Scientia Horticulturae 328 (2024) 112928

Contents lists available at ScienceDirect

Scientia Horticulturae

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

Research Paper

Potential application of hyperspectral imaging and FT-NIR spectroscopy for discrimination of soilless tomato according to growing techniques, water use efficiency and fertilizer productivity

Hassan Fazayeli [a,*], Maria Luisa Amodio [a], Danial Fatchurrahman [a], Francesco Serio [b], Francesco Fabiano Montesano [c], Ingunn Burud [d], Andrea Peruzzi [e], Giancarlo Colelli [a]

[a] Dipartimento di Scienze Agrari, Alimenti, Risorse Naturali, e Ingegneria, Università di Foggia, Via Napoli 25, Foggia 71122, Italy
[b] National Research Council of Italy, Institute of Sciences of Food Production, Via G. Amendola, 122/O, Bari 70126, Italy
[c] Department of Soil, Plant and Food Sciences, University of Bari Aldo Moro, via G. Amendola 165/A, Bari 70126, Italy
[d] Faculty of Science and Technology, Norwegian University of Life Sciences, PB 5003, Aas 1430, Norway
[e] Dipartimento di Scienze Agrarie, Alimentari e Agro-ambientali, Università di Pisa, Via del Borghetto 80, Pisa 56124, Italy

ARTICLE INFO

Keywords:
Classification
Cultural practices
FT-NIR spectroscopy
Hyperspectral imaging
PLS-DA

ABSTRACT

Background: This study aimed to test the potential of nondestructive optical techniques for classifying sustainable-produced tomatoes according to a) growing practices; b) levels of water use (WUE) and partial factor productivity of nutrients (PFP). Three distinct hydroponic growing techniques for water and fertilizer use were applied over two cultivation cycles for two varieties (cv 'Carminio', and cv 'Mose'): i) free drain open cycle cultivation (OPEN); ii) open cycle cultivation with on-demand sensor-based fertigation (SMART); iii) closed cycle cultivation (CLOSED). Hyperspectral images (HSI) in the Vis-NIR and NIR range, as well as reflectance spectra obtained through Fourier Transform (FT)-NIR spectroscopy, were acquired throughout the harvesting period, with approximately 300 fully ripe tomatoes obtained per variety. For each variety, partial least squares discriminant analysis (PLS-DA) was initially employed to discriminate the three cultivation systems and subsequently distinguish two levels of WUE and PFP, per variety. Finally, the data obtained from both varieties were combined and PLS-DA utilized to categorize three levels of WUE: LOW, MEDIUM, and HIGH.
Results: The PLS-DA models applied in external prediction for discriminating tomatoes of each variety according the three cultural practices achieved accuracy higher than 79.55 % for data obtained with FT-NIR and HSI in the Vis-NIR range, and lower for HSI (NIR). The performance increased when considering only two distinct classes (based on WUE and nutrients PFP) with accuracy, specificity, and sensitivity higher than 86 %. General models based on spectra obtained with HSI Vis-NIR, using data over the two varieties to classify tomato according to three levels of WUE, yielded accuracy and specificity of 89.8 % and 91.7 %, respectively, utilizing only 20 variables.
Conclusions: Results of this study indicated the effectiveness of FT-NIR and HSI (Vis-NIR) techniques for discrimination of tomato fruits cultivated with varying levels of water and fertilizers. Further validation of these methods is needed before their widespread application to support the adoption of low input growing techniques.

1. Introduction

With the progression of agricultural science and technology in last decades, sustainable agricultural practices such as precision agriculture, sensor-based management of agronomic inputs, and organic farming have gained increasing popularity due to their benefits in reducing environmental impact as well as input costs, which bring higher profits to growers (Lubell et al., 2011). With raising efforts to increase the amount of sustainably-produced products, the availability of tools that make it possible to distinguish products also on the basis of the agricultural practices adopted for their production is becoming critical. Sustainability certification and labelling have become a must-have to show the authenticity of the product, with specific reference to the compliance of the production process with sustainable agricultural

* Corresponding author.
E-mail address: hassan.fazayeli@unifg.it (H. Fazayeli).

https://doi.org/10.1016/j.scienta.2024.112928
Received 30 August 2023; Received in revised form 3 December 2023; Accepted 22 January 2024
Available online 30 January 2024

practices associated, to anyone worried about unfair market competition, including consumers and processors (Reid et al., 2006). In fact, sustainability is becoming an important market leverage, and every tool capable of attributing a clear sustainability rating to products is crucial to the diffusion of sustainable practices and to prevent unfair commercialization of non-sustainable products as sustainable. In this perspective, the higher prices paid for labelled sustainable products shoot up the risk of frauds. Thus, there is a clear need of additional tools to prove the authenticity of the product (Brooks et al., 2021). Conventional approaches such as molecular markers and other biochemical methods along with morphological tests can accurately discriminate products grown with different cultural practices. Even so, these methods are time, cost, and labor consuming because they need sample preparation, chemical analyses, expensive instruments, and expert people (Amodio et al., 2017b).

Nowadays, among non-destructive analytical approaches in food science, hyperspectral image analysis (HSI) and near infrared (NIR) spectroscopy have been employed for different applications. Both techniques allow obtaining the spectral fingerprint on the base of the interaction between molecular structure of samples and the incident light (Liu et al., 2017; Nicolaï et al., 2007; Roberts et al., 2004; van de Voort, 1992; Wu and Sun, 2013). NIR spectroscopy, and particularly Fourier Transform-NIR (FT-NIR) spectroscopy, have been applied for food authentication purposes (Roberts et al., 2004). The upside of FT-NIR over regular IR spectroscopy is that it would perform the entire source spectrum rather than individual wavelengths and apply a Fourier transformation to convert the digital interferogram to a typical IR emittance spectrum (van de Voort, 1992). HSI, on the other hand, is an integration of imaging and spectroscopy, which provides spectral data for each pixel. The superiority of this technique over spectroscopy is its capability of generating maps and visualize the chemical composition of a sample (Jiang et al., 2021; Liu et al., 2017; Pourdarbani et al., 2021; Wu and Sun, 2013). Generally, for food application, detectors that are sensitive in the visible-near infrared (Vis-NIR) and NIR are used. Several studies have reported on the viability of using Vis-NIR and NIR-based analysis to discriminate and assess the quality of fresh fruits (Amodio et al., 2017b, 2017a; Ding et al., 2015; Feng et al., 2019; Kusumiyati et al., 2019; Morellos et al., 2020).

Tomato (*Solanum lycopersicum* L.) is an internationally cultivated crop that needs a lot of water to grow and provide high yield performance (Cuartero and Fernández-Muñoz, 1998; Quinet et al., 2019). In recent years, there was an increase in the soilless production of tomato crops in greenhouses due to the scarcity of high-quality water for irrigation in many parts of the world (Savvas et al., 2013). The cultivation of soilless tomatoes using various cultivation systems has been the subject of numerous studies, outlining closed cycle cultivation and use of smart technologies (i.e., sensor-based management of the cultivation process) as promising approaches to increase the sustainability of the production according to inputs use efficiency (Buttaro et al., 2015; Djurović et al., 2016; Montesano et al., 2010, 2015). In fact, not all management approaches in soilless cultivation are efficient enough according to water and fertilizer use efficiency and productivity, and/or environmental effects. For example, it was shown that with a poor quality water and empirical irrigation management, the free drain open cycle cultivation management jeopardizes the overall sustainability of soilless cultivation, since 30–50 % of the total NS supplied to the crop is commonly discharged into the environment as leaching fraction. Nonetheless, in the Mediterranean area, characterized by low-tech greenhouses, the closed-cycle nutrient solution recirculation is still poorly adopted, and in most cases only semi-closed cycles (i.e. providing periodic discharge of the recirculating NS when specific salts accumulation set-points in the NS are reached) are actually achievable (Bouchaaba et al., 2015; Massa et al., 2020; Montesano et al., 2015). From a regulatory perspective, open-cycle free drain cultivation is generally permitted in most countries, with a few exceptions and limitations (e.g., the Netherlands). Even though regulations may not specifically target soilless cultivation, they impose certain limits that must be taken into consideration in soilless cropping. In this perspective, it is worth highlighting that also in the case of the open free-drain cultivation, different sustainability levels can be achieved, since water and nutrient use can be improved, and their losses minimized, by the application of innovative approaches, as the use of sensing technologies for sensor-based agronomic inputs management, aiming at “zero emissions” (Massa et al., 2020).

The constraints of producing tomatoes with reduced input of water and fertilizers, as well as the importance of increasing the information on the sustainability of the production, point out the need for a reliable and fast non-destructive authentication method.

