

Article

Non-Destructive Quality Evaluation of 80 Tomato Varieties Using Vis-NIR Spectroscopy

Lilija Duckena ^{1,*} , Reinis Alksnis ², Ieva Erdberga ¹, Ina Alsina ¹, Laila Dubova ¹ and Mara Duma ³

¹ Faculty of Agriculture, Institute of Soil and Plant Science, Latvia University of Life Sciences and Technologies, 2 Liela Street, LV-3001 Jelgava, Latvia; laila.dubova@lbtu.lv (L.D.); ieva.erdberga@lbtu.lv (I.E.); ina.alsina@lbtu.lv (I.A.)

² Department of Mathematics, Faculty of Information Technologies, Latvia University of Life Sciences and Technologies, 2 Liela Street, LV-3001 Jelgava, Latvia; ralksnis@llu.lv

³ Department of Chemistry, Faculty of Food Technology, Latvia University of Life Sciences and Technologies, 2 Liela Street, LV-3001 Jelgava, Latvia; mara.duma@lbtu.lv

* Correspondence: lilija.duckena@lbtu.lv

Abstract: Traditional biochemical methods are resource- and time-consuming; therefore, there is a need for cost-effective alternatives. A spectral analysis is one of the non-destructive techniques that are more widely used for fruit quality determination; however, references are needed for traditional methods. In this study, visible and near-infrared (Vis-NIR) spectroscopy was used to analyze the internal quality attributes of tomatoes. For the first time, 80 varieties with large differences in fruit size, shape, color, and internal structure were used for an analysis. The aim of this study was to develop models suitable to predict a taste index, as well as the content of lycopene, flavonoids, β -carotene, total phenols, and dry matter of intact tomatoes based on Vis-NIR reflectance spectra. The content of phytochemicals was determined in 80 varieties of tomatoes. A total of 140 Vis-NIR reflectance spectra were obtained using the portable spectroradiometer RS-3500 (Spectral Evolution Inc.). Partial least squares regression (PLS) and multiple scatter correction (MSC) were used to develop calibration models. Our results indicated that PLS models with good prediction accuracies were obtained. The present study showed the high capability of Vis-NIR spectroscopy to determine the content of lycopene and dry matter of intact tomatoes with a determination coefficient of 0.90 for both parameters. A regression fit of $R^2 = 0.86$, $R^2 = 0.84$, $R^2 = 0.82$, and $R^2 = 0.73$ was also achieved for the taste index, flavonoids, β -carotene, and total phenols, respectively.

Keywords: *Solanum lycopersicum* L.; Vis-NIR spectroscopy; partial least squares regression (PLS); phytochemicals



Citation: Duckena, L.; Alksnis, R.; Erdberga, I.; Alsina, I.; Dubova, L.; Duma, M. Non-Destructive Quality Evaluation of 80 Tomato Varieties Using Vis-NIR Spectroscopy. *Foods* **2023**, *12*, 1990. <https://doi.org/10.3390/foods12101990>

Academic Editor: Zhengjun Qiu

Received: 17 April 2023

Revised: 7 May 2023

Accepted: 12 May 2023

Published: 14 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The tomato (*Solanum lycopersicum* L.) is an economically important vegetable crop. Its global production quantity has increased 87% in the past six decades, from 32.4 million tons in 1961 to 251.7 million tons in 2020. China, India, Turkey, the United States of America, Egypt, and Italy are leading global tomato producers [1]. The fresh tomato and tomato products are an essential source of antioxidants, and consumption of them is characterized by various health benefits, especially reduced risk of chronic diseases [2–5].

The tomato is a climacteric fruit that differs in size, shape, color, and physical and biochemical traits [6]. Breeding for high yield, long shelf life, and good external fruit quality has impacted fruit sensory quality attributes. Therefore, during the last three decades, taste and nutritional quality have received more attention in tomato breeding programs [7]. Various factors affect consumers' preferences and perceptions, including tomato fruit size and color, and most importantly, internal sensory quality attributes and the health value of varieties [8,9].

Nutritional and health-beneficial components of a tomato fruit traditionally have been determined by destructive, time-consuming, and expensive biochemical methods that use toxic solvents. Moreover, a small number of fruits are examined using destructive testing methods. Therefore, there is a need for quick and non-destructive alternatives that could provide a large-scale analysis [10,11].

Different imaging techniques can be used in fruit quality assessment, more specifically, Raman imaging, fluorescence imaging, magnetic resonance imaging, a soft X-ray, thermal imaging, infrared thermography, laser backscattering imaging, microwave imaging, and hyperspectral imaging. Among these, nuclear magnetic base techniques have high equipment prices; however, infrared base techniques are unsuitable for a large surface analysis, and a precise interpretation of results is difficult in the case of unpredictable temperatures. Moreover, permittivity and conductivity base techniques have slow detection speeds [12].

Therefore, spectral base methods combined with mathematical techniques are more dominant in fruit and vegetable quality control and prediction due to non-destructiveness and high-speed detection [13,14]. Vis-NIR spectroscopy is one of the techniques that effectively predicts multiple quality parameters in a single scan, providing a resource- and time-saving analysis [15]. This technology has been demonstrated to characterize internal quality attributes in various fruits, including apples [16], pears [17], kiwifruit [18], bananas [19], melons [20], and tomatoes [14]. Moreover, Vis-NIR spectroscopy allows for following the biochemical changes of the fruits during the ripening process, harvesting fruits at the optimum time, and assessing postharvest ripeness of climacteric fruits cost-effectively [21,22].

Clément et al. [23] reported simultaneously measuring tomatoes' quality parameters in a non-destructive manner using Vis-NIR reflectance spectroscopy and chemometrics. Results showed a well-predicted tomato color index. Borba et al. [24] used a portable NIR spectrometer (Neospectra-Module, Si-Ware Systems, Egypt) based on a micro-electro-mechanical system spectrometer at a wavelength range of 1295–2611 nm. They reported the simultaneous quantification of various quality parameters in Roma, round, grape, and cherry tomatoes. Good prediction models were obtained for soluble solids, fructose, glucose, citric acid, and ascorbic acid content, and the coefficient of determination (R^2) was higher than 0.80 for all compounds. Excellent prediction accuracy of $R^2 = 0.94$ was obtained for titratable acidity. Moreover, the high prediction accuracy of an independent prediction set measured in the field was obtained, indicating the good robustness of obtained models.

Although Vis-NIR spectroscopy has been widely used to determine fruit quality parameters, the research has been mainly focused on one or few cultivars, which have a small variation in color and size [25–27]. In our study, 80 varieties of tomato were used for the first time with large differences in fruit size, shape, color, and internal structure, which has not been previously reported.

The aim of this study was to develop models suitable to predict six internal quality traits (i.e., a taste index, and the content of lycopene, flavonoids, β -carotene, total phenols, and dry matter) of intact tomatoes based on visible and near-infrared reflectance spectra.

