Test of a LED fully integrated pre-prototype for rapid evaluation of table tomato (Solanum lycopersicum L., Marinda F1) quality

Journal:	*Journal of Near Infrared Spectroscopy*
Manuscript ID	JNS-21-0036.R3
Manuscript Type:	SI: NIRItalia2021
Date Submitted by the Author:	n/a
Complete List of Authors:	Tugnolo, Alessio; Universita degli Studi di Milano Facolta di Scienze Agrarie e Alimentari, Department of Agricultural and Environmental Sciences - Production, Landscape, Agroenergy Pampuri, Alessia; Universita degli Studi di Milano Facolta di Scienze Agrarie e Alimentari, Department of Agricultural and Environmental Sciences - Production, Landscape, Agroenergy GIOVENZANA, VALENTINA; Universita degli Studi di Milano Facolta di Scienze Agrarie e Alimentari, Department of Agricultural and Environmental Sciences - Production, Landscape, Agroenergy Casson, Andrea; Universita degli Studi di Milano Facolta di Scienze Agrarie e Alimentari, Department of Agricultural and Environmental Sciences - Production, Landscape, Agroenergy Guidetti, Riccardo; Universita degli Studi di Milano Facolta di Scienze Agrarie e Alimentari, Department of Agricultural and Environmental Sciences Beghi, Roberto; Universita degli Studi di Milano Facolta di Scienze Agrarie e Alimentari, Department of Agricultural and Environmental Sciences – Production, Landscape, Agroenergy
Keywords:	simplified optical device, ripening, vis/NIR and NIR spectroscopy, instrumental comparison, chemometrics, preharvest, postharvest, moisture content, total soluble solids, passing-bablok
Abstract:	The present research aims to evaluate the performance of an optical pre-prototype (TRL3) based on LED technology (light emitting diode, 450 - 860 nm) to quantify table tomatoes' quality features in a rapid and non-destructive way (Solanum lycopersicum L., Marinda F1). A total of 200 samples were analysed. Performances related to the pure near-infrared (NIR, 960 - 1650 nm) and visible/near-infrared (VIS/NIR, 400 - 1000 nm) commercial spectrophotometers to estimate the main tomato quality parameters, i.e. moisture content (MC) and total soluble solids (TSS), were calculated by using PLS regression method. Since no substantial differences were highlighted between the two commercial devices, to reduce the complexity keeping the performance of the model built using the whole spectra (1647 variables for VIS/NIR), a cost-effective pre-prototype was designed and built by using 12 bands in the VIS/NIR optical range. The pre-prototype shows slightly lower performance, displaying in prediction a RMSEP equal to 2% and 1.45 °Brix for MC and TSS respectively, respect to an RMSEP equal to 1% and 1.19 °Brix for VIS/SW-NIR device (using the entire spectrum). Moreover, no significant

	differences at 95% were highlighted by using Passing-Bablok regression. In conclusion, the pre-prototype performance can be considered enough accurate to allow an initial field screening of the trend of the analysed parameters (MC and TSS) using a new generation of simplified optical sensors.

Title:

Test of a LED fully integrated pre-prototype for rapid evaluation of table tomato (Solanum lycopersicum L., Marinda F1) quality

Short title:

Test of a simply optical pre-prototype to evaluate tomato quality

Authors:

A. Tugnolo, A. Pampuri, V. Giovenzana*, A. Casson, R. Guidetti & R. Beghi

Department of Agricultural and Environmental Sciences (DiSAA), Università degli Studi di Milano, via Celoria 2, 20133 Milano, Italy

* valentina.giovenzana@unimi.it

Abstract

The present research aims to evaluate the performance of an optical pre-prototype (TRL3) based on LED technology (light emitting diode, 450 - 860 nm) to quantify table tomatoes' quality features in a rapid and non-destructive way (Solanum lycopersicum L., Marinda F1). A total of 200 samples were analysed. Performances related to the pure near-infrared (NIR, 960 - 1650 nm) and visible/near-infrared (VIS/NIR, 400 - 1000 nm) commercial spectrophotometers to estimate the main tomato quality parameters, i.e. moisture content (MC) and total soluble solids (TSS), were calculated by using PLS regression method. Since no substantial differences were highlighted between the two commercial devices, to reduce the complexity keeping the performance of the model built using the whole spectra (1647 variables for VIS/NIR), a cost-effective pre-prototype was designed and built by using 12 bands in the VIS/NIR optical range.

The pre-prototype shows slightly lower performance, displaying in prediction a RMSEP equal to 2% and 1.45 °Brix for MC and TSS respectively, respect to an RMSEP equal to 1% and 1.19 °Brix for VIS/SW-NIR device (using the entire spectrum). Moreover, no significant differences at 95% were highlighted by using Passing-Bablok regression. In conclusion, the pre-prototype performance can be considered enough accurate to allow an initial field screening of the trend of the analysed parameters (MC and TSS) using a new generation of simplified optical sensors.

Keywords: simplified optical device, ripening, VIS/NIR and NIR spectroscopy, instrumental comparison, chemometrics

INTRODUCTION

Tomato (Solanum lycopersicum) is the second most-produced and consumed vegetable in the world and its global production in 2019 reached almost 197 million tons (1).

Concerning tomato production, in 2016 Italy has been the third larger producer (5.2 million tons of processed tomato) in the world behind California and China. Half of the Italian tomatoes are grown and processed in Northern Italy, in particular in Lombardy, Veneto, Piedmont and Emilia-Romagna (2).

Due to the great importance of this cultivation, it is essential to identify innovative solutions that maximize crop production and at the same time reduce waste. Technology-based solutions may be confounded by the large number of species and the complexity of plant-environment interactions within crop production systems. Therefore, the development of new approaches for improving understanding of crop biology to maximize production, minimize losses and to improve pre and post-harvest production and utilization is a crucial aspect (3).

Among the new approaches, it is well known that in recent years research has been focusing on non-destructive spectroscopic methods capable of exploring many samples and providing a complete and rapid overview of the maturation of fruit and vegetable products. Furthermore, spectroscopy, combined with multivariate management of the data, is a powerful analytical method that doesn't require any treatment to the sample and can be integrated on existing machinery for an increasingly automated quality control. (4).

