LWT - Food Science and Technology 182 (2023) 114831

Contents lists available at ScienceDirect

LWT

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

Green, multivariate approach for obtaining a fingerprint of quality of watermelons at supermarket level using near infrared spectroscopy

Miguel Vega-Castellote [a], Dolores Pérez-Marín [b,**], Irina Torres-Rodríguez [b], José-Manuel Moreno-Rojas [c,d], José-Luis Ordoñez-Díaz [c], María-Teresa Sánchez [a,*]

[a] Department of Bromatology and Food Technology, University of Cordoba, Rabanales Campus, 14071, Córdoba, Spain
[b] Department of Animal Production, University of Cordoba, Rabanales Campus, 14071, Córdoba, Spain
[c] Department of Agroindustry and Food Quality, Andalusian Institute of Agricultural and Fisheries Research and Training (IFAPA), Avenida Menéndez-Pidal, SN, 14004, Córdoba, Spain
[d] Foods for Health Group, Instituto Maimónides de Investigación Biomédica de Córdoba (IMIBIC), Córdoba, Spain

ARTICLE INFO

Keywords:
Portable NIR sensor
In situ analysis
Non-destructive quality assessment
Fruit sweetness
Antioxidant activity

ABSTRACT

Currently, one of the main demands of consumers - especially in large fruits such as watermelon - is for supermarkets to use techniques of non-invasive analysis to enable them to measure the sweetness of the fruits at the time of purchase, and thus avoid having to base the choice exclusively on external appearance. In addition, increasing interest is being shown by consumers in knowing the nutritional quality of healthy foods, such as watermelon. Near infrared spectroscopy (NIRS) was used to assess *in situ* the physicochemical and nutritional quality of half-watermelons, which is the format in which they are usually sold in supermarkets at the beginning of the season, due to their high price. A handheld, new-generation spectrophotometer was used for this purpose, and two modes of analysis, static and dynamic, were studied. The results obtained show the viability of using NIRS technology in dynamic mode at the supermarket level to obtain a reliable measurement of the sweetness of the half-fruits, thus meeting the consumers' demand for sweet-tasting fruits. Promising results were also obtained for measuring the antioxidant activity of the half-watermelons, thus paving the way for the nutritional labelling of this healthy food at the supermarket level.

1. Introduction

Watermelon is a fruit with a very high water content (90% w/w), a sweet flavour and a juicy texture, and it is considered refreshing by consumers (Aimpoint Research, 2020). The fall in the average size of households and the fact that watermelons are very large fruits has meant that supermarkets usually sell half-fruits, or even quarter-fruits, especially at the beginning of the season, when prices are extremely high.

The quality of watermelons is closely linked to their appearance (size, shape, defects), internal colour, and sweetness (Ali, Hashim, Bejo, & Shamsudin, 2017; Sun, Huang, Xu, & Ying, 2010). In watermelons, although the appearance and colour can influence consumer first purchase intention, the decision for subsequent purchases is highly dependent upon consumer satisfaction based on sweetness (Opara & Pathare, 2014). It is also important to take into account that the cultivars vary widely in soluble solid content (SSC) when mature, and an SSC of at least 10% in the flesh, near the centre of the watermelon, can be considered a good indicator of proper maturity (Jie, Zhou, & Wei, 2019; Suslow, 1997).

Moreover, watermelons are also considered a healthy food (Davis et al., 2011), due to their high content in bioactive phytochemicals, and their consumption is linked to a reduction in the risk of cardiovascular diseases, certain types of cancers and other age-related pathologies (Pool-Zobel, Bub, Muller, Wollowski, & Rechkemmer, 1997; Rao & Agarwal, 2000; Rubenstein, 2000). It is therefore of particular interest to consumers to have information on the nutritional quality of these fruits available at the time of purchase. The main bioactive phytochemicals contained in watermelons are carotenoids (lycopene, β-carotene and lutein), vitamins (A, B, C and E) and amino acids (citrulline and arginine) (Guo et al., 2011; Tlili et al., 2011), and the concentration of these compounds depends largely on the watermelon cultivar. Thus, Tarazona-Díaz, Viegas, Maldonado-Martins, and Aguayo (2011) who

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

https://doi.org/10.1016/j.lwt.2023.114831
Received 30 December 2022; Received in revised form 19 April 2023; Accepted 30 April 2023
Available online 1 May 2023

evaluated 5 different watermelon cultivars, reported significant differences in terms of the lycopene concentrations among them, ranging from 11.62 to 14.26 mg kg^{-1}. These authors also found differences in the citrulline concentrations among the watermelon cultivars studied, with a range from 1100 to 4700 mg kg^{-1}. Furthermore, Tlili et al. (2011) found wide variations in vitamin C and total phenol concentrations between 5 watermelons cultivars, with a variation in parameters of 119.7–204.0 mg kg^{-1} and 137.2–260.2 mg kg^{-1}, respectively.

The advent of handheld near infrared (NIR) spectrophotometers over the past few decades has now made it possible to use near infrared spectroscopy (NIRS) technology throughout the different steps of the food supply chain, i.e., in the field, at reception in the industry, delivery to the supermarket, sale, and consumption. Nevertheless, the NIRS applications developed featuring new-generation sensors have been intended in most cases for use in the field or in the agri-food industry (Bec, Grabska, Siesler, & Huck, 2020; Pasquini, 2018). In watermelons, NIRS applications – mainly used to measure SSC – have been used with intact products, using different bench-top NIRS instruments, which are, however, not ideally suited for their incorporation at supermarket level (Abebe, 2006; Ibrahim, Daood, Égei, Takács, & Helyes, 2022; Jie, Xie, Fu, Rao, & Ying, 2013; Jie, Xie, Rao, & Ying, 2014; Qi et al., 2014; Tamburini, Costa, Rugiero, Pedrini, & Marchetti, 2017; Tian, Ying, Lu, Fu, & Yu, 2007). Only Flores et al. (2008) have carried out SSC measurements on slices of the fruit, analysing a slice 6 mm thick cut from the equatorial region and also using a bench-top instrument, while Vega--Castellote, Sánchez, Torres, De la Haba, and Pérez-Marín (2022) determined SSC in intact watermelons using a portable handheld instrument, taking spectra at the equator of the fruits. NIRS analysis with handheld instruments can be performed not only in static mode, taking point spectra readings on the surface of the watermelons analysed, without the instrument moving during the measurement, but also in dynamic mode, with the sensor being moved during the spectral measurements all along the area of the watermelons analysed**.** It is therefore highly suitable for cut fruits, since one characteristic of watermelons is the gradation in sweetness from the rind to the centre of the fruit (Magwaza & Opara, 2015).

