given the low representativeness of these samples in the calibration set after removing the outliers when setting up the calibration model ($N = 180$) (see distribution in Fig. 3). Thus, a sample belonging to the 'Boston' cultivar with a reference value of SSC = 11.3% was predicted as SSC = 8.6%, which turned out to be the highest residual predicted value of the samples in the validation set. The slope, SEP(c) and bias fell within the recommended limits of the Windham et al. (1989) protocol for the LVF instrument (Fig. 4). In addition, although the r^2_p value did not meet the validation requirements established in this protocol, this equation would enable to carry out an initial screening for watermelon maturity in the field, which can be considered as a first step in the development of an *in situ* application to measure the internal quality of intact

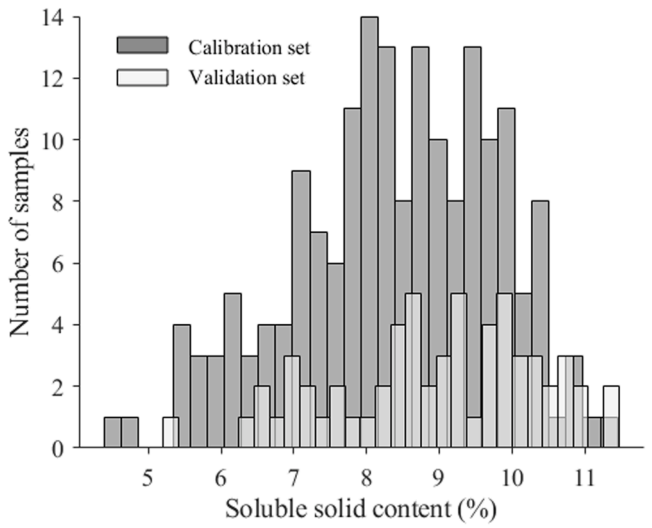

Fig. 3. Histogram of frequencies for the soluble solid content (%) for the calibration* ($N = 180$) and validation sets ($N = 61$). *Calibration set: samples used to develop the model after removing outliers ($N = 180$).

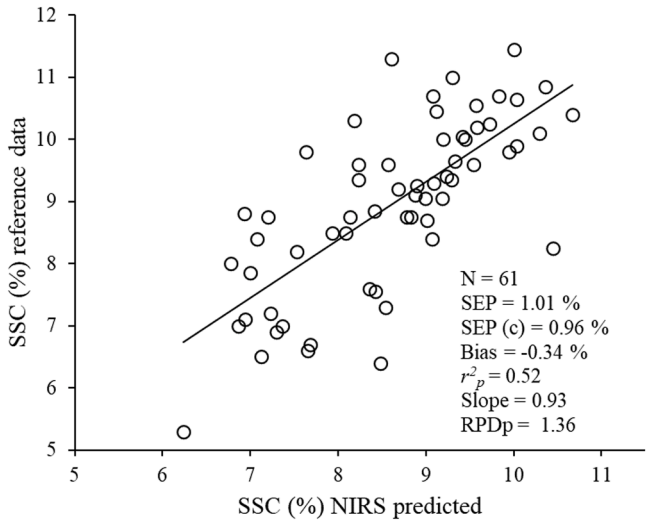

Fig. 4. Actual *versus* predicted values for the validation of the best models to predict the soluble solid content (SSC,%) of intact watermelons analysed using the linear variable filters instrument.

watermelons. On the contrary, the validation results obtained for the assay carried out using the DA instrument showed a very low predictive capacity (data not shown). We can therefore confirm that the LVF instrument performed better when measuring *in situ* the SSC parameter in intact watermelons, and that this instrument would be the one used from that moment onwards in this study.

