Application NIR Spectroscopy for Prediction Soluble Solids Content and Classification of Tomatoes During Storage

Raisal Maulana, Andasuryani*, Dinah Cherie

Department of Agriculture and Biosystem Engineering, Faculty of Agricultural Engineering and Technology, Andalas University, Jl. Dr. Mohammad Hatta, Limau Manis, Kec. Pauh, Kota Padang, Sumatera Barat 25163.

*Corresponding author, email: andasuryani@ae.unand.ac.id

Article Info	Abstract
	Tomatoes are a horticultural commodity that is highly susceptible to quality degradation after harvest; therefore, appropriate postharvest handling is required to maintain quality. This study aims to evaluate the potential of near-infrared (NIR) spectroscopy for assessing tomato quality by applying partial least squares (PLS) to predict soluble solids content (SSC) and linear discriminant analysis (LDA) for classification based on storage temperature and ripeness level, with SNV pretreatment. Tomato samples were stored at 10 °C and 28 °C and observed at the breaker and pink ripeness stages. The best PLS model was obtained with SNV pretreatment and 10 latent variables, yielding R² calibration = 0.89, RMSEC = 0.19°Brix, R² prediction = 0.80, and RMSEP = 0.26 °Brix. The RPD value of 2.04 and the RER of 8.08 indicate that the model has a good predictive ability for evaluating tomato SSC. Meanwhile, LDA distinguished storage temperature better (accuracy 89.13%) than ripeness level (accuracy 65.21%). These results demonstrate that NIR spectroscopy can be used as an effective nondestructive method for analyzing the SSC of tomatoes during storage, reflecting the levels of sugars, organic acids, and other soluble compounds that contribute to the taste and overall fruit quality.

1. Introduction

Tomatoes (*Solanum lycopersicon* L.) are horticultural commodities that play an essential role in the farming and food industries (Collins et al., 2022). One of the most commonly used quality parameters for assessing tomato quality is the Soluble Solids Content (SSC) (Pathare dan Al-Dairi, 2021; Bapary et al., 2024). The SSC value reflects the content of simple sugars, organic acids, and other soluble compounds that contribute to the tomato flavor (Amr & Raie, 2022). Changes in SSC values during storage can indicate metabolic processes within the fruit, such as the conversion of starch into sugars during the early ripening stage or the degradation of sugars in later stages, leading to a decline in the organoleptic quality (Quinet et al., 2019).

Tomato quality is influenced by environmental factors, particularly storage temperature, which determines the rate of physiological and chemical changes that occur during storage. In addition, tomato quality strongly depends on the ripeness stage, as each stage presents distinct physical and chemical characteristics, including SSC levels (Priyankara et al., 2017). Storage at low temperatures, such as 10 °C, can maintain the SSC content for a longer period (Pathare & Al-Dairi, 2021). Conversely, storage at ambient temperatures tends to accelerate respiration, transpiration, and enzymatic reactions, resulting in quality deterioration, including a decrease in SSC during the later stages of storage (Thole et al., 2021).

To date, SSC quality has generally been assessed using a destructive method: crushing the tomato to obtain its juice, which is then tested with a refractometer (Najjar & Abu-khalaf, 2021). However, this method is challenging to apply and unsuitable for continuous monitoring (Huang et al., 2025). Therefore, nondestructive approaches based on near-infrared (NIR) spectroscopy have been increasingly developed as efficient, rapid, and environmentally friendly solutions to this problem (Cortés et al., 2019). NIR spectroscopy operates based on the interaction of light with molecular bonds in the material, particularly C–H, O–H, and N–H bonds, which are closely related to quality parameters such as sugar, water, and texture content (Beć & Huck., 2019).

NIR technology not only enables quality prediction without damaging the fruit but also provides opportunities for quantitative modelling using multivariate approaches such as partial least squares (PLS) (Evi dan Rachbini, 2022). In addition to quantitative analysis, NIR can be used for qualitative analysis using approaches such as linear discriminant analysis (LDA) (Ribeiro et al., 2021). LDA works by finding a linear combination of spectra that maximizes the separation between groups while minimizing the variation within groups (Silalahi et al., 2016). Previous studies have shown that the LDA method was successfully applied to the NIR spectral data of mangoes grouped based on harvest time, achieving a classification accuracy of over 70% (Polinar et al., 2019).