2. Materials and Methods

2.1. Sampling

Tomato fruits were harvested from 2018 to 2022 in different seasons from plants grown in plastic film, polycarbonate, and glass greenhouses with or without additional lighting, including store-bought fruits. Tomatoes were tested at the full ripeness stage.

Measurements were done with 80 varieties of tomatoes:

- red fruit varieties 'Amaneta F1', 'Aurea F1', 'Bellastar F1', 'Berberana F1', 'Conchita F1', 'Elegance F1', 'Gardener's Delight F1', 'Gourmandia F1', 'Lancelot F1', 'Pozano F1', 'Sunstream F1', 'Bolstar Gimli F1', 'Encore F1', 'Delizia F1', 'Belfast F1', 'Nectar F1', 'Cristal F1', 'Forticia F1', 'Nectar F1', 'Borsalina F1', 'Sandoline F1', 'Panekra F1', 'Sakura F1', 'Cocktail Crush F1', 'De Barao Krasny', 'Hector F1', 'Strabena', 'Cornue des Andes F1', and 'Brooklyn F1',

- pink ‘Cipars F1’, ‘Dimerosa F1’, ‘DRK936 F1’, ‘Fuji Pink F1’, ‘Rosa Star F1’, ‘Rosetta F1’, ‘Pink Wonder F1’, ‘Pink Oxheart’, ‘Rhianna F1’, ‘Kongo F1’, ‘Gusto Pink F1’, ‘Pink Rock F1’, ‘Honey Moon F1’, and ‘Cassarosa F1’,
- orange—‘Apressa F1’, ‘Beorange F1’, ‘Organza F1’, ‘Santorange F1’, ‘Orange Wellington F1’, ‘Taijo F1’, ‘Orange Star F1’, ‘Kinkanstar F1’, and ‘Orange Queen’,
- yellow—‘Bolzano F1’, ‘Gualdino F1’, ‘Citrusovij Sad’, ‘Taiko’, and ‘Lemon Tree’,
- light green—‘Green Envy F1’ and ‘Limetto F1’,
- multicolored—‘Ananas’, ‘Gargamel’, ‘Green Tiger’, ‘Pink Tiger’, ‘Blush Tiger’, and ‘Zebrino F1’,
- brown—‘Black Cherry F1’, ‘Bucanero F1’, ‘Chocostar F1’, ‘Chocomate F1’, ‘De Barao Chornij’, ‘Noire De Crimée’, and ‘Sacher F1’,
- and others (‘Ruskije Kolokola’, ‘Timenta’, ‘Sunstream’, ‘Toila’, ‘Pinedo’, ‘Luciestar’, ‘Delikates Moskvi’, and ‘Indigo Rose’) were included.

The average tomato mass varied from 8.6 g (‘Bellastar F1’) to 212.0 g (‘Pink Wonder F1’).

2.2. Non-Destructive Methods and Reference Analysis

Reflectance spectra of tomato fruits were obtained with a portable spectroradiometer RS-3500 (Spectral Evolution Inc., Haverhill, MA, USA). The RS-3500 Remote Sensing Bundle includes a fiber optic field portable remote sensing spectroradiometer, a spectral range of 350–2500 nm, and a resolution of 1 nm. The spectroradiometer RS-3500 was tuned and calibrated with a built-in white light reference before data acquisition. A total of 12–16 reflectance spectra for each variety were obtained. Spectra were represented as averages for each variety. After recording the spectra, the same fruits were used immediately for the reference measurement of quality parameters.

The extraction and determination of phytochemicals were performed as reported in the study by Alsina et al. [28]. The following internal quality attributes were determined (Table 1): the content of lycopene (expressed as mg 100 g⁻¹), β-carotene (mg 100 g⁻¹), total phenols (gallic acid equivalent (GAE) per 100 g of fresh tomato mass), total flavonoids (amount of quercetin equivalents (QE) per 100 g of fresh tomato mass), dry matter (mg g⁻¹), soluble solids (°Brix), titratable acidity (mg of citric acid per 100 g of fresh tomato weight), and a taste index [29].

Table 1. Methods used to determine tomatoes’ quality parameters.

Parameters	Methods of Analyses	Reference
Lycopene	Solvent—tetrahydrofuran, spectrophotometry	[28]
β-carotene	Solvent—tetrahydrofuran, spectrophotometry	[28]
Total phenols	Solvent—methanol, spectrophotometry	[30]
Total flavonoids	Solvent—ethanol, spectrophotometry	[31]
Dry matter	Gravimetry	[28]
Soluble solids	Refractometry	[28]
Titratable acidity	Alkaline titration	[28]
Taste index	Calculations using soluble solids and titratable acidity	[29]

2.3. Mathematical Methods and Data Analysis

In total, 158 tomato samples were used in this study. To obtain models that could predict the internal attributes listed above, partial least squares regression (PLS) was implemented, which is a common approach in the related literature [32,33]. It is applied when the number of predictors is larger than observations, as was the case in this study since there were 2250 predictors— a reflectance percentage for each wavelength. The goal of PLS regression is to compress the main information from all predictors into some relatively small amount of PLS components that could be used for the prediction [34]. The approach is similar to principal component analysis (PCA); however, this algorithm searches for components that would be correlated as much as possible with the dependent variable.

Before model training, pre-processing transformations were applied to improve the prediction performance. Some common methods, such as centering and smoothing as well as taking derivatives of the reflectance spectra, were investigated. Finally, the optimal performance was reached when multiple scatter correction (MSC) was applied, which reduced the light scatter that was present in the obtained spectrum data [35]. Afterwards, the Savitzky–Golay filter with third-order polynomials and window size 5 was used for smoothing, thus improving the model performance even further. Original and smoothed spectra are shown in Figure 1.

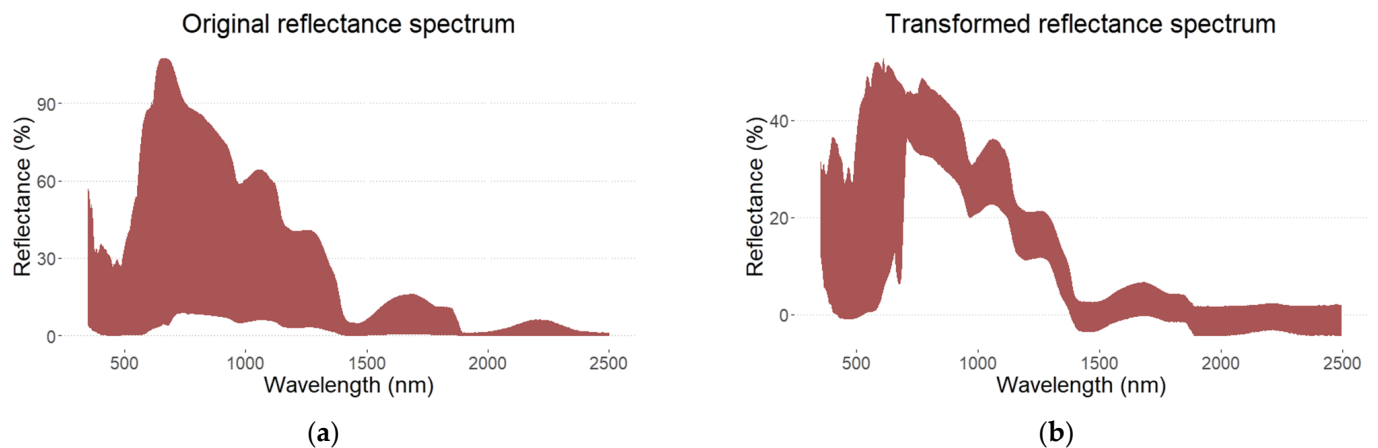


Figure 1. Original (a) and multiplicative scatter corrected and smoothed (b) visible and near-infrared (Vis-NIR) reflectance spectrum of intact tomato at a wavelength range of 350–2500 nm.