In this study, the utility of FT-NIR spectroscopy and hyperspectral image analysis (HSI) was evaluated for discriminating tomatoes according to i) cultural practices adopted in the production process, and ii) their degree of sustainability in terms of water use efficiency (WUE) and partial factor productivity of nutrients (PFP).

2. Materials and methods

2.1. Cultural practices and plant material

Aiming to obtain tomato fruits differing for the environmental impact as an effect of different growing techniques, two cultivation cycles, autumn-winter (August-December 2020) and spring-summer (February-June 2021), were implemented at the experimental farm ‘‘La Noria’’ of the Institute of Sciences of Food Production (CNR - ISPA) in Mola di Bari (Southern Italy). Soilless cultivation was adopted, using perlite as growing substrate and a nutrient solution (NS) complete of macro and micro nutrients for plant fertigation obtained by mixing fertilizers to irrigation water. Three different cultivation management strategies were set: i) free drain open cycle cultivation (OPEN); ii) improved open cycle cultivation providing on-demand sensor-based fertigation for optimal control of NS supply and leaching according to real plant needs (SMART); iii) semi-closed cycle cultivation, providing NS recirculation and discharge at specific electrical conductivity (EC) set-points (CLOSED). For the OPEN, NS was supplied in excess with respect to plants consumption based on an empirically prefixed fertigation schedule operated by a timer, according to the most common practice in low tech greenhouse industry (e.g. typically in Mediterranean area); the excess NS drained from substrates (typically ranging from 30 to 50 % of the supplied NS in the common practice, with even higher values in case of saline water usage) was not recycled for subsequent fertigation events and accounted as discharged into the environment; this practice generally is connected with the highest environmental impact in terms of resources use (namely water and fertilizers) and pollution implications. For the SMART treatment, NS supply was managed by an automatic smart irrigation system (Canaj et al., 2022; Montesano et al., 2018), based on real-time measurements of growing substrate moisture (volumetric water content) and EC performed by a wireless sensor network (GS3 sensor, Decagon Devices, Pullman - WA, USA); occurrence of fertigation events and volume of NS supplied per fertigation event was automatically adjusted based on volumetric water content and EC variations in the growing substrate, according to set-point values for both parameters, in order to maintain acceptable water availability and salinity levels with minimum water and fertilizers use and leaching into the environment. Finally, for CLOSED system, NS supplied in excess and drained from substrates was collected and recirculated for subsequent fertigation; according to a generally adopted practice, the NS is recirculated until a predetermined level of salinity considered detrimental for plants is reached, after which the recirculating NS is discharged into the environment, thus resulting in a semi-closed cycle. Closed cycle management is virtually the one providing the lowest environmental impact, although implying several drawbacks (nutritional imbalances linked to changes in the composition of the NS over time, rapid detrimental fungal infestation via the recycled

NS). In case of low-quality water use (typically brackish water) it may be necessary to open the cycle frequently, with potential decrease of the sustainability of this agronomical strategy.

In the autumn-winter cycle (August-December 2020) a cherry tomato variety (cv 'Carminio', Seminis-Bayer, Italy) was grown, and brackish irrigation water containing 1 g/L NaCl was used to prepare the NS, with a final EC=3.7 dS/m, in order to simulate a growing condition typical in Mediterranean coastal regions where greenhouse tomato cultivation is more widespread. In the spring-summer cultivation cycle (February-June 2021), an intermediate tomato variety (cv 'Mose', Syngenta, Italy) and good quality irrigation water (collected rain water) were used (final NS EC=2.2 dS/m).

The cultivation set-up consisted in a completely randomized block design, with 3 replications per cultivation treatment, each one constituted by 15 plants providing tomato samples for subsequent analysis.

Irrigation water use efficiency (WUE = total marketable yield mass / total irrigation water used volume) and partial factor productivity of nutrients (N_PFP, K_PFP, P_PFP, Ca_PFP, Mg_PFP = total marketable yield mass / total nitrogen, potassium, phosphorus, calcium, magnesium mass supplied as fertilizer in the NS, respectively, (Dobermann, 2007)) were used as indexes to discriminate the level of sustainability achieved in the different production processes. A summary of WUE and nutrients PFP performances resulting from the different growing conditions (in terms of tomato variety and cultivation techniques) is reported in Table 1. Within each variety, mean values of cultivation treatments were subjected to the ANOVA analysis followed by Tukey test for means separation (with STATISTICA 10.0 software, StatSoft, Tulsa, OK, USA). This analysis allowed to discriminate on a statistical basis cultivation systems according to the level of WUE and nutrients PFP, giving information about the sustainability of each treatment.

After harvesting, tomatoes were transferred within 2 h to the Postharvest facilities of the University of Foggia. In order to standardize for the maturity stage, only full ripe red tomatoes were used, with a total of about 300 fruits for each variety. Due to time limitation for sampling, not all fruit could be scanned in the same day. In this case, samples were stored in controlled humidity conditions, at 15 °C for one or two days and then scanned after leaving them at the room temperature /about 25 °C) for 3 h.

2.2. Obtaining hyperspectral images

The HS images were captured using a hyperspectral line-scan scanner (Version 1.4, DV srl, Padova, Italy) outfitted with two spectrographs, one in the Vis-NIR range and the other in the NIR region. The Vis-NIR spectrograph (400–1000 nm) had a spatial resolution of 1000 × 2000 pixels with a spectral resolution of 5 nm and was connected to a CCD camera. As for the NIR spectrograph (900–1700 nm), the spatial resolution was 600 × 320 pixels with a spectral resolution of 5 nm and it was connected to a CMOS (Specim Spectral Imaging Ltd., Oulu, Finland) with 50 frames per second, equipped with C-mount lenses. The device was equipped with 150 W Halogen lamps through fiber optic bundles and two light lines. The GigE vision was used as the interface with a 37° field of view (FOV). Depending on the size of tomatoes, 6 to 15 tomatoes were acquired in each scan, first with Vis-NIR and then with NIR scanners. For calibration purposes, dark and white references were taken before switching between the two spectrographs. Closing the shutter of camera allowed for the acquisition of the black reference (0 % reflectance), while a 99 % Spectralon reflectance standard was employed for the white reference. Using Eq. (1), it was possible to determine the reflectance of tomatoes (Tsouvaltzis et al., 2020)

$$R = \frac{H - B}{W - B} \tag{1}$$

where H stands for the raw hyperspectral image; B and W stand for the acquired signals for the black and white references, respectively. The surface characteristics of tomato fruit, as well as the existence of directional radiations, may result in glare pixels. As a result, there is spectrum distortion, which affects the quality of obtained information, and therefore must be eliminated before processing the data. The Region of Interest (ROI) consists of pixels that correspond to each individual fruit, whereas glare pixels are deleted. A self-developed MATLAB code was employed to detect ROI and retrieve the corresponding reflectance data. Two grayscale images at two different wavelengths that showed a higher contrast between the background and tomato fruit, and between the background and the glare regions, were chosen for each tomato fruit. The Otsu approach was then used to threshold and binarize the selected images. After identifying the background and the glare pixels, with a 0 value for the background and glare regions and a 1 value for the pixels of fruit skin which are not glare, a 2-D binary picture was obtained as a mask. This mask later applied to all wavelength images to keep only the ROI in each image. Finally, the reflectance values for each pixel of ROI were extracted to generate a mean spectra for each individual fruit and collected as a raw data in the dataset, where the columns correspond to the number of wavelengths.

Table 1
Irrigation water use efficiency (WUE = total marketable yield/total water supplied) and partial factor productivity of nutrients (N PFP, K PFP, P PFP, Ca PFP, Mg PFP = total marketable yield/total nitrogen, potassium, phosphorus, calcium, magnesium supplied as fertilizer in the NS, respectively) achieved in 'Carminio' and 'Mose' cv under different cultivation conditions. Different letters among the rows indicate statistical difference at the Tukey test ($p < 0.05$).

Autumn-winter growing cycle (Cherry tomato cv 'Carminio', brackish irrigation water)						
	WUE	N PFP	K PFP	P PFP	Ca PFP	Mg PFP
	(kg MY/m^3 WS)	(kg MY / kg nutrient supplied as fertilizer)				
OPEN	15 b	116 b	64 b	250 b	125 b	313 b
SMART	26 a	290 a	161 a	629 a	315 a	787 a
CLOSED	16 b	124 b	68 b	268 b	134 b	335 b
Spring-summer growing cycle (tomato Cv 'Mose', good quality irrigation water)						
	WUE	N PFP	K PFP	P PFP	Ca PFP	Mg PFP
	(kg MY/m^3 WS)	(kg MY / kg nutrient supplied as fertilizer)				
OPEN	12 b	80 b	40 b	239 b	80 b	239 b
SMART	18 a	121 a	60 a	361 a	121 a	361 a
CLOSED	17 a	128 a	64 a	383 a	128 a	383 a

2.3. Obtaining FT-NIR spectra

Tomato fruits were subjected to Fourier transformed-near infrared (FT-NIR) spectroscopy at room temperature (25 °C). Three scans of different parts of each tomato fruit were collected by manually turning the fruit (MPA Multi-Purpose FT-NIR Analyzer, Bruker Optics, Ettlingen, Germany). A representative spectrum for each fruit was obtained by averaging these three spectra. In reflectance mode, the device was set to obtain spectrum across the 800–2777 nm range (sphere macrosample resolution 1.71 nm, scanner velocity 10 kHz, sample and background scan time 64 scans.) The instrument included a permanently aligned, extremely stable ROCKSOLID interferometer and a high-energy air-cooled NIR source (20 W tungsten-halogen lamp) (Bruker).