However, most of the studies reported on tomatoes involve the use of expensive benchtop instruments that cover the entire spectral range of visible and near infrared (between 400 nm and 2500 nm approximately). Some recent studies have reported the possibility of using NIR spectroscopy for the determination of total soluble solids (TSS), titratable acidity and carotenoid compounds in salad tomatoes (5–7), while hyperspectral imaging has been applied for the estimation of moisture content, pH and TSS (8). A portable spectrometer was used by Sheng et al. (2019) (9) to predict soluble solids and lycopene in cherry tomatoes at different temperatures and by Arruda de Brito et al. (2021) (10) to develop models for intact tomatoes' TSS.

Over the last three decades, researchers have looked into the possibility of developing simplified optical devices for specific applications (11, 12), and focusing on the latest years the research has been moving towards portable, inexpensive, easy and quick to use instruments that do not need the entire spectral range to provide useful information (13).

Internet of things, big data and artificial intelligence and their disruptive role in shaping the future of agri-food systems (e.g. greenhouse monitoring, intelligent farm machines, drone-based crop imaging, food quality assessment using spectral methods), are making an impact only in very recent times, thanks to the advent of industry 4.0. In the IoT framework, proximal-remote sensors gather information generated by machines to increase efficiency, promote better decision-making and build competitive advantages, regardless of industry or company size (14).

Miniaturized sensors enabled a new and previously unattainable spectrum of applications of NIR spectroscopy, agriculture and the food sector, materials science, industry and environmental studies, having an impact on operational characteristics, marking a significant turning point in the evolution of the practical applications of NIR spectroscopy (15). In contrast to a mature benchtop spectrometer sector, the handheld devices are much less uniform and incorporate various novel technologies resulting in different performance, with narrower spectral regions, lower resolution, leading applicability limits and lower analytical performance (16).

In the agri-food sector there is still a need to develop customized cost-effective solutions for specific applications and the set of characteristics required must be merged together in devices and applications that are not actually already available. Also considering complementary but fundamental aspects as (i) the development of specific multivariate calibrations already on board the devices, and (ii) the optimization of the interconnection of the devices, e.g. cloud data storage and cloud computing.

The maturity of the tomato, in particular for small companies, is normally evaluated basing on the experience and on the color of the surface perceived by the human eye, in this way the fruits could undergo overripe. On the contrary, if the tomatoes are picked too early, they will not reach the desired ripeness, also causing an economic and technological (in the case of tomatoes intended for processing) damage.

The case of tomatoes is even more particular than other fruits because there are varieties that reach maturity in the absence of the color change towards red: detection of mature-green and immature-green tomatoes has been a challenge for researchers since there is no difference in terms of external appearance of the fruit (17).

The accurate and objective judgment of the maturity and harvest time is a critical prerequisite to maintain the quality of tomatoes. Additionally, supply chain operators may also need tomatoes with different maturities to meet various commercial purposes, so the harvest standard of tomatoes is various.

However, in terms of the producers, it may not be able to accurately determine the optimum harvest time due to the large subjective error. The quality grading for the picked tomatoes through the use of innovative optical sensors could be an effective method to improve the agricultural output value and reduce the economic loss (18).

Therefore, the aim of this work was to test a miniaturized LED fully integrated pre-prototype for rapid evaluation of MC and TSS of table tomatoes (Solanum lycopersicum L., Marinda F1) in a rapid and non-destructive way. Moreover, a statistical comparison with a commercial portable VIS/NIR device was carried out to verify the effectiveness and reliability of this new generation of simplified optical sensors.

MATERIALS AND METHODS

2. Materials and methods

2.1 Sampling

The experimentation was performed on table tomatoes Solanum lycopersicum L., Marinda F1, a variety easily available on the market at different stages of maturation. In order to represent the ripening process, in May 2019, a total of 200 tomato samples were bought and analyzed over 3 weeks. Three arbitrary ripening classes (figure 1) were created according with the skin color: totally green (extremely unripe, class t1, 50 samples), green/yellow/orange/red surface (medium ripe, class t2, 100 samples due to the tomatoes variability), totally red surface (ripe, class t3, 50 samples).

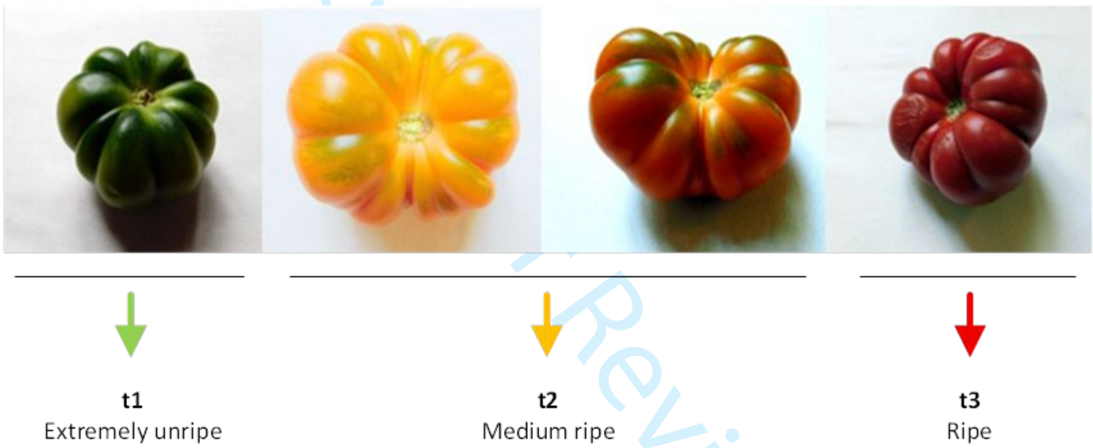

Figure 1: Ripening tomatoes classes (from left, class t1, class t2, and class t3)

Samples were processed in a laboratory within a few hours of purchase to acquire optical spectra and to perform moisture content MC (%) and TSS (°Brix) used as wet-chem reference parameters.

2.2 Optical analysis

Optical analyses were performed (before the wet-chem analyses) on tomato without any sample preparation. Each tomato sample was analyzed using two commercial portable spectrophotometers: NIR (Aurora NIR, Grainit, Italy) and a portable VIS/NIR (Jaz Modular Optical Sensing Suite, OceanOptics, Inc., Dunedin, FL, USA). The NIR spectrophotometer is equipped with a halogen light source, an InGaAs sensor in the NIR module (960 - 1650 nm, spectral resolution 10 nm) and it is designed for diffuse reflectance acquisition with automatic internal calibration. The

VIS/NIR spectrophotometer (400 - 1000 nm) works using a bifurcated fiber that conveys the light from the halogen lamp to the sample and back to the detector. The tip of the fiber was equipped with a cap that standardized the analysis distance (about 2 mm) and reduced the environmental light interference. The background consists of white (100% of light reflected) and black (0% of light reflected) standards. An integration time of 50 ms was set in order to collect the best spectral dynamics using a light intensity of 3000 lumens. Spectral measurements were taken from three points on each sample, apical, central, and basal area, and the values were averaged to obtain a mean optical spectrum for each sample.