After the pre-processing stage, models were trained on 70% of the samples using k-fold cross-validation with 10 folds repeated five times. As hyperparameters, the number of PLS components was searched in the range from 1 to 40. The best-performing model was chosen based on the root mean squared error (RMSE), which is calculated as the square root of the mean of the sum of squared differences between predicted and measured values. Finally, the obtained model was validated on the hold-out sample of 30% of tomatoes. The coefficient of determination R^2 was reported for calibration and validation sets as well as the root mean squared error (RMSE), mean absolute error (MAE), and ratio of the standard error of prediction to standard deviation (RPD). All reported analyses were conducted using the software R (version 4.2.2; R Core Team, 2022) on Windows 10 × 64 (build 19045). More specifically, the package *prospectr* (version 0.2.6; Stevens A, Ramirez-Lopez L, 2022) (version 0.2.4) was used in the pre-processing stage and *caret* (version 6.0.93; Kuhn M, 2022) was used to train and validate the PLS models.

3. Results

3.1. Lycopene

The measured level of lycopene content in the sample was 3.56 ± 2.34 (sample mean \pm sample standard deviation) $\text{mg } 100 \text{ g}^{-1}$ with a total range from 0 to $8.38 \text{ mg } 100 \text{ g}^{-1}$. The sample distribution was somewhat skewed and bimodal (skewness was -0.06 and kurtosis 2.03) with no apparent outliers.

The final calibration model with the root mean squared error (RMSE) of $0.72 \text{ mg } 100 \text{ g}^{-1}$ showed a good fit with $R^2 = 0.90$. When applied to a hold-out validation set, $R^2 = 0.85$ was obtained with the RMSE of $0.95 \text{ mg } 100 \text{ g}^{-1}$. The ratio of the standard error of prediction to standard deviation (RPD) for the calibration sample was 3.16, indicating a good prediction accuracy, while RPD for the validation sample was 2.59. Hence, although the model performed worse on the new samples, its predictions were still satisfactory (Figure 2). The optimal number of PLS components for the model was 10, which is relatively small compared to other models in this study.

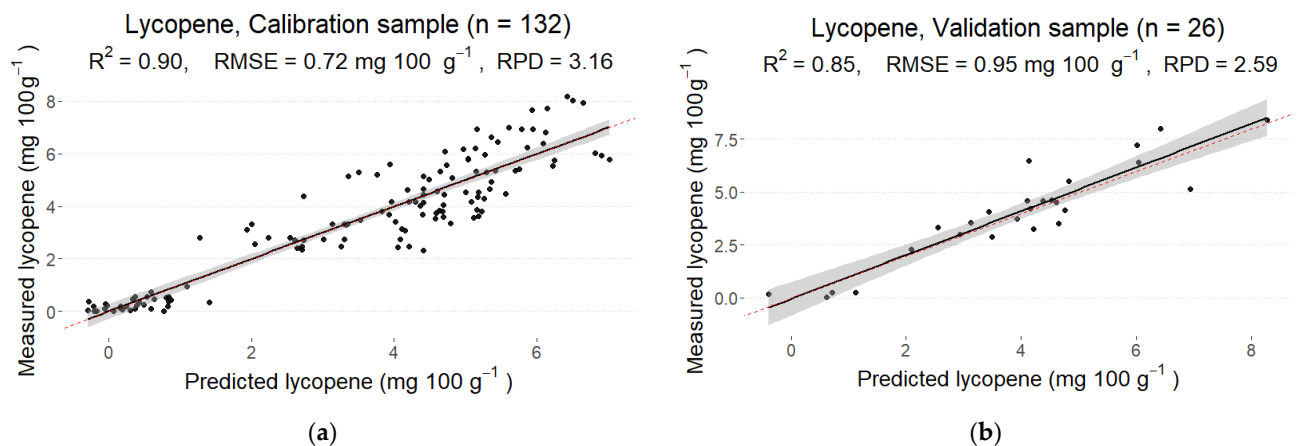


Figure 2. Predicted and measured values of lycopene in the calibration (a) and validation (b) sets derived from the best partial least squares regression (PLS) model from the reflectance spectrum with MSC pre-processing.

3.2. β -Carotene

The level of measured content of β -carotene in the sample was $5.68 \pm 2.60 \text{ mg } 100 \text{ g}^{-1}$, and it varied in the range from $0.04 \text{ mg } 100 \text{ g}^{-1}$ to $11.73 \text{ mg } 100 \text{ g}^{-1}$. The sample distribution was negatively skewed and bimodal (sample skewness was -0.18 and kurtosis 2.68).

The final calibration model with a root mean squared error of $RMSE = 1.07 \text{ mg } 100 \text{ g}^{-1}$ showed a good linear fit ($R^2 = 0.82$). The coefficient of determination for the validation sample was $R^2 = 0.85$ with $RMSE = 1.01 \text{ mg } 100 \text{ g}^{-1}$. Thus, the obtained model performed equally well for calibration and validation samples (Figure 3). The ratios of the standard error of prediction to standard deviation (RPD) for calibration and validation samples were 2.09 and 2.37 , respectively, thus showing poor prediction accuracy. The number of PLS components for the final calibration model was 10 .