2.4. Maturity indexes and composition

Total soluble solid content (SSC) of the juice from each fruit was measured using a digital refractometer (Atago N1, PR32-Palette, Tokyo, Japan). A 5 g sample of juice was used to determine the titratable acidity (TA) and pH using an automatic titrator (TitroMatic CRISON, Barcelona, Spain). The samples were titrated up to a final pH of 8.1 against a 0.1 mol l-1 NaOH solution, and the citric acid content per 100 g of sample was recorded. With slight adjustments, the levels of vitamin C (Vit-C), l-ascorbic acid (AA), and l-dehydroascorbic acid (DHA) were determined as per (Zapata and Dufour, 1992).

2.5. Data handling and analytics

The image and spectral data handling were carried out using MATLAB R2020b (version 9.9.0.1467703, MathWorks, Natick, MA, USA) and PLS_TOOLBOX 8.6 (Eigenvector Research Inc., Manson, WA, USA). Spectral preprocessing, feature extraction, and model parameter optimization are the most used functions for handling complex data (Vidal and Amigo, 2012). The samples were first divided into calibration and prediction (test) sets. The model was trained using the calibration set, and it was then tested using the prediction set which was not used for training purpose. The data splitting was carried out utilizing the Kennard-Stone approach (Kennard and Stone, 1969): the calibration set and the prediction set were each given 70 % and 30 % of the samples, respectively. Then, in order to improve the performance or interpretation of discrimination, FT-NIR spectral and HSI data were preprocessed to remove measurement artifacts. After that, defective spectral outliers were identified and eliminated using 'robustpls' algorithm, which uses the function 'rsimpls' from the LIBRA Toolbox (kuleuven.be) (Hubert and Vanden Branden, 2003).

Partial Least Squares Discriminant Analysis (PLS-DA) through SIMPLS algorithm (Barker and Rayens, 2003; Sijmen Jong, 1993) was used to discriminate tomatoes according to cultural practices, and to WUE and nutrients PFP indexes. In order to choose the appropriate number of latent variables (LVs) and prevent the calibration model from being overfit, a venetian blinds cross-validation with 10 splits was utilized. By maximizing the explained variance and optimizing the number of variables, the best number of LVs was determined as described by Ballabio and Consonni (2013). As for the first aim, PLS-DA was implemented to discriminate the three treatments of cultivation approaches -OPEN, CLOSED, and SMART- for each variety (Table 1). This involved identifying the fundamental relationships (i.e. latent variables) between the matrix of spectra and the vector of classes, and maximizing the covariance between these two spaces. The second objective was to test the capability of classifiers to differentiate tomatoes according to WUE and nutrients PFP. With this aim for each experiment, different binary classifiers were calibrated using PLS-DA to discriminate tomatoes according to the two levels of WUE, resulting from the statistical analysis, as reported in Table 1. For the 'Carminio' variety, OPEN and CLOSED were not statistically different and differentiated from SMART, whereases for 'Mose', SMART and CLOSED resulted in the same WUE differing from OPEN. In order to create a more robust model with clear distinctions, the datasets were combined, and the classes were redefined into three levels: HIGH, LOW, and MEDIUM. The different classes used for the discrimination models, in each experiment and in the general model, are summarized in Table 2.

Finally, for the last classification model on three levels of WUE and nutrients PFP, the interval partial least square (IPLS) variable selection method was applied to find and select the best features or combination of them (Nørgaard et al., 2000) in order to simplify the model, enhance the interpretability and its robustness, by eliminating non-relevant information, according to the fact that a model using a limited number of wavelengths is more desirable, due to its high potential for industrial applications and on-line control by replacing expensive and time-consuming instruments to an efficient filter-based one. A repeated double cross-validation (rDCV) technique, as described by Filzmoser et al. (2009), was used to optimize the complexity of model, as well as to carefully estimate the range of prediction errors to be expected for new cases and to assess the robustness of model.

The effectiveness of PLS-DA models for classification was assessed. in terms of Sensitivity (percentage of correctly identified positive instances), Specificity (percentage of correctly identified negative cases), and Accuracy (proportion of samples which were correctly classified); and reported for calibration (Cl), cross-validation (CV), and Prediction (Pr).

$$Sensitivity\,(SENS) = \frac{true\,positives}{true\,positives + false\,negatives} \tag{2}$$

$$Specificity\,(SPEC) = \frac{true\,negatives}{true\,negatives + false\,positives} \tag{3}$$

$$Accuracy = \frac{true\,positives + true\,negatives}{true\,positives + true\,negatives + false\,positives + false\,negatives} \tag{4}$$

Table 2
Assigned classes (i.e., 1, 2, and 3) for each discrimination objective on individual dataset for 'Carminio' tomatoes and 'Mose' cv, and for combined datasets (combining the two varieties).

	Discrimination Classes		
	First Objective	Second Objective	
'Carminio', autumn-winter, brackish water	'Carminio'	'Carminio'	Combined Datasets
OPEN	1	2	2 or "MEDIUM"
SMART	2	1	3 or "HIGH"
CLOSED	3	2	2 or "MEDIUM"
'Mose', spring-summer, good water	'Mose'	'Mose'	–
OPEN	1	1	1 or "LOW"
SMART	2	2	2 or "MEDIUM"
CLOSED	3	2	2 or "MEDIUM"

3. Results

3.1. WUE, nutrient PFP, maturity indexes and composition

For 'Carminio' cv tomatoes cultivated in the autumn-winter cycle, fruits collected from CLOSED system showed value of WUE similar to OPEN, due to the frequent renewal of the recirculating nutrient solution (NS) caused by NaCl accumulation; SMART exhibited higher WUE and nutrient partial factor productivity (PFP) than OPEN and CLOSED (Table 1). For 'Mose' cv tomatoes cultivated in the spring-summer cycle, SMART and CLOSED showed similar performances, with significantly higher use efficiency and productivity of inputs compared to open free-drain growing conditions (Table 1).

Table 3
Quality attributes of 'Carminio' and 'Mose' cv tomatoes under different cultivation conditions. Data are shown as Mean ± Standard Error. SSC (%); TA (g citric acid /100 mL); AA, DHAA, Vit-C (g kg-1). Different letters indicate significant differences by the ANOVA/Tukey-HSD ($p < 0.05$).

Variety	Quality Attribute	Mean ± SE OPEN	CLOSED	SMART
'Carminio'	pH	4.1 ± 0.03 ns	4.17 ± 0.03 ns	4.11 ± 0.03 ns
	SSC	6.8 ± 0.15 b	7.5 ± 0.19 a	7.7 ± 0.15 a
	TA	0.65 ± 0.01 b	0.62 ± 0.01 c	0.68 ± 0.01 a
	AA	8.40 ± 0.79 ns	6.85 ± 0.76 ns	7.17 ± 0.81 ns
	DHAA	16.82 ± 3.02 ns	12.99 ± 2.89 ns	22.22 ± 3.07ns
	Vit-C	25.23 ± 2.8 ab	19.83 ± 2.68 b	29.39 ± 2.85 a
'Mose'	pH	4.36 ± 0.03 a	4.26 ± 0.02 b	4.29 ± 0.02 b
	SSC	5.5 ± 0.19 b	6.5 ± 0.30 a	6.1 ± 0.10 a
	TA	0.67 ± 0.02 b	0.83 ± 0.03 a	0.75 ± 0.03 ab
	AA	8.10± 0.59 ns	10.03 ± 0.57 ns	9.66 ± 0.58 ns
	DHAA	6.69 ± 0.60 ns	5.4 ± 0.58 ns	4.78 ± 0.6 ns
	Vit-C	14.56 ± 0.49 ns	15.43 ± 0.47 ns	14.44 ± 0.49 ns

A summary of the main maturity indexes and Vit-C content of tomatoes at harvest is presented in Table 3. Although only fully ripe tomatoes were used for the experiment, significant differences in maturity indexes and composition were observed in terms of SSC and TA for both varieties, as well as Vit-C content for 'Carminio' tomatoes, and pH for 'Mose' tomatoes.

