Journal of Agriculture and Food Research 21 (2025) 101974

Contents lists available at ScienceDirect

Journal of Agriculture and Food Research

journal homepage: www.sciencedirect.com/journal/journal-of-agriculture-and-food-research

Application of handheld near infrared device for in-plant quality assessment of tomato paste samples

Shreya Madhav Nuguri [a], Silvia de Lamo Castellvi [a,b], Didem Peren Aykas [c], Mustafa Mortas [d], Luis Rodriguez-Saona [a,*]

[a] Department of Food Science and Technology, The Ohio State University, Parker Food Science and Technology Building, 2015 Fyffe Road, Columbus, OH, USA
[b] Departament d'Enginyeria Química, Universitat Rovira i Virgili, Av. Països Catalans 26, 43007, Tarragona, Spain
[c] Department of Food Engineering, Faculty of Engineering, Adnan Menderes University, Aydin, 09100, Turkey
[d] EngineeringFaculty, Food Engineering Department, Ondokuz Mayıs University, Samsun, Turkey

ARTICLE INFO

Keywords:
Tomato paste
Handheld NIR
Rheological properties
Color
Titratable acidity

ABSTRACT

The current study aims to evaluate the applicability of a novel handheld NIR scanner and an NIR sensor in transflectance mode for non-destructive and rapid *in-situ* analysis of important quality parameters in tomato paste samples. The predictive variables included five key quality traits-natural tomato soluble solids (NTSS), titratable acidity (TA), Bostwick consistency, serum viscosity, and the a/b ratio. Reference levels of these parameters were determined using conventional analytical techniques. A total of 224 tomato pastes samples, supplied by a tomato processing industry from 2015 to 2020 and in 2022, were considered for this study. The samples provided a unique range of concentration for individual parameter (NTSS = 25.7–37.4 ^{0}Brix, TA = 1.01–1.95 %, Bostwick consistency = 1.0–10.4 cm, serum viscosity = 190.0–452.0 cSt and a/b ratio = 1.9–4.4). A transflectance approach was employed to collect NIR spectra using a 0.50 mm pathlength reflector. Partial least squares regression (PLSR) was used to analyze the multivariate data and develop predictive models for the quality traits. Both, the NIR scanner (1350 nm–2500 nm, R^2pre = 0.83 to 0.98 and RMSEP = 0.03 to 0.63) and the NIR sensor (1100 nm–2550 nm, R^2pre = 0.85 to 0.98 and RMSEP = 0.05 to 0.43) exhibited comparable performance with good figures of merit (RER >9 and SEP/SECV = 0.8–1.1), emphasizing their suitability for quality assessment of tomato paste samples. Additionally, model transfer from the NIR scanner to the NIR sensor was investigated, and the performance of the resulting multivariate calibration transfer was compared with that of the NIR sensor. Models generated using the NIR sensor performed better than the calibration transfer models, however, the validation performance of the calibration transfer models (R^2pre = 0.55 to 0.94 and RMSEP = 0.06 to 0.68) suggested their suitability for screening calibration. Overall, the results underscore the reliability of miniaturized NIR systems to streamline quality analysis of tomato paste samples in the tomato processing industry, providing a cost-effective, high-throughput and multicomponent monitoring technique.

1. Introduction

Tomato (*Lycopersicon esculentum)*, being rich in beneficial nutrients, is among the most widely consumed and largest produced fruit in the world. Tomatoes are frequently consumed in the form of processed products such as canned vegetables, juice, soup, ketchup, and sauces. California produces around 10–12 million metric tons of processed tomatoes each year, contributing to approximately 96 % of the total US processing tomato production [1]. Most of these tomatoes are thermally processed and concentrated to form pastes, which are stored year-round and diluted to produce ketchup, sauces, and other value-added products [2]. Monitoring quality traits is crucial for the tomato industry, and it is widely known that tomato paste is the main ingredient in many tomato products, therefore, quality assessment is essential during the production [3,4].

For tomato paste production, tomatoes are washed, sorted according to USDA guidelines [5], and homogenized, followed by pulping, filtering, evaporation, and pasteurization [6]. Homogenization activates

* Corresponding author.
E-mail addresses: nuguri.2@osu.edu (S.M. Nuguri), delamocastellvi.1@osu.edu (S. de Lamo Castellvi), didem.cinkilic@adu.edu.tr (D.P. Aykas), mustafa.mortas@omu.edu.tr (M. Mortas), rodriguez-saona.1@osu.edu (L. Rodriguez-Saona).

https://doi.org/10.1016/j.jafr.2025.101974
Received 11 March 2025; Received in revised form 29 April 2025; Accepted 30 April 2025
Available online 6 May 2025

enzymes, specifically pectin methylesterase (PME) and polygalacturonase (PG), which causes the rapid breakdown of pectin. These homogenized tomatoes undergo thermal treatment, referred to as the "break". In the hot break process, they are heated to a temperature of at least 82 °C (mostly >90 °C), while in the cold break process, the temperature is lower (~65 °C). The activity of pectic enzymes is enhanced at temperatures between 60-66 °C but is gradually inactivated at about 82 °C [7]. The choice of hot or cold treatment is crucial for processors, as it determines important quality parameters, including titratable acidity (TA), pH, Bostwick consistency, serum viscosity, citric acid, and acetic acid, among others. In the cold break, juice retains fresh flavor and color, while high temperatures in the hot break juice accelerate color loss [8]. High temperatures lead to the inactivation of pectic enzymes and efficient extraction of pectic substances in hot break juice, resulting in a homogeneous paste [8]. Hydrolysis of pectin in cold break juice causes a loss of pectin, reducing its viscosity. Additionally, the methyl ester hydrolysis of pectin releases polygalacturonic acid residues, increasing the TA and pH of the tomato juice [2]. Furthermore, the consistency of the paste varies during evaporation; the end point of this step is determined based on soluble solids [9].

The rheological properties of the paste are important for formulating products like ketchup and sauces, as they determine the consistency of both intermediate and final products [9,10]. The Bostwick consistency of the whole juice is influenced by insoluble materials, while serum viscosity results from soluble solutes in the tomato juice [10]. Soluble solids and TA play significant roles in influencing the flavor of the product. Generally, a high sugar content is associated with a high soluble solid content. The major sugars present in tomato paste include glucose and fructose, along with smaller quantities of raffinose, arabinose, xylose, and galactose [11]. Soluble solids, TA, and pH are influenced by several factors, such as cultivar, crop yield, horticulture practices, seasonal variations, ripening stage at the time of harvesting, processing conditions (hot vs. cold break processing), and growing location [2,12–14]. Tomatoes have sufficient acidity to maintain a pH below 4.6, so they do not require the severe thermal treatment needed to ensure food safety for low-acid food [2]. Citric acid is a major contributor to the acid content, along with ascorbic acid, malic acid, and glutamic acid, and greatly affects the titratable acidity (TA) of tomato products [2]. A colorimetric measurement parameter, the a/b ratio, which indicates the relationship between redness and yellowness in the sample, was used to assess the redness of the tomato paste product [15]. Conventional laboratory-based techniques used to analyze these important parameters are time-consuming, require specialized expertise, involve hazardous solvents, and are not suited for routine in-line quality analysis [3]. The significance of the proposed research lies in providing the tomato processing industry with predictive models using a novel, field deployable Near infrared (NIR) device that enables rapid and reliable quantitative analysis of multiple quality variables simultaneously with minimal sample preparation and personnel training.

Our lab has been intensively working on characterizing industrially important quality traits in processed tomato products using handheld and benchtop mid-infrared (MIR) spectrometers [3,16–18]. In this study, we evaluated two different near-infrared (NIR) spectrometers (a NIR scanner and a NIR sensor) using a transflectance approach to determine quality traits of tomato paste samples, including Brix, TA, Bostwick consistency, serum viscosity, and the a/b ratio. Compared to MIR, NIR offers a greater penetration depth (2 mm-to 3 mm) [19] and a more miniaturized system (reduced size, wight, and cost), making it convenient for on-site analysis. NIR also represents overtones and combinations of the major X-H bonds, enabling the analysis of organic traits [20]. The NIR scanner and the NIR sensor encompass the long-wave NIR region (LWNIR, 1100–2500 nm). Short-wave NIR region (SWNIR, 700–1100 nm) focuses on the 3rd or 4th overtones of vibrations, whereas LWNIR addresses the 1st or 2nd overtones and combination vibrations, which exhibit stronger absorptivity [21].

Furthermore, calibration transform was performed on data obtained using the NIR scanner (primary instrument) to another NIR sensor (secondary instrument). Subsequently, the effectiveness of the calibration transfer was examined by comparing its performance to that of models developed using the NIR sensor (secondary instrument). Generally, calibration model developed with one NIR spectrometer introduce large predictive errors when applied to a different NIR spectrometer, which cannot be addressed using spectral pre-treatment methods [22]. Calibration transfer employs chemometric techniques to correct for instrumental differences, thereby enabling the transfer of multivariate calibration models to the secondary instrument with minimal recalibration [23]. Thus, our aim was to evaluate the applicability of a novel handheld NIR scanner and an NIR sensor in transflectance mode-using spectra acquired with a 0.50 mm pathlength reflector, for non-destructive and rapid *in-situ* analysis of key quality parameters (Brix, TA, Bostwick consistency, serum viscosity and the a/b ratio) in tomato paste samples.