Then, based on the results, a LED fully integrated stand-alone pre-prototype (technology readiness level estimated equal to 3, figure 2), designed by the Department of Agriculture and Environmental Sciences of the Università degli Studi di Milano, were tested on tomatoes.

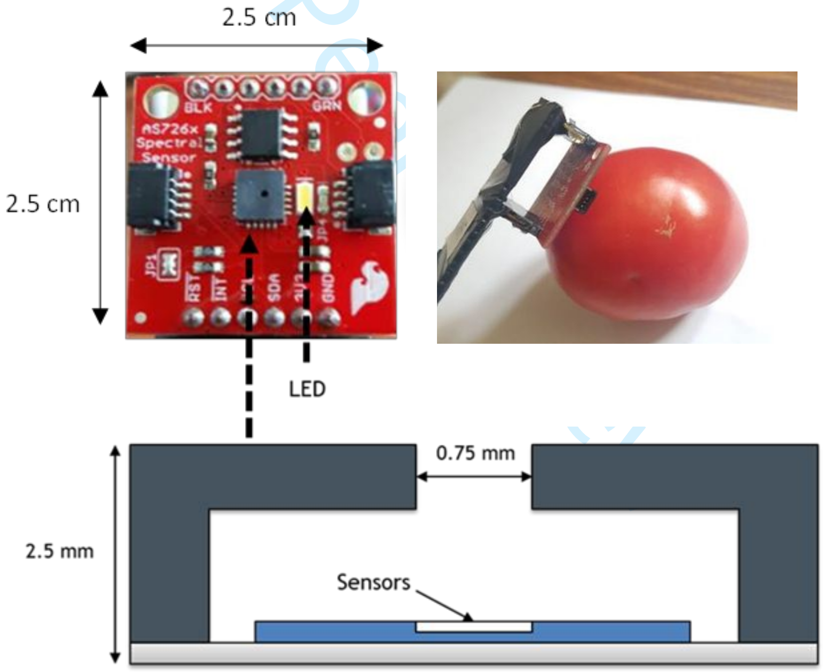

Figure 2. LED fully integrated stand-alone pre-prototype (TRL 3) managed by using a portable computer

The proposed prototype is composed by tuned photodiode arrays, interference filters, LEDs and optics. In detail, the device incorporates two sensitive spectrometers (6 optical channels each one, dimensions indicated in Figure 2) available in the form of breakout boards (AMS, models AS7262 visible and AS7263 NIR, Premstaetten, Austria-Europe), which include sensors and auxiliary electronic components for spectral measurement in the visible (VIS) and Short Wave Near-infrared (SW-NIR) region, have been used (Fletcher & Fisher, 2018). The sensors are 4.5 × 4.4 mm in size and are classified as ultra-low power consumption sensors. They have a 16-bit radiometric resolution and 12 independent on-device optical filters from 450 nm to 860 nm as summarized in table 1. The prototype enables chip-scale spectral analysis by integrating Gaussian filters into standard complementary metal-oxide-semiconductor (CMOS) silicon via nano-optic deposited interference filter technology. The six-channel VIS sensor (VIS module) is sensitive to the 400 - 700 nm spectral range with center wavelengths of 450, 500, 550, 570, 600 and 650 nm (interesting to get also color information, a crucial feature to get indications regarding the ripening progress). While, the six-channel SW-NIR sensor (NIR module) is sensitive to the 600 - 900 nm spectral range with center wavelengths at 610, 680, 730, 760, 810 and 860 nm. Each module is composed of 6 independent optical filters whose spectral bandpass is defined with full-width half-max (FWHM) of 40 nm for the VIS module and 20 nm for NIR module.

The sensors can read the intensity of light at the 12 wavelengths (6 for each module) and give digital output (readout) corresponding to the intensity of light falling on it. The light source is a white super bright LED illumination (5700K) with an irradiance of ~600µW/cm^2. The light detection position is in contact with the fruits, while the LED emission position is on the side at about 2 mm from the surface of the tomato.

The pre-prototype has been configurated using Arduino to perform an average of 10 scans for each acquisition point in order to reduce the experimental noise.

Table 1. Sensor's wavelengths.

	Wavelengths (nm)					
Sensor 1	450	500	550	570	600	650
Sensor 2	610	680	730	760	810	860

2.3 Chemical analyses

After optical measures, the wet-chemical analyses (reference parameters) were also performed on the same samples. Each tomato was centrifugated and the juice analyzed for a representative measure of whole tomato.

The analytical methods to determine the moisture content (MC) recommend a temperature of 103°C for 24 h, until constant weight of product (19). The samples were weighed with a balance (LAZ 30P, Sartorius Lab Holding GmbH, Goettingen, Germany) and dried with a laboratory oven

(UNB400, Memmert GmbH & Co, Schwabach, Germany) in order to determine the dry weight of samples collected. Once reached room temperature, samples are re-weighed (dry weight) for the MC determination (Eq. 1):

$$MC\ (\%) = \frac{(m_{gw} - m_{dw}) * 100}{m_{gw}} \qquad \text{Eq. 1}$$

where:

MC (%) = percentage of moisture content.

m_{gw} = gross weight.

m_{dw} = dry weight.

Instead, TSS content was evaluated by using a digital refractometer (PAL-1 ATAGO, Tokyo, Japan, accuracy refractive index ±0.2 °Brix) measuring the refractive index of the juice (°Brix).

2.4 Data processing

The entire data analysis was performed in the Matlab® environment, version 2019b (The MathWorks, Inc., Natick, MA, USA) using PLS-Toolbox package (Eigenvector Research, Inc. Manson, Washington) and Passing and Bablok regression by Andrea Padoan (Jan 16, 2010).

The reference parameters (MC and TSS), representative of tomato ripening process, and the optical diffuse reflectance data obtained from NIR and VIS/NIR commercial spectrophotometers were analyzed in a multivariate way to (i) qualitatively understand the relationships among all variables and among variables and samples, i.e. principal component analysis (PCA), and to highlight outliers based on a detection procedure applied on PCA scores using the 'Hotelling T^2 computation' function (α value was set to 0.05), (ii) quantitative predict the reference parameters, i.e. partial least square regression (PLS). The logic scheme of data processing is presented in figure 3.