Specifically, for 'Carminio' tomatoes, while AA and DHAA were not significantly affected by the treatment, their summation (total Vit-C) was higher in the SMART treatment than in the CLOSED, with an intermediate content observed in the OPEN treatment. Regarding SSC, elevated concentrations were observed in the CLOSED and SMART treatments in comparison to the OPEN treatment. This may suggest a potential outcome of stress induced by these two techniques on the plants. However, the effect on TA varied, suggesting an interaction with the variety (Table 1).

3.2. Hyperspectral data

The reflectance spectra in the Vis-NIR (400–1000 nm) and NIR (900–1700 nm) ranges are shown in Fig. 1.

Looking at the raw data of Vis-NIR camera (Fig. 1A), for 'Carminio' cv tomatoes, a very high difference was observed between the SMART class and two other classes (i.e. OPEN and CLOSED), while for 'Mose' cv tomatoes, SMART and CLOSED were more similar and differed from OPEN. These findings were in agreement with the WUE and nutrients PFP indexes of each class (Table 1), where the SMART class and the OPEN class showed different values compared to the other two classes, respectively for 'Carminio' and 'Mose' tomatoes. In addition, it is possible to realize that differences in WUE and nutrients PFP for 'Carminio' cv tomatoes were higher than in 'Mose' cv tomatoes (Table 1, and Fig. 1A).

For NIR Region (Fig. 1B), although the aforementioned finding was generally applicable, the spectral data of different classes were more overlapped in comparison to Vis-NIR region and therefore lower performance of discrimination could be expected.

3.3. FT-NIR spectral data

The raw data obtained from FT-NIR instrument in range of 800–2777 nm is represented in Fig. 2. Similar to HSI data, the spectral data of FT-NIR for the range of 800–2000 nm (Fig. 2) indicated the same group of classes, allowing the same considerations regarding WUE and nutrients PFP, and being as for NIR spectra in HSI less separated if compared to VIS-NIR.

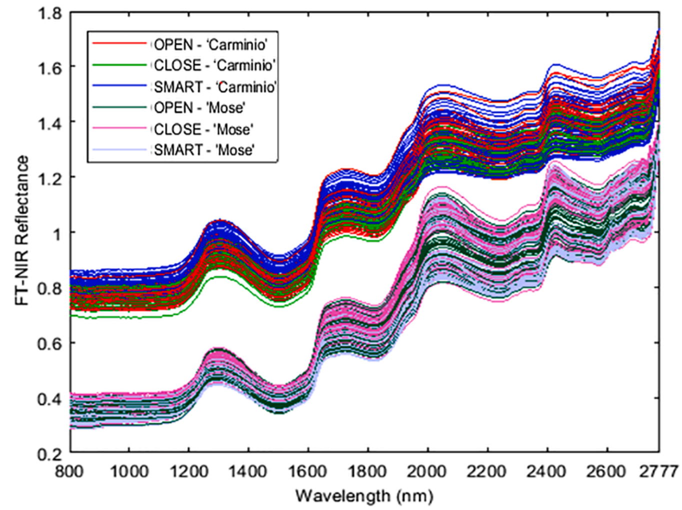

Fig. 2. Raw data obtained from FT-NIR instrument for both 'Carminio and 'Mose' cv tomatoes (colored according to the cultural practices).

3.4. Classification models

Data cleaning is required before any interpretation since data may contain inaccurate, mistaken, and redundant values. It will contribute to higher performance in terms of decision-making. Therefore, before applying the discrimination algorithms, 'robustpls' algorithm was applied to eliminate spectral outliers (Barker and Rayens, 2003) selecting 286 'Carminio' and 295 'Mose' cv tomatoes for Vis-NIR range, and 285 'Carminio' and 291 'Mose' cv tomatoes for NIR range. The number of FT-NIR spectra chosen for discrimination analysis, after outlier elimination, was equal to 283 for 'Carminio' cv and 289 for 'Mose' cv tomatoes. Then, the raw data were initially preprocessed by different combination of spectral pre-treatments inorder to guarantee the model reliability. The most effective pre-treatments, allowing the highest accuracy and the lowest error in classification, were selected. It was found that SNV followed by MC, and SNV followed by 2nd Derivative and then MC were the best pre-processing treatments, for Vis-NIR and NIR ranges, respectively. For data acquired from FT-NIR device, the best pre-treatment was a combination of smoothing, Detrend, and MC.

As for the first objective, for both the two tomato varieties, PLS-DA algorithm was applied to develop calibration models aimed to discriminate tomatoes based on three compared cultural practices: OPEN,

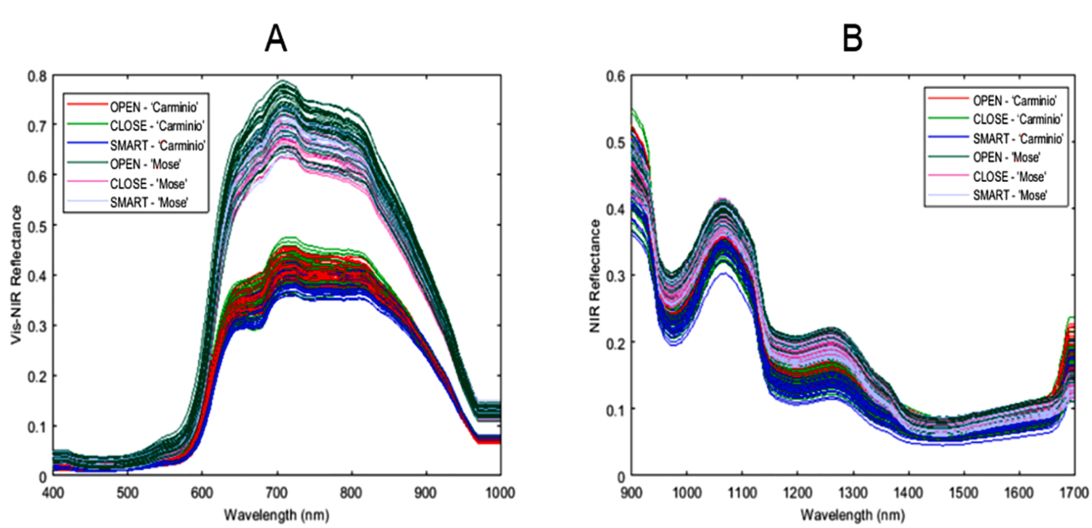

Fig. 1. Raw data obtained from (A) Vis-NIR and (B) NIR hyperspectral camera for both 'Carminio' and 'Mose' cv tomatoes (colored according to the cultural practices).

CLOSED, and SMART. Confusion Tables 4 and 5 summarize the number of samples that were correctly classified in their actual respective classes, and the performance metrics of each class for 'Carminio' and 'Mose' tomatoes, respectively, for FT-NIR and HSI instruments in Vis-NIR and NIR ranges. All of the classification models successfully discriminated tomatoes into three different classes, with the accuracy and the specificity of prediction higher than 75 % and 81 %, respectively. As for sensitivity, lowest values were observed, being as low as 62 % for HSI NIR, and slightly higher than 71 % for FT-NIR.

Although the performance results of models made with FT-NIR data seem to be the highest among all three instruments (around 81 % and 86 % for the accuracy and the specificity of prediction set, respectively), the model calibrated on Vis-NIR range is more preferable. This is due to the fact that the model complexity in terms of the number of latent variables utilized must also be addressed. Indeed, the model in the VIS-NIR uses less latent variables (LV=5) than FT-NIR (LV=8). Thus, the best model for discriminating tomatoes into three different classes could be defined by the model generated with Vis-NIR HSI spectra, yielding accuracy of 80.6 % and 79.5 %, and specificity of 86.6 % and 84.7 % on prediction set for 'Carminio' and 'Mose' cv tomatoes, respectively.

The second goal was to examine classifiers' capacity to identify treatments that had comparable levels of WUE and nutrients PFP. As resulted from the ANOVA and mean separation (Table 1), only two levels of WUE and nutrients PFP resulted for each experiment, since two classes of OPEN and CLOSED for 'Carminio' cv tomatoes showed similar indexes (low efficiency) compared to SMART (high efficiency); whereas, for 'Mose' cv tomatoes OPEN had lower efficiency than SMART and CLOSED. The results showed that discrimination based on WUE yielded specificity and sensitivity of 84.5 % and 86.3 % on prediction set for 'Carminio' cv and 'Mose' cv tomatoes, respectively (Table 6).