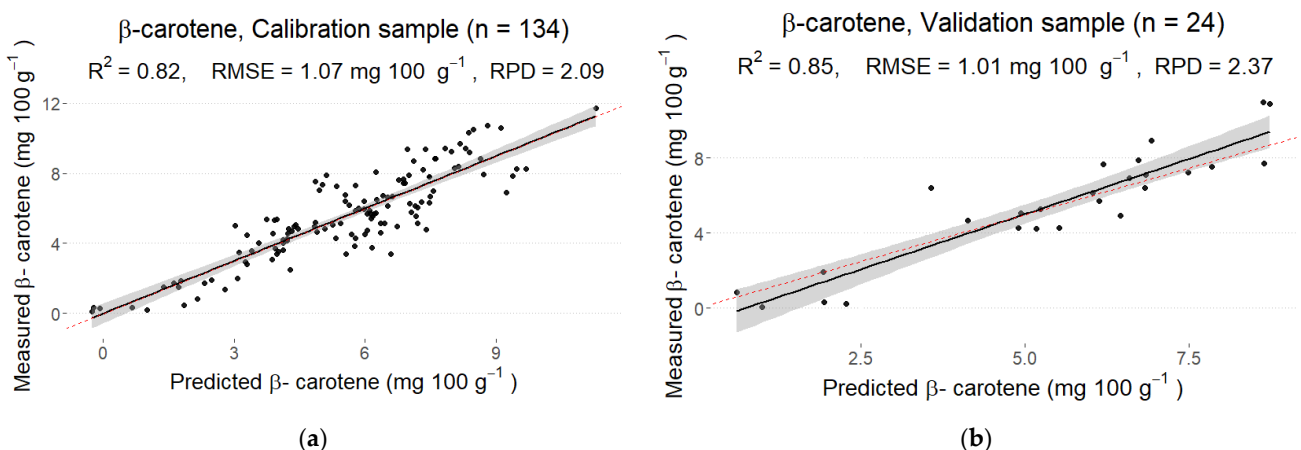


Figure 3. Predicted and measured values of β -carotene in the calibration (a) and validation (b) sets derived from the best partial least squares regression (PLS) model from the reflectance spectrum with MSC pre-processing.

3.3. Total Phenols

The level of the measured total phenol content in the sample was $34.68 \pm 8.67 \text{ mg GAE } 100 \text{ g}^{-1}$, ranging from 16.77 to $60.91 \text{ mg GAE } 100 \text{ g}^{-1}$. The values of skewness and kurtosis were 0.68 and 3.06 , respectively, indicating a positive skew in the sample.

The determination coefficient for the calibration sample was $R^2 = 0.73$ with $RMSE = 4.57 \text{ mg GAE } 100 \text{ g}^{-1}$, indicating a somewhat good fit; however, $R^2 = 0.50$ with $RMSE = 6.33 \text{ mg GAE } 100 \text{ g}^{-1}$ was obtained when applied to the validation sample, show-

ing a poor fit for samples unseen in the calibration process (Figure 4). The ratios of the standard error of prediction to standard deviation (RPD) for calibration and validation samples were 1.91 and 1.41, respectively, thus showing a very poor prediction accuracy. The number of PLS components was 21.

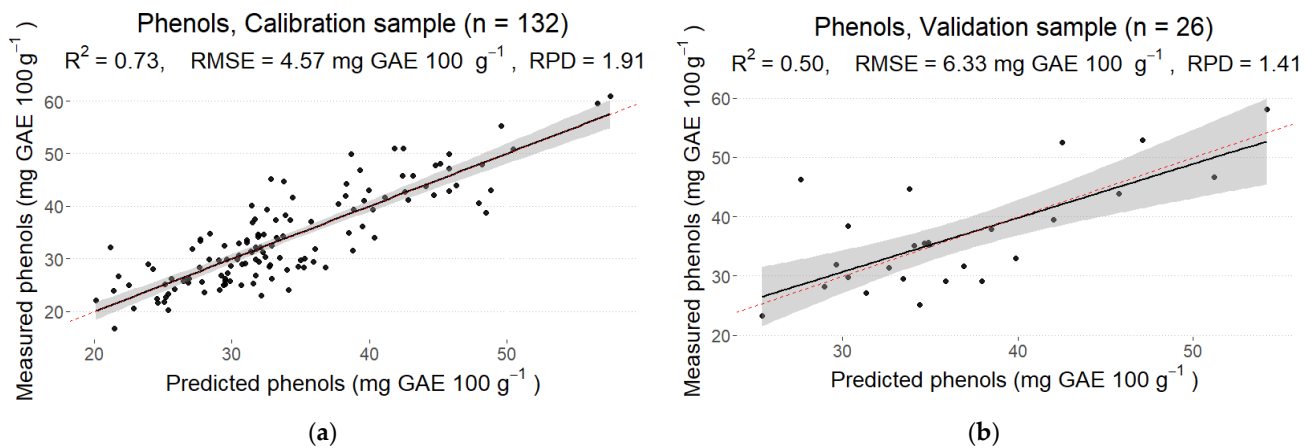


Figure 4. Predicted and measured values of phenols in the calibration (a) and prediction (b) sets with partial least squares regression (PLS) models.

3.4. Flavonoids

The level of measured flavonoid content in the sample was 4.73 ± 2.04 mg QE 100 g⁻¹, and it varied from 1.09 to 11.02 mg QE 100 g⁻¹. The sample skewness and kurtosis were 0.38 and 2.98, respectively, indicating a positive skewness with some relatively large sample values; however, not necessarily considered outliers.

The determination coefficient for the calibration sample was $R^2 = 0.84$ with $RMSE = 1.08$ mg QE 100 g⁻¹ while $R^2 = 0.80$ with $RMSE = 1.31$ mg QE 100 g⁻¹ was obtained for the validation sample, indicating a good fit for both calibration and validation samples (Figure 5). The ratios of the standard error of prediction to standard deviation (RPD) for calibration and validation samples were 2.51 and 2.15, respectively, indicating a fair prediction accuracy. The optimal number of PLS components in the final model was 21.

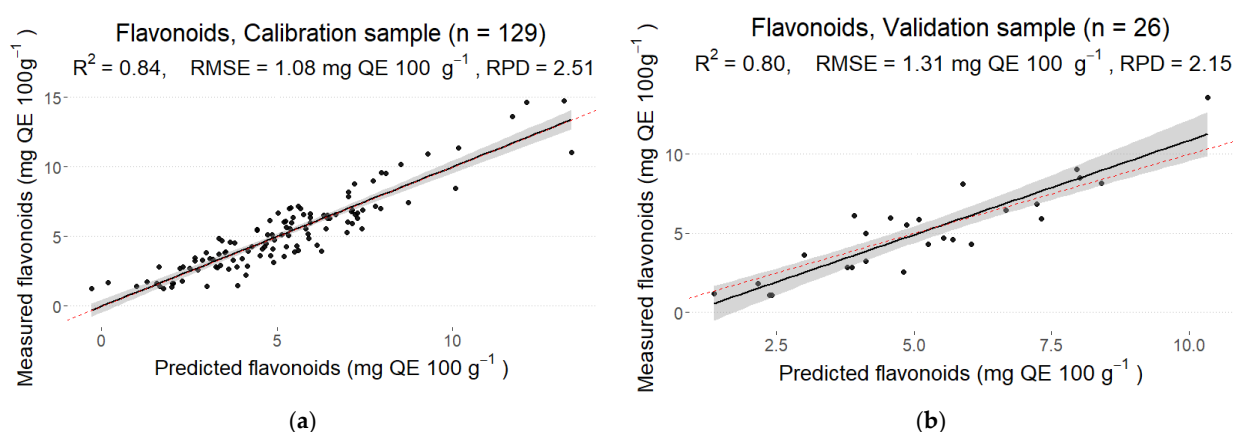


Figure 5. Predicted and measured values of flavonoids in the calibration (a) and prediction (b) sets with PLS models.