Moreover, in order to reduce the instrumentation complexity keeping the performance of the model built using the whole spectra (1647 variables for VIS/NIR), a cost-effective pre-prototype, characterize from 12 bands (table 1) was tested and the same data processing approach (figure 3) carried out. Afterwards, for a better understanding of the practical applicability of the proposed LED technology for the quality parameters prediction (MC and TSS), the Passing-Bablok regression method was applied on the PLS prediction outcome deriving from VIS/NIR instrumentation and pre-prototype device.

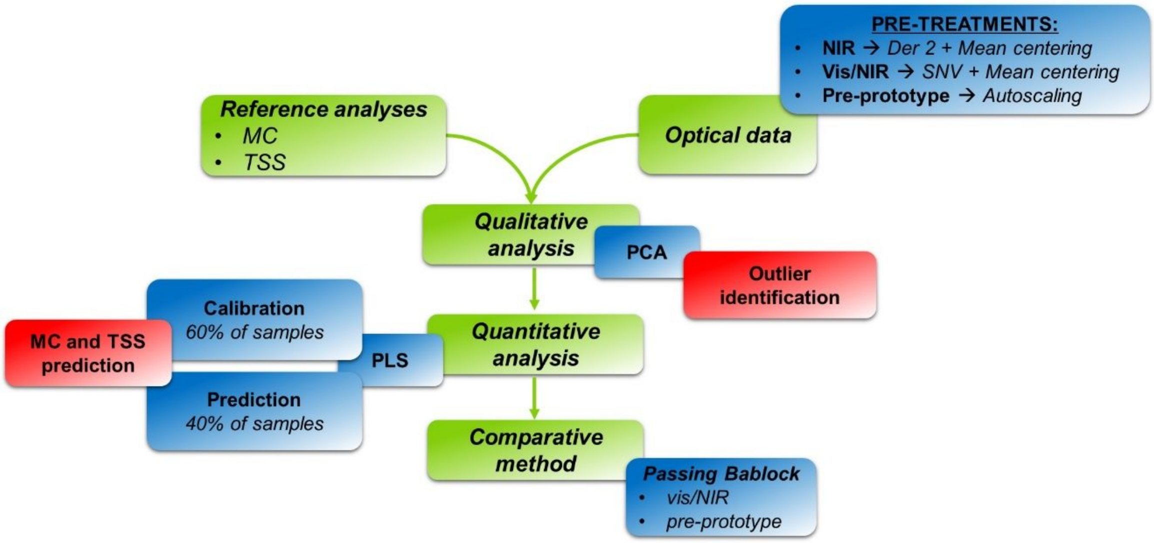

Figure 3. The logic scheme of data processing (green), methods applying (blue) to obtain results (red)

Before applying multivariate analysis, the optical data were pre-treated to reduce instrumental noise and the scattering effects (due to the inhomogeneous physical structure of the tomato's samples, size and shape, and the various positions and distances where the sample can interact with the sensor).

On NIR spectra, second derivative (Der 2) and mean centering were applied. Der 2 performed by using Savitzky-Golay algorithm (filter width = 15, polynomial order = 2) which enhances the separation of overlapping peaks allowing more specific identification of small and nearby lying absorption peaks which are not resolved in the original spectrum, thereby offering means to increase the selectivity of absorption peaks for certain molecules of matrix. Der 2 was applied paying attention to avoid suppression of broad bands and enhancement of noise. Mean centering ensures that all results will be interpretable in terms of variation around the mean. For VIS/NIR optical data a correction of the baseline vertical shifts (offsets) and of the global intensity effects (typically arising from unwanted light scattering) was performed applying the Standard Normal Variate (SNV) transform and mean centering.

The data obtained from the two sensors of pre-prototype were processed and analyzed as data obtained from one single pre-prototype in order to take advantage of all wavelengths considered. A data scaling phase was applied to the readouts of pre-prototype, in order to make the different variables comparable in importance before applying scale-dependent multivariate analysis methods (such as PCA or PLS). For this purpose, the unit variance scaling (or autoscaling), variables are divided by their respective standard deviations, was applied. The method is commonly used to data sets containing variables with different units and scales in order to impose equal weights in the analysis (20,21).

Then, the reference parameters were used for the calculation of PLS predictive models. Kennard–Stone algorithm (22) was applied to select a representative subset to ensure training samples spread evenly throughout the sample space. In this paper a 200 samples set was partitioned into training (60% for calibration set, 120 samples) and test (40% for prediction set, 80 samples) sets.

To identify the most suitable pre-treatments and the number of latent variables (LV), the models were evaluated using an internal validation (Cross-validation) through the Venetian Blinds splitting method with five cancellation groups.

To measure the PLS models accuracy, the statistical parameters used were the RMSE (root mean square error, RMSEC, in calibration and RMSEP in prediction), as well as latent variables, bias, and R^2 (determination coefficient).

2.5 Methods comparison

In order to compare the practical applicability of the pre-prototype device to predict quality parameters correlated to the tomato ripening, the Passing-Bablok regression (23) method was applied on the MC and TSS values obtained by pre-prototype and the commercial VIS/NIR spectrometer on the same dataset, i.e. the prediction set. Since the Passing-Bablok regression method is a symmetrical non-parametric technique, which can build regression models also when both variables (independent and dependent) have a non-negligible experimental error, the regression method results particularly suitable for method comparison. For statistically evaluating the similarity/diversity between these two independent estimations, slope and intercept of the fitted line were calculated, and a significance test was conducted. The null hypothesis (H_0) was verified when the slope was not significantly different from 1 and, simultaneously, the intercept was not significantly different from 0, meaning that there are no significant differences between the two methods, at a 95% confidence level (24). Hence, the Passing-Bablok regression allows to evaluate if the performance deriving from pre-prototype are equally comparable to performance from commercial VIS/NIR instrumentation for the quantification of tomato quality parameters.

RESULTS AND DISCUSSIONS

Reference data analysis

The boxplots in figure 4 provide a visualization of summary statistics for MC (figure 4a) and TSS (figure 4b) obtained using the destructive reference analysis on the tomato's samples at each sampling time. The mean, median, the interquartile range, and the data range were represented into the graphs. Moreover, the potential outliers (observations beyond the data range whisker length) were also reported. By default, a potential outlier is a value that is more than 1.5 times the interquartile range away from the bottom or top of the box.

Firstly, no apparent differences were highlighted for MC which ranged about 83% to 93% during the three sampling times (median of 90% at t1, 87% at t2 and 88% at t3).