2. Materials and methods

2.1. Tomato paste samples

A total of 224 tomato paste samples were supplied by a tomato processing industry in California, USA, over the years 2015–2020 and 2022. The tomato variety used to obtain the paste samples was mainly Roma. The major differences among the samples were primarily contributed by variations in processing parameters including cold break and hot break. Each processing line was designed to produce tomato paste tailored to specific tomato products or end-users. Based on their requirements, the processing parameters were adjusted which led to the variability in the quality attributes of the tomato paste. The tomato paste samples were thermally processed with no additives and aseptically packaged.

2.2. Reference analysis

The reference values for the parameters of interest were provided by the quality assurance department of the tomato processor. These included natural tomato soluble solids (NTSS), TA, Bostwick consistency, serum viscosity and the a/b ratio. The samples were shipped overnight and were stored in a freezer (−20 °C) until analyzed by NIR systems.

For NTSS measurements, the tomato paste was centrifuged at 140,000 rpm for 10 min using a Sorvall-Discovery M150 SE microcentrifuge (Thermo Fisher, Waltham, MA, USA). The NTSS readings were obtained from the supernatant using a temperature-controlled refractometer (RX 5000i ATAGO, Bellevue, WA, USA) at 25 °C. Paste samples were diluted with deionized water to a Brix value of 12^0 and homogenized using a stomacher (Seward™ Stomacher™ Model 400, West Sussex, UK). The diluted blend was placed in a Bostwick consistometer, and the consistency was determined by the distance covered in centimeters after 30 s. Similarly, the paste aliquots were diluted to 6^0 Brix and centrifuged at 2000 rpm for 15 min (Thermo Fisher Scientific, Waltham, MA, USA) to prepare samples for serum viscosity determination. The supernatant was filtered through Whatman No.1 filter paper (Whatman PLC, Maidstone, UK), and the filtrate was introduced into a Cannon-Fenske viscometer. The viscometer was placed at 30 ± 1 °C for 5 min to equilibrate the temperature. Serum viscosity, recorded in centistokes (cSt), was calculated using the absolute viscosity of distilled water under the same conditions. TA and a/b ratio were determined using manual titration and a lab colorimeter respectively. All reference values were obtained in duplicates.

2.3. Near-infrared analysis

NIR analysis was performed using two different systems: the NeoSpectra Scanner developed by Si-Ware (Cairo, Egypt) and a compact FT-

NIR sensor developed by Hamamatsu (Hamamatsu Photonics Co., Hamamatsu, Japan). The Neospectra scanner (Si-Ware, Cairo, Egypt) was connected to a mobile application via Bluetooth, while the spectrum was measured using the NIR sensor by connecting it to a laptop via USB.

Fig. 1a shows the schematic of the scanner, which features seven tungsten light sources arranged concavely to effectively illuminate the sample, along with a centrally located detector that captures the scattered light. The scanner is equipped with a single-chip Michelson interferometer and integrates an uncooled indium-gallium-arsenide (InGaAs) photodetector. It has a spectral window of approximately **0.01 m** in diameter, covering a wavelength range of 1350–2500 nm with a resolution of 16 nm at 1550 nm and a signal-to-noise ratio (S/N) of 2000:1 at 2350 nm.

The NIR sensor (Fig. 1b) consists of a light source that integrates three tungsten lamps and a bundle of optical fibers (Diffuse reflection light source L164621-01; Hamamatsu Photonics Co., Hamamatsu, Japan). The lamps irradiate the sample, and the diffused and reflected light is guided into the optical fibers for spectroscopic measurements [24] (Fig. 1b). The transmitted photons enter the Michelson optical interferometer (C15511-01 FTIR; Hamamatsu Photonics Co., Hamamatsu, Japan), which is equipped with a beam splitter, a fixed mirror, a moving mirror and an InGaAs PIN photodiode detector. The typical spectral response range extends from 1100 to 2550 nm, with a spectral resolution of 5.7 nm ($\lambda = 1533$ nm) and a S/N of 10000.

A transflectance accessory kit comprising Petri dish bases and aluminum diffuse reflectors (0.50 mm pathlength) from PerkinElmer (Waltham, MA, USA) was used to introduce the sample to the NIR spectrometers. A small amount of paste was placed on the reflecting surface of the transflector and pressed against the Petri dish base. The excess paste from the sides of the transflector was cleaned using cotton swabs (Q-tips, Missouri, US). Care was taken to ensure that the space within the transflector was homogeneously covered with the sample,

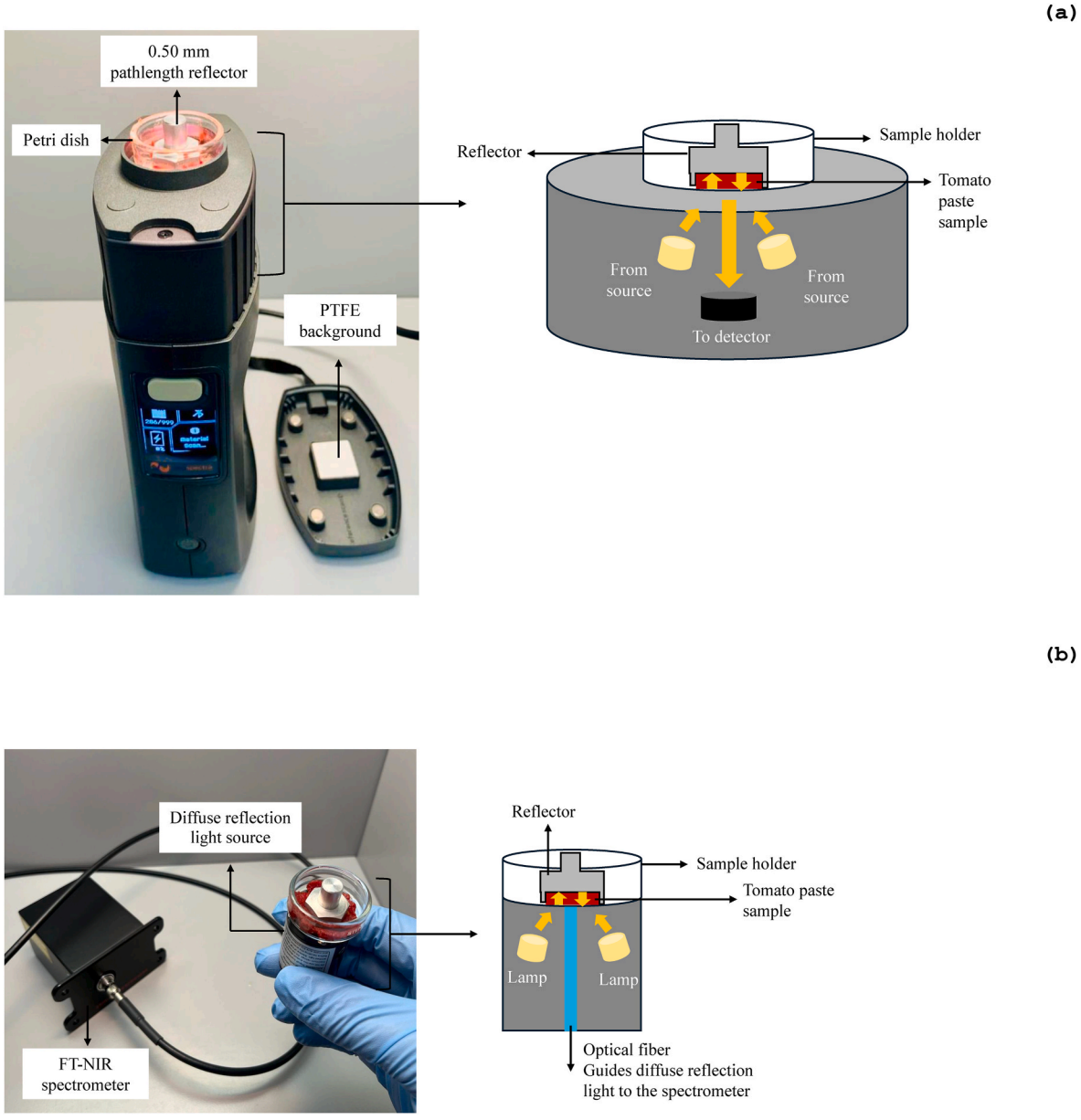

Fig. 1. (a) Image illustrating spectral acquisition using the NIR scanner and the schematics of its transflectance approach. (b) Image illustrating spectral acquisition using the NIR sensor and the schematics of its transflectance approach. The transflector is set on top of the sample, which is placed in a Petri dish (sample holder) positioned on top of the device. The yellow arrow indicates the light path. The incident light travels through the sample enclosed within the transflector, reflects off the reflecting surface, and transmits back through the sample.

with no gaps or seeds present. The Petri dish with the sample and the transflector was positioned on the optical window of the NIR spectrometers for spectral acquisition. In transflectance mode, the light travels through the sample enclosed within the transflector, reflects off the reflecting surface, and transmits back through the sample before reaching the detector (Fig. 1). Thus, the pathlength is roughly twice the width of the transflector. Background spectra were acquired using a polytetrafluoroethylene (PTFE) reflector. Spectra for each sample were collected in duplicates and averaged during data analysis.