3.5. Taste Index

The taste index is a dimensionless quantity and the measured level of this index in the sample was 1.15 ± 0.16 with a total range between 0.86 and 1.52. The sample skewness and kurtosis were 0.37 and 2.24, respectively, indicating a positive skewness in the sample; however, no outliers were detected.

The final calibration model with $RMSE = 0.06$ resulted in $R^2 = 0.86$. The determination coefficient for the validation sample was $R^2 = 0.77$ with $RMSE = 0.10$, showing a worse but still acceptable fit compared to the calibration sample (Figure 6). According to the RPD metric, a fair model ($RPD = 2.68$) was obtained for the calibration sample, but a poor prediction model ($RPD = 1.93$) was obtained for the validation sample. The optimal number of PLS components was 25—the largest of all the models considered in this study.

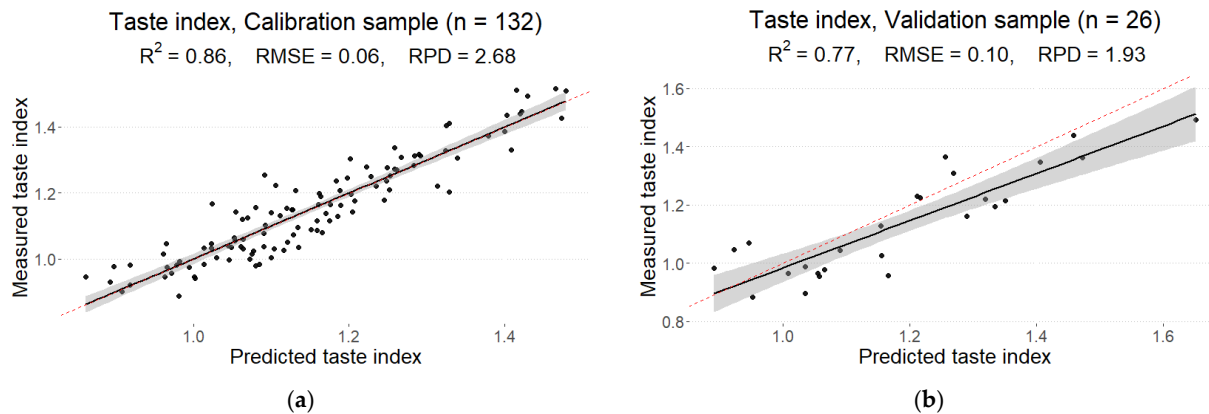


Figure 6. Measured and predicted values of taste index in the calibration (a) and prediction (b) sets derived from the best partial least squares regression (PLS) model from the reflectance spectrum with MSC pre-processing.

3.6. Dry Matter

The level of measured dry matter content in the sample was $7.12 \pm 1.82\%$, and it varied in the range between 4.55% and 13.15%. The sample skewness and kurtosis were 0.88 and 3.07, respectively, showing a positive skewness in the sample. Some more extreme values were observed compared to the rest of the sample; hence, care must be taken so that the validation sample does not contain them all.

The final calibration model with $RMSE = 0.42\%$ showed $R^2 = 0.90$, while the determination coefficient for the validation sample was $R^2 = 0.83$ with $RMSE = 0.98\%$. Thus, a good performance was observed for both calibration and validation samples. According to the ratios of the standard error of prediction to standard deviation (RPD) metric, a good model ($RPD = 3.23$) was obtained for the calibration sample, but a poor prediction model ($RPD = 2.03$) was obtained for the validation sample. The optimal number of PLS components in the final model was 24 (Figure 7).

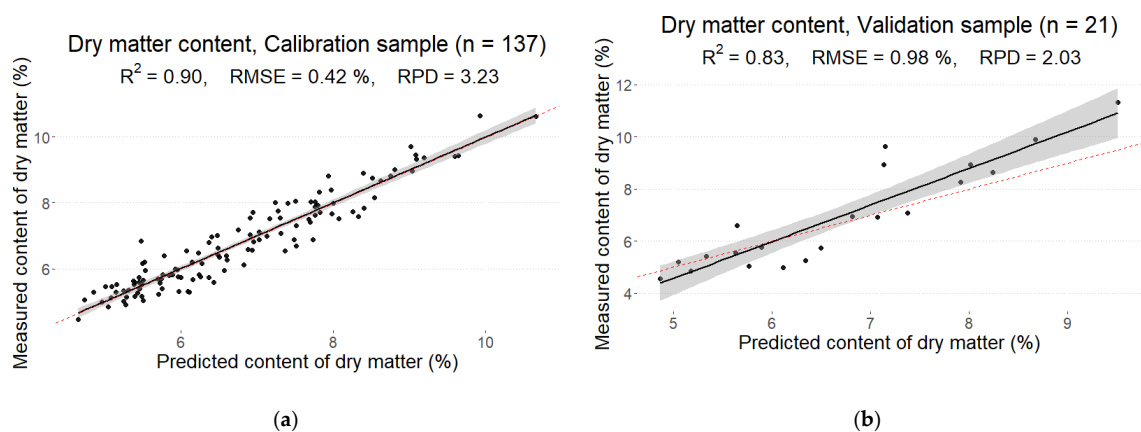


Figure 7. Measured and predicted values of dry matter in the calibration (a) and prediction (b) sets derived from the best partial least squares regression (PLS) model from the reflectance spectrum with MSC pre-processing.

4. Discussion

Visible and near-infrared (Vis-NIR) spectroscopy is one of the methods used to identify various compounds in fruits and estimate their quality in a non-destructive manner [23,36]. The possibility of determining tomato quality attributes using Vis-NIR spectroscopy has been investigated in various studies [23,33,37,38]. Although this method can be applied in various sensing modes, the reflectance technique is one of the best ways to determine internal fruit quality attributes [39]. In our study, intact tomato fruits' Vis-NIR reflectance spectra were used to develop partial least squares regression (PLS) models and predict the taste index, as well as the content of lycopene, flavonoids, β -carotene, dry matter, and total phenols.

In the present study, the linear regression model of lycopene ($R^2 = 0.90$) produced a better fit than the content of β -carotene ($R^2 = 0.82$). Similar results were reported by Tilahun et al. [40]. They obtained good lycopene determination accuracy in the calibration ($R^2 = 0.89$) and prediction ($R^2 = 0.85$) sets; however, the content of β -carotene had lower prediction accuracy in both the calibration ($R^2 = 0.88$) and prediction ($R^2 = 0.77$) sets. Moreover, linear regression results for lycopene content obtained by Li et al. [41] also demonstrated similar performance to that observed in our study. They obtained a model with high accuracy using the genetic algorithm partial least square regression method (the correlation coefficient was 0.907, and the root mean squared errors of cross-validation and prediction were 8.76 and 8.93, respectively).