However, when the combined dataset was used to test the discrimination model including the 2 varieties and considering 3 levels of WUE and nutrients PFP (see Table 2) the performance of the model was really excellent as discrimination of tomatoes based on three levels of water and fertilizers used efficiency reached accuracy and specificity in prediction equal to 92.1 % and 94.1 %, respectively (Table 6). This restructuring allowed for the establishment of more meaningful differences between the classes, providing a solid foundation for subsequent analyses. The decision to combine the datasets was driven by the aim to improve the generalizability of the model, as tested on a larger dataset featuring more distinct classes. This approach ensures a more comprehensive evaluation of the model performance across a broader range of scenarios. Finally, Using the IPLS variable selection method, the top twenty significant wavelengths were chosen in order to streamline the final model and increase its generalizability and robustness. The selected 20 wavelengths, as shown in Fig. 3, were nm [425, 470:490, 515, 535, 545:555, 565, 575, 635, 765, 775, 785, 925, 930, 970]. A repeated double cross validation (rDCV) technique (50 runs) yielded accuracy and specificity of 89.8 % ± 1.8 and 91.7 % ± 2.0, respectively, on the outer loop procedure. In comparison to employing the full-range wavelengths, the model performance after variable selection was marginally lower.

4. Discussion

Several rapid and nondestructive discrimination studies have tested various approaches for discrimination of tomato fruits, plants, leaves, and crops using both HSI and FT-NIR spectroscopy (Lu et al., 2018;

Table 4
Confusion matrices for discrimination of 'Carminio' cv tomatoes according to cultural practices using three instruments A. FT-NIR, B. HSI (Vis-NIR), C. HSI (NIR). Cal: calibration set; CV: cross validation set; Pr. prediction set. LV: latent variables.

Ins.	LV		Confusion Matrix					Performance Metrics (%)					
				Actual				Sensitivity		Specificity		Accuracy	
FT-NIR	**8**	**Cal**	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	50	0	14	75.76	77.37	89.39	88.61	84.85	84.85
				SMART	1	57	7	87.7		93.99		91.92	
				CLOSED	15	8	46	68.66		82.44		77.78	
		CV	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	48	1	15	72.73	72.30	87.88	86.08	82.83	81.48
				SMART	2	52	9	80.00		91.73		87.88	
				CLOSED	16	12	43	64.18		78.63		73.74	
		Pr.	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	19	0	6	67.86	**71.95**	89.47	**85.96**	82.35	**81.18**
				SMART	2	25	4	83.33		89.09		87.06	
				CLOSED	7	5	17	62.96		79.31		74.12	
HSI (Vis-NIR) ()	**5**	**Cal**	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	49	1	16	75.39	76.12	87.41	87.96	83.50	84.00
				SMART	2	56	6	84.85		94.03		91.00	
				CLOSED	14	9	47	68.12		82.44		77.50	
		CV	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	47	2	18	72.31	72.15	85.18	85.96	81.00	81.33
				SMART	2	54	8	81.82		92.54		89.00	
				CLOSED	16	10	43	62.32		80.15		74.00	
		Pr.	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	20	1	8	71.43	**70.45**	84.48	**85.67**	80.23	**80.62**
				SMART	0	25	3	80.65		94.55		89.54	
				CLOSED	8	5	16	59.26		77.97		72.09	
HSI (NIR) ()	**6**	**Cal**	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	44	4	18	68.76	67.98	83.70	83.87	78.90	78.56
				SMART	3	50	10	75.76		90.23		85.43	
				CLOSED	17	12	41	59.42		77.69		71.36	
		CV	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	42	4	18	63.64	64.57	83.58	82.21	77.00	76.33
				SMART	4	48	11	72.73		88.81		83.50	
				CLOSED	20	14	39	57.35		74.24		68.50	
		Pr.	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	17	3	8	60.71	**62.41**	81.04	**81.57**	74.42	**75.20**
				SMART	2	22	4	70.97		89.09		82.56	
				CLOSED	9	6	15	55.56		74.58		68.61	

Table 5
Confusion matrices for discrimination of 'Mose' cv tomatoes according to cultural practices using three instruments A. FT-NIR, B. HSI (Vis-NIR), C. HSI (NIR). Cal: calibration set; CV: cross validation set; Pr. prediction set. LV: latent variable.

Ins.	LV		Confusion Matrix					Performance Metrics (%)					
				Actual				Sensitivity		Specificity		Accuracy	
FT-NIR	**8**	**Cal**	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	61	5	4	84.72	76.07	93.13	88.25	90.15	84.23
				SMART	6	45	13	69.23		86.23		80.79	
				CLOSED	5	15	49	74.24		85.40		81.77	
		CV	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	59	7	7	81.94	70.07	89.31	85.29	86.70	80.30
				SMART	8	42	17	64.62		81.89		76.36	
				CLOSED	5	16	42	63.64		84.67		77.83	
		Pr.	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	21	2	1	75.00	**71.22**	94.83	**85.68**	88.37	**80.62**
				SMART	5	18	10	72.00		75.41		74.42	
				CLOSED	2	5	22	66.67		86.79		79.07	
HSI (Vis-NIR) ()	**5**	**Cal**	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	58	8	4	80.56	73.27	91.11	86.78	87.44	82.29
				SMART	8	46	17	69.70		82.27		78.26	
				CLOSED	6	12	48	69.57		86.96		81.16	
		CV	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	58	8	5	80.56	72.28	90.37	86.29	86.96	81.64
				SMART	8	45	17	68.18		82.27		77.78	
				CLOSED	6	13	47	68.12		86.23		80.19	
		Pr.	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	24	4	3	75.00	**69.19**	87.50	**84.71**	82.96	**79.55**
				SMART	5	19	5	63.33		82.76		86.14	
				CLOSED	3	7	18	69.23		83.87		79.55	
HSI (NIR) ()	**6**	**Cal**	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	57	8	7	78.08	70.15	88.89	85.23	85.10	80.13
				SMART	9	43	20	69.36		80.14		76.92	
				CLOSED	7	11	46	63.01		86.67		78.37	
		CV	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	54	10	9	73.97	66.25	85.93	83.29	81.73	77.56
				SMART	11	40	20	64.52		78.77		74.52	
				CLOSED	8	12	44	60.27		85.19		76.44	
		Pr.	**Predicted**		**OPEN**	**SMART**	**CLOSED**		Average		Average		Average
				OPEN	18	4	3	64.29	**62.63**	87.27	**81.30**	79.52	**75.10**
				SMART	6	18	7	62.07		75.93		71.09	
				CLOSED	4	7	16	61.54		80.70		74.70	

Shrestha et al., 2016; Xia et al., 2021; Xie et al., 2010). In this study, the possibility of the aforementioned techniques was tested for discriminating soilless tomato fruits according to cultural practices using different approaches to water use, resulting in different levels of WUE and PFP The results demonstrated that different cultural practices can affect tomato fruit fingerprints by altering their chemical components (Table 3), and/or the cell and tissue morphology, providing the possibility of non-destructive tomato discrimination. Particularly, results showed that the observed spectral changes were not solely just a direct consequent of SSC, which represents the primary constituents of fruits. This was evident as the mean separation for SSC did not align with the same distribution of WUE groups, especially for the 'Carminio' cv, and nor did it correspond with TA. Additionally, the significant difference in Vit-C content observed only for 'Carminio' cv tomatoes, did not account for the variations attributed to WUE.

On the other side, when considering the combined dataset, the concurrent variation in the main constituents—SSC, TA, and ascorbic acids—correlated with both WUE and PFP, suggesting efficient utilization of water and fertilizers for mass production and ripening mechanisms, including the accumulation of SSC and Vit-C.

Additionally, it is possible that cultural practices based on low input agricultural practices (LIP) may significantly impact the spectral fingerprint of tomatoes, overcoming differences attributed to maturity and main constituents, as they were consistent across the differences of the two used varieties and of the water quality. The spectral signal may, in fact, also be affected by other tissue characteristics such as firmness, cell size, cell wall characteristics, and other morphological factors, inducing difference in absorption and scattering (Franceschini et al., 1999; Gono et al., 2004; Wilson et al., 2015).

This brings to the conclusion that cultural practice based on low input agricultural practices (LIP) may affect the spectral fingerprint of tomatoes in a significant way, overcoming differences due to maturity, and being consistent over the differences of the two used varieties and of the water quality.

Calibrating several classification models on data from both HSI and FT-NIR devices demonstrated that both the optical non-destructive approaches utilized in this study may be used for the purpose of tomato classification (with great accuracy) based on their cultural practices and the amount of water and fertilizer used. Models calibrated on HSI in the Vis-NIR band, however, are preferable due to their lesser complexity and smaller number of latent variables (Ballabio and Consonni, 2013), which give a larger potential for distinguishing tomato fruits.