2.4. Regression analysis

Predictive models for the reference values were developed using the partial least squares regression (PLSR) algorithm in Pirouette 4.5 software (Pirouette® version 4.0, Infometrix Inc., Woodville, WA, USA). Briefly, PLSR extracts important variables (latent variables) in spectral data (X) that best predicts the response parameter (Y), developing a correlation between two matrices comprising of the independent variables (X) and the dependent variables (Y). PLSR is a versatile regression technique when the predicting datasets are multi-collinear, noisy, have high dimensionality and less dependent variables (Y) than spectral variables (X) [25].

NIR data with their corresponding reference values were imported into the chemometrics software, Pirouette® version 5.0 (Woodville, WA, USA). The dataset was divided into a training set, which comprised 80 % of the dataset used to train the calibration model, and a test set representing the remaining 20 %, which was used as an unseen dataset to validate the trained models. "Random" algorithm provided by the Pirouette software was applied to select the calibration and validation datasets [26]. The algorithm uses a random process to generate the subsets. The 80:20 ratio was used as it provides more data points for model training, which helps capture most of the variation in the response variable. The test set comprising 20 % of the dataset, ensured that enough datapoints were included for reliable validation.

Pre-processing including normalization, Savitzky-Golay (SG) second derivative and mean centering were applied to ensure that the data better aligned with Beer's law [27], and additionally enhancing outliers' detection [28]. Leave-one-out cross validation was used to internally validate the calibration models. It is a resampling method where each data point is left out and predicted by the model generated using the remaining datapoints estimating the calibration performance. A number of factors providing a low variance were selected to avoid overfitting. Correlation coefficient of cross-validation (Rcv), correlation coefficient of external prediction (Rpre), root mean square error of cross validation (SECV), root mean square error of prediction (SEP), range error ratio (RER), and the ratio of SEP to SECV (SEP/SECV) were used to evaluate regression performance.

2.5. Calibration transfer

Based on the established literature [29,30], an informative subset of 15 samples containing key variations in the reference value, particularly the Brix (NTSS) was selected to generate the transformation matrix. The spectral data of these 15 samples, collected on both the NIR systems were uploaded to the calibration transfer interface of PLS_toolbox (SOLO ©, Eigenvector Research Inc., Manson, Washington). Piecewise direct standardization (PDS), one of the most popular and well-established methods for standardizing spectral response was applied [31]. In PDS, a transformation matrix is constructed using a moving window procedure that corrects for non-linearities in the data, providing a significant advantage [31]. It develops a structured transformation matrix based on a few samples measured on both the instruments to correct for instrumental differences, making the spectra transferrable and eliminating the need for full recalibration on the new instrument [29,30,32]. The generated transformation matrix was then applied to the spectral data from the NIR scanner, transforming it into the corresponding spectra as if they had been collected using the NIR sensor.

3. Results and discussion

3.1. Reference values in tomato paste samples

Table 1 summarizes the reference data for the four major quality parameters (NTSS, citric acid, Bostwick consistency, and a/b ratio) in tomato paste samples. The number of samples (n) slightly varied among the parameters of NTSS (n = 224), TA (n = 220) and Bostwick consistency (n = 216) due to differences in reference value availability of these parameters. For serum viscosity, only the hot break samples (n = 130) were used to build the model, as the cold break samples did not have serum viscosity values, and the samples with a/b ratio information were limited (n = 109). The wide range of reference values can be attributed to the inclusion of samples from multiple years and processing facilities. Furthermore, the high variation in the reported data could result from the analysis performed at different locations using distinct instruments measured by various individuals [3]. The US Food and Drug Administration requires the food to contain not less than 24 % of NTSS to be labeled as tomato paste [33]. The NTSS of all samples complied with the regulation, exceeding 24 ^{0}Brix (25.7 ^{0}Brix to 37.4 ^{0}Brix) with a mean value 29.5 ± 2.8 ^{0}Brix. The values of Bostwick consistency ranged from 1.0 to 10.4 cm (mean = 4.0 ± 2.2 cm) and the values of serum viscosity ranged from 190.0 to 452.0 cSt (mean = 313.8 ± 57.5 cSt). The logarithmic transformed data of serum viscosity (range = 2.3–2.7 and mean = 2.5 ± 0.1) was considered while developing its regression model to linearize the relationship, improve the normal distribution, and reduce the influence of outliers [34]. The data of rheological parameters obtained in this study aligned with the reference values obtained previously [3]. The a/b ratio relates to the darkness of the product, it ranged from 1.9 to 4.4 (mean = 2.1 ± 0.2) consistent with the results reported by other literature sources [35–37].

3.2. Spectral characterization

Fig. 2 shows a typical NIR spectra obtained from NIR scanner (Fig. 2a) and NIR sensor (Fig. 2b) system collected using 0.50 mm pathlength transflector. The broad IR bands in the NIR spectra can be attributed to the high-water content in the sample and overlapping of numerous bands. The region around 2000, 1800 and 1430 nm can be correlated with the first overtone of O-H, first overtone of C-H and O-H stretching respectively. The high-water content of the sample causes saturation of the O-H first overtone at 1980–1900 nm leading to a peculiar pattern [38]. The absorption characteristics of tomato paste samples resembled the spectral pattern of fruit juice samples including apple juice [38], bayberry juice [39], strawberry juice [40] and commercial fruit juices [41] in the 1350–2500 nm spectral region. The spectra collected from NIR sensor (Fig. 2b) extended in the SWNIR region capturing information from carbohydrates (combination of CH vibrations) [42,43].

Table 1
Summary of reference values for the five major quality parameters (NTSS, TA, Bostwick consistency, serum viscosity, and a/b ratio) in tomato paste samples.

Parameter	N	Range	mean ± SD
NTSS (^{0}Brix)	224	25.7–37.4	29.5 ± 2.8
TA (%)	220	1.01–1.95	1.4 ± 0.2
Bostwick consistency (cm)	216	1.0–10.4	4.0 ± 2.2
Serum viscosity (cSt)	130	190.0–452.0	313.8 ± 57.5
a/b ratio	109	1.9–4.4	2.1 ± 0.2

N: Number of samples, SD: Standard deviation.

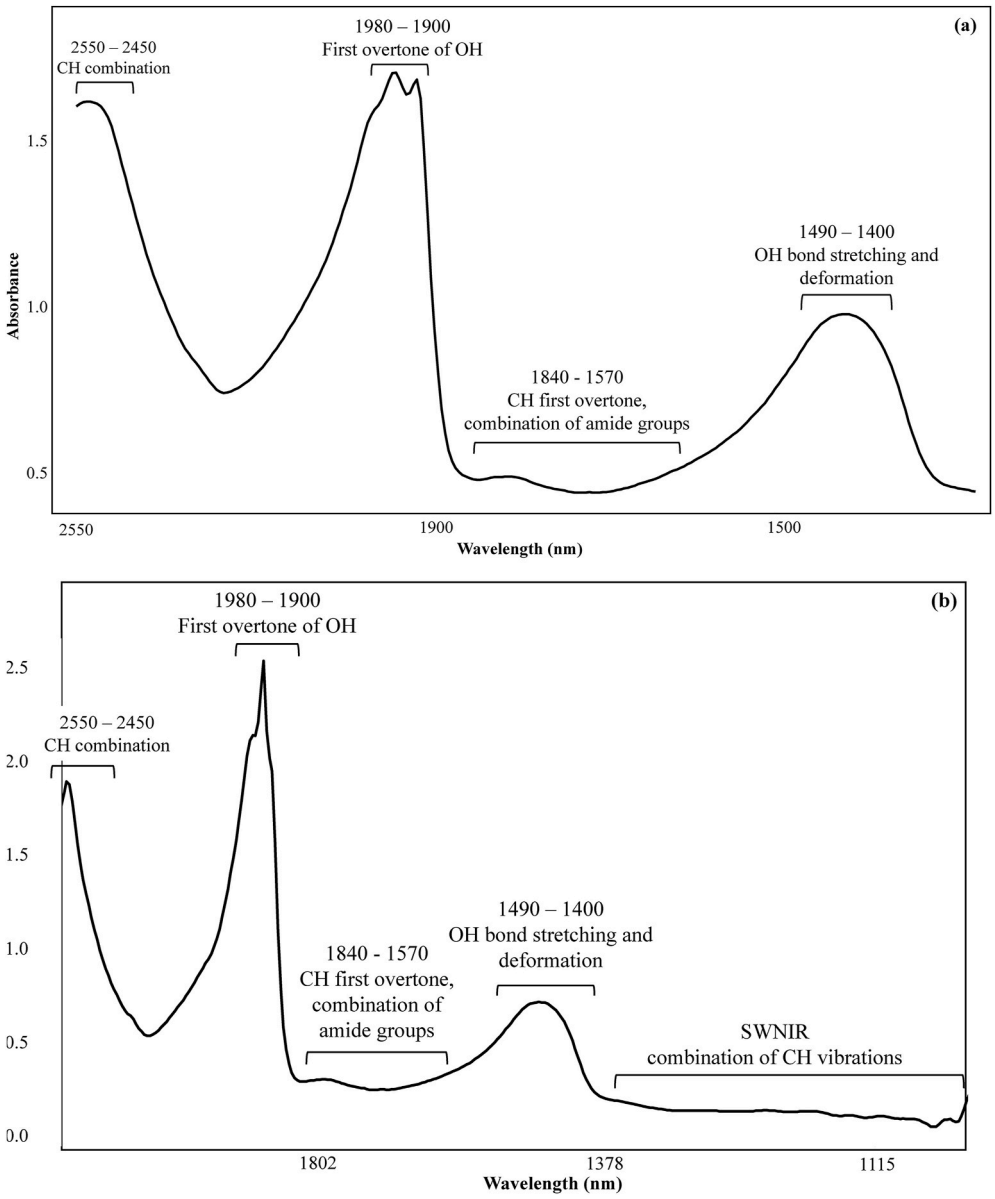

Fig. 2. Typical NIR spectrum obtained from NIR scanner (a) and NIR sensor (b) collected using 0.50 mm pathlength transflector.