Other similar studies showed poorer performance for the PLS linear regression of carotenes. Baranska et al. [42] reported lower prediction accuracy of lycopene and β -carotene using NIR spectroscopy (regression coefficient was 0.85 and 0.80, respectively). However, the study performed by Goisser et al. [27] showed more accurate performance for the PLS linear regression of lycopene content. They used three portable NIR spectrometers to predict lycopene content in the intact tomato 'Avalantino'. A good linear fit was obtained using an F-750 Produce Quality meter (Portland, OH, USA) followed by H-100F (Incheon, Republic of Korea) (coefficient of determination was 0.96 and 0.95, respectively). For CiO™ (Hod HaSharon, Israel), good performance was observed, and the coefficient of determination was 0.92. In the present study, fruits were tested only at the full ripeness stage. However, all ripening stages (green, breaker, turning, light red, and red) tested by Goisser et al. [27] resulted in better prediction capabilities, indicating the importance of data variation for the accurate prediction of lycopene.

In our study, good flavonoid content determination accuracy was observed in calibration ($R^2 = 0.84$) and validation sets ($R^2 = 0.80$); however, the content of total phenols had lower prediction accuracy in both calibration ($R^2 = 0.73$) and validation ($R^2 = 0.50$) sets. A superior prediction performance was obtained in other studies. Alenazi et al. [26] used a near infra-red (NIR) spectrometer (F-750, Produce Quality Mater, Felix Instruments, Camas, WA, USA) at a wavelength range of 285–1200 nm to evaluate the tomato 'Red Rose' quality traits. They reported a higher coefficient of determination for total phenolic content in calibration ($R^2 = 0.966$) and validation sets ($R^2 = 0.834$). Moreover, higher prediction accuracy was observed for total phenolic content in calibration ($R^2 = 0.984$) and validation sets ($R^2 = 0.790$). In their study, five different ripening stages of tomatoes were used for the analysis, indicating the importance of data variation to obtain more precise linear regression models.

Applying linear regression to quantify dry matter content resulted in a fit of $R^2 = 0.90$ and $R^2 = 0.83$ for calibration and validation sets, respectively. Similar results were obtained in previous studies. Radzevičius et al. [43] used a NIR Case NCS001A (Sacmi Imola S.C., Bologna, Italy) spectrometer at a wavelength range of 600–1000 nm to evaluate the main quality attributes in tomatoes. They observed a high correlation for dry matter (regression coefficient was 0.91), indicating good prediction accuracy. Acharya et al. [44] reported high prediction accuracy of dry matter using a portable "Nirvana" SWNIR spectrometer (the coefficient of determination for the prediction set ranged from 0.86 to 0.92). In comparison to our study, the poorer prediction accuracy of dry matter content was obtained by Arruda de Brito et al. [33]. They used a portable Vis-NIR spectrometer (Felix Instruments, model F-

750, Camas, WA, USA) and reported a coefficient of determination for dry matter prediction of $R^2 = 0.62$ and 0.59 for calibration and validation sets, respectively.

In the presented study, a good prediction model was obtained for the taste index ($R^2 = 0.86$ and 0.77 for the calibration and validation sets, respectively). This index is calculated based on soluble solids content (SSC) and titratable acidity (TA), as indicated by Soare et al. [29]. Oliveira et al. [32] used a multi-purpose analyzer (MPA) spectrometer (Bruker Optics) in a reflectance mode to determine internal sensory attributes in various fruit species, including tomatoes. Excellent results were obtained for the apricot, and the regression fit of ($R^2 = 0.95$ and 0.93) was obtained for TA and SSC, respectively. In contrast, the notably poorer fit of ($R^2 = 0.51$ and 0.52) was observed for TA and SSC in tomatoes, respectively. Such differences were mainly attributed to the structural differences between the fruits. A superior regression fit for taste-related compounds was obtained in the study by Najjar and Abu-Khalaf [25]. They used Vis-NIR spectroscopy with a USB 2000+ miniature fiber optic spectrometer (Ocean Optics, Orlando, FL, USA) at a wavelength range of 550–1100 nm to model quality parameters in three varieties of tomato (i.e., 'Ekram', 'Harver', and 'Izmer') at a light red stage. A good prediction of soluble solids content (SSC), titratable acidity (TA), and taste (SSC/TA) was obtained (the coefficient of determination for the validation set was 0.98 , 0.89 , and 0.94 , respectively).

Nevertheless, obtained results in our study are satisfactory since the error range in predicting the taste index's reference values was lower than 1% (root mean squared error of calibration and validation was RMSE = 0.05% and 0.10% , respectively). Furthermore, 80 varieties that vary in color, shape, size, and internal structure were used, which could affect prediction performance as indicated by previous studies. Research performed by de Oliveira et al. [32] proved that internal fruit heterogeneity is one of the major limiting factors for the precise prediction of taste-related compounds in tomatoes. Ibáñez et al. [45] reported that calibrations are necessary for each assay to obtain NIR spectra partial least square regression (PLS) models with a precise prediction of tomato taste-related compounds. Borba et al. [24] reported that separated calibration models are necessary for large and small varieties of tomatoes to obtain models with a higher prediction performance of soluble solids content and titratable acidity.

According to Williams et al. [46], the determination coefficient of $R^2 = 0.92$ or higher indicates that the obtained model is suitable for fruit quality assessment; however, lower values ($R^2 = 0.83$ to 0.90) indicate that the obtained model can be used with caution for most applications, including research. Moreover, the ratios of the standard error of prediction (SEP) to standard deviation (SD) or RPD values must be higher than three to use the models for quality control [46]. Therefore, in our study, good results were obtained for lycopene and dry matter content, and the ratios of the standard error of prediction to standard deviation were RPD = 3.23 and RPD = 3.16 , respectively. However, there is a need for further research to increase the robustness of the obtained models for the taste index, flavonoids, β -carotene, and total phenols for the application in fruit quality control.

5. Conclusions

The results of the present study confirmed that visible and near-infrared (Vis-NIR) spectroscopy is a precise non-destructive method to determine tomato fruit quality in a single scan. The portable Vis-NIR spectroradiometer RS-3500 (Spectral Evolution Inc., Haverhill, MA, USA) allowed the quick and simultaneous quantification of multiple internal quality traits of the tomato. The models suitable to predict the taste index and the lycopene, flavonoid, β -carotene, total phenol, and dry matter content were developed using the partial least squares regression (PLS) algorithm. Although 80 varieties of tomato were used that have large differences in fruit size, shape, color, and internal structure, a strong correlation was observed between reflectance spectra data and the reference values obtained using a biochemical analysis for lycopene and dry matter. The results of the present study are in-line with previous research, indicating a high potential for non-destructive measurements in fruit quality control. Fast and cost-effective measurements are essential

to ensure the high quality of fruits desired by breeders, producers, and consumers. Therefore, further research is required to increase the robustness of the current models for the prediction of a taste index, flavonoids, β -carotene, and total phenols.