Taking into account that the specificity metric is defined as ability of a model to identify negative cases, and the initial goal of the study was to avoid having products with fake labels that are not truly sustainable-produced, a model that is more capable of identifying negative cases is preferable. Consequently, in this research, the higher specificity means the more reliable model. In the concluded trials, three distinct production lines can be distinguished with the specificity of 86 %, while discrimination of three degree of WUE and nutrients PFP yielded a specificity of 94 % (Tables 4, and 5).

Thus, it is possible to conclude that the higher differences between cultural practices in terms of WUE and nutrients PFP result in a better discrimination with lower error and higher specificity/accuracy. Furthermore, although the results showed that the performances were slightly worse after feature selection (down to 20 features) when compared to using the full-range wavelengths, the obtained method is viable to be implemented for on-line control and industrial applications

Table 6
Different classification models for additional elaboration of HSI in the Vis-NIR range. The first 2 models are binary classifiers discriminating 2 classes of WUE and nutrients PFP. The last model is a three-class classifier according to three different levels of WUE and nutrients PFP, using full range spectra and selected wavelengths.

Model		Sensitivity						Specificity						Accuracy					
		'Carminio' cv			'Mose' cv			'Carminio' cv			'Mose' cv			'Carminio' cv			'Mose' cv		
		Cl	CV	Pr.	Cl	CV	Pr.	Cl	CV	Pr.	Cl	CV	Pr.	Cl	CV	Pr.	Cl	CV	Pr.
OPEN Vs. (CLOSED and SMART)		–			88.1	86.8	**86.3**	–			88.1	85.6	**86.3**	–			87.9	86.7	**86.1**
SMART Vs. (OPEN and CLOSED)		88.1	86.5	**84.5**	–			88.1	86.5	**84.5**	–			88.4	86.6	**84.7**	–		
"LOW" Vs. "MEDIUM" Vs. "HIGH"		Merged Experiments						Merged Experiments						Merged Experiments					
	Full Range	**Cl**		**CV**		**Pr.**		**Cl**		**CV**		**Pr.**		**Cl**		**CV**		**Pr.**	
		92.8		91.6		**88**		97		95.8		**94.1**		96.1		94.4		**92.1**	
	Selected Variables	**Inner rDCV loop**			**Outer rDCV loop**			**Inner rDCV loop**			**Outer rDCV loop**			**Inner rDCV loop**			**Outer rDCV loop**		
		89.7 ± 2.1			**87.3 ± 1.3**			94.9 ± 1.8			**91.7 ± 2.0**			92.1 ± 1.4			**89.8 ± 1.8**		

by replacing expensive and time-consuming instruments to an efficient filter-based one, or even for consumer use at the distribution market by making the elaboration faster and the acquisition simpler, as for instance by using equipment with an efficient light filters.

Several literary studies have looked into the suitability of non-destructive technologies for classification of tomato fruits in different varieties (Najafian et al., 2021); maturity stages (Jiang et al., 2021); levels of firmness (Rahman et al., 2018); and production methods (organic and conventional technologies) (Araújo et al., 2019). Nonetheless, to the best of our knowledge, there was not any research on non-destructive discrimination of tomato fruits according to cultural practices and/or their degree of WUE and nutrients PFP reported. Some researchers reported in literature, aimed to apply these technologies for early detection of tomato diseases (Abdulridha et al., 2020; Morellos et al., 2020), discrimination of drought stress origin in tomatoes (Žibrat et al., 2019), and assessing the water stress severity in soilless tomato crop (Elvanidi et al., 2018). A study looked into how useful HSI could be for classification of water-stressed tomato plants in a field experiment with different irrigation treatments (Rinaldi et al., 2015). They demonstrated that a clear distinction between the two differently irrigated areas could only be made by employing the entire wavelength range from 400 to 1000 nm. They were able to differentiate between the two areas irrigated with different amounts of water with a mean accuracy of roughly 77 % (Rinaldi et al., 2015). In comparison to this study, which had the goal of discriminating tomato plants rather than tomato fruits, the results obtained in this research showed that the reached performances were qualitatively better, achieving an accuracy of more than 80 % for discriminating three cultural practices and around 90 % for three levels of water use efficiency despite using a limited number of wavelengths.

The aforementioned non-destructive technologies were also applied for the dentification of the cultural practices of rocket leaves (Palumbo et al., 2021), diagnosis of nitrogen nutrition of tea plants under field conditions (Wang et al., 2020), discrimination of harvest time in fennels (Amodio et al., 2017a) and discrimination of tropical grasses grown under different nitrogen fertilizer regimes (Naicker et al., 2020). In the case of Rocket leaves (Palumbo et al., 2021), authors showed that about 70 % of samples were accurately discriminated using computer vision systems with digital imaging and concluded that the obtained performance is not relevant for practical application. Moreover, they claimed that the properties of the product were not considerably impacted by the various cultivation techniques. In comparison to this study, the results obtained in the conducted research were showed a superior performance (Accuracy = 81 %; Specificity = 86 %) in discriminating tomatoes produced adopting three different cultural practices. The obtained results in the present research also showed that the cultivation methods significantly affected some of quality attributes of tomatoes (Table 3). For the tea plants study (Wang et al., 2020), HSI with PLS-DA and least squares–support vector machines (LS-SVM) algorithms were employed to classify different nitrogen status. The findings obtained in this research demonstrated that the LS-SVM model produced superior performance, with correct classification rates of 82 % in prediction sets for the diagnosis of various N application rates and 92 % for N status (Wang et al., 2020). In a study on tropical grasses (Naicker et al., 2020), the effectiveness of using remotely sensed data and multivariate discrimination ensembles to characterize grasses grown in various nitrogen concentrations was examined by the authors and achieved an accuracy of 91 %. However, compared to the tea plant and the tropical grasses studies, the results of the present study demonstrated higher performance for classifying tomatoes in two levels of WUE and nutrients PFP (accuracy = specificity = sensitivity = 95.3 %).

To broaden the scope of this study, two tomato varieties ('Carminio' and 'Mose') and different water quality were used. The obtained results confirmed the fact that the WUE and nutrients PFP indexes, as well as the chemical compounds may differ from one tomato variety to another. Nonetheless, given the constraints of this research, we proved that these

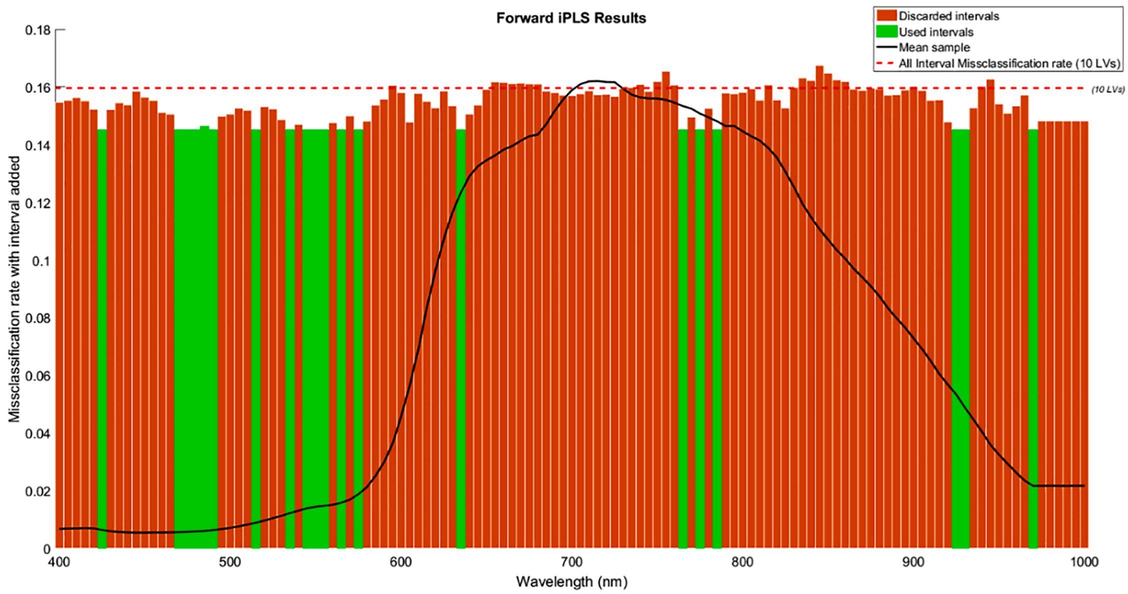

Fig. 3. Misclassified rate with wavelength interval added using IPLS technique, red regions: Discarded intervals, green regions: Selected intervals: [425 470:490 515 535 545:555 565 575 635 765 775 785 925 930 970].

non-destructive methods were capable of distinguishing tomato fruits obtained with low agronomic input. As a result, the obtained findings emphasize the need of further study to validate the potential of HSI analysis and NIR spectroscopy in ensuring the authenticity of soilless tomato fruits. This include the performance of the models through additional studies and refinements.