As well as measuring SSC, Tamburini et al. (2017) determined the content of phytonutrients (lycopene and β-carotene) in intact 'Minirossa' watermelon, a cultivar with a very thin rind (<0.5 mm), using a bench-top instrument working in reflectance mode. They positioned the fruits on stationary and moving belts, and proved the suitability of using NIRS technology for evaluation by the industry of watermelon as a functional food. Likewise, Ibrahim et al. (2022) determined the phytonutrient content (lycopene, total carotenoids, vitamin C, β-carotene, γ-carotene) using a bench-top diode-array instrument working in the 950–1650 nm range in reflectance mode, placing the portion of watermelon flesh in a dish to acquire the spectral information. Although these works indicate the suitability of NIRS technology for measuring the nutritional value of the product, the type of instrumentation used by these authors is not suitable for monitoring the fruits at the points of dispatch and sale, since the equipment must be used in controlled laboratory conditions. Similarly, it should be noted that the NIRS predictive models of phytonutrients were carried out either using a single cultivar of watermelon with a very thin rind, or using a small portion of watermelon flesh.

The aim of this research was to fine-tune and optimize the methodology to measure *in situ* the sweetness, colour and antioxidant activity of half-watermelons belonging to different cultivars using a portable NIR instrument based on linear variable filter (LVF) technology (MicroNIR™ Pro 1700), suitable for use at supermarket level.

2. Materials and methods

2.1. Watermelon sampling and reference analysis

A total of 249 watermelons were analysed in this study: *Citrullus lanatus* (Thunb.) cv. 'Bazman' (N = 5), 'Bengala' (N = 26), 'Boston' (N = 33), 'Fashion' (N = 54), 'Fenway' (N = 38), 'Jamaica' (N = 5), 'Premium' (N = 68), 'Premium Frilly' (N = 17) and 'Red Jasper' (N = 3), all harvested between June and October 2021 and provided by an agricultural company from Almería (Spain). The intact fruits were transported to the laboratory in the Faculty of the Agricultural and Forestry Engineering at the University of Cordoba (Spain) and kept at a room temperature (20 °C) until the following day, when laboratory testing was performed. Prior to measurement, the watermelons were cut in half along the equatorial axis.

In order to obtain an objective indication of the internal colour of the fruits, the colour parameters (L* and a* values and colour index (CI) (1000 [a*/(L*b*)]), which can be related to the organoleptic quality of foods (Tarazona-Díaz et al., 2011), were measured at three equidistant points (close to the rind, close to the heart and in the middle of these two points) and then averaged using a Minolta Chroma Meter CR-400 (Minolta Corporation, Ramsay, NJ, USA), using the standard illuminant D65 and 2° observer (Pardo, Gómez, Tardaguila, Amo, & Varón, 1997). After that, the flesh was cut and liquefied to determine the SSC (%) in samples of freshly squeezed juice by refractometry, using a temperature compensated WYA-1S Digital Abbe refractometer (Xi'an Yima Optoelec Co., Ltd., Shanghai, China). Additionally, the flesh was lyophilized in order to measure the antioxidant activity using the DPPH methodology (free radical DPPH 1,1-diphenyl-2-picryl-hydrazyl scavenging capacity). Next, the lyophilized samples were ground and stored at −80 °C until analysis. The bioactive compounds were extracted following the method described by Moreno-Rojas et al. (2018), using a mixed solution of methanol and deionized water (80:20, v/v), acidified with formic acid. Thus, 0.2 mg of lyophilized sample was extracted with the solvent solution before being sonicated and centrifuged. The supernatant was collected, and pellets were re-extracted using the same protocol. The extracted samples were transferred into vials and stored at −80 °C until their analysis. Antioxidant activity was quantified using DPPH radical by measuring the absorption at 515 nm after 50 min of reaction (Ordóñez-Díaz et al., 2020). The results were expressed as g Trolox equivalents (TE) per kg of dried weight.

All 249 watermelons were used to measure the SSC and internal colour, while only 106 out of the 249 samples were used for the antioxidant activity analysis. On each day of analysis, either one out of every two or one out of every three watermelons were chosen – depending on the number of samples analysed per day – and we tried to avoid those fruits showing any kind of external or internal defects. These N = 106 samples belonged to the 'Bazman' (N = 2), 'Bengala' (N = 8), 'Boston' (N = 16), 'Fashion' (N = 26), 'Fenway' (N = 19), 'Jamaica' (N = 2), 'Premium' (N = 25), 'Premium Frilly' (N = 7) and 'Red Jasper' (N = 1) cultivars.

2.2. Instrumentation and NIR spectra acquisition

The spectral information of the half-watermelons was read using a portable linear variable filter (LVF)-NIRS instrument (MicroNIR™ Pro 1700, VIAVI Solutions, Inc., San Jose, California, USA), working in reflectance mode in the spectral range 908–1676 nm with a constant interval of 6.2 nm. The illumination source of this handheld spectrophotometer were two integrated vacuum tungsten lamps, and the optical window was around 227 mm^2. The spectra were acquired using the VIAVI MicroNIR software Pro version 2.2 (VIAVI Solutions, Inc., San Jose, California, USA). The performance of the instrument was checked every 10 min using white and dark reference measurements – for the former, Spectralon™, 99% diffuse reflectance, and for the latter, by placing a black cover over the analysis window. The half-watermelons were placed centrally on the fruit holder and two modes of analysis were used for analysing each fruit: static and dynamic. For static mode, the sensor was placed at 2 different points in the fruit; first, close to the rind and then in the central point close to the heart. In dynamic mode, two spectra were taken for each sample by moving the sensor over the

entire surface of the cut watermelon; for the first, a straight line was traced near the area of the watermelon rind, and for the second, the sensor was moved in a straight line parallel to the previous one, passing through the centre of the fruit. In both cases, the 2 spectra were averaged to provide a mean spectrum per half-fruit in the two analysis modes.