Instead, concerning TSS, an increasing trend from a median of 6.7 to 10 °Brix was observed along sampling, representing the ripening process. Moreover, a very high variance, strictly related to the higher number of samples (100 samples) and to the different features of the medium ripe samples, was captured for the samples analysed at t2 (from 4.8 to 14.8 °Brix).

Finally, three and two potential outliers were identified in the reference data for MC and TSS, respectively.

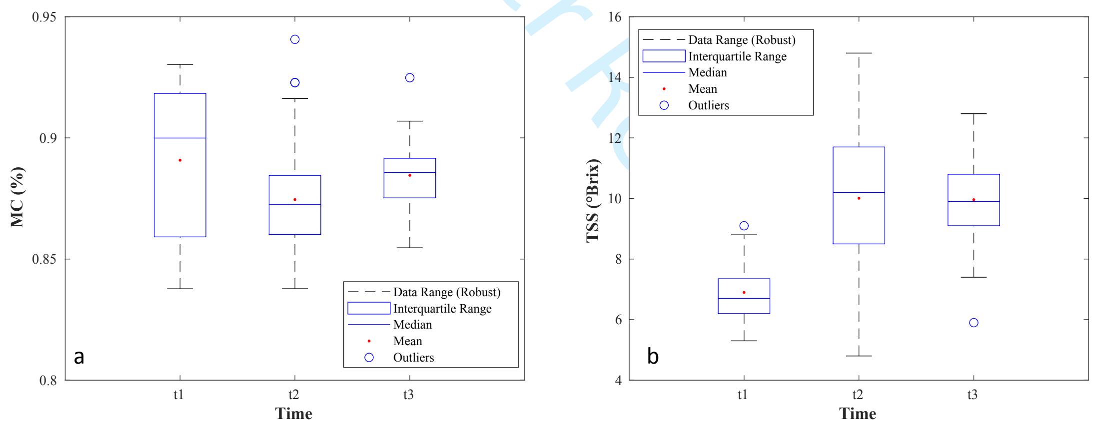

Figure 4. Descriptive statistics of the reference analysis for MC (a) and TSS (b) obtained from the tomato's samples.

Spectra exploration and regression

Figure 5 shows raw and pre-treated spectra obtained from the two commercial VIS/NIR (450-950 nm, Figure 5a) and NIR (950-1650 nm, Figure 5c) spectrophotometers. A visual analysis of the spectra highlights the main absorption bands related to the main constituents in tomato's samples. The visible range exhibits clear differences between the different tomato maturity groups around 550 nm and 675 nm, which are mainly related to the variation of anthocyanin and chlorophylls. As the maturity of tomato advances from green to red, its chlorophyll content decreases, while anthocyanin increase (13). Moreover, in the short-wave near-infrared (SWNIR), an absorption band at 760 nm (caused by the third overtone of OH stretching) is noticeable (23). While, in the pure NIR region (from the NIR spectrophotometer, figure 5c), the main absorption bands are related to the stretching of the CO bonds of aliphatic esters, to the second overtone of CH stretching vibrations of alkyl groups and alkenes (1212 and 1245 nm) and the water absorption related to the OH stretch first overtone around 1440 nm (25–27).

However, due to the nature of the samples and to the use of portable devices, a scattering effect was highlighted in the raw spectra. Therefore, two different pre-treatments (SNV for VIS/NIR spectra and second derivative for the NIR spectra) were applied in order to ensure a correction of the global intensity effect – scattering at the sample surface (Figure 5b) and the baseline drift, typically, ascribable to the use of portable devices (Figure 5d).

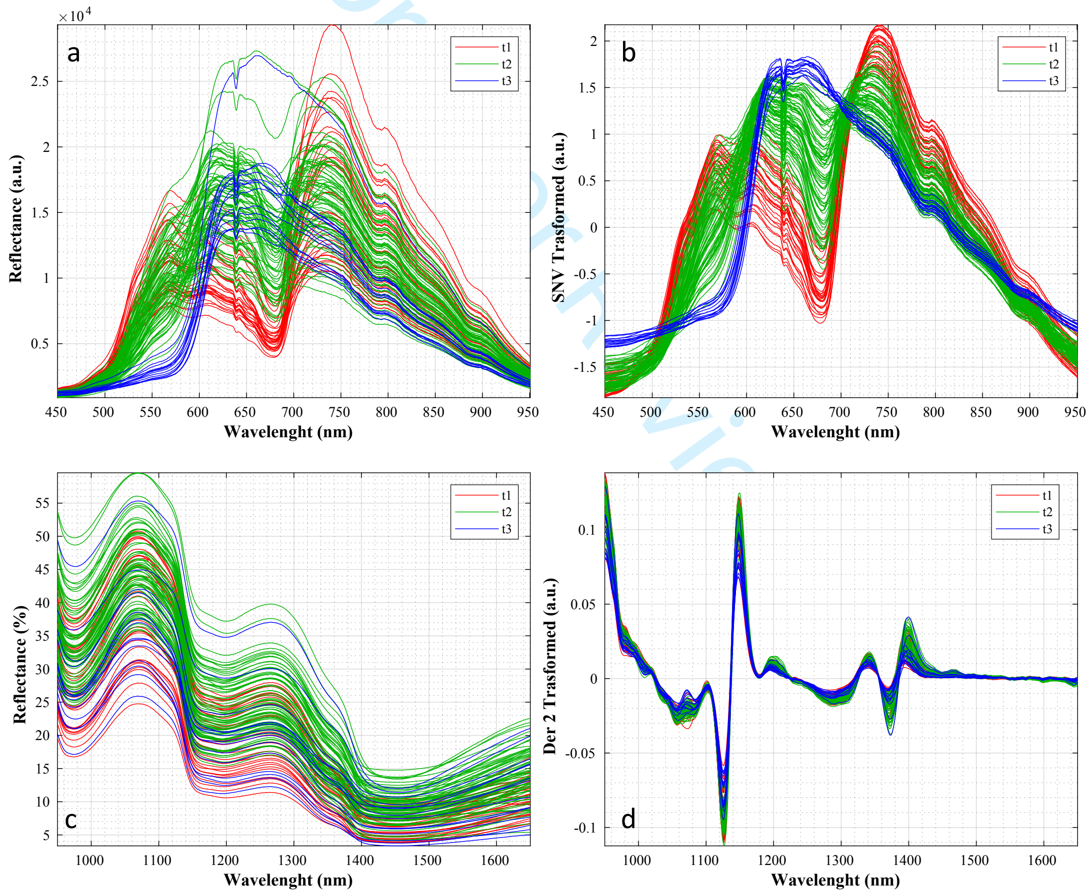

Figure 5. Raw (5a, 5c) and pre-treated spectra (SNV for VIS/NIR spectra and second derivative for NIR spectra) obtained from intact tomato's samples.