5. Conclusion

This research study confirms the ability of hyperspectral imaging and FT-NIR spectroscopy for differentiating soilless tomatoes based on their cultural practices, as well as their degree of water and fertilizer usage efficiency. Indeed, the obtained results demonstrated that various cultural practices may vary the spectral information of tomato fruits, providing predictive value for accurately identifying the majority of samples. The findings showed that the higher the variations in agronomic input efficiency indices between samples, the greater the likelihood of non-destructive discrimination. However, further studies are needed to extend the applicability of the obtained results. Furthermore, the outcome of the feature selection technique showed that the technology is appropriate for technology transfer, to be used online or at the market, as a quick supplement to increase the information about the sustainability of the production. In conclusion, the obtained results emphasize the need of further studies on spectral techniques for the discrimination of tomato fruits and other crops based on water used efficiency, in order to make available new tools to help the diffusion of good sustainable practices, which may represent a good market leverage in the future.

CRediT authorship contribution statement

Hassan Fazayeli: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Maria Luisa Amodio:** Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Supervision, Writing – original draft, Writing – review & editing. **Danial Fatchurrahman:** Data curation, Methodology. **Francesco Serio:** Conceptualization, Data curation, Methodology, Investigation. **Francesco Fabiano Montesano:** Conceptualization, Data curation, Investigation, Methodology, Writing – review & editing. **Ingunn Burud:** Conceptualization, Methodology, Writing – review & editing. **Andrea Peruzzi:** Conceptualization, Methodology, Writing – review & editing. **Giancarlo Colelli:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing.

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

No data was used for the research described in the article.

Acknowledgment

This research was funded by the project Prin 2017 “SUS&LOW-Sustaining low-impact practices in horticulture through non-destructive approach to provide more information on fresh produce history and quality” (grant number: 201785Z5H9) from the Italian Ministry of Education University. We also acknowledge the Italian Ministry of Education University who funded the scholarship for Hassan Fazayeli through the PON RICERCA e INNOVAZIONE 2014–2020- Dottorati innovativi a caratterizzazione industrial a.a. 2018/2019 (codice = DOT13YISJ8).

References

Abdulridha, J., Ampatzidis, Y., Kakarla, S.C., Roberts, P., 2020. Detection of target spot and bacterial spot diseases in tomato using UAV-based and benchtop-based hyperspectral imaging techniques. Precis. Agric. 21, 955–978. https://doi.org/10.1007/s11119-019-09703-4.

Amodio, M.L., Capotorto, I., Chaudhry, M.M.A., Colelli, G., 2017a. The use of hyperspectral imaging to predict the distribution of internal constituents and to classify edible fennel heads based on the harvest time. Comput. Electron. Agric. 134, 1–10. https://doi.org/10.1016/j.compag.2017.01.005.

Amodio, M.L., Ceglie, F., Chaudhry, M.M.A., Piazzolla, F., Colelli, G., 2017b. Potential of NIR spectroscopy for predicting internal quality and discriminating among strawberry fruits from different production systems. Postharvest Biol. Technol. 125, 112–121. https://doi.org/10.1016/j.postharvbio.2016.11.013.

Araújo, E.M., de Lima, M.D., Barbosa, R., Alleoni, L.R.F., 2019. Using machine learning and multi-element analysis to evaluate the authenticity of organic and conventional