2.3. Exploratory spectral analysis of the watermelon population and definition of calibration and internal validation sets

Data pre-processing and chemometric treatments were performed using the Matlab software version 2019a (The Mathworks, Inc., Natick, MA, USA). Prior to the development of NIR calibrations, a Principal Components Analysis (PCA) was applied to the four spectral data sets: the 249 watermelons used to study the SSC and colour parameters and the 106 watermelons used to assess the antioxidant activity, in both cases in static and dynamic modes of analysis. The spectra were pre-treated using a combination of standard normal variate (SNV) and de-trending (DT) for scatter correction (Barnes, Dhanoa, & Lister, 1989), together with the first derivative treatment. Spectral outlier samples were studied using the Mahalanobis Global distance (GH) –calculated according to Walsh, Guthrie, and Burney (2000) as $GH = D^2/f$, where 'D' is the Mahalanobis distance calculated on principal components and 'f' the number of factors used in the PCA. Those samples showing GH values over 4 were studied in detail and eliminated from their respective sets if their removal was justified. Once the spectral outliers were removed from these sets, a new PCA was carried out using the same spectral pre-treatments, and the loadings plot was studied in detail using the set of watermelons used for studying the SSC and colour parameters analysed in static mode. Prior to the comparison of both modes of analysis, the authors decided to use the set of data analysed in static mode to interpret the loadings, since this analysis mode was considered simpler than the dynamic one.

A structured selection of the calibration and internal validation groups was carried out by ordering the samples by their GH distance to the centre of the population following the recommendations of Shenk and Westerhaus (1991). In order to match the calibration and internal validation sets with the same samples and, consequently, to be able to compare the prediction accuracy of the quantitative models, in the case of the sets used to measure SSC and colour, one in every five samples were then selected from the set of watermelons analysed in static mode to be part of the internal validation set, and the remaining samples were reserved for calibration. For both mode of analysis, static and dynamic, the calibration sets were finally made up of the same samples. The same procedure was followed for the sets used to determine the antioxidant activity parameter, although, in this case, given the smaller number of watermelons included in these groups, only one on every seven samples was selected to be included in the internal validation set.

The calibration models for the SSC (%), L*, a* and CI colour parameters and the antioxidant activity (g TE kg^{-1}) studied in half-watermelons were developed using the Partial Least Squared (PLS) regression method. For each of the physicochemical and nutritional quality parameters analysed, the first and the second derivative treatments in combination with SNV and DT for scatter correction were tested. 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. For each of the parameters analysed, a confidence limit for the Hotelling T^2 statistic was calculated at 95% and the 'Y' Studentized Residuals were studied. To identify the most suitable mode of analysis for predicting the internal physicochemical and nutritional quality in half-watermelons *in situ*, tests were run to identify potential differences between the models developed. To achieve this, the SECV values for the best models previously selected for the parameters studied were compared using Fisher's F test (Mark & Workman, 2003), with $P = 0.05$.

The best models for the SSC, colour and antioxidant activity parameters were subjected to evaluation following the protocols outlined by Shenk, Westerhaus, and Abrams (1989) and Windham, Mertens, & Barton II (1989) and using the samples included in the internal validation set.

3. Results and discussion

3.1. Characterization of the research population

The average spectrum for the sets of samples analysed in static and dynamic modes used to predict the SSC and colour parameters were plotted (Fig. 1). As it can be observed, the average spectrum for the samples analysed in static and dynamic modes of analysis did not show great absorbance differences between them. Two spectral bands related to the absorption of water were identified at 970 nm and 1450 nm. In addition, two peaks at around 1200 nm and 1440 nm could be observed which could be related to sugar absorption bands (Osborne, Fearn, & Hindle, 1993; Shenk, Workman, & Westerhaus, 2008).

The analysis of spectrally anomalous samples by means of a PCA highlighted three samples in the sets of watermelons used to predict the SSC and colour parameters. The same three samples were identified as anomalous for both static and dynamic modes of analysis. These samples showed GH values of 4.36–13.40, and belonged to the 'Fashion', 'Fenway' and 'Premium' cultivars. In all cases, the fruits were overripe, and the flesh had a soft texture. These three samples were removed from the sets of samples analysed in static and dynamic modes, used for measuring SSC and colour. In the case of the samples used for measuring antioxidant activity, the PCA highlighted two samples belonging to the 'Fashion' and 'Fenway' cultivars showing high GH of 4.08–4.14, again, for both modes of analysis. These two samples were, in fact, the same ones identified from the sets used for the SSC and colour determination and were also removed. The third sample identified in the sets used to measure the physicochemical parameters belonging to the 'Premium' cultivar was not originally included in the N = 106 samples used to measure antioxidant activity. As a result, after discarding the outlier samples, the sets analysed in static and dynamic modes to measure SSC and colour contained a total of N = 246 samples, and the sets analysed in both modes of analysis to measure antioxidant activity were made up of N = 104 samples.

The loadings plot obtained from the PCA applied to the set of N = 246 watermelons analysed in static mode to measure SSC and colour (Fig. 2) highlighted three peaks in the 1100–1250 region, which could

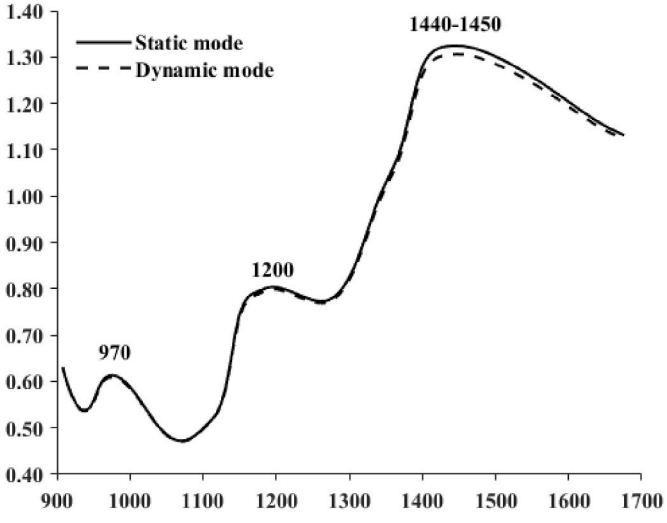


Fig. 1. Mean NIR spectra for the half-watermelons analysed using the linear variable filters (LVF) instrument in static and dynamic modes.