This study aimed to develop a calibration model using NIR spectroscopy to predict SSC values in tomatoes. In addition, it aims to develop a classification model for tomatoes based on storage temperature (10 °C and 28 °C) and ripeness level (breaker and pink). The results are expected to contribute to the development of a faster, more accurate, and nondestructive postharvest quality analysis method, thereby improving supply chain efficiency and reducing yield losses caused by quality deterioration during storage.

2. Material and Methods

2.1 Sample Preparation

Tomatoes of the TM Blazzer variety were harvested at the breaker and pink ripeness stages according to the USDA standards. Breaker-stage tomatoes had 10–30% of their surface turning yellowish or pink, whereas pink-stage tomatoes had more than 60% of the surface colored pink to

reddish. A total of 146 uniformly sized tomatoes were randomly collected from local farmers in Alahan Panjang, Lembah Gumanti, Solok, West Sumatra (Latitude: -1°46′0.01″ S; Longitude: 101°22′1.20″ E). Tomatoes were stored at 10°C and 28°C. NIR spectra and reference data were collected on days 0, 7, 14, 21, and 28 for storage at 10°C, and on days 0, 2, 4, and 7 for 28°C. The observation intervals were designed to account for the differing rates of quality degradation at each temperature, ensuring that the data collected were representative for building calibration and classification models using NIR spectroscopy.

2.2 NIR Spectral Data Acquisition

Spectral data acquisition was carried out using a Buchi NIRFlex N-500 near-infrared spectrometer, which operates within a wavelength range of 1000–2500 nm at 4 nm intervals. Spectral measurements were performed at a single point on the surface of each tomato sample.

2.3 SSC Measurement

SSC measurements were performed using a digital refractometer (HANNA HI 96801). The samples were homogenized, and a few drops of the extract were placed on the prism surface of the instrument for analysis. Each measurement was performed in triplicate per sample to improve data reliability, and the SSC results were expressed in °Brix.

2.4 Spectral Pretreatment

Raw spectra obtained from NIR instruments often contain noise and irrelevant variations, such as light scattering caused by particle size or surface structure variations. Therefore, preprocessing is necessary to eliminate these interferences and enhance the signals related to the chemical characteristics of the sample. Common techniques include the Standard Normal Variate (SNV), Gap Segment First Derivative (GS 1stD), and their combination (Rinnan et al., 2009). The SNV method corrects spectral intensity fluctuations caused by scattering, whereas the Gap Segment Derivative enhances spectral information by emphasizing small changes and reducing baseline effects. Combining both methods yields a more optimal correction by simultaneously addressing scattering and highlighting the relevant spectral differences associated with chemical features (Zeaiter & Rutledge, 2009).

2.5 Partial Least Square (PLS)

The use of this method enables the analysis of the relationship between spectral data and various quality parameters of the oil. A good calibration model was built by selecting an adequate number of PLS factors. The selection of factors in PLS was performed to determine the optimal number of components in the calibration model. This process generally employs the cross-validation method to evaluate the model performance across different factor numbers and select the factor that yields the lowest prediction error. Selecting too many factors may lead to overfitting, causing the model to fail

to make optimal predictions (Andasuryani et al., 2014). All spectral data pretreatment and model development processes were performed using The Unscrambler X version 10.4 software (CAMO Software AS, Oslo, Norway).

Model accuracy was evaluated using several statistical parameters: coefficient of determination (R^2), Root Mean Square Error (RMSE), Ratio of Performance to Deviation (RPD), and Range Error Ratio (RER). These parameters assess the model's ability to predict and classify based on the spectral and reference data. A higher R^2 value indicates better precision, classified as excellent for $R^2 \geq 0.90$, good for 0.70–0.89, and poor for < 0.30. A smaller RMSE indicates better model performance (Martín, 2022). An RPD greater than 3 indicates an excellent model, whereas values below 1.5 are unsuitable for prediction (Maurel et al., 1994). An RER value greater than 15 indicates that the model has excellent performance and is suitable for use in quantitative prediction (Yarce & Rojas, 2012).