PCA was performed on both types of optical data (data not shown) allowing to detect one and two possible outliers in the VIS/NIR and NIR spectra, respectively. Afterwards, regression models (by means of the PLS method) were developed for the prediction of MC and TSS starting from the data coming from the two different type of portable spectrophotometers. The parameters used for evaluating the model goodness are presented in Table 2. Overall, regarding R^2 and RMSEC and RMSEP, relative slightly low differences between the performance of the NIR models in calibration (R^2_{Cal} = 0.82 and RMSEC = 0.01 for MC and R^2_{Cal} = 0.84 and RMSEC = 0.83 °Brix for TSS) and in prediction (R^2_{Pred} = 0.75 and RMSEP = 0.01 for MC and R^2_{Pred} =0.75 and RMSEP = 1.05 °Brix for TSS) were obtained showing the good capability of the NIR models to accurately predict the MC and TSS in tomatoes samples. Concerning the VIS/NIR models a lower performance was obtained for the optical estimation of MC and TSS in calibration (R^2_{Cal} = 0.67 and RMSEC = 0.01 for MC and R^2_{Cal} = 0.72 and RMSEC = 1.10 °Brix for TSS) and in prediction (R^2_{Pred} = 0.53 and RMSEP = 0.01 for MC and R^2_{Pred} = 0.65 and RMSEP = 1.19 °Brix for TSS).

However, considering the very low error difference obtained from the models built using NIR and VIS/NIR (from the two different devices) optical data, the use of fewer LVs and less invasive pre-treatments, the prototype has been developed using wavelengths coming from the VIS/NIR optical range (from 450 to 1000 nm). Moreover, this decision was also taken considering the capability of the VIS/NIR optical detectors to be less expensive in order to develop cost-effective optical sensors able to get qualitative information from tomato's samples.

Table 2. Figures of merit of the NIR and VIS/NIR PLS models obtained from the two commercial spectrophotometers for the estimation of MC and TSS.

						Calibration			Prediction		
Optical range	Qualitative parameter	N° of Cal samples	N° of Pred samples	LVs	Pre-pro	R^2_{Cal}	RMSEC	$Bias_{Cal}$	R^2_{Pred}	RMSEP	$Bias_{Pred}$
NIR (950-1650 nm)	MC (%)	116	80	6	Der 2 + Mean centering	0.82	1	-2.2 e-16	0.75	1	6.2 e-4
	TSS (°Brix)	117		10		0.84	0.83	1.1 e-14	0.75	1.05	-0.2
VIS/NIR (450-950 nm)	MC (%)	114	80	3	SNV + Mean centering	0.67	1	1.1 e-16	0.53	1	2.2 e-3
	TSS (°Brix)	115		3		0.72	1.10	-1.7 e-15	0.65	1.19	-0.24

Pre-pro=pre-processing, LVs = latent variables, Cal = calibration, Pred = prediction

Sensor readouts exploration and regression

Figure 6 shows the pre-prototype readouts (12 wavelengths, table 1). The data have been handled by merging both sensors in one integrated pre-prototype. However, even though the data do not come from a single sensor, the final reflectance output provide a complete profile which is highly comparable in terms of trend and reflectance of the tomato's samples acquired with the VIS/NIR commercial spectrophotometer.

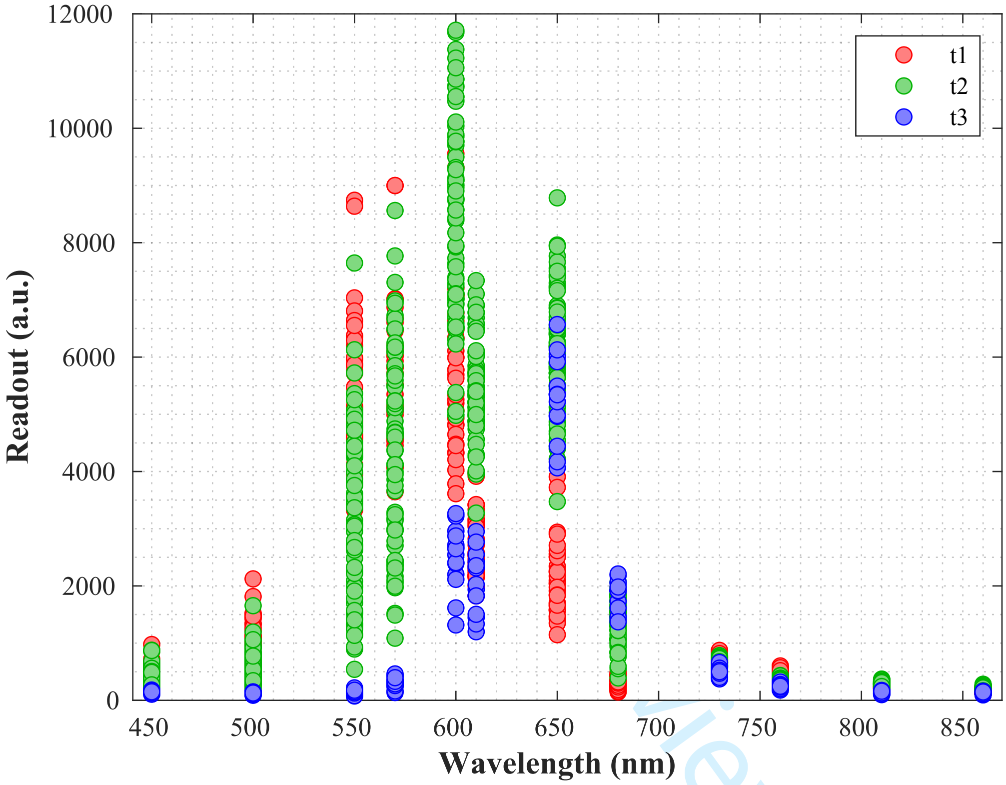

Figure 6. Sensors' readouts obtained from intact tomato's samples.

Table 3 shows the figure of merit of the PLS models obtained by the integrated prototype system for the evaluation of MC and TSS. Overall, a pour coefficient of determination was obtained for the prediction of both qualitative parameters (R_{Pred}^2 = 0.52) using 4 latent variables. However, a slightly higher RMSEP was obtained respect to the model developed using the VIS/NIR spectrophotometer (RMSEP = 0.02 for MC and RMSEP = 1.45 for TSS). These results suggest the capability of the new integrated prototype to be able to obtain better results using more tomato's samples in order to boost the predictive performance of the PLS model.