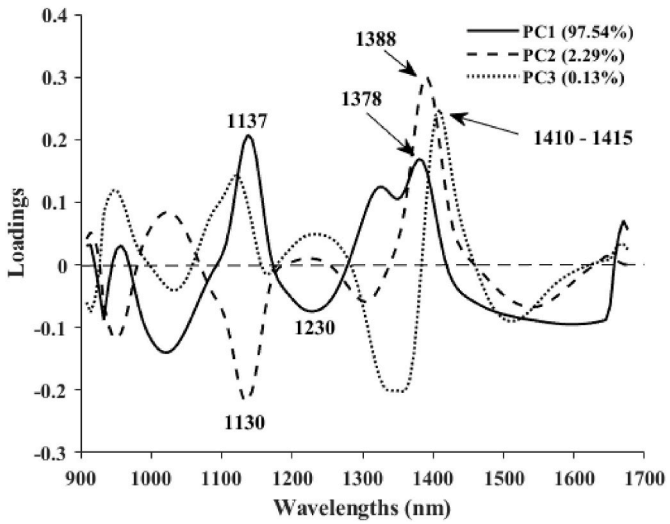


Fig. 2. PC1, PC2 and PC3 loadings plot for the N = 246 watermelons used to predict the SSC and colour parameters analysed in static mode of analysis.

be associated with the third overtone of the C–H groups which, in turn, could be linked to carotenoids (Osborne et al., 1993; Shenk et al., 2008; Tamburini et al., 2017). Furthermore, the peaks at around 1378, 1388 and 1410 nm have been linked to sugar-related absorption bands (Deák, Szigedi, Pék, Baranowski, & Helyes, 2015; Osborne et al., 1993), and the peak at around 1415 nm to bands characteristic of phenolics (Miller, 2001). This spectral variability identified by means of the PCA associated to SSC and antioxidant compounds (mainly carotenoids and phenolic compounds) can be related to the wide variability in the samples in terms of maturity stage and cultivars. These same factors affected the concentration of SSC and antioxidant compounds (Ibrahim et al., 2022; Tlili et al., 2011) in the watermelons received in the laboratory, with a range of fruits harvested from early to full maturity.

Finally, prior to model development, the sets used for measuring SSC, colour and antioxidant activity were divided into the calibration and internal validation groups, which displayed close values for all the parameters studied, with the internal validation range values lying within those of the calibration sets in all cases (Table 1). The mean and standard deviation (SD) values for the SSC and colour parameters were very similar to those obtained in other NIR studies in the literature working with cut watermelons (Flores et al., 2008; Ibrahim et al., 2022). No studies were found in the literature for the prediction of antioxidant activity in half-watermelons using NIR spectroscopy.

3.2. Model development for the prediction of SSC, colour and antioxidant activity parameters using PLS regression

In all cases, except for the a* parameter of colour, the SECV values obtained were lower in those models calculated to predict the SSC, colour and antioxidant activity parameters using data acquired in the dynamic analysis mode (Table 2). Furthermore, significant differences were found ($P < 0.05$) between the two modes of analysis for the L* and CI parameters. These differences point out the importance of moving the sensor along the cut watermelon surface during the analysis, given the characteristic heterogeneity of watermelon flesh, with irregular concentrations of sugars and phytochemical compounds with antioxidant activity (Ibrahim et al., 2022; Magwaza & Opara, 2015). In this way, greater spectral variability, and thus more representative information, can be obtained from the analysis. Consequently, the dynamic mode was considered as the optimal way to perform the analysis in this study, and it was the mode chosen to assess the results from this point in the study onwards.

The cross-validation results for the best models to measure SSC in cut watermelons (Table 2) reported prediction results that could be considered as good (Shenk & Westerhaus, 1996; Williams, 2001). These results are of particular interest, since these models could be used to measure whether the fruits are sufficiently mature once cut and ready to be placed on the supermarket shelves, thus ensuring the loyalty of consumers of cut watermelons in their subsequent purchases.

For the colour parameter L*, the model developed could be used according to Shenk and Westerhaus (1996) to classify watermelons into high, medium and low values of this parameter. Moreover, for the a* and CI* parameters, these models could discriminate samples showing low and high values for these parameters. One issue to take into account was that the portable NIR instrument used did not work within the visible range, which hindered the prediction of these parameters. Nevertheless, these models for colour prediction in half-watermelons could be considered as an interesting support tool to measure the degree of maturity of the watermelon and, consequently, the quality of the fruit.

Finally, the models used to predict antioxidant activity would enable

Table 1
Characterization of the calibration and internal validation sets used for measuring soluble solid content, colour, and antioxidant activity.

Parameter	Mode of analysis	Set	[a] N	Range	Mean	[b] SD	[c] CV (%)
Soluble solid content (%)	Static	Calibration	197	4.40–12.05	8.42	1.48	17.58
		Validation	49	4.80–11.45	8.79	1.55	17.63
	Dynamic	Calibration	197	4.40–12.05	8.42	1.48	17.58
		Validation	49	4.80–11.45	8.79	1.55	17.63
L*	Static	Calibration	197	29.03–59.98	41.76	5.70	13.65
		Validation	49	32.13–53.80	42.01	5.58	13.28
	Dynamic	Calibration	197	29.03–59.98	41.76	5.70	13.65
		Validation	49	32.13–53.80	42.01	5.58	13.28
a*	Static	Calibration	197	13.83–35.16	27.21	4.14	15.22
		Validation	49	19.78–33.97	27.89	3.26	11.69
	Dynamic	Calibration	197	13.83–35.16	27.21	4.14	15.22
		Validation	49	19.78–33.97	27.89	3.26	11.69
Colour index	Static	Calibration	197	20.54–62.31	36.68	8.00	21.81
		Validation	49	25.60–58.67	37.49	8.03	21.42
	Dynamic	Calibration	197	20.54–62.31	36.68	8.00	21.81
		Validation	49	25.60–58.67	37.49	8.03	21.42
Antioxidant activity (g TE kg^{-1})	Static	Calibration	89	56.68–137.43	83.37	19.32	23.17
		Validation	15	57.88–126.98	82.09	19.77	24.08
	Dynamic	Calibration	89	56.68–137.43	83.37	19.32	23.17
		Validation	15	57.88–126.98	82.09	19.77	24.08

[a] N: number of samples.
[b] SD: standard deviation.
[c] CV: coefficient of variation.

Table 2
Calibration statistics for the best equations obtained for the prediction of the soluble solid content, colour, and antioxidant activity parameters in half-watermelons.