Author Contributions: Conceptualization, I.A. and L.D. (Laila Dubova); methodology, I.A. and M.D.; software, R.A.; validation, R.A. and I.A.; formal analysis, R.A.; investigation, M.D. and I.E.; resources, I.E. and I.A.; data curation, R.A.; data interpretation, L.D. (Lilija Duckena); writing—original draft preparation, L.D. (Lilija Duckena); writing—review and editing, L.D. (Lilija Duckena), R.A. and I.A.; visualization, R.A.; supervision, I.A.; project administration, I.A. and L.D. (Laila Dubova); funding acquisition, I.A. and L.D. (Laila Dubova). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Latvian Rural Development Program 2014–2020 Cooperation, call 16.1 project Nr. 19-00-A01612-000010 Investigation of innovative solutions and new method development for efficiency and quality increase in Latvian greenhouse sector (IRIS).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. FAOSTAT. Available online: <https://www.fao.org/faostat/en/#data/QCL> (accessed on 1 February 2023).
2. Bauerfeind, J.; Hintze, V.; Kschonsek, J.; Killenberg, M.; Böhm, V. Use of photochemiluminescence for the determination of antioxidant activities of carotenoids and antioxidant capacities of selected tomato products. *J. Agric. Food Chem.* **2014**, *62*, 7452–7459. [[CrossRef](#)]
3. Chaudhary, P.; Sharma, A.; Singh, B.; Nagpal, A.K. Bioactivities of phytochemicals present in tomato. *J. Food Sci. Technol.* **2018**, *55*, 2833–2849. [[CrossRef](#)]
4. Puah, B.P.; Jalil, J.; Attiq, A.; Kamisah, Y. New Insights into Molecular Mechanism behind Anti-Cancer Activities of Lycopene. *Molecules* **2021**, *26*, 3888. [[CrossRef](#)]
5. Tilesi, F.; Lombardi, A.; Mazzucato, A. Scientometric and Methodological Analysis of the Recent Literature on the Health-Related Effects of Tomato and Tomato Products. *Foods* **2021**, *10*, 1905. [[CrossRef](#)]
6. Ercolano, M.R.; Carli, P.; Soria, A.; Cascone, A.; Fogliano, V.; Frusciante, L.; Barone, A. Biochemical, sensorial and genomic profiling of traditional Italian tomato varieties. *Euphytica* **2008**, *164*, 571–582. [[CrossRef](#)]
7. Bai, Y.; Lindhout, P. Domestication and breeding of tomatoes: What have we gained and what can we gain in the future? *Ann. Bot.* **2007**, *100*, 1085–1094. [[CrossRef](#)]
8. Bertin, N.; Génard, M. Tomato quality as influenced by preharvest factors. *Sci. Hortic.* **2018**, *233*, 264–276. [[CrossRef](#)]
9. Sinesio, F.; Cammareri, M.; Cottet, V.; Fontanet, L.; Jost, M.; Moneta, E.; Palombieri, S.; Peparai, M.; del Castillo, R.R.; Civitelli, E.S.; et al. Sensory Traits and Consumer's Perceived Quality of Traditional and Modern Fresh Market Tomato Varieties: A Study in Three European Countries. *Foods* **2021**, *10*, 2521. [[CrossRef](#)]
10. Davis, A.R.; Fish, W.W.; Perkins-Weazie, P. A rapid spectrophotometric method for analyzing lycopene content in tomato and tomato products. *Postharvest. Biol. Technol.* **2003**, *28*, 425–430. [[CrossRef](#)]
11. Pedro, A.M.K.; Ferreira, M.M.C. Nondestructive Determination of Solids and Carotenoids in Tomato Products by Near-Infrared Spectroscopy and Multivariate Calibration. *Anal. Chem.* **2005**, *77*, 2505–2511. [[CrossRef](#)]
12. Hussain, A.; Pu, H.; Sun, D.W. Innovative nondestructive imaging techniques for ripening and maturity of fruits—a review of recent applications. *Trends Food Sci. Technol.* **2018**, *72*, 144–152. [[CrossRef](#)]
13. Wang, H.; Peng, J.; Xie, C.; Bao, Y.; He, Y. Fruit Quality Evaluation Using Spectroscopy Technology: A Review. *Sensors* **2015**, *15*, 11889–11927. [[CrossRef](#)]
14. Zhang, D.; Yang, Y.; Chen, G.; Tian, X.; Wang, Z.; Fan, S.; Xin, Z. Nondestructive evaluation of soluble solids content in tomato with different stage by using Vis/NIR technology and multivariate algorithms. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2021**, *248*, 119139. [[CrossRef](#)]
15. Cortés, V.; Blasco, J.; Aleixos, N.; Cubero, S.; Talens, P. Monitoring strategies for quality control of agricultural products using visible and near-infrared spectroscopy: A review. *Trends Food Sci. Technol.* **2019**, *85*, 138–148. [[CrossRef](#)]
16. Ignat, T.; Lurie, S.; Nyasordzi, J.; Ostrovsky, V.; Egozi, H.; Hoffman, A.; Friedman, H.; Weksler, A.; Schmilovitch, Z. Forecast of Apple Internal Quality Indices at Harvest and During Storage by VIS-NIR Spectroscopy. *Food Bioprocess Technol.* **2014**, *7*, 2951–2961. [[CrossRef](#)]
17. Li, J.; Huang, W.; Zhao, C.; Zhang, B. A comparative study for the quantitative determination of soluble solids content, pH and firmness of pears by Vis/NIR spectroscopy. *J. Food Eng.* **2013**, *116*, 324–332. [[CrossRef](#)]