3.3. Regression models

The absorption bands in the NIR region are widely overlapping, characterized by broad, weak and non-specific spectra [44], requiring multivariate calibration methods, such as PLSR to extract relevant variations and relate them to chemical data [45]. Additionally, parameters such as mathematical pre-treatments, transformations, wavelength selection, and the number of factors help minimize the complicating effects of NIR spectra and correct for non-linearity [46]. The parameters (pre-treatments, wavelength selection, and the number of factors) were optimized to achieve a high correlation coefficient, low error and a good fit along the regression line. During model development, the spectral data were normalized, SG second derivatized (11-point window), and mean centered. Normalization reduces signal fluctuations and enhances the predictive ability of the regression models [47]. Vector length normalization was applied and scaled to 100. The SG second derivative deconvolutes important peaks, enhances resolution, and improves accuracy of quantitative analysis. A second derivative with an 11-point window size and a polynomial order of 2 was utilized in this study [48]. Mean-centering subtracts the mean value of an independent variable from each of its data point, evaluating inter-variable relationships by considering differences relative to the mean [26,49]. The transformed spectra required fewer factors and yielded better performance than the models developed using the original spectra (Table S1). The removal of non-informative wavelengths has been shown to increase predictive ability and reduce model complexity [50]. The loadings plot of the corresponding parameter was used to identify the important wavelength range. The peculiar pattern around 1960–1860 nm was excluded from both spectra to avoid random variations caused by the saturation effect of the water. Outliers are abnormal data points that strongly influence the model, and their detection is crucial for generating a high-quality multivariate calibration model [51]. Observations having extreme values in the studentized versus leverage plot were identified as outliers and eliminated from the regression analysis.

3.3.1. Regression models developed using NIR systems

Tables 2 and 3 present the mean and standard derivation of calibration and validation datasets, along with the performance (number of factors, R^2cv, RMSECV, R^2pre, RMSEP, RPD, RER, SEP/SECV) of each attribute, for NIR scanner and NIR sensor respectively. The NIR scanner

Table 2
Chemical composition of the calibration and validation sets, and statistical performance of the regression models obtained using the portable NIR scanner.

NIR scanner													
Parameter	**F**	**Spectral region**	**Calibration**					**Validation**					
			N	**Mean**	**SD**	**R2cv**	**SECV**	**N**	**Mean**	**SD**	**R2pre**	**SEP**	**RER**	**SEP/SECV**
NTSS (0Brix)	5	2550–1950 1730–1350	177	29.6	2.9	0.98	0.44	45	29.0	2.3	0.97	0.39	26.4	0.9
TA (%)	5	2550–1950 1550–1350	168	1.4	0.2	0.79	0.09	43	1.4	0.2	0.82	0.08	9.5	0.9
Bostwick consistency (cm)	6	2550–1950	164	3.8	1.9	0.91	0.57	41	3.8	2.0	0.91	0.63	10.4	1.1
Serum viscosity (cSt)	4	2550–2070	100	311.1	55.5	0.67	33.28	26	306.9	51.9	0.71	30.53	8.7	0.9
a/b ratio	7	2500–1950 1870–1350	87	2.1	0.1	0.91	0.03	22	2.1	0.1	0.94	0.03	13.6	1.0

N: Number of samples, F: Factors, SD: Standard deviation, R^2cv: correlation coefficient of cross validation, SECV: Root mean square error of cross-validation, R^2pre: correlation coefficient of prediction, SEP: Root mean square error of prediction, RER: Range error ratio.

Table 3
Chemical composition of the calibration and validation sets, and statistical performance of the regression models obtained using the NIR sensor.

NIR sensor													
Parameter	**F**	**Spectral region**	**Calibration**					**Validation**					
			N	**Mean**	**SD**	**R2cv**	**SECV**	**N**	**Mean**	**SD**	**R2pre**	**SEP**	**RER**	**SEP/SECV**
NTSS (0Brix)	7	2500–2100 1800–1340	89	29.2	2.7	0.97	0.44	27	29.4	2.6	0.98	0.36	29.9	0.8
TA (%)	6	2450–2050 1740–1350	85	1.4	0.2	0.90	0.06	24	1.4	0.2	0.94	0.06	13.2	0.9
Bostwick consistency (cm)	4	2450–2050 1740–1375	76	3.4	1.3	0.85	0.50	21	3.4	1.3	0.90	0.43	11.7	0.9
a/b ratio	8	2400–2000	85	2.1	0.1	0.80	0.05	25	2.1	0.1	0.84	0.05	9.2	1.0

N: Number of samples, F: Factors, D: Standard deviation, R^2cv: correlation coefficient of cross validation, SECV: Root mean square error of cross validation, R^2pre: correlation coefficient of prediction, SEP: Root mean square error of prediction, RER: Range error ratio.

was released a few years earlier than the NIR sensor, which was recently marketed in 2024. Hence, the samples that retained freshness were used for developing the NIR sensor model, resulting in a smaller number of samples for the sensor. A number of factors corresponding to the first local minimum of the factor selection line plot (RMSECV against the number of factors) were considered for model development. The number of factors in the cross-validated NIR scanner models varied from 5 to 7, describing 83–95 % of the variability. The serum viscosity model of the NIR scanner performed better than the NIR sensor (data not shown). The number of factors was comparatively higher for the NIR sensor models (4–8), covering a higher cumulative variation of 94–98 %. Selecting an appropriate number of factors is essential to balance model complexity. Considering too few factors may not model important spectral variation leading to under-fitting, while a higher number of factors may introduce random noise resulting in an over-fitted model [17,52]. The mean and standard deviation of the calibration and validation datasets were comparable. A higher correlation coefficient indicates enhanced predictive accuracy [17] while a low RMSE implies good precision and predictive ability [52] of the regression model. A good correlation was obtained between the measure values and the predicted values with R^2cv $\geq$ 0.67 and R^2pre $\geq$0.70 for NIR scanner models and R^2cv $\geq$ 0.81 and R^2pre $\geq$085 for NIR sensor. Furthermore, the results of calibration (Rcv, RMSECV) and validation (Rpre, RMSEP) models were comparable indicating the external prediction with an unseen dataset performs as good as the internal cross-validation. In terms of correlation coefficient, the NTSS parameter gave the best performance in NIR scanner models followed by a/b ratio, Bostwick consistency and TA. While in NIR sensor, the NTSS gave the best performance followed by TA, Bostwick consistency and a/b ratio.

According to the feasibility criteria based on the correlation coefficient, as detailed by Ruiz-Roso et al. (2022) [53], NTSS gave an excellent performance, TA gave a good performance, Bostwick consistency and a/b ratio demonstrated a very good performance while Serum viscosity model can be considered for prediction in NIR scanner models. In the NIR sensor, NTSS showed excellent performance, TA and Bostwick consistency gave very good results while a/b ratio showed good performance. NTSS, Bostwick consistency, and TA showed superior performance to the study by Ruiz-Roso et al. (2022) [53] that analyzed tomato paste (n = 50) by NIR reflectance spectroscopy (correlation coefficient of NTSS (0.96) and Bostwick consistency (0.79)). However, the correlation coefficient of a/b ratio parameter was comparatively less than the result obtained by Ruiz-Roso et al. (2022) [53] (R^2cv = 0.94) which can be attributed to the changes during prolonged sample storage.

Additional metrics including RER and SEP/SECV were calculated to assess the prediction performance of the regression models. RER is the ratio of the range of reference values in the validation dataset to its RMSEP. The RER values of all regression models exceeded 4, indicating their suitability for screening calibration [54]. Based on the categorization of the RER values by Alves et al. (2012) [54] NTSS model was appropriate for quantification while TA, Bostwick consistency, serum viscosity, and a/b ratio models can be used for calibration for quality control. The good model typically has an SEP/SECV ratio close to 1.0 [54]; our models approximately met this guideline with SEP/SECV ranging from 0.8 to 1.1.

Figs. 3 and 4 represents the correlation plots which marks the predicted values against the reference values, obtained from NIR scanner and NIR sensor respectively. The data points scatter along the regression line (diagonal line) indicating a good fit. Moreover, the lighter circles representing validation data points lie within the calibration data points (depicted by darker circles), suggesting that the calibration models capture the variability, practically expected in the tomato paste processing facilities.