Parameter	Mode of analysis	Math treatment	[A] N	Range	Mean	[B] SD	[C] LV	[D] SECV	[E] R^2_{cv}	[F] RPD_{cv}
Soluble solid content (%)	Static	SNV + DT + Derivate 2,5,5	191	4.40–12.05	8.42	1.43	16	0.67[a]*	0.78	2.13
	Dynamic	SNV + DT + Derivate 1,5,5	193	4.40–12.05	8.41	1.47	16	0.70[a]	0.78	2.10
L*	Static	SNV + DT + Derivate 1,5,5	193	29.03–56.41	41.64	5.43	6	4.04[a]	0.45	1.34
	Dynamic	SNV + DT + Derivate 1,5,5	194	29.03–58.11	41.61	5.48	4	3.48[b]	0.60	1.57
a*	Static	SNV + DT + Derivate 1,5,5	186	19.27–35.16	27.78	3.44	5	2.74[a]	0.37	1.26
	Dynamic	SNV + DT + Derivate 2,5,5	188	19.27–35.16	27.71	3.48	4	2.79[a]	0.36	1.25
Colour index	Static	SNV + DT + Derivate 2,5,5	195	20.54–62.31	36.57	7.94	5	7.15[a]	0.20	1.11
	Dynamic	SNV + DT + Derivate 2,5,5	194	20.54–62.31	36.45	7.78	14	6.09[b]	0.40	1.28
Antioxidant activity (g TE kg^{-1})	Static	SNV + DT + Derivate 2,3,3	88	0.57–1.36	0.83	0.19	13	0.15[a]	0.33	1.27
	Dynamic	SNV + DT + Derivate 2,3,3	87	0.57–1.32	0.82	0.18	11	0.13[a]	0.44	1.38

[A] N: number of samples; [B] SD: standard deviation; [C] LVs: latent variables; [D] SECV: standard error for cross-validation; [E] R^2_{cv}: coefficient of determination for cross-validation; [F] RPD_{cv}: residual predictive deviation for cross-validation.
* SECV values for one parameter in static and dynamic modes of analysis with different letters indicates significant differences ($P < 0.05$) between them.

us to distinguish samples with high and low values of this parameter. The *in-situ* prediction of antioxidant activity in watermelons and the possibility of making this information available to the consumer on the product label could be used as a way of boosting the consumption of this fruit, given the health benefits that have been associated to the intake of antioxidant compounds present in watermelon flesh (Perkins-Veazie, Collins, & Clevidence, 2007; Rao & Agarwal, 2000). No studies can be found in the literature which predict antioxidant activity in watermelons using portable NIR spectroscopy sensors. Only Tamburini et al. (2017) and Ibrahim et al. (2022) developed models to predict different antioxidant compounds in intact watermelons and in homogenized samples of flesh, respectively, although none of them used portable sensors. This, therefore, highlights the importance of our research aimed at predicting the antioxidant activity of watermelons on the supermarket shelves.

3.3. Internal validation

The models developed to predict the SSC, colour and antioxidant activity parameters using samples analysed in dynamic mode were subjected to a validation procedure using those samples included in the internal validation set.

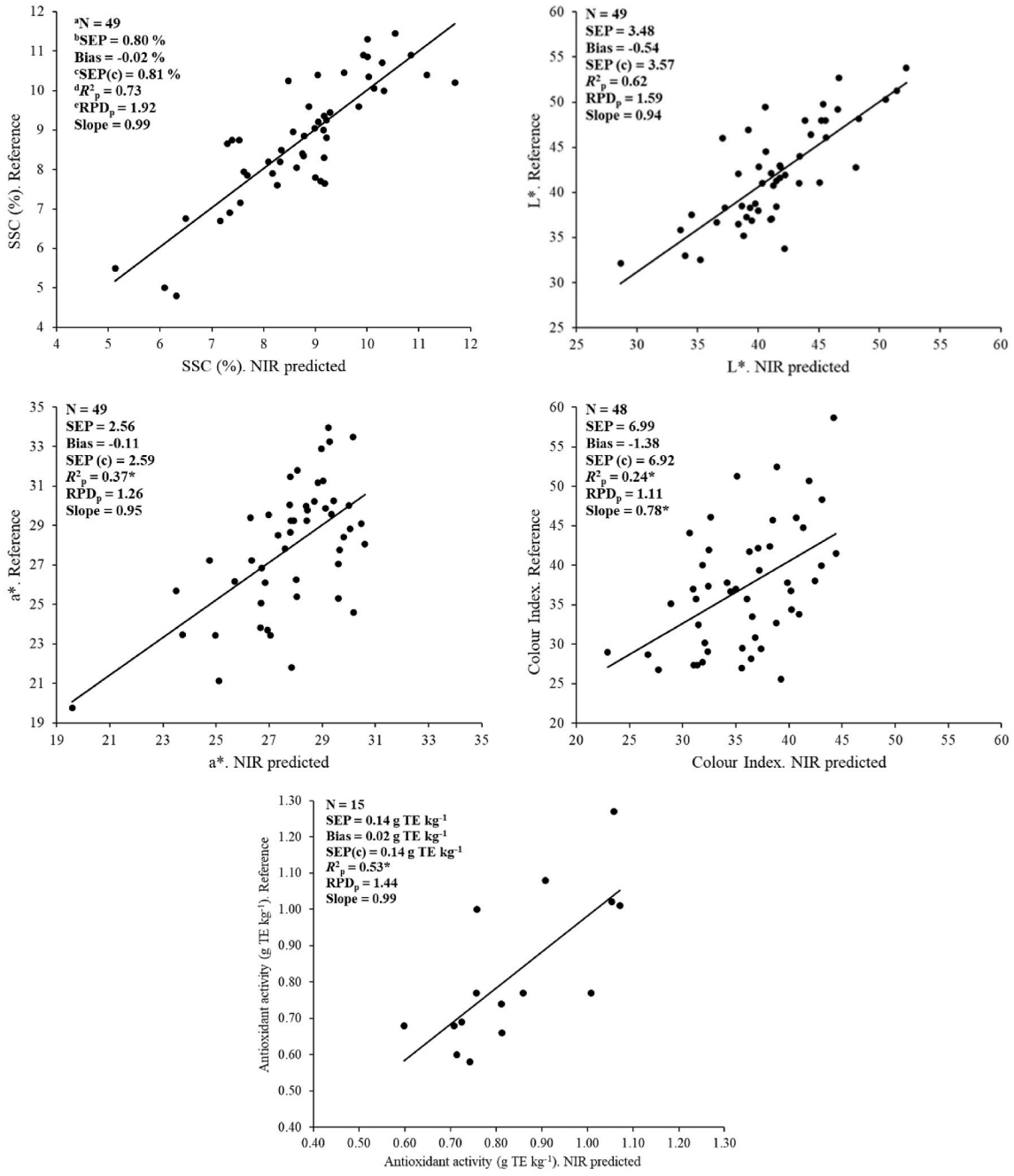
For the SSC, very good prediction results were obtained (Fig. 3), given the complexity of the model in which 9 different watermelon cultivars were included. One sample belonging to the 'Bengala' cultivar showed a 'Y' Studentized Residual value of −2.83. It is important to note that although the Hotelling T^2 value for this sample was of 12.20 (Hotelling T^2_{limit} for the SSC parameter, 29.52), this sample presented an external bruise that could have influenced the texture of the watermelon flesh. The R^2_p value ($R^2_p > 0.6$), slope (0.9 < slope <1.1), the SEP(c) (SEP (c) < 1.30 × SEC) and the bias (bias < ± 0.6 × SEC) met the requirements stablished in the protocol of Windham, Mertens, & Barton II (1989). Consequently, this equation could be used to evaluate the maturity and sweetness of the half-watermelons *in situ* on the supermarket shelves.