Table 3. Figures of merit of the VIS/NIR PLS-models obtained from the sensors' readouts for the estimation of MC (%) and TSS (°Brix).

						Calibration			Prediction		
Optical range	Qualitative parameter	N° of Cal samples	N° of Pred samples	LVs	Pre-pro	R^2	RMSEC	Bias	R^2	RMSEP	Bias
VIS/NIR (12 wavelengths)	MC (%)	113	74	4	Autoscaling	0.60	2	-2.2 e-16	0.52	2	-4.5 e-05
	TSS (°Brix)					0.63	1.27	1.7 e-15	0.52	1.45	0.02

Pre-pro=pre-processing, LVs = latent variables, Cal = calibration, Pred = prediction

Methods comparison

To evaluate whether significant differences of the performance between the VIS/NIR commercial spectrophotometer and the VIS/NIR prototype in determining the MC and TSS exist, the Passing–Bablok regression was performed on the same data used as external validation of the PLS models built using the commercial spectrophotometer and the prototype. Applying a joint test on slopes and intercepts, the devices were compared in pairs analysing the differences between the prediction values for MC and TSS obtained by the models developed using the two instruments. No statistical differences between the instruments were highlighted from the Passing–Bablok tests, at a confidence level of 95%. Therefore, the null hypothesis (slope not significantly different from 1 and intercept not significantly different from 0) was accepted for all the paired comparisons: MC-sensor prediction vs. MC-commercial VIS/NIR spectrophotometer prediction and TSS- sensor prediction vs. TSS-commercial VIS/NIR spectrophotometer prediction. In figure 7, the Passing–Bablok regression lines (solid blue lines) and the bisector of the quadrants (ideal lines) represented as dotted red lines were reported for comparison.

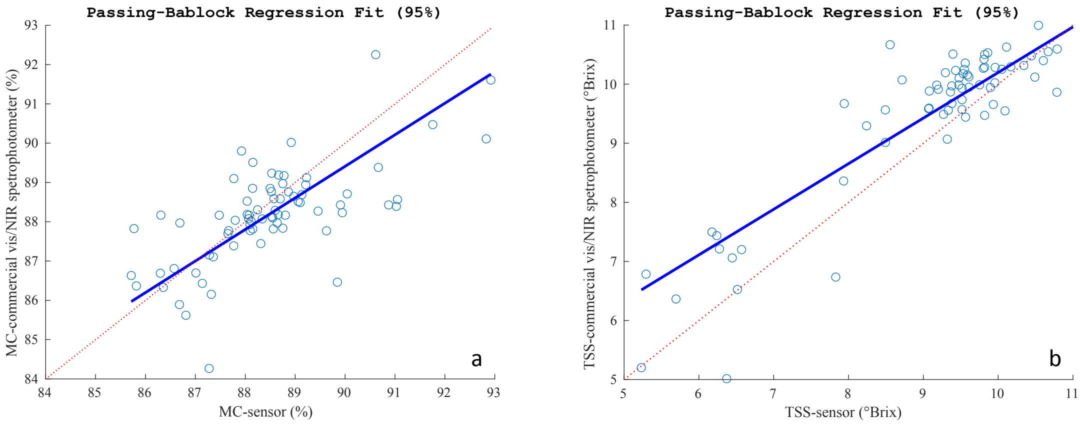

Figure 7. Passing–Bablok regression outcomes: a) comparison between sensor and commercial VIS/NIR spectrophotometer for MC prediction; b) comparison between sensor and commercial VIS/NIR spectrophotometer for TSS prediction

CONCLUSIONS

The interest in small-sized, cost-effective and simplified portable devices to quantify quality parameters in a rapid and non-destructive way in pre and post-harvest steps of agri-food chain is desirable. Moreover, the commercial availability of optical components, highly miniaturized, extremely low-cost and robust, given the opportunity to develop an optical device based on LED technology (light emitting diode), which will be tested in this work to monitor the ripening of table tomatoes (Solanum lycopersicum L., Marinda F1).

In this work, a pre-prototype equipped by two visible and SW-NIR sensors for spectral acquisition based on LED technology was tested for rapid estimation of tomato’s quality parameters. The overall results of the PLS models from the pre-prototype were encouraging considering the initial development stage of the device. In these terms, a larger dataset in order to improve the robustness of the prediction models is needed. The instrument should be able to acquire and predict moisture content and total soluble solids related to tomato. The integration of simple processing algorithms derived from the PLS models in the microcontroller software would easy calculate and visualize the real-time values of the predicted parameters on the device dashboard. In the envisaged future updated version of the prototype, it is planned that the shape of the device will be able to fully embrace the sample (in this case a tomato, but it can also be used on other vegetable matrices) to minimize and keep under control the incidence of sunlight for in field use. Moreover, the optical core of the device is a commercial one, as stated, and it is equipped with a standard white LED that showed little power at longer wavelengths. The next evolutional step is to reinforce the lighting system by using one or two additional LEDs with emission peaks around 750 nm and 840 nm.

The potential device deriving from the pre-prototype analysed in this work could result as user-friendly devices to support small-scale growers in determining the optimal harvest date according to tomato ripening degree or applied in post-harvest to classify the fruits based on quality parameters objectively measured. Moreover, the development of interconnected optical systems (creation of remote storage of optical databases) would also allow the updating of the prediction performance of the mathematical models integrated in the prototype for the parameters control. Finally, thanks to the low cost of the components, comparable to consumer electronics, the pre-prototype could be also hypothesized for applications requiring a high number of distributed sensors (sensors network approach), for example, to control agri-food products directly in the field in defined sentinel parcels, or along the supply chains where extreme miniaturization, simplification and low cost can be crucial aspects.

Acknowledgements

The authors wish to thank “Officina delle Soluzioni” (Magliano Alfieri, Cuneo, Italy) for the technical support for the prototype design.