A regression vector is a weighted sum of the loadings plot corresponding to the factors considered during model development [26]. It reflects the loading weights of each independent variable with high variations reflecting wavelengths of important contributions [45,55]. The region around 2350 nm appears with a high coefficient in regression vector of NIR scanner models (Fig. 5). The absorption in the region

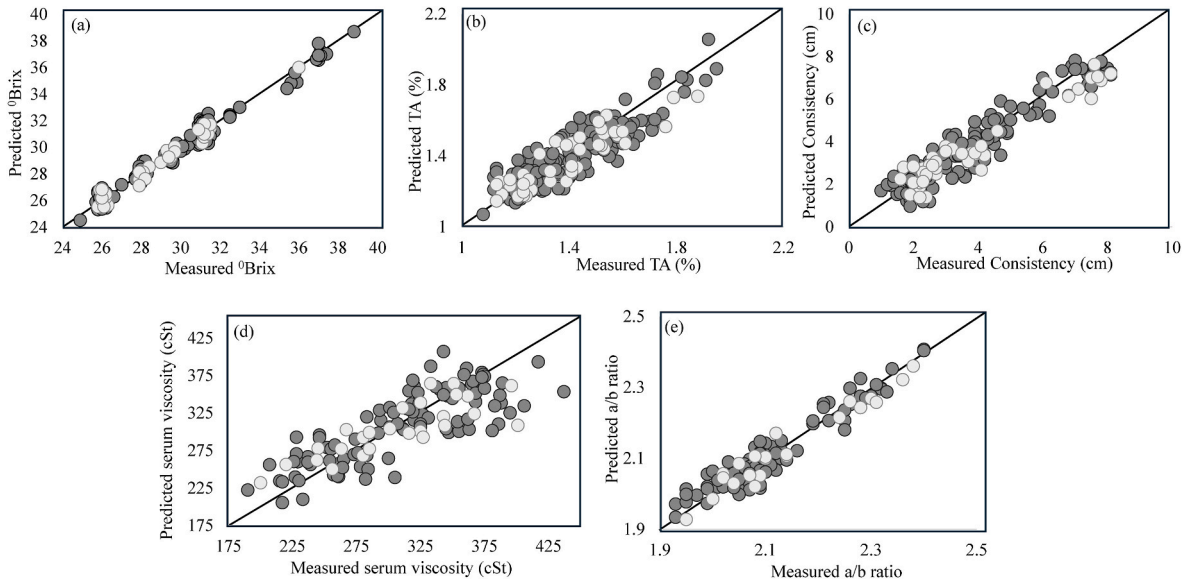

Fig. 3. PLSR correlation plots of NTSS (a), TA (b), Bostwick consistency (c), serum viscosity (d) and a/b ratio (e) obtained using NIR scanner (● and ○ denote calibration and validation samples, respectively).

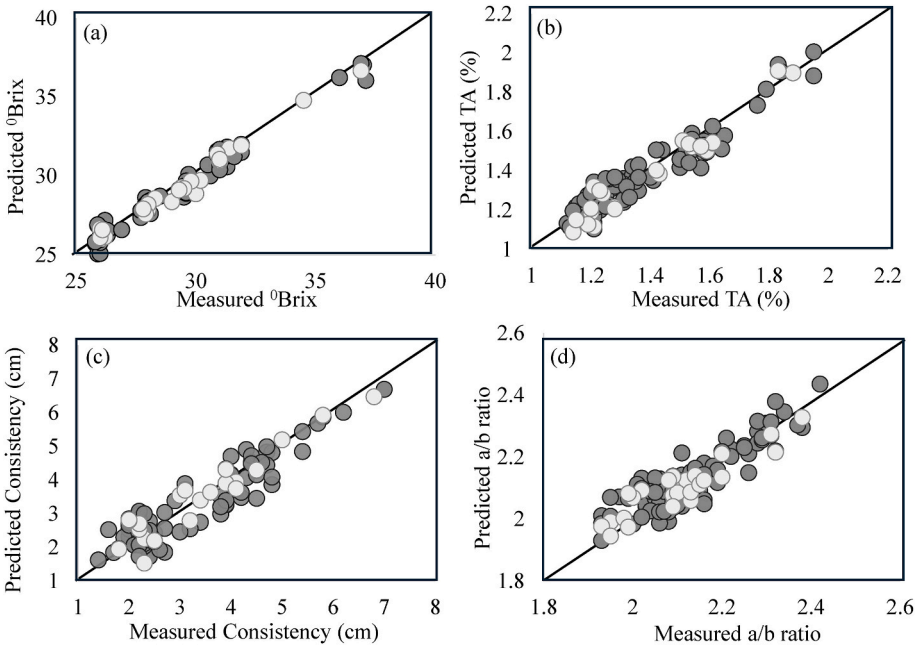

Fig. 4. PLSR correlation plots of NTSS (a), TA (b), Bostwick consistency (c) and a/b ratio (d) obtained using NIR sensor (● and ○ denote calibration and validation samples, respectively).

ranging from 2272 nm to 2480 nm has been associated to C-H groups of carbohydrates [45]. NIR wavenumbers at 2219 nm, 2232 nm, 2225 nm and 2246 nm in regression vector plots of NTSS, TA, Bostwick consistency, serum viscosity and a/b ratio respectively correspond to the combination bands of C-H and O-H [56,57]. The region between 2100 nm and 2200 nm in regression vectors of all quality parameters relate to the first combination bands of C-H bonds, primarily associated to sugars and acids [56,58]. Lan et al. (2022) [56] reported that the wavelength around 1400 nm was related to the soluble solids variation and was the major contributor for describing the rheological behavior of apple juice. The region around 1440 nm in the TA regression vector plot was attributed to the first overtone of the stretching vibrations of free OH groups [59]. Examination of the regression vectors showed that the absorption peaks in the LWNIR region (2550-1950 nm) largely influenced the spectral variations, emphasizing the contribution from combination bands. The regression vector plots of NIR sensor models (Fig. 6) showed similar patterns with high coefficients around 2227 nm, 2270 nm, 2107 nm and 1400 nm. Chang et al. (1998) observed the absorbance at 2270 nm (C-H stretch and deformation [60]) combination to be the most important factor in estimating soluble solids in fruits juices [61].

3.3.2. *Regression models of calibration transferred data*

The ability to transfer models from one NIR device (NIR scanner) to another (NIR sensor) was evaluated by performing calibration transfer and comparing the resulting performance of multivariate calibration transfer with models generated using the NIR sensor. The calibration dataset for the transfer models included both the original spectral data

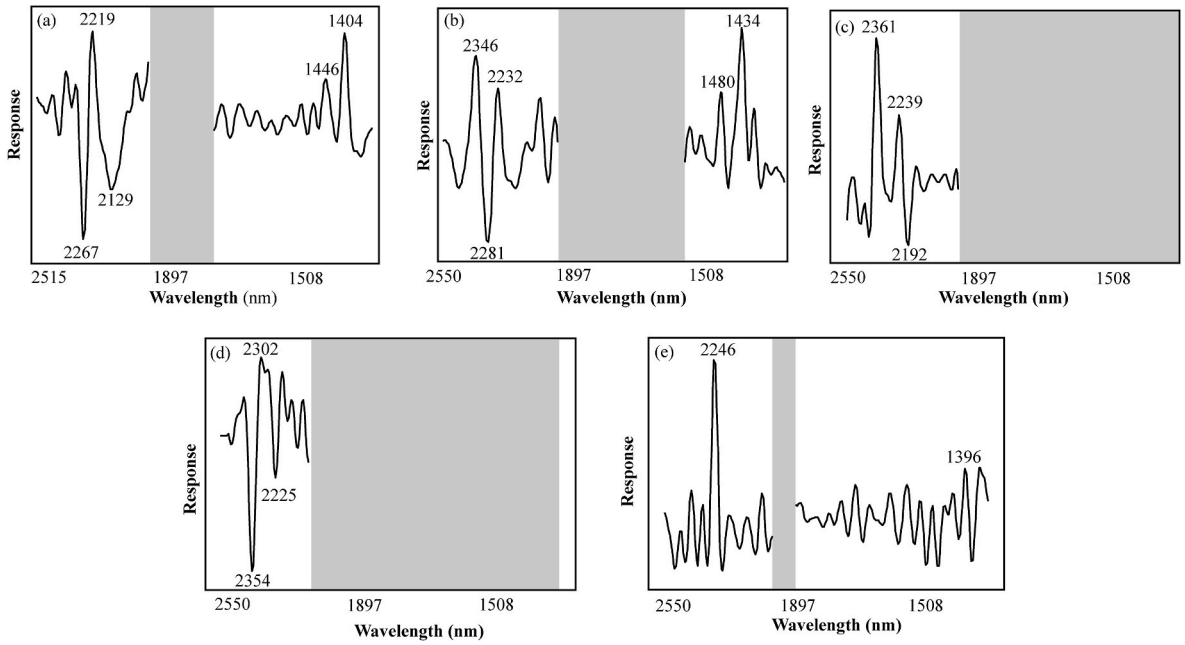

Fig. 5. Regression vector line plot of quality parameters NTSS (a), TA (b), Bostwick consistency (c), serum viscosity (d) and a/b ratio (e), corresponding to the models of NIR scanner.