Given the important differences that can be observed between the two types of watermelons studied (striped light green and solid dark green), in terms of external color, brightness, rind thickness, etc., which could affect the light interaction with the product and the results obtained, we decided to create new, separate models for striped light green and solid dark green rind watermelons to improve their performance. The comparison of rind thickness (mm) and SSC (%) between the cultivars emphasised the need to distinguish between the two types of fruit, since significant differences ($P < 0.05$) were found between the 'Bazman', 'Boston', 'Premium', and 'Premium-Frilly' cultivars (striped light green rind) and at least one of the solid dark green rind cultivars, both in terms of the SSC (%) and in rind thickness (Table 1). Furthermore, the second derivative plot of the raw mean spectra for the striped light green and solid dark green rind watermelons showed differences in absorption peaks at around 1378 and 1447 nm (Fig. 5), which have been linked to sugar-related absorption peaks (Osborne et al., 1993).

The set of striped light green rind watermelons included a greater number of samples and showed a wider variability in terms of the CV value for the SSC, compared to the one made up for solid dark green rind watermelons (Table 5). The prediction capacity in terms of RPD_{cv} of the models developed, after separating the watermelons into the two available types (Table 6), were very similar to that obtained for all the samples without differentiating between types. Nevertheless, the number of samples is a key factor in developing NIRS models, especially when the size of the sets is around 100–150 samples. It must therefore be noted that, despite the lower number of samples of solid dark green rind watermelons used to set up the model ($N = 94$), the results obtained were slightly better in terms of RPD_{cv} than those obtained with the set of samples including the two types, in which a total of $N = 182$ samples were used to make the model. These results confirm the importance of differentiating between the striped light green and solid dark green rind watermelons when creating NIRS prediction models.

3.3. Discriminant analysis of watermelons according to maturity

According to the criterion established by Suslow (1997), the number of mature and immature striped light green watermelons analysed with the LVF instrument used to develop the classification models was 21 and 131, respectively. Likewise, the number of mature and immature solid dark green samples available was 21 and 75, respectively. The best results for both types of watermelons were obtained using the 2nd derivative treatment, along with the SNV + DT for scatter correction. The optimum threshold values were 0.82 and 0.65 for the assays carried out using striped light green and solid dark green rind watermelons, respectively. In the former case, the difference in terms of number of samples per class –mature and immature watermelons– was greater and, consequently, the optimum threshold value was further from the mean value of the discriminant variable (Naes et al., 2002; Brereton, 2009).

The model created for the striped light green rind watermelons correctly classified 66.4% (101/152) of the samples (Table 7), 71.4% (15/21) as mature and 65.6% (86/131) as immature. Among the poorly classified samples, 7, 2, 13, 4, and 1 samples belonging to the 'Bengala', 'Boston', 'Premium', 'Premium-Frilly' and 'Red Jasper' cultivars, respectively, presented a rind thickness 18.25% greater, on average, than the mean value of its corresponding cultivar, which can negatively impact on ability of the NIRS to predict the quality characteristics of the fruit (De Oliveira et al., 2014). The immature not correctly classified samples presented a mean and SD SSC values of 8.53% and 1.26%,

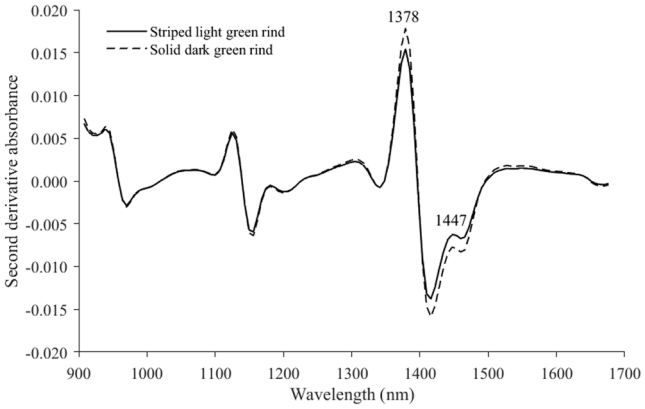

Fig. 5. Second derivative of the absorbance spectra of the striped light green and solid dark green rind watermelons analysed using the linear variable filters instrument.

Table 5
Number of samples (N), range, mean, standard deviation (SD) and coefficient of variation (CV) for the soluble solid content (%) of the striped light green and solid dark green rind samples analysed using the linear variable filters instrument.