18. Moghimi, A.; Aghkhani, M.H.; Sazgarnia, A.; Sarmad, M. Vis/NIR spectroscopy and chemometrics for the prediction of soluble solids content and acidity (pH) of kiwifruit. *Biosyst. Eng.* **2010**, *106*, 295–302. [[CrossRef](#)]
19. Zude, M. Non-destructive prediction of banana fruit quality using VIS/NIR spectroscopy. *Fruits* **2003**, *58*, 135–142. [[CrossRef](#)]
20. Hu, R.; Zhang, L.; Yu, Z.; Zhai, Z.; Zhang, R. Optimization of soluble solids content prediction models in ‘Hami’ melons by means of Vis-NIR spectroscopy and chemometric tools. *Infrared Phys. Technol.* **2019**, *102*, 102999. [[CrossRef](#)]
21. Costa, G.; Noferini, M.; Fiori, G.; Torrigiani, P. Use of Vis/NIR Spectroscopy to Assess Fruit Ripening Stage and Improve Management in Post-Harvest Chain. *Fresh Prod.* **2009**, *1*, 35–41.
22. Melado-Herreros, A.; Nieto-Ortega, S.; Olabarrieta, I.; Gutiérrez, M.; Villar, A.; Zufia, J.; Gorretta, N.; Roger, J.M. Postharvest ripeness assessment of ‘Hass’ avocado based on development of a new ripening index and Vis-NIR spectroscopy. *Postharvest. Biol. Technol.* **2021**, *181*, 111683. [[CrossRef](#)]
23. Clément, A.; Dorais, M.; Vernon, M. Nondestructive Measurement of Fresh Tomato Lycopene Content and Other Physicochemical Characteristics Using Visible–NIR Spectroscopy. *J. Agric. Food Chem.* **2008**, *56*, 9813–9818. [[CrossRef](#)]
24. Borba, K.R.; Aykas, D.P.; Milani, M.I.; Colnago, L.A.; Ferreira, M.D.; Rodriguez-Saona, L.E. Portable near Infrared Spectroscopy as a Tool for Fresh Tomato Quality Control Analysis in the Field. *Appl. Sci.* **2021**, *11*, 3209. [[CrossRef](#)]
25. Najjar, K.; Abu-Khalaf, N. Non-destructive quality measurement for three varieties of tomato using Vis/NIR spectroscopy. *Sustainability* **2021**, *13*, 10747. [[CrossRef](#)]
26. Alenazi, M.M.; Shafiq, M.; Alsadon, A.A.; Alhelal, I.M.; Alhamdan, A.M.; Solieman, T.H.I.; Ibrahim, A.A.; Shady, M.R.; Saad, M.A.O. Non-destructive assessment of flesh firmness and dietary antioxidants of greenhouse-grown tomato (*Solanum lycopersicum* L.) at different fruit maturity stages. *Saudi J. Biol. Sci.* **2020**, *27*, 2839–2846. [[CrossRef](#)] [[PubMed](#)]
27. Goisser, S.; Wittmann, S.; Fernandes, M.; Mempel, H.; Ulrichs, C. Comparison of colorimeter and different portable food-scanners for non-destructive prediction of lycopene content in tomato fruit. *Postharvest. Biol. Technol.* **2020**, 167. [[CrossRef](#)]
28. Alsina, I.; Erdberga, I.; Duma, M.; Alksnis, R.; Dubova, L. Changes in Greenhouse Grown Tomatoes Metabolite Content Depending on Supplemental Light Quality. *Front. Nutr.* **2022**, *9*, 830186. [[CrossRef](#)]
29. Soare, R.; Maria, D.; Alexandru-Ioan, A.; Soare, M. The evolution of some nutritional parameters of the tomato fruit during the harvesting stages. *Hortic. Sci.* **2019**, *46*, 132–137. [[CrossRef](#)]
30. Singleton, V.L.; Orthofer, R.; Lamuela-Raventós, R.M. Analysis of total phenols and other oxidation substrates and antioxidants by means of Folin-Ciocalteu reagent. *Methods Enzymol.* **1999**, *299*, 152–178. [[CrossRef](#)]
31. Kim, D.; Jeond, S.; Lee, C. Antioxidant capacity of phenolic phytochemicals from various cultivars of plums. *Food Chem.* **2003**, *81*, 321–326. [[CrossRef](#)]
32. Oliveira, G.A.; Bureau, S.; Renard, C.M.G.C.; Pereira-Netto, A.B.; de Castilhos, F. Comparison of NIRS approach for prediction of internal quality traits in three fruit species. *Food Chem.* **2014**, *143*, 223–230. [[CrossRef](#)] [[PubMed](#)]
33. Aruda de Brito, A.; Campos, F.; dos Reis Nascimento, A.; Damiani, C.; Alves da Silva, F.; de Almeida Teixeira, G.H.; Cunha Júnior, L.C. Non-destructive determination of color, titratable acidity, and dry matter in intact tomatoes using a portable Vis-NIR spectrometer. *J. Food Compos. Anal.* **2022**, *107*, 104288. [[CrossRef](#)]
34. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [[CrossRef](#)]
35. Windig, W.; Shaver, J.; Bro, R. Loopy MSC: A simple way to improve multiplicative scatter correction. *Appl. Spectrosc.* **2008**, *62*, 1153–1159. [[CrossRef](#)]
36. Lan, W.; Jaillais, B.; Chen, S.; Renard, C.M.G.C.; Leca, A.; Bureau, S. Fruit variability impacts puree quality: Assessment on individually processed apples using the visible and near infrared spectroscopy. *Food Chem.* **2022**, *390*, 133088. [[CrossRef](#)]
37. Alsina, I.; Dubova, L.; Duma, M.; Erdberga, I.; Avotínš, A.; Rakutko, S. Comparison of lycopene and β -carotene content in tomatoes determined with chemical and non-destructive methods. *Agron. Res.* **2019**, *17*, 343–348. [[CrossRef](#)]
38. Li, X.; Tsuta, M.; Hayakawa, F.; Nakano, Y.; Kazami, Y.; Ikehata, A. Estimating the sensory qualities of tomatoes using visible and near-infrared spectroscopy and interpretation based on gas chromatography–mass spectrometry metabolomics. *Food Chem.* **2021**, *343*, 128470. [[CrossRef](#)]
39. Khodabakhshian, R.; Emadi, B.; Khojastehpour, M.; Golzarian, M.R. A comparative study of reflectance and transmittance modes of Vis/NIR spectroscopy used in determining internal quality attributes in pomegranate fruits. *J. Food Meas. Charact.* **2019**, *13*, 3130–3139. [[CrossRef](#)]
40. Tilahun, S.; Seo, M.H.; Hwang, I.G.; Kim, S.H.; Choi, H.R.; Jeong, C.S. Prediction of lycopene and β -carotene in tomatoes by portable chroma-meter and VIS/NIR spectra. *Postharvest. Biol. Technol.* **2018**, *136*, 50–56. [[CrossRef](#)]
41. Li, T.; Zhong, C.; Lou, W.; Wei, M.; Hou, J. Optimization of Characteristic Wavelengths in Prediction of Lycopene in Tomatoes Using Near-Infrared Spectroscopy. *J. Food Process. Eng.* **2017**, *40*, e12266. [[CrossRef](#)]
42. Baranska, M.; Schütze, W.; Schulz, H. Determination of lycopene and beta-carotene content in tomato fruits and related products: Comparison of FT-Raman, ATR-IR, and NIR spectroscopy. *Anal. Chem.* **2006**, *78*, 8456–8461. [[CrossRef](#)] [[PubMed](#)]
43. Radzevičius, A.; Viškelis, J.; Karklelienė, R.; Juškevičienė, D.; Viškelis, P. Determination of tomato quality attributes using near infrared spectroscopy and reference analysis. *Zemdirb.-Agric.* **2016**, *103*, 91–98. [[CrossRef](#)]
44. Acharya, U.K.; Subedi, P.P.; Walsh, K.B. Robustness of Tomato Quality Evaluation Using a Portable Vis-SWNIRS for Dry Matter and Colour. *Int. J. Anal. Chem.* **2017**, *2017*, 1–8. [[CrossRef](#)] [[PubMed](#)]

45. Ibáñez, G.; Cebolla-Cornejo, J.; Martí, R.; Roselló, S.; Valcárcel, M. Non-destructive determination of taste-related compounds in tomato using NIR spectra. *J. Food Eng.* **2019**, *263*, 237–242. [[CrossRef](#)]
46. Williams, P.; Manley, M.; Antoniszyn, J. Statistical terms for evaluation of accuracy and prediction. In *Near Infrared Technology: Getting the Best out of Light*; African Sun Media: Stellenbosch, South Africa, 2019; pp. 69–103.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.