The internal validation of the L* parameter showed good prediction results (Fig. 3) and met all the requirements established in the validation protocol used, ensuring accurate predictions for this parameter. The model developed for the a* parameter also met all the requirements established in the protocol, except for the coefficient of determination for the internal validation ($R^2_p < 0.6$) (Fig. 3). In this case, one sample belonging to the 'Bengala' cultivar showed high residual predicted value for the a* parameter ('Y' Studentized Residual value = 2.20) and a Hotelling T^2 value of 1.25 which was under the calculated limit (Hotelling T^2_{limit} for the a* parameter, 9.84). In addition, another sample belonging to the 'Premium' cultivar showed a Hotelling T^2_{limit} value of 20.62, which was over the calculated limit, and a 'Y' Studentized Residual value of −1.44. These particular samples, which presented a low and a high reference value for the a* parameter, respectively, of 21.81 and 32.90 - when the average and standard deviation values for the calibration set of this parameter were 27.21 and 4.14, respectively - were not a good representation of the set of samples included in the model developed to predict the a* parameter (Fig. 4). In addition, the results obtained for the CI parameter did not meet the requirements established in the protocol in terms of the R^2_p and the slope (Fig. 3), with one sample belonging to the 'Fashion' cultivar showing a 'Y' Studentized Residual value of 4.64 and Hotelling T^2 of 159.95 (Hotelling T^2_{limit} for the CI parameter, 26.23). The extremely high residual and Hotelling T^2 values obtained were due to the over-ripeness of the sample, which was therefore eliminated from the internal validation set for the CI parameter. In addition, this sample showed an extreme value for the colour parameter (52.04), which was not a good representation of the set of samples included in the model (Fig. 4). The poor results for the CI parameter indicate that this equation cannot be applied in routine operations. This may be due to the limited range of the LVF instrument, which did not work in the visible range, which would facilitate the prediction of these colour parameters.

The model developed to predict the antioxidant activity parameter met all the requirements established by Windham, Mertens, & Barton II (1989), except for the R^2_p (Fig. 3). The sample presenting the highest 'Y' Studentized Residual value (−2.08) belonged to the 'Premium' cultivar and showed a Hotelling T^2 value of 19.76 under the limit (Hotelling T^2_{limit} for the antioxidant activity parameter, 23.86). This watermelon showed abrasions in the rind and an extreme value for the L* parameter (51.79). These results can be considered as a promising initial approximation for predicting antioxidant activity in half-fruits using portable NIR sensors. Further studies will be needed which include a greater number of samples both in the calibration and internal validation sets belonging to a wider range of watermelon cultivars and growing seasons.

4. Conclusions

The results obtained showed that NIRS technology can be successfully used *in situ* at the supermarket level, to inform consumers of the degree of sweetness of the half-watermelons they are purchasing, thus ensuring their loyalty for subsequent purchases of these types of large half-fruits, which are sold in this way due to the reduction in the size of households and the high prices of watermelons. Likewise, the incorporation of this non-destructive and multiparameter analysis technique permits to know the antioxidant activity of the fruits on sale, thus enabling them to be labelled with nutritional information, which is especially important since they are considered a healthy food due to their high content in bioactive phytochemicals. Furthermore, it is recommendable to carry out the NIRS analysis in dynamic mode, in order to evaluate the entire half of the fruit to be sold, given the heterogeneity in the distribution of sugars and phytochemical substances, characteristic of watermelons. Thus, this research can be considered as a first step towards the selective purchase of half-watermelons based on



[a] N: number of samples; [b] SEP: standard error of prediction; [c] SEP(c): standard error of prediction with bias corrected; [d] R^2_p: coefficient of determination for the prediction; [e] RPD$_p$: residual predicted deviation for the prediction.
*Values with an asterisk did not meet the protocol established by Windham et al. (1989).

Fig. 3. Reference *vs* NIR predicted values for the prediction of the soluble solid content (SSC), colour and antioxidant activity parameters using half-fruits analysed in dynamic mode
[a] N: number of samples; [b] SEP: standard error of prediction; [c] SEP(c): standard error of prediction with bias corrected; [d] R^2_p: coefficient of determination for the prediction; [e] RPD$_p$: residual predicted deviation for the prediction. *Values with an asterisk did not meet the protocol established by Windham, Mertens, & Barton II (1989).

the taste of consumers for sweet or very sweet fruits and their demand for a high content of nutraceuticals. A larger database including more samples for each of the cultivars used in this study will be needed in the future to develop more robust prediction models for all the parameters tested and, in particular, for antioxidant activity.

Funding

This research was carried out as part of the research project P-12021056 'NIRS technology for pre and postharvest quality control of watermelons', funded by Agrícola Navarro de Haro S.L. (Cuevas de Almanzora, Almería, Spain). The funding source was not involved in the preparation of this article.