Fruit rind	N	Range	Mean	SD	CV (%)
Striped light green	152	4.65–11.45	8.28	1.54	18.60
Solid dark green	96	4.40–12.05	8.81	1.34	15.21

Table 6
Calibration statistics for the best equations obtained to predict the soluble solid content (%) in intact striped light green and solid dark green rind watermelons using the linear variable filters instrument.

Fruit rind	[a] N	Range	[b] Mean	[c] SD	[d] r^2_{cv}	[e] SECV	[f] RPD_{cv}
Striped light green	147	4.80–11.45	8.38	1.46	0.48	1.06	1.38
Solid dark green	94	4.40–11.45	8.81	1.27	0.51	0.90	1.41

[a] Number of samples.
[b] Mean of the calibration set.
[c] Standard deviation of the calibration set.
[d] Coefficient of determination of cross validation.
[e] Standard error of cross validation.
[f] Residual predictive deviation for cross validation.

Table 7
Percentage of samples correctly classified (%) calculated as number of samples per cultivar correctly classified (N) in the PLS-DA models divided by total number of samples available per cultivar (N_{total}).

Rind type	Cultivar	N/N_{total} (%)
Striped light green	Bazman	5/5 (100.0%)
	Bengala	13/26 (50.0%)
	Boston	23/33 (69.7%)
	Premium	49/68 (72.1%)
	Premium-Frilly	10/17 (58.8%)
	Red Jasper	1/3 (33.3%)
	Total	**101/152 (66.4%)**
Solid dark green	Fashion	43/53 (81.1%)
	Fenway	30/38 (78.9%)
	Jamaica	4/5 (80.0%)
	Total	**77/97 (82.2%)**

respectively. Likewise, the mature but not correctly classified samples showed a mean and SD SSC values of 10.60% and 0.45%, respectively. In addition, 6 samples showed an SSC value within the 10 ± 2 * SEL (SEL = 0.1%) range, which could be considered difficult to discriminate, since the error obtained could be attributed to the error of the reference method.

It should also be noted that those cultivars with the lowest number of samples available showed the greatest relative number of misclassified samples (number of misclassified samples of a cultivar/total number of samples of this cultivar available), which highlights the importance of including a proper representation of each cultivar in the training set. For example, a total of 2 out of the 3 samples belonging to the 'Red Jasper' cultivar were not correctly classified (Table 7). It is clear that, for complex applications, such as the one proposed here, the number of samples needed to cover the variability of the product must be higher.

The number of correctly classified fruits for the solid dark green rind type was 77/96 (82.2%) (Table 7), of which 15/21 (71.4%) belonged to the mature category and 62/75 (82.7%) to the immature set. In this case, of the poorly classified samples, 4, 5, and 1 watermelons belonging to the 'Fashion', 'Fenway' and 'Jamaica' cultivars, respectively, presented a rind thickness 15.70% greater, on average, than the mean value of its corresponding variety. From these samples, the immature but not correctly classified samples presented mean and SD SSC values of 9.24% and 0.68%, respectively. Similarly, the mature but not correctly classified samples showed mean and SD SSC values of 10.15% and 0.17%, respectively, and these values were very close to SSC = 10%. Furthermore, 5 samples showed an SSC value within the 10 ± 2 * SEL range.

The difference in terms of correctly classified fruits for both types of watermelons –66.4% for the striped light green and 82.2% for the solid dark green rind– could possibly be attributed to the greater number of cultivars included in the striped light green rind set of watermelons ('Bazman', 'Bengala', 'Boston', 'Premium', 'Premium-Frilly', and 'Red Jasper') compared to the solid dark green rind set ('Fashion', 'Fenway', and 'Jamaica'), and also to the differences in the brightness and hardness of the product that hinders analysis and light penetration in the case of the striped watermelons more than for the dark green ones. In addition, the number of fruits included per cultivar was not the same, especially in the case of the striped light green rind watermelons, which could compromise the cross-validation results, since it may imply that these cultivars were not well represented in some of the cross-validation tests, where it is of the utmost importance to work with equitable cross-validation splits (Shenk and Westerhaus, 1991) when aiming to obtain representative cross-validation results. Therefore, given the importance of including different cultivars in the training set, which should include a wider variability in future watermelon fruits to be predicted, further studies should focus on adding more samples to the training set in order to work with a similar number of watermelon fruits per cultivar.