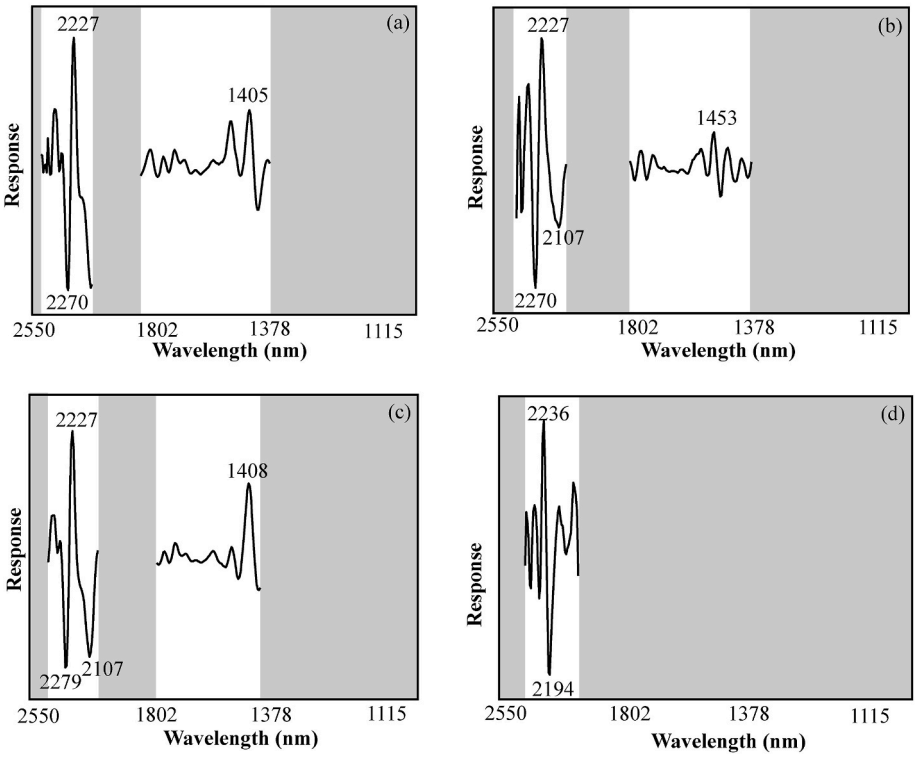

Fig. 6. Regression vector line plot of quality parameters NTSS (a), TA (b), Bostwick consistency (c), and a/b ratio (d), corresponding to the models of NIR sensor.

collected using NIR sensor and the transferred spectra from the NIR scanner. The external prediction set comprised the validation dataset used in the NIR sensor model. The performance of the external prediction was evaluated by incorporating different numbers of original spectra (20, 30, 40, 50, 60, 70, 80) into the calibration model. Including original spectra is essential to prevent the model from being dominated by the transformed spectra [31] and to account for the expected variation when acquiring the spectra in the new setting. The calibration transfer models of NTSS, TA, Bostwick consistency, and the a/b ratio demonstrated the best prediction performances when 21 %, 15 %, 33 %, and 70 % of the dataset consisted of original spectra from the NIR sensor. Table 4 shows the chemical composition of the calibration and validation datasets used in the calibration transfer models, along with their regression performance.

Table 4
Chemical composition of the calibration and validation sets, and statistical performance of the regression models obtained using the calibration transferred data.

Parameter	F	Calibration					Validation						
		N	Mean	SD	R2cv	SECV	N	Mean	SD	R2pre	SEP	RER	SEP/SECV
NTSS (0Brix)	7	219	29.2	2.6	0.94	0.65	27	29.4	2.6	0.81	1.22	8.9	1.9
TA (%)	5	224	1.4	0.2	0.75	0.09	25	1.4	0.2	0.83	0.09	8.2	1.0
Consistency (cm)	6	203	3.1	1.2	0.78	0.58	21	3.3	1.3	0.70	0.77	6.5	1.3
a/b ratio	7	135	2.1	0.1	0.56	0.07	24	2.1	0.1	0.62	0.07	6.5	1.0

F: Factors, SD: Standard deviation, R2cv: correlation coefficient of cross validation, SECV: Root mean square error of prediction of cross validation, R2pre: correlation coefficient of prediction, SEP: Root mean square error of prediction, RER: Range error ratio.

The calibration transfer models (Table 4) demonstrated acceptable performance with good R2pre (0.62–0.83) and considerable RMSEP. The RER values for all parameters were greater than 4, indicating their applicability for screening calibration [54]. The SEP/SEC ratio was close to 1.0, except for NTSS, which demonstrated a value of 1.9. The first paper published by Wang et al., 1992 which reported the application of PDS for calibration transfer among NIR instruments found that the transferred calibration model resulted in an average standard error twice that of the original model [29]. It has been reported that the standardization methods cannot provide quality transfer results for very complex systems [32] and there is always a certain amount of natural variation expected in the prediction [62]. Tomato paste samples are quite complex due to their chemical composition and moisture content. Although the transfer model is not as good as the original model, it provides a basic model to work with instead of requiring full recalibration on the new instrument [62]. Correlation scatter plots of predicted versus measured values for all parameters obtained through calibration transfer are shown in Fig. 7, providing a visual understanding of the variation and regression fit of the data points.

Fig. 8 displays the regression vector plots for all parameters obtained using the calibration transferred data. The plot shows higher regression coefficients around 2130 nm in the LWNIR region (C-H stretch and deformation [60]) and the informative region from 1574 nm to 1674 nm was observed in the SWNIR region (first overtones of O-H stretching modes, primarily from soluble solids) [63].

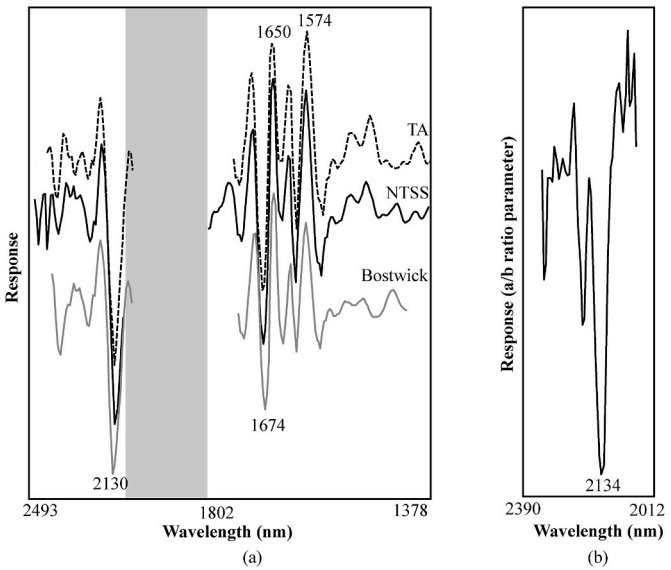

Fig. 8. Regression vector line plot of quality parameters (a) NTSS, TA, Bostwick consistency, and (b) a/b ratio, corresponding to the calibration transfer models.

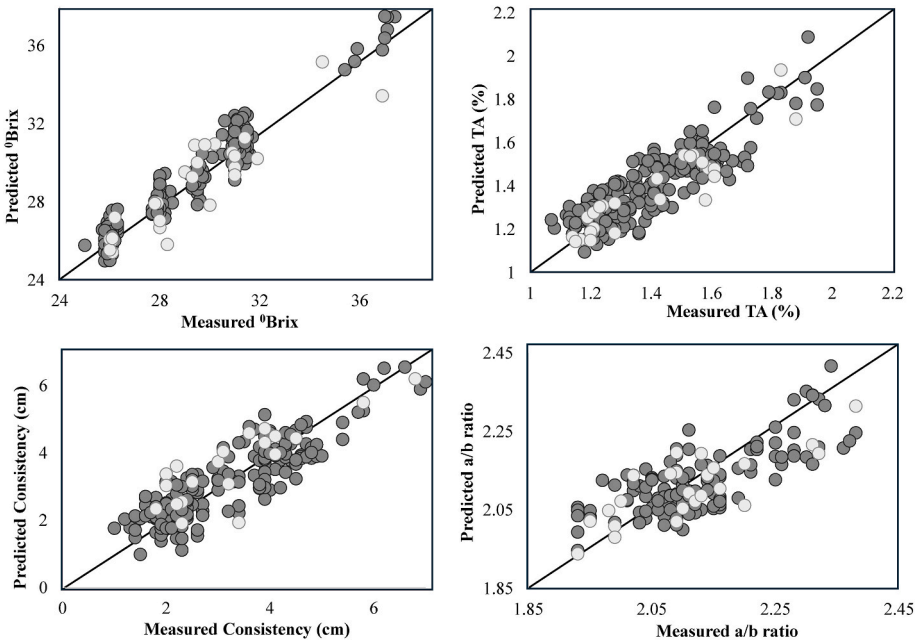

Fig. 7. PLSR correlation plots of Brix (a), TA (b), Bostwick consistency (c), and a/b ratio (d) regression models obtained using calibration transferred data (● and ○ denote calibration and validation set samples, respectively).