Literature cited

1. FAO Statistical Programme of Work 2020–2021. FAO Statistical Programme of Work 2020–2021. FAO; 2020.
2. Donati M, Guareschi M, Veneziani M. Organic tomatoes in Italy. In: Sustainability of European Food Quality Schemes: Multi-Performance, Structure, and Governance of PDO, PGI, and Organic Agri-Food Systems. Springer International Publishing; 2019 [cited 2021 May 31]. p. 171–89. Available from: https://link.springer.com/chapter/10.1007/978-3-030-27508-2_9
3. Skolik, P., Morais, C. L., Martin, F. L., & McAinsh, M. R. (2019). Determination of developmental and ripening stages of whole tomato fruit using portable infrared spectroscopy and Chemometrics. BMC plant biology, 19(1), 1-15.
4. Casson, A., Beghi, R., Giovenzana, V., Fiorindo, I., Tugnolo, A., & Guidetti, R. (2020). Environmental advantages of visible and near infrared spectroscopy for the prediction of intact olive ripeness. Biosystems Engineering, 189, 1-10.

5. Acharya, U. K., Subedi, P. P., & Walsh, K. B. (2017). Robustness of tomato quality evaluation using a portable Vis-SWNIRS for dry matter and colour. International journal of analytical chemistry, 2017.

6. Feng, L., Zhang, M., Adhikari, B., & Guo, Z. (2019). Nondestructive detection of postharvest quality of cherry tomatoes using a portable NIR spectrometer and chemometric algorithms. Food Analytical Methods, 12(4), 914-925.

7. Ibáñez, G., Cebolla-Cornejo, J., Martí, R., Roselló, S., & Valcárcel, M. (2019). Non-destructive determination of taste-related compounds in tomato using NIR spectra. Journal of Food Engineering, 263, 237-242.

8. Rahman, A., Kandpal, L. M., Lohumi, S., Kim, M. S., Lee, H., Mo, C., & Cho, B. K. (2017). Nondestructive estimation of moisture content, pH and soluble solid contents in intact tomatoes using hyperspectral imaging. Applied Sciences, 7(1), 109.

9. Sheng, R., Cheng, W., Li, H., Ali, S., Agyekum, A. A., & Chen, Q. (2019). Model development for soluble solids and lycopene contents of cherry tomato at different temperatures using near-infrared spectroscopy. Postharvest Biology and Technology, 156, 110952.

10. de Brito, A. A., Campos, F., dos Reis Nascimento, A., de Carvalho Corrêa, G., da Silva, F. A., de Almeida Teixeira, G. H., & Júnior, L. C. C. (2021). Determination of soluble solid content in market tomatoes using near-infrared spectroscopy. Food Control, 126, 108068.

11. Kawano, S., Watanabe, H., & Iwamoto, M. (1992). Determination of sugar content in intact peaches by near infrared spectroscopy with fiber optics in interactance mode. Journal of the Japanese Society for Horticultural Science, 61(2), 445-451.

12. Civelli, R., Giovenzana, V., Beghi, R., Naldi, E., Guidetti, R., & Oberti, R. (2015). A simplified, light emitting diode (LED) based, modular system to be used for the rapid evaluation of fruit and vegetable quality: Development and validation on dye solutions. Sensors, 15(9), 22705-22723.

13. Huang, Y., Lu, R., & Chen, K. (2018). Prediction of firmness parameters of tomatoes by portable visible and near-infrared spectroscopy. *Journal of Food Engineering*, *222*, 185-198.

14. Misra, N. N., Dixit, Y., Al-Mallahi, A., Bhullar, M. S., Upadhyay, R., & Martynenko, A. (2020). IoT, big data and artificial intelligence in agriculture and food industry. IEEE Internet of Things Journal

15. Beć, K. B., Grabska, J., & Huck, C. W. (2021). Principles and applications of miniaturized near-infrared (NIR) spectrometers. Chemistry (Weinheim an der Bergstrasse, Germany), 27(5), 1514
16. Huck, C. W. (2021). New Trend in Instrumentation of NIR Spectroscopy—Miniaturization. In Near-Infrared Spectroscopy (pp. 193-210). Springer, Singapore
17. Fatchurrahman, D., Amodio, M. L., de Chiara, M. L. V., Chaudhry, M. M. A., & Colelli, G. (2020). Early discrimination of mature-and immature-green tomatoes (Solanum lycopersicum L.) using fluorescence imaging method. Postharvest Biology and Technology, 169, 111287.
18. Huang, Y., Dong, W., Chen, Y., Wang, X., Luo, W., Zhan, B., ... & Zhang, H. (2021). Online detection of soluble solids content and maturity of tomatoes using Vis/NIR full transmittance spectra. Chemometrics and Intelligent Laboratory Systems, 210, 104243.
19. ASAE Standards for Moisture Determination. Available from: https://engineering.purdue.edu/~abe305/moisture/html/page13.htm
20. Tugnolo, A., Beghi, R., Giovenzana, V., & Guidetti, R. (2019). Characterization of green, roasted beans, and ground coffee using near infrared spectroscopy: A comparison of two devices. Journal of Near Infrared Spectroscopy, 27(1), 93-104.
21. Marini, F., de Beer, D., Walters, N. A., de Villiers, A., Joubert, E., & Walczak, B. (2017). Multivariate analysis of variance of designed chromatographic data. A case study involving fermentation of rooibos tea. Journal of Chromatography A, 1489, 115-125.
22. Morais, C. L., Santos, M. C., Lima, K. M., & Martin, F. L. (2019). Improving data splitting for classification applications in spectrochemical analyses employing a random-mutation Kennard-Stone algorithm approach. Bioinformatics, 35(24), 5257-5263.
23. Passing, H., & Bablok, W. (1983). A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, Part I.
24. Mustorgi, E., Malegori, C., Oliveri, P., Hooshyary, M., Bounneche, H., Mondello, L., ... & Casale, M. (2020). A chemometric strategy to evaluate the comparability of PLS models obtained from quartz cuvettes and disposable glass vials in the determination of extra virgin olive oil quality parameters by NIR spectroscopy. Chemometrics and Intelligent Laboratory Systems, 199, 103974.

25. Clément, A., Dorais, M., & Vernon, M. (2008). Nondestructive measurement of fresh tomato lycopene content and other physicochemical characteristics using visible− NIR spectroscopy. Journal of Agricultural and Food Chemistry, 56(21), 9813-9818.

26. Fernández-Espinosa, A. J. (2016). Combining PLS regression with portable NIR spectroscopy to on-line monitor quality parameters in intact olives for determining optimal harvesting time. Talanta, 148, 216-228.

27. Trapani, S., Migliorini, M., Cecchi, L., Giovenzana, V., Beghi, R., Canuti, V., Fia, G.,& Zanoni, B. (2017). Feasibility of filter-based NIR spectroscopy for the routine measurement of olive oil fruit ripening indices. European Journal of Lipid Science and Technology, 119(6), 1600239.