The results obtained demonstrate the possibility of developing applications for the *in situ* differentiation between mature and immature intact watermelons using portable NIRS handheld instruments, which is of great interest since it could be used as a support tool for decision-making about the optimum harvest time for watermelons in the field.

No studies in the literature were found for the classification of intact watermelon samples based on their stage of maturity using near infrared reflectance spectroscopy techniques.

4. Conclusions

The results showed the potentiality of NIRS technology to be used as a tool to study the stage of maturity of intact watermelons during their development on the vine based on the prediction of the SSC of the fruits analysed. This covers the demands of the growing watermelon sector for the incorporation at the field level of non-destructive and objective technologies to follow the development of the fruit and to establish the best time for harvest to meet the consumers' demand for sweet, tasty fruits. Moreover, due to the differences between the striped light green and solid dark green rind watermelons, the results suggest that the two types of fruits must be treated separately, and better results for SSC prediction were obtained with this approach. Additionally, successful results were obtained when NIRS technology was used to classify the striped light green and solid dark green rind watermelons as fully mature or immature, respectively. Thus, this research can be considered as a successful first step towards the *in situ* implementation of NIRS technology in this kind of fruit with a thick rind and large size, using a handheld new generation spectrophotometer working in reflectance mode. A larger database will be needed in the future to develop more robust prediction models and to test the equations using samples belonging to different harvesting seasons and considering the influence of outdoors environmental factors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper in any way.

Data availability

Data will be made available on request.

Ethics approval and consent to participate

Not applicable.

Acknowledgments

The authors are grateful to Mrs. María Carmen Fernández from the Animal Production Department for her technical assistance. Furthermore, the authors wish to express their gratitude to the Spanish Ministry of Universities for the support offered to Miguel Vega-Castellote through the Training Programme for Academic Staff (FPU).

References

Abbaszadeh, R., Rajabipour, A., Ahmadi, H., Delshad, M., Mahjoob, M., 2011. Assessment of watermelon quality using vibration spectra. In: Pichappan, P., Ahmadi, H., Ariwa, E. (Eds.), Innovative Computing Technology. INCT 2011. Communications in Computer and Information Science, 241. Springer, Berlin, Germany, pp. 21–29.

Abebe, A.T., 2006. Total sugar and maturity evaluation of intact watermelon using near infrared spectroscopy. J. Near Infrared Spectrosc. 14, 67–70. https://doi.org/10.1255/jnirs.588.

Arendse, E., Fawole, O.A., Magwaza, L.S., Opara, U.L., 2018. Non-destructive prediction of internal and external quality attributes of fruit with thick rind: a review. J. Food Eng. 217, 11–23. https://doi.org/10.1016/j.jfoodeng.2017.08.009.

Barnes, R.J., Dhanoa, M.S., Lister, S.J., 1989. Standard normal variate transformation and de-trending of near infrared diffuse reflectance spectra. Appl. Spectrosc. 43, 772–777.

Bec, K.B., Grabska, J., Siesler, H.W., Huck, C.W., 2020. Handheld near-infrared spectrometers: where are we heading? NIR News 31, 28–35. https://doi.org/10.1177/0960336020916815.

Bianchi, G., Rizzolo, A., Grassi, M., Provenzi, L., Lo Scalzo, R., 2018. External maturity indicators, carotenoid and sugar compositions and volatile patterns in 'Cuoredolce®' and 'Rugby' mini-watermelon (*Citrullus lanatus* (Thunb) Matsumura & Nakai) varieties in relation of ripening degree at harvest. Postharvest Biol. Technol. 136, 1–11. https://doi.org/10.1016/j.postharvbio.2017.09.009.

Brereton, R.G., 2009. Chemometrics For Pattern Recognition. John Wiley and Sons, Chichester, West Sussex, UK.