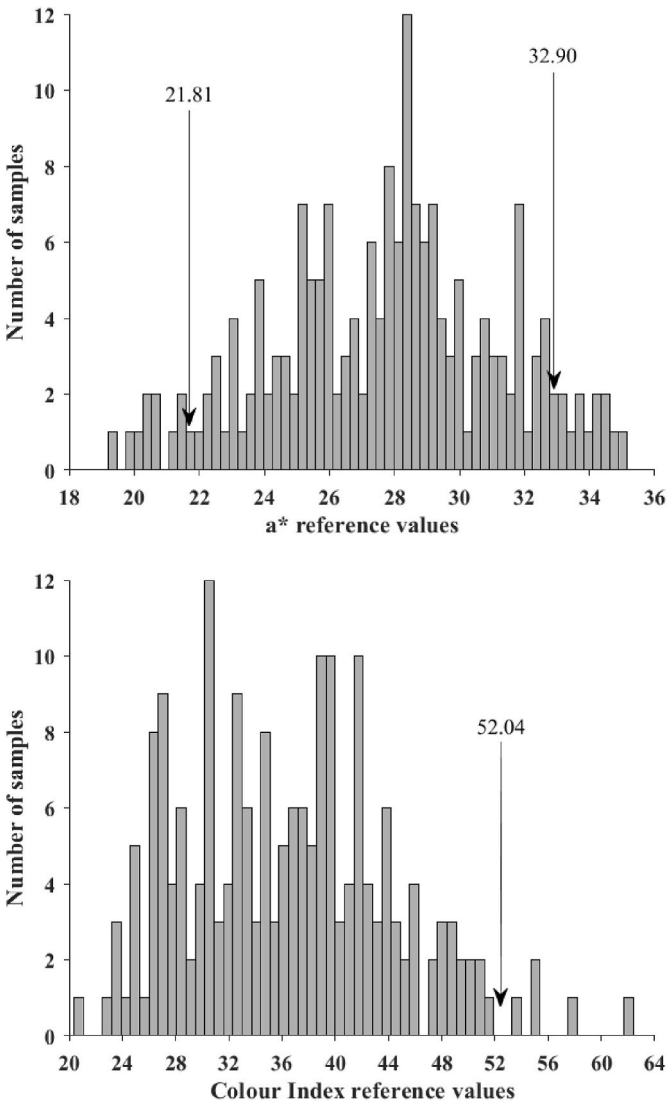


Fig. 4. Distribution of the watermelons included in the models developed for parameters a* and colour index throughout their respective ranges.

CRediT authorship contribution statement

Miguel Vega-Castellote: Funding acquisition, Data acquisition, Methodology, Formal analysis, Investigation, Software, Data curation, Validation, Writing – original draft, Writing – review & editing, Visualization. **Dolores Pérez-Marín:** Conceptualization, Methodology, Validation, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Irina Torres-Rodríguez:** Funding acquisition, Data acquisition, Formal analysis, Investigation, Software, Data curation, Writing – original draft, Writing – review & editing, Visualization. **José-Manuel Moreno-Rojas:** Funding acquisition, Data acquisition, Methodology, Investigation, Writing – original draft. **José-Luis Ordoñez-Díaz:** Funding acquisition, Data acquisition, Methodology, Investigation, Writing – original draft. **María-Teresa Sánchez:** Conceptualization, Methodology, Validation, Investigation, Resources, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

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

Data availability

Data will be made available on request.

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).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.lwt.2023.114831.

References

Abebe, T. (2006). Total sugar and maturity evaluation of intact watermelon using near infrared spectroscopy. *Journal of Near Infrared Spectroscopy, 14*, 67–70. https://doi.org/10.1255/jnirs.588

Aimpoint Research. (2020). *Annual consumer research findings [Fact Sheet]. National watermelon promotion board.* accessible https://www.watermelon.org/wp-content/uploads/2020/04/NWPBWinter2020CONSUMERFINALFORWEB.pdf. (Accessed 26 December 2022).

Ali, M. M., Hashim, N., Bejo, S. K., & Shamsudin, R. (2017). Rapid and nondestructive techniques for internal and external quality evaluation of watermelons: A review. *Scientia Horticulturae, 225*, 689–699. https://doi.org/10.1016/j.scienta.2017.08.012

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard normal variate transformation and de-trending of near infrared diffuse reflectance spectra. *Applied Spectroscopy, 43*, 772–777. https://doi.org/10.1366/0003702894202201

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

Davis, A. R., Webber, C. L., III, Fish, W. W., Wehner, T. C., King, S., & Perkins-Veazie, P. (2011). L-citrulline levels in watermelon cultigens tested in two environments. *HortScience, 46*, 1572–1575. https://doi.org/10.21273/HORTSCI.46.12.1572

Deák, K., Szigedi, T., Pék, Z., Baranowski, P., & Helyes, L. (2015). Carotenoid determination in tomato juice using near infrared spectroscopy. *International Agrophysics, 29*, 275–282. https://doi.org/10.1515/intag-2015-0032

Flores, K., Sánchez, M. T., Pérez-Marín, D. C., López, M. D., Guerrero, J. E., & Garrido-Varo, A. (2008). Prediction of total soluble solid content in intact and cut melon and watermelons using near infrared spectroscopy. *Journal of Near Infrared Spectroscopy, 16*, 91–98. https://doi.org/10.1255/jnirs.771

Guo, S., Liu, J., Zheng, Y., Huang, M., Zhang, H., Gong, G., et al. (2011). Characterization of transcriptome dynamics during watermelon fruit development: Sequencing, assembly, annotation and gene expression profiles. *BMC Genomics, 12*, 454. https://doi.org/10.1186/1471-2164-12-454, 1–13.

Ibrahim, A., Daood, H. G., Égei, M., Takács, S., & Helyes, L. (2022). A comparative study between Vis/NIR spectroradiometer and NIR spectroscopy for the non-destructive quality assay of different watermelon cultivars. *Horticulturae, 8*(509), 1–17. https://doi.org/10.3390/horticulturae8060509

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. *Journal of Food Engineering, 118*, 387–392. https://doi.org/10.1016/j.jfoodeng.2013.04.027

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 Biology and Technology, 90*, 1–6. https://doi.org/10.1016/j.postharvbio.2013.11.009

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

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

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

Miller, C. E. (2001). Chemical principles of near-infrared technology. In P. C. Williams, & K. H. Norris (Eds.), *Near-infrared technology in the agricultural and food industries* (pp. 19–37). St. Paul, MN, USA: American Association of Cereal Chemists.