4. Conclusion

In this study, application of a novel handheld NIR scanner and a NIR sensor in transflectance mode were evaluated for non-destructive and rapid *in-situ* analysis of important quality parameters in tomato paste samples. The predictive algorithms developed using both the NIR systems yielded excellent performance (R^2pre = 0.83 to 0.98) and low standard error, demonstrating their feasibility for quality assessment of tomato paste samples in processing facilities. Furthermore, model transfer from the NIR scanner to the NIR sensor was investigated, and the performance of the resulting multivariate calibration transfer was compared with that of the NIR sensor. Models generated using the NIR sensor performed better than the calibration transfer models, however, the validation performance of the calibration transfer models suggested their suitability for screening calibration. Overall, the study presents the potential of miniaturized NIR systems as a cost-effective, high-throughput, multi-component monitoring technique for tomato paste samples in tomato processing industries, optimizing processing parameters, improving production rates, and increasing the quality of the products.

CRediT authorship contribution statement

Shreya Madhav Nuguri: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Formal analysis, Data curation. **Silvia de Lamo Castellvi:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Didem Peren Aykas:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Mustafa Mortas:** Formal analysis, Data curation. **Luis Rodriguez-Saona:** Writing – review & editing, Visualization, Supervision, Software, Resources, Project administration, Methodology, Investigation, Conceptualization.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

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.

Acknowledgements

The authors would like to acknowledge Marcal Plans for his guidance on the calibration transfer analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jafr.2025.101974.

Data availability

Data will be made available on request.

References

[1] K. Winans, S. Brodt, A. Kendall, Life cycle assessment of California processing tomato: an evaluation of the effects of evolving practices and technologies over a 10-year (2005–2015) timeframe, Int. J. Life Cycle Assess. 25 (2020) 538–547, https://doi.org/10.1007/s11367-019-01688-6.

[2] G.E. Anthon, D.M. Barrett, Pectin methylesterase activity and other factors affecting pH and titratable acidity in processing tomatoes, Food Chem. 132 (2012) 915–920, https://doi.org/10.1016/j.foodchem.2011.11.066.

[3] D.P. Aykas, K.R. Borba, L.E. Rodriguez-Saona, Non-destructive quality assessment of tomato paste by using portable mid-infrared spectroscopy and multivariate analysis, Foods 9 (2020), https://doi.org/10.3390/foods9091300.

[4] L.E. Rodriguez-Saona, D.P. Aykas, New approaches for rapid tomato quality control, Food Chemistry, Function and Analysis 2019-January (2019) 85–113, https://doi.org/10.1039/9781788016247-00085.

[5] United States Standards for Grades of Fresh Tomatoes, 1991.

[6] E. Koh, S. Charoenprasert, A.E. Mitchell, Effects of industrial tomato paste processing on ascorbic acid, flavonoids and carotenoids and their stability over one-year storage, J. Sci. Food Agric. 92 (2012) 23–28, https://doi.org/10.1002/jsfa.4580.

[7] A complete course in canning and related processes | WorldCat.org, (n.d.). https://www.worldcat.org/title/complete-course-in-canning-and-related-processes/oclc/476217331?referer=di&ht=edition (accessed September 6, 2022).

[8] K. Sridhar, H.A. Makroo, B. Srivastava, Effect of cold- and hot-break heat treatments on the physicochemical characteristics of currant tomato (solanum pimpinellifolium) pulp and paste, Foods 11 (2022) 1730, https://doi.org/10.3390/FOODS11121730.

[9] L. Zhang, M.A. Schultz, R. Cash, D.M. Barrett, M.J. McCarthy, Determination of quality parameters of tomato paste using guided microwave spectroscopy, Food Control 40 (2014) 214–223, https://doi.org/10.1016/j.foodcont.2013.12.008.

[10] G.E. Anthon, D.M. Barrett, Changes in tomato paste during storage and the effects of heating on consistency of reconstituted tomato paste, J. Texture Stud. 41 (2010) 262–278, https://doi.org/10.1111/j.1745-4603.2010.00225.x.

[11] B. Congcong Zhang, L.E. Rodriguez-Saona, Lynn Knipe, Monica Giusti, Rapid assessment of sugars and organic acids in tomato paste using a portable mid-infrared spectrometer and multivariate analysis, 2016, pp. 9–11.

[12] B.R. Thakur, R.K. Singh, P.E. Nelson, Quality attributes of processed tomato products: a review, Food Rev. Int. 12 (1996) 375–401, https://doi.org/10.1080/87559129609541085.

[13] D.P. Aykas, K.R. Borba, L.E. Rodriguez-Saona, Non-destructive quality assessment of tomato paste by using portable mid-infrared spectroscopy and multivariate analysis, Foods 9 (2020), https://doi.org/10.3390/foods9091300.

[14] Tomato Production, Processing and Technology - WA Gould - Google Books, (n.d.). https://books.google.com/books?hl=en&lr=&id=cIVwAgAAQBAJ&oi=fnd&pg=PP1&dq=W.+V.+Gould,+%22Tomato+Production,+Processing+and+Technology,%22+CTI+Publ.,+Baltimore,+1992.&ots=DqONyk-9yG&sig=0_UlQ_bQnZ2pF-Y2hsQorPBGz2c#v=onepage&q&f=false (accessed September 10, 2022).

[15] M. Ganje, S.M. Jafari, V. Farzaneh, N. Malekjani, Kinetics modelling of color deterioration during thermal processing of tomato paste with the use of response surface methodology, Heat and Mass Transfer/Waerme- Und Stoffuebertragung 54 (2018) 3663–3671, https://doi.org/10.1007/S00231-018-2394-3/FIGURES/3.

[16] D.E. Rubio-Diaz, A. Santos, D.M. Francis, L.E. Rodriguez-Saona, Carotenoid stability during production and storage of tomato juice made from tomatoes with diverse pigment profiles measured by infrared spectroscopy, J. Agric. Food Chem. 58 (2010) 8692–8698, https://doi.org/10.1021/JF1012665.

[17] H. Ayvaz, A. Sierra-Cadavid, D.P. Aykas, B. Mulqueeney, S. Sullivan, L. E. Rodriguez-Saona, Monitoring multicomponent quality traits in tomato juice using portable mid-infrared (MIR) spectroscopy and multivariate analysis, Food Control 66 (2016) 79–86, https://doi.org/10.1016/J.FOODCONT.2016.01.031.

[18] E.D. Wilkerson, G.E. Anthon, D.M. Barrett, G.F.G. Sayajon, A.M. Santos, L. E. Rodriguez-Saona, Rapid assessment of quality parameters in processing tomatoes using hand-held and benchtop infrared spectrometers and multivariate analysis, J. Agric. Food Chem. 61 (2013) 2088–2095, https://doi.org/10.1021/jf304968f.

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

[20] M. Manley, Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials, Chem. Soc. Rev. 43 (2014) 8200–8214, https://doi.org/10.1039/c4cs00062e.

[21] Z. Guo, W. Huang, Y. Peng, Q. Chen, Q. Ouyang, J. Zhao, Color compensation and comparison of shortwave near infrared and long wave near infrared spectroscopy for determination of soluble solids content of 'Fuji' apple, Postharvest Biol. Technol. 115 (2016) 81–90, https://doi.org/10.1016/J.POSTHARVBIO.2015.12.027.

[22] Y. Liu, H. Xu, Z. Xia, Z. Gong, Multi-spectrometer calibration transfer based on independent component analysis, Analyst 143 (2018) 1274–1280, https://doi.org/10.1039/C7AN01555K.

[23] R.N. Feudale, N.A. Woody, H. Tan, A.J. Myles, S.D. Brown, J. Ferré, Transfer of multivariate calibration models: a review, n.d. www.elsevier.com/locate/chemometrics.

[24] Diffuse reflection light source, (n.d.). www.hamamatsu.com (accessed October 8, 2024).

[25] H. Abdi, L.J. Williams, Partial least squares methods: partial least squares correlation and partial least square regression, Methods Mol. Biol. 930 (2013) 549–579, https://doi.org/10.1007/978-1-62703-059-5_23/COVER.

[26] Infometrix Inc., Pirouette Multivariate Data Analysis Software, 2011, p. 506. http://www.infometrix.com/.

[27] S.I. Sohn, S. Pandian, Y.J. Oh, J.L.Z. Zaukuu, H.J. Kang, T.H. Ryu, W.S. Cho, Y. S. Cho, E.K. Shin, B.K. Cho, An overview of near infrared spectroscopy and its applications in the detection of genetically modified organisms, Int. J. Mol. Sci. 22 (2021) 9940, https://doi.org/10.3390/IJMS22189940.