Cattaneo, T.M.P., Stellari, A., 2019. Review: NIR spectroscopy as a suitable tool for the investigation of the horticultural field. Agronomy 9 (503), 1–20. https://doi.org/10.3390/agronomy9090503.

De Oliveira, G.A., Bureau, S., Claire Renard, C.M-G., Pereira-Netto, A.B., De Castilhos, F., 2014. Comparison of NIRS approach for prediction of internal quality traits in three fruit species. Food Chem. 143, 223–230. https://doi.org/10.1016/j.foodchem.2013.07.122.

Flores, K., Sánchez, M.T., Pérez-Marín, D.C., López, M.D., Guerrero, J.E., Garrido, A., 2008. Prediction of total soluble solid content in intact and cut melons and watermelons using NIR spectroscopy. J. Near Infrared Spectrosc. 16, 91–98. https://doi.org/10.1255/jnirs.771.

ISI, 2000. The Complete Software Solution Using a Single Screen For Routine Analysis, Robust Calibrations, and Networking. Manual, FOSS NIRSystems/TECATOR. Infrasoft International, LLC, Silver Spring, MD, USA.

Jie, D., Zhou, W., Wei, X., 2019. Nondestructive detection of maturity of watermelon by spectral characteristic using NIR diffuse transmittance technique. Sci. Hortic. 257 (108718), 1–7. https://doi.org/10.1016/j.scienta.2019.108718.

Jie, D., Xie, L., Rao, X., Ying, Y., 2014. Using visible and near infrared diffuse transmittance technique to predict soluble solids content of watermelon in an online detection system. Postharvest Biol. Technol. 90, 1–6. https://doi.org/10.1016/j.postharvbio.2013.11.009.

Jie, D., Xie, L., Fu, X., Rao, X., Ying, Y., 2013. Variable selection for partial least squares analysis of soluble solids content in watermelon using near-infrared diffuse transmission technique. J. Food Eng. 118, 387–392. https://doi.org/10.1016/j.jfoodeng.2013.04.027.

Lammertyn, J., Peirs, J., De Baerdemaeker, J., Nicolaï, B.M., 2000. Light penetration properties of NIR radiation in fruit with respect to non-destructive quality assessment. Postharvest Biol. Technol. 18, 121–132. https://doi.org/10.1016/S0925-5214(99)00071-X.

Long, R., Walsh, K., 2006. Limitations to the measurement of intact melon total soluble solids using near infrared spectroscopy. Aust. J. Agric. Res. 57, 403–410. https://doi.org/10.1071/AR05285.

Magwaza, L.S., Opara, U.L., 2015. Analytical methods for determination of sugars and sweetness of horticultural products—a review. Sci. Hortic. 184, 179–192. https://doi.org/10.1016/j.scienta.2015.01.001.

Mark, H., Workman, J., 2003. The F statistic. In: Mark, H., Workman, J. (Eds.). Statistics in Spectroscopy, 2nd ed. Academic Press, San Diego, CA, USA, pp. 205–211.

Naes, T., Isaksson, T., Fearn, T., Davies, A., 2002. A User-Friendly Guide to Multivariate Calibration and Classification. NIR Publications, Chichester, UK.

Osborne, B.G., Fearn, T., Hindle, P.H., 1993. Practical NIR Spectroscopy with Applications in Food and Beverage Analysis. Longman Scientific and Technical, Harlow, Essex, England.

Pasquini, C., 2018. Near infrared spectroscopy: a mature analytical technique with new perspectives – a review. Anal. Chim. Acta 1026, 8–36. https://doi.org/10.1016/j.aca.2018.04.004.

Qi, S., Song, S., Jiang, S., Chen, Y., Li, W., Han, D., 2014. Establishment of a comprehensive indicator to nondestructively analyze watermelon quality at different ripening stages. J. Innov. Opt. Health Sci. 7 (1350034), 1–11. https://doi.org/10.1142/S179354581350034X.

Saranwong, S., Kawano, S., 2007. Fruits and vegetables. In: Ozaki, Y., McClure, W.F., Christy, A.A. (Eds.), Near-Infrared Spectroscopy in Food Science and Technology. John Wiley & Sons, Inc., Hoboken, New Jersey, USA, p. 219.