Moreno-Rojas, J. M., Moreno-Ortega, A., Ordóñez, J. L., Moreno-Rojas, R., Pérez-Aparicio, J., & Pereira-Caro, G. (2018). Development and validation of UHPLC-HRMS methodology for the determination of flavonoids, amino acids and organosulfur compounds in black onion, a novel derived product from fresh shallot

onions (*Allium* cepa var. *aggregatum*). *Lebensmittel-Wissenschaft & Technologie, 97*, 376–383. https://doi.org/10.1016/j.lwt.2018.07.032

Opara, U. L., & Pathare, P. B. (2014). Bruise damage measurement and analysis of freshhorticultural produce–A review. *Postharvest Biology and Technology, 91*, 9–24. https://doi.org/10.1016/j.postharvbio.2013.12.009

Ordóñez-Díaz, J. L., Hervalejo, A., Pereira-Caro, G., Muñoz-Redondo, J. M., Romero-Rodríguez, E., Arenas-Arenas, F. J., et al. (2020). Effect of rootstock and harvesting period on the bioactive compounds and antioxidant activity of two orange cultivars ('Salustiana' and 'Sanguinelli') widely used in juice industry. *Processes, 8*, 1212. https://doi.org/10.3390/pr8101212, 1–19.

Osborne, B. G., Fearn, T., & Hindle, P. H. (1993). *Practical NIR spectroscopy with application in food and beverage analysis*. Harlow, Essex, UK: Longman Scientific and Technical.

Pardo, J. E., Gómez, R., Tardaguila, J., Amo, M., & Varón, R. (1997). Quality evaluation of watermelon varieties (*Citrullus vulgaris* S.). *Journal of Food Quality, 20*, 547–557. https://doi.org/10.1111/j.1745-4557.1997.tb00495.x

Pasquini, C. (2018). Near infrared spectroscopy: A mature analytical technique with new perspectives – a review. *Analytica Chimica Acta, 1026*, 8–36. https://doi.org/10.1016/j.aca.2018.04.004

Perkins-Veazie, P., Collins, J. K., & Clevidence, B. (2007). Watermelons and health. Acta. *Horticulturae, 731*, 121–128. https://doi.org/10.17660/ActaHortic.2007.731.17

Pool-Zobel, B. L., Bub, A., Muller, H., Wollowski, I., & Rechkemmer, G. (1997). Consumption of vegetables reduces genetic damage in humans: First result of a human intervention trial with carotenoid-rich foods. *Carcinogenesis, 18*, 1847–1850. https://doi.org/10.1093/carcin/18.9.1847

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. *Journal of Innovative Optical Health Sciences, 7*. https://doi.org/10.1142/S179354581350034X, 1350034-1–1350034-11.

Rao, A. V., & Agarwal, S. (2000). Role of antioxidant lycopene in cancer and heart disease. *Journal of the American College of Nutrition, 19*, 563–569. https://doi.org/10.1080/07315724.2000.10718953

Rubenstein, E. (2000). Biologic effects of and clinical disorders caused by nonprotein amino acids. *Medicine, 79*, 80–89. https://doi.org/10.1097/00005792-200003000-00002

Shenk, J. S., & Westerhaus, M. O. (1991). Population structuring of near infrared spectra and modified partial least squares regression. *Crop Science, 31*, 1548–1555. https://doi.org/10.2135/cropsci1991.0011183X003100060034x

Shenk, J. S., & Westerhaus, M. O. (1996). Calibration the ISI way. In A. M. C. Davies, & P. C. Williams (Eds.), *Near infrared spectroscopy: The future waves* (pp. 198–202). Chichester, UK: NIR Publications.

Shenk, J. S., Westerhaus, M. O., & Abrams, S. M. (1989). Protocol for NIRS calibration: Monitoring. Analysis results and recalibration. In G. C. Martens, J. S. Shenk, & F. E. Barton II (Eds.), *Near infrared spectroscopy (NIRS): Analysis of forage quality, Agriculture handbook, n° 643* (pp. 106–110). Washington D. C., USA: USDA-ARS, US Government Printing Office.

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

Sun, T., Huang, K., Xu, H., & Ying, Y. (2010). Research advances in nondestructive determination of internal quality in watermelon/melon: A review. *Journal of Food Engineering, 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 26 December 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. https://doi.org/10.3390/s17040746, 1–12.

Tarazona-Díaz, M. P., Viegas, J., Maldonado-Martins, M., & Aguayo, E. (2011). Bioactive compounds from flesh and by-product of fresh-cut watermelon cultivars. *Journal of the Science of Food and Agriculture, 91*, 805–812. https://doi.org/10.1002/jsfa.4250

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. *Journal of Zhejiang University - Science B, 8*, 105–110. https://doi.org/10.1631/jzus.2007.B0105

Tlili, I., Hdider, C., Lenucci, M. S., Riadh, I., Hager, J., & Dalessandro, G. (2011). Bioactive compounds and antioxidant activities during fruit ripening of watermelon cultivars. *Journal of Food Composition and Analysis, 24*, 923–928. https://doi.org/10.1016/j.jfca.2011.03.016

Vega-Castellote, M., Sánchez, M. T., Torres, I., De la Haba, M. J., & Pérez-Marín, D. (2022). Assessment of watermelon maturity using portable new generation NIR spectrophotometers. *Scientia Horticulturae, 304*, Article 111328. https://doi.org/10.1016/j.scienta.2022.111328, 1–7.

Walsh, K. B., Guthrie, J. A., & Burney, J. W. (2000). Application of commercially available, low-cost, miniaturized NIR spectrometers to the assessment of the sugar content of intact fruit. *Australian Journal of Plant Physiology, 27*, 1175–1186. https://doi.org/10.1071/PP99111

Williams, P. C. (2001). Implementation of near-infrared technology. In P. C. Williams, & K. H. Norris (Eds.), *Near-infrared technology in the agricultural and food industries* (pp. 145–169). St. Paul, MN, USA: AACC Inc.

Windham, W. R., Mertens, D. R., & Barton, F. E., II (1989). Protocol for NIRS calibration: Sample selection and equation development and validation. In G. C. Martens, J. S. Shenk, & F. E. Barton, II (Eds.), *Near infrared spectroscopy (NIRS): Analysis of forage quality. Agriculture handbook n°643* (pp. 96–103). Washington, D. C., USA: USDA-ARS, Government Printing Office.