vegetables. Food Anal. Methods 12, 2542–2554. https://doi.org/10.1007/s12161-019-01597-2.
Ballabio, D., Consonni, V., 2013. Classification tools in chemistry. Part 1: linear models. PLS-DA. Anal. Methods 5, 3790–3798. https://doi.org/10.1039/c3ay40582f.
Barker, M., Rayens, W., 2003. Partial least squares for discrimination. J. Chemom. 17, 166–173. https://doi.org/10.1002/cem.785.
Bouchaaba, Z., Santamaria, P., Choukr-Allah, R., Lamaddalena, N., Montesano, F.F., 2015. Open-cycle drip vs closed-cycle subirrigation: effects on growth and yield of greenhouse soilless green bean. Sci. Hortic. 182, 77–85. https://doi.org/10.1016/j.scienta.2014.11.007 (Amsterdam).
Brooks, C., Parr, L., Smith, J.M., Buchanan, D., Snioch, D., Hebishy, E., 2021. A review of food fraud and food authenticity across the food supply chain, with an examination of the impact of the COVID-19 pandemic and Brexit on food industry. Food Control 130, 108171.
Buttaro, D., Santamaria, P., Signore, A., Cantore, V., Boari, F., Montesano, F.F., Parente, A., 2015. Irrigation management of greenhouse tomato and cucumber using tensiometer: effects on yield, quality and water use. Agric. Agric. Sci. Procedia 4, 440–444. https://doi.org/10.1016/j.aaspro.2015.03.050.
Canaj, K., Parente, A., D'imperio, M., Boari, F., Buono, V., Toriello, M., Mehmeti, A., Montesano, F.F., 2022. Can precise irrigation support the sustainability of protected cultivation? A life-cycle assessment and life-cycle cost analysis. Water 14. https://doi.org/10.3390/w14010006 (Switzerland).
Cuartero, J., Fernández-Muñoz, R., 1998. Tomato and salinity. Sci. Hortic. 78, 83–125. https://doi.org/10.1016/S0304-4238(98)00191-5 (Amsterdam).
Ding, X., Ni, Y., Kokot, S., 2015. NIR spectroscopy and chemometrics for the discrimination of pure, powdered, purple sweet potatoes and their samples adulterated with the white sweet potato flour. Chemom. Intell. Lab. Syst. 144, 17–23. https://doi.org/10.1016/j.chemolab.2015.03.004.
Djurović, N., Ćosić, M., Stričević, R., Savić, S., Domazet, M., 2016. Effect of irrigation regime and application of kaolin on yield, quality and water use efficiency of tomato. Sci. Hortic. 201, 271–278. https://doi.org/10.1016/j.scienta.2016.02.017 (Amsterdam).
Dobermann, A., 2007. Nutrient use efficiency – measurement and management. IFA International Workshop on Fertilizer Best Management Practices. Brussels, Belgium, pp. 1–28.
Elvanidi, A., Katsoulas, N., Ferentinos, K.P., Bartzanas, T., Kittas, C., 2018. Hyperspectral machine vision as a tool for water stress severity assessment in soilless tomato crop. Biosyst. Eng. 165, 25–35. https://doi.org/10.1016/j.biosystemseng.2017.11.002.
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 Anal. Methods 12, 914–925. https://doi.org/10.1007/s12161-018-01429-9.
Filzmoser, P., Liebmann, B., Varmuza, K., 2009. Repeated double cross validation. J. Chemom. 23, 160–171. https://doi.org/10.1002/cem.1225.
Franceschini, M.A., Gratton, E., Hueber, D.M., Fantini, S., 1999. Near-infrared absorption and scattering spectra of tissues *in vivo*. Optical Tomography and Spectroscopy of Tissue III. SPIE, pp. 526–531.
Gono, K., Igarashi, M., Obi, T., Yamaguchi, M., Ohyama, N., 2004. Enhancement of spectral change from change of cell nucleus size distribution in epithelium using multiclass linear discriminant. Optical Biopsy V. SPIE, pp. 45–53.
Hubert, M., Vanden Branden, K., 2003. Robust methods for partial least squares regression. J. Chemom. 17, 537–549. https://doi.org/10.1002/cem.822.
Jiang, Y., Chen, S., Bian, B., Li, Y., Sun, Y., Wang, X., 2021. Discrimination of tomato maturity using hyperspectral imaging combined with graph-based semi-supervised method considering class probability information. Food Anal. Methods 14, 968–983. https://doi.org/10.1007/s12161-020-01955-5.
Kennard, R.W., Stone, L.A., 1969. Computer aided design of experiments. Technometrics 11, 137–148. https://doi.org/10.1080/00401706.1969.10490666.
Kusumiyati, K., Hadiwijaya, Y., Putri, I.E., 2019. Non-destructive classification of fruits based on vis-nir spectroscopy and principal component analysis. J. Biodjati 4, 89–95. https://doi.org/10.15575/biodjati.v4i1.4389.
Liu, Y., Pu, H., Sun, D.W., 2017. Hyperspectral imaging technique for evaluating food quality and safety during various processes: a review of recent applications. Trends Food Sci. Technol. 69, 25–35. https://doi.org/10.1016/j.tifs.2017.08.013.
Lu, J., Zhou, M., Gao, Y., Jiang, H., 2018. Using hyperspectral imaging to discriminate yellow leaf curl disease in tomato leaves. Precis. Agric. 19, 379–394. https://doi.org/10.1007/s11119-017-9524-7.
Lubell, M., Hillis, V., Hoffman, M., 2011. Innovation, cooperation, and the perceived benefits and costs of sustainable agriculture practices. Ecol. Soc. 16 https://doi.org/10.5751/ES-04389-160423.
Massa, D., Magán, J.J., Montesano, F.F., Tzortzakis, N., 2020. Minimizing water and nutrient losses from soilless cropping in southern Europe. Agric. Water Manag. 241, 106395 https://doi.org/10.1016/j.agwat.2020.106395.
Montesano, F., Parente, A., Santamaria, P., 2010. Closed cycle subirrigation with low concentration nutrient solution can be used for soilless tomato production in saline conditions. Sci. Hortic. 124, 338–344. https://doi.org/10.1016/j.scienta.2010.01.017 (Amsterdam).
Montesano, F.F., Serio, F., Mininni, C., Signore, A., Parente, A., Santamaria, P., 2015. Tensiometer-based irrigation management of subirrigated soilless tomato: effects of substrate matric potential control on crop performance. Front. Plant Sci. 6, 1–11. https://doi.org/10.3389/fpls.2015.01150.
Montesano, F.F., van Iersel, M.W., Boari, F., Cantore, V., D'Amato, G., Parente, A., 2018. Sensor-based irrigation management of soilless basil using a new smart irrigation system: effects of set-point on plant physiological responses and crop performance. Agric. Water Manag. 203, 20–29. https://doi.org/10.1016/j.agwat.2018.02.019.
Morellos, A., Tziotzios, G., Orfanidou, C., Pantazi, X.E., Sarantaris, C., Maliogka, V., Alexandridis, T.K., Moshou, D., 2020. Non-destructive early detection and quantitative severity stage classification of tomato chlorosis virus (ToCV) infection in young tomato plants using Vis-NIR spectroscopy. Remote Sens. 12 https://doi.org/10.3390/rs12121920.
Naicker, R., Mutanga, O., Sibanda, M., Peerbhay, K., 2020. Discriminating Tropical Grasses Grown Under Different Nitrogen Fertilizer Regimes in KwaZulu-Natal, South Africa. Hyperspectral Remote Sensing: Theory and Applications. LTD. https://doi.org/10.1016/B978-0-08-102894-0.00014-0.
Najafian, K., Ghanbari, A., Stavness, I., Jin, L., Hassan Shirdel, G., Maleki, F., 2021. A Semi-self-supervised learning approach for wheat head detection using extremely small number of labeled samples. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1342–1351. https://doi.org/10.1109/ICCVW54120.2021.00155, 2021-Octob.
Nicolaï, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I., Lammertyn, J., 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. Postharvest Biol. Technol. 46, 99–118. https://doi.org/10.1016/j.postharvbio.2007.06.024.
Nørgaard, L., Saudland, A., Wagner, J., Nielsen, J.P., Munck, L., Engelsen, S.B., 2000. Interval partial least-squares regression (iPLS): a comparative chemometric study with an example from near-infrared spectroscopy. Appl. Spectrosc. 54, 413–419. https://doi.org/10.1366/0003702001949500.
Palumbo, M., Pace, B., Cefola, M., Montesano, F.F., Serio, F., Colelli, G., Attolico, G., 2021. Self-configuring cvs to discriminate rocket leaves according to cultivation practices and to correctly attribute visual quality level. Agronomy 11. https://doi.org/10.3390/agronomy11071353.
Pourdarbani, R., Sabzi, S., Rohban, M.H., García-Mateos, G., Arribas, J.I., 2021. Nondestructive nitrogen content estimation in tomato plant leaves by Vis-NIR hyperspectral imaging and regression data models. Appl. Opt. 60, 9560. https://doi.org/10.1364/ao.431886.
Quinet, M., Angosto, T., Yuste-Lisbona, F.J., Blanchard-Gros, R., Bigot, S., Martinez, J.P., Lutts, S., 2019. Tomato fruit development and metabolism. Front. Plant Sci. 10 https://doi.org/10.3389/fpls.2019.01554.
Rahman, A., Park, E., Bae, H., Cho, B.K., 2018. Hyperspectral imaging technique to evaluate the firmness and the sweetness index of tomatoes. Agrcultural Sci. Korean J. Agric. Sci. 45, 823–837. https://doi.org/10.7744/kjoas.20180075.
Reid, L.M., O'Donnell, C.P., Downey, G., 2006. Recent technological advances for the determination of food authenticity. Trends Food Sci. Technol. 17, 344–353. https://doi.org/10.1016/j.tifs.2006.01.006.
Rinaldi, M., Castrignanò, A., De Benedetto, D., Sollitto, D., Ruggieri, S., Garofalo, P., Santoro, F., Figorito, B., Gualano, S., Tamborrino, R., 2015. Discrimination of tomato plants under different irrigation regimes: analysis of hyperspectral sensor data. Environmetrics 26, 77–88. https://doi.org/10.1002/env.2297.
Roberts, C.A., Workman, J., Reeves, J.B., 2004. Near-Infrared Spectroscopy in Agriculture. Madison, Wisconsin, USA: American Society of Agronomy, Crop Science Society of America, Soil Science Society of America.
Savvas, D., Gianquinto, G.P., Tüzel, Y., Gruda, N., 2013. Soilless culture. Good Agricultural Practices for Greenhouse Vegetable Crops. Plant Production and Protection Paper 217. Food and Agriculture Organization of the United Nations, Rome, pp. 303–354. http://www.fao.org/3/a-i3284e.pdf.
Shrestha, S., Deleuran, L.C., Gislum, R., 2016. Classification of different tomato seed cultivars by multispectral visible-near infrared spectroscopy and chemometrics. J. Spectr. Imaging 5, 1–9. https://doi.org/10.1255/jsi.2016.a1.
Sijmen, J., 1993. SIMPLS : an alternative approach to partial least squares regression. Chemom. Intell. Lab. Syst. 18, 251–263.
Tsouvaltzis, P., Babellahi, F., Amodio, M.L., Colelli, G., 2020. Early detection of eggplant fruit stored at chilling temperature using different non-destructive optical techniques and supervised classification algorithms. Postharvest Biol. Technol. 159, 111001 https://doi.org/10.1016/j.postharvbio.2019.111001.
van de Voort, F.R., 1992. Fourier transform infrared spectroscopy applied to food analysis. Food Res. Int. 25, 397–403. https://doi.org/10.1016/0963-9969(92)90115-L.
Vidal, M., Amigo, J.M., 2012. Pre-processing of hyperspectral images. Essential steps before image analysis. Chemom. Intell. Lab. Syst. 117, 138–148. https://doi.org/10.1016/j.chemolab.2012.05.009.
Wang, Y.J., Li, T.H., Jin, G., Wei, Y.M., Li, L.Q., Kalkhajeh, Y.K., Ning, J.M., Zhang, Z.Z., 2020. Qualitative and quantitative diagnosis of nitrogen nutrition of tea plants under field condition using hyperspectral imaging coupled with chemometrics. J. Sci. Food Agric. 100, 161–167. https://doi.org/10.1002/jsfa.10009.
Wilson, R.H., Nadeau, K.P., Jaworski, F.B., Tromberg, B.J., Durkin, A.J., 2015. Review of short-wave infrared spectroscopy and imaging methods for biological tissue characterization. J. Biomed. Opt. 20, 30901.
Wu, D., Sun, D.W., 2013. Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: a review - Part II: applications. Innov. Food Sci. Emerg. Technol. 19, 15–28. https://doi.org/10.1016/j.ifset.2013.04.016.
Xia, J.A., Zhang, W.Y., Zhang, W.X., Yang, Y.W., Hu, G.Y., Ge, D.K., Liu, H., Cao, H.X., 2021. A cloud computing-based approach using the visible near-infrared spectrum to classify greenhouse tomato plants under water stress. Comput. Electron. Agric. 181, 105966 https://doi.org/10.1016/j.compag.2020.105966.
Xie, L.J., Ying, Y.B., Chen, M.L., Ying, T.J., 2010. Detection of transgenic tomato leaf with LeETR1 antisense gene by near-infrared spectroscopy. Trans. ASABE 53, 313–318.
Zapata, S., Dufour, J., 1992. Interaction HPLC. J. Food Sci. 57, 506–511.
Žibrat, U., Susič, N., Knapič, M., Širca, S., Strajnar, P., Razinger, J., Vončina, A., Urek, G., Gerič Stare, B., 2019. Pipeline for imaging, extraction, pre-processing, and

processing of time-series hyperspectral data for discriminating drought stress origin in tomatoes. MethodsX 6, 399–408. https://doi.org/10.1016/j.mex.2019.02.022.