Shenk, J.S., Westerhaus, M.O., 1991. Population structuring of near infrared spectra and modified partial least squares regression. Crop Sci. 31, 1548–1555.

Shenk, J.S., Westerhaus, M.O., 1995. Analysis of Agriculture and Food Products By Near Infrared Reflectance Spectroscopy. Monograph. NIRSystems, Inc., Silver Spring, MD, USA.

Shenk, J.S., Westerhaus, M.O., 1996. Calibration the ISI way. In: Davis, A.M.C., Williams P.C. (Eds.), Near Infrared Spectroscopy: The Future Waves. NIR Publications, Chichester, UK, pp. 198–202.

Shenk, J.S., Workman, J., Westerhaus, M., 2008. Application of NIR spectroscopy to agricultural products. In: Burns, D.A., Ciurczac, E.W. (Eds.), Handbook of Near Infrared Analysis, 3rd ed. Marcel Dekker, New York, NY, USA, pp. 347–386.

Sun, T., Huang, K., Xu, H., Ying, Y., 2010. Research advances in nondestructive determination of internal quality in watermelon/melon: a review. J. Food Eng. 100, 569–577. https://doi.org/10.1016/j.jfoodeng.2010.05.019.

Suslow, T. 1997. Watermelon: Recommendations for Maintaining Postharvest Quality. http://postharvest.ucdavis.edu/Commodity_Resources/Fact_Sheets/Datastores/Fruit_English/?uid=60&ds=798. (Accessed March 24, 2022).

Tamburini, E., Costa, E., Rugiero, I., Pedrini, P., Marchetti, M.G., 2017. Quantification of lycopene, β-carotene, and total soluble solids in intact red-flesh watermelon (*Citrullus lanatus*) using on-line near-infrared spectroscopy. Sensors 17 (746), 1–12. https://doi.org/10.3390/s17040746.

Tian, H.Q., Ying, Y.B., Lu, H.S., Fu, X.P., Yu, H.Y., 2007. Measurement of soluble solids content in watermelon by Vis/NIR diffuse transmittance technique. J. Zhejiang Univ. Sci. B 8, 105–110. https://doi.org/10.1631/jzus.2007.B0105.

Torres, I., Sánchez, M.T., Entrenas, J.A., Garrido-Varo, A., Pérez-Marín, D., 2019. Monitoring quality and safety assessment of summer squashes along the food supply chain using near infrared sensors. Postharvest Biol. Technol. 154, 21–30. https://doi.org/10.1016/j.postharvbio.2019.04.015.

Walsh, K.B., McGlone, V.A., Han, D.H., 2020. The uses of near infra-red spectroscopy in postharvest decision support: a review. Postharvest Biol. Technol. 163 (111139), 1–12. https://doi.org/10.1016/j.postharvbio.2020.111139.

Williams, P.C., 2001. Implementation of near-infrared technology. Williams, P.C., Norris, K.H. (Eds.). Near Infrared Technology in the Agricultural and Food Industries. AACC Inc., St. Paul, Minnesota, pp. 145–169.

Windham, W.R., Mertens, D.R., Barton II, F.E., 1989. Protocol for NIRS calibration: sample selection and equation development and validation. Martens, G.C., Shenk, J.S., Barton II, F.E. (Eds.). In: Near Infrared Spectroscopy (NIRS): Analysis of Forage Quality. Agriculture Handbook, 643. USDA-ARS, Government Printing Office, Washington DC, USA, pp. 96–103.

Yan, H., Siesler, H.W., 2018. Hand-held near-infrared spectrometers: state-of-the-art instrumentation and practical applications. NIR News 29, 8–12. https://doi.org/10.1177/0960336018796391.

Yativ, M., Harary, I., Wolf, S., 2010. Sucrose accumulation in watermelon fruits: genetic variation and biochemical analysis. J. Plant Physiol. 167, 589–596. https://doi.org/10.1016/j.jplph.2009.11.009.