[28] T. Wang, L.E. Rodriguez-Saona, Rapid determination of sugar level in snack products using infrared spectroscopy, J. Food Sci. 77 (2012) C874–C879, https://doi.org/10.1111/J.1750-3841.2012.02824.X.
[29] Calibration Transfer and Measurement Stability of Near-Infrared Spectrometers, 1992.
[30] C.-S. Chen, C.W. Brown, S.-C. Lo, Calibration Transfer from Sample Cell to Fiber-Optic Probe, 1997.
[31] A.J. Parrott, A.C. Mcintyre, M. Holden, G. Colquhoun, Z.-P. Chen, D. Littlejohn, A. Nordon, Calibration model transfer in mid-infrared process analysis with in situ attenuated total reflectance immersion probes. https://doi.org/10.15129/42fb0eac-a54a-4b63-85f5-4ed6c965144f, 2022.
[32] R.N. Feudale, N.A. Woody, H. Tan, A.J. Myles, S.D. Brown, J. Ferré, Transfer of multivariate calibration models: a review, n.d. www.elsevier.com/locate/chemometrics.
[33] CFR - Code of Federal Regulations Title 21, (n.d.).
[34] K. Benoit, Linear Regression Models with Logarithmic Transformations, 2011.
[35] J.A. Barreiro, M. Milan, A.J. Sandoval, Kinetics of Colour Change of Double Concentrated Tomato Paste during Thermal Treatment, n.d.
[36] H.A. Makroo, N.K. Rastogi, B. Srivastava, Enzyme inactivation of tomato juice by ohmic heating and its effects on physico-chemical characteristics of concentrated tomato paste, J. Food Process. Eng. 40 (2017) e12464, https://doi.org/10.1111/JFPE.12464.
[37] M. Ganje, S.M. Jafari, A. Dusti, D. Dehnad, M. Amanjani, V. Ghanbari, Modeling quality changes in tomato paste containing microencapsulated olive leaf extract by accelerated shelf life testing, Food Bioprod. Process. 97 (2016) 12–19, https://doi.org/10.1016/J.FBP.2015.10.002.
[38] K. Włodarska, I. Khmelinskii, E. Sikorska, Evaluation of quality parameters of apple juices using near-infrared spectroscopy and chemometrics, Journal of Spectroscopy 2018 (2018), https://doi.org/10.1155/2018/5191283.
[39] L. Xie, X. Ye, D. Liu, Y. Ying, Quantification of glucose, fructose and sucrose in bayberry juice by NIR and PLS, Food Chem. 114 (2009) 1135–1140, https://doi.org/10.1016/j.foodchem.2008.10.076.
[40] K. Włodarska, J. Szulc, I. Khmelinskii, E. Sikorska, Non-destructive determination of strawberry fruit and juice quality parameters using ultraviolet, visible, and near-infrared spectroscopy, J. Sci. Food Agric. 99 (2019) 5953–5961, https://doi.org/10.1002/jsfa.9870.
[41] F. Mabood, J. Hussain, F. Jabeen, G. Abbas, B. Allaham, M. Albroumi, S. Alghawi, S. Alameri, S.A. Gilani, A. Al-Harrasi, Q.M.I. Haq, S. Farooq, Applications of FT-NIRS combined with PLS multivariate methods for the detection & quantification of saccharin adulteration in commercial fruit juices, Food Addit. Contam. Part A Chem Anal Control Expo Risk Assess 35 (2018) 1052–1060, https://doi.org/10.1080/19440049.2018.1457802.
[42] A. Zeb, W.S. Qureshi, A. Ghafoor, A. Malik, M. Imran, A. Mirza, M.I. Tiwana, E. Alanazi, Towards sweetness classification of orange cultivars using short-wave NIR spectroscopy, Scientific Reports 2023 13 (1 13) (2023) 1–12, https://doi.org/10.1038/s41598-022-27297-2.
[43] M.M. Giusti, A. Atnip, C. Sweeney, L.E. Rodriguez-Saona, Rapid authentication of fruit juices by infrared spectroscopic techniques, ACS (Am. Chem. Soc.) Symp. Ser. 1081 (2011) 275–299, https://doi.org/10.1021/BK-2011-1081.CH020.
[44] X. Dong, X. Sun, A case study of characteristic bands selection in near-infrared spectroscopy: nondestructive detection of ash and moisture in wheat flour, J. Food Meas. Char. 7 (2013) 141–148, https://doi.org/10.1007/S11694-013-9149-0/FIGURES/9.
[45] L.E. Rodriguez-Saona, F.S. Fry, E.M. Calvey, Use of fourier transform near-infrared reflectance spectroscopy for rapid quantification of castor bean meal in a selection of flour-based products, J. Agric. Food Chem. 48 (2000) 5169–5177, https://doi.org/10.1021/jf000604m.
[46] J. Wan, Y.C. Chen, A.J. Morris, S.N. Thennadil, A comparative investigation of the combined effects of pre-processing, wavelength selection, and regression methods on near-infrared calibration model performance, Appl. Spectrosc. 71 (2017) 1432–1446, https://doi.org/10.1177/0003702817694623/SUPPL_FILE/ASP694623_SUPPLEMENTARY_MATERIAL.PDF.
[47] J. Guezenoc, A. Gallet-Budynek, B. Bousquet, Critical review and advices on spectral-based normalization methods for LIBS quantitative analysis, Spectrochim. Acta Part B At. Spectrosc. 160 (2019) 105688, https://doi.org/10.1016/J.SAB.2019.105688.
[48] M.A. Czarnecki, Resolution enhancement in second-derivative spectra, Appl. Spectrosc. 69 (2015) 67–74, https://doi.org/10.1366/14-07568/ASSET/IMAGES/LARGE/10.1366_14-07568-FIG8.JPEG.
[49] A.A. Munawar, Multivariate analysis and artificial neural network approaches of near infrared spectroscopic data for non-destructive quality attributes prediction of Mango (Mangifera indica L.), Dissertation (2014) 4–6. https://ediss.uni-goettingen.de/handle/11858/00-1735-0000-0022-5E52-8.
[50] Z. Xiaobo, Z. Jiewen, M.J.W. Povey, M. Holmes, M. Hanpin, Variables selection methods in near-infrared spectroscopy, Anal. Chim. Acta 667 (2010) 14–32, https://doi.org/10.1016/j.aca.2010.03.048.
[51] Y.D. Liu, Y.B. Ying, H.Y. Jiang, Rapid determination of maturity in apple using outlier detection and calibration model optimization, Trans. ASABE (Am. Soc. Agric. Biol. Eng.) 49 (2006) 91–95, https://doi.org/10.13031/2013.20215.
[52] A.O. Aptula, N.G. Jeliazkova, T.W. Schultz, M.T.D. Cronin, The better predictive model: high q2 for the training set or low root mean square error of prediction for the test set? QSAR Comb. Sci. 24 (2005) 385–396, https://doi.org/10.1002/qsar.200430909.
[53] M.J.R. Ruiz-Roso, R. Rodríguez-Corchado, M. Gómez-Cardoso, M.C. Carrillo, A. Crespo, P. Ramayo, J.L. Llerena-Ruiz, M.J. Redondo, A. Roldán, E. Valdezate, The potential of near infrared reflectance spectroscopy (NIRS) for the estimation of quality parameters in tomato paste, in: Acta Hortic, International Society for Horticultural Science, 2022, pp. 223–230, https://doi.org/10.17660/ActaHortic.2022.1351.33.
[54] A. Alves, A. Santos, P. Rozenberg, L.E. Pâques, J.P. Charpentier, M. Schwanninger, J. Rodrigues, A common near infrared-based partial least squares regression model for the prediction of wood density of Pinus pinaster and Larix × eurolepis, Wood Sci. Technol. 46 (2012) 157–175, https://doi.org/10.1007/S00226-010-0383-X/TABLES/3.
[55] A.A. Gowen, G. Downey, C. Esquerre, C.P. O'Donnell, Preventing over-fitting in PLS calibration models of near-infrared (NIR) spectroscopy data using regression coefficients, J. Chemom. 25 (2011) 375–381, https://doi.org/10.1002/CEM.1349.
[56] W. Lan, V. Baeten, B. Jaillais, C.M.G.C. Renard, Q. Arnould, S. Chen, A. Leca, S. Bureau, Comparison of near-infrared, mid-infrared, Raman spectroscopy and near-infrared hyperspectral imaging to determine chemical, structural and rheological properties of apple purees, J. Food Eng. 323 (2022) 111002, https://doi.org/10.1016/J.JFOODENG.2022.111002.
[57] H. Wang, J. Peng, C. Xie, Y. Bao, Y. He, Fruit quality evaluation using spectroscopy technology: a review, Sensors 15 (2015) 11889–11927, https://doi.org/10.3390/S150511889, 11889-11927 15 (2015).
[58] W. Lan, B. Jaillais, A. Leca, C.M.G.C. Renard, S. Bureau, A new application of NIR spectroscopy to describe and predict purees quality from the non-destructive apple measurements, Food Chem. 310 (2020), https://doi.org/10.1016/J.FOODCHEM.2019.125944.
[59] S.N. Jha, Near infrared spectroscopy, nondestructive evaluation of food quality, Theory and Practice (2010) 141–212, https://doi.org/10.1007/978-3-642-15796-7_6/TABLES/10.
[60] L.J. Xie, X.Q. Ye, D.H. Liu, Y. Bin Ying, Application of principal component-radial basis function neural networks (PC-RBFNN) for the detection of water-adulterated bayberry juice by near-infrared spectroscopy, J. Zhejiang Univ. - Sci. B 9 (2008) 982–989, https://doi.org/10.1631/JZUS.B0820057/METRICS.
[61] W.H. Chang, S. Chen, C.C. Tsai, Development of a universal algorithm for use of nir in estimation of soluble solids in fruit juices, Transactions of the ASAE 41 (1998) 1739–1745, https://doi.org/10.13031/2013.17317.
[62] J.J. Workman, A review of calibration transfer practices and instrument differences in spectroscopy, Appl. Spectrosc. 72 (2018) 340–365, https://doi.org/10.1177/0003702817736064.
[63] Y. Ozaki, Near-infrared spectroscopy-its versatility in analytical chemistry, Anal. Sci. 28 (2012) 545–563, https://doi.org/10.2116/ANALSCI.28.545/METRICS.