The mathematical formulas for these statistical parameters are as follows:

$$R^2 \text{calibration or prediction} = 1 - \frac{\sum_{i=1}^{n} (X_n - Y_n)^2}{\sum_{i=1}^{n} (\bar{X} - Y_n)^2} \tag{1}$$

$$\text{RMSE (°Brix)} = \sqrt{\frac{\sum (Y_n - X_n)^2}{n}} \tag{2}$$

$$\text{RPD} = \frac{SD}{RMSEP} \tag{3}$$

$$\text{RER} = \frac{Y_{max} - Y_{min}}{RMSEP} \tag{4}$$

Where R^2 = Coefficient of Determination, RMSEP = Root Mean Square Error of Prediction, RPD = Ratio of Performance to Deviation, RER = Range Error Ratio, SD = standard deviation of SSC prediction values, X_n = reference SSC value, Y_n = predicted SSC value from NIR, Y_{max} = maximum predicted SSC value from NIR, Y_{min} = minimum predicted SSC value from NIR, and $\bar{X}$ = mean of the reference SSC data.

2.6 Linear Discriminant Analysis (LDA)

Qualitative spectral analysis was conducted using LDA to distinguish tomato NIR spectra based on storage temperature (10 °C and 28 °C) and ripeness stage (breaker and pink). An Independent Sample T-Test was performed on the SSC values to verify whether the classification results aligned with the differences in SSC between the treatments. Data from days 0 and 7 were used, representing equivalent observation points under both storage conditions, ensuring a valid comparison and confirming the effect of temperature and ripeness on tomato SSC changes. The classification model was developed using random calibration data with cross-validation, and internal validation data were used to assess the model performance in correctly classifying the samples.

Model evaluation was conducted using a classification table comparing the actual and predicted classes, and the number of true and misclassified samples was identified. The average accuracy value

served as the leading indicator, reflecting the model's ability to distinguish classes during calibration and prediction and its reliability in classifying new data.

3. Results and Discussion

3.1 Characteristics of Tomato Absorbance Spectra

The NIR spectra (Figure 1) show absorption peaks at specific wavelengths. According to Mechram et al. (2024), this phenomenon occurs because of the interaction of specific chemical bonds with NIR light. Figure 1 shows the original NIR spectra of the tomato samples. The absorption peak located at approximately 1450 nm was identified as the characteristic absorption of water, associated with the vibration of the O–H bond (Sun et al., 2021). Furthermore, a broader and stronger absorption band was observed in the 1900–2100 nm range, corresponding to the O–H bond overtone band. This band reflects the combined contribution of water and carbohydrate compounds, such as simple sugars, which are the main components of SSC (Cao et al., 2025).
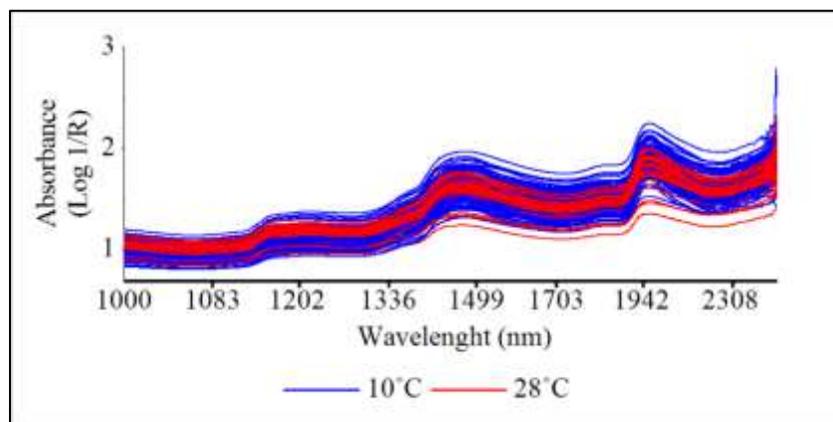

Figure 1. Characteristics of NIR Spectra

3.2 Outlier Data

A total of 146 samples were divided using the Kennard–Stone algorithm into 2/3 for calibration and 1/3 for prediction. Outlier detection was performed using PCA-based ellipse projection to identify samples outside the confidence boundary (Figure 2). Six outliers were removed following the general guideline of 1–5% to maintain data stability. A comparison was also made by processing the entire data set with the filtered data. If the resulting model performed better using the filtered data, it was confirmed that the previously removed data were outliers. If not, the data deemed outliers were not included in that category and were used for model development. After filtering, descriptive statistics were analyzed for the calibration and prediction datasets. The results showed that the prediction data remained within the representative range of the calibration data, ensuring the accuracy and reliability of the model in prediction (Andasuryani et al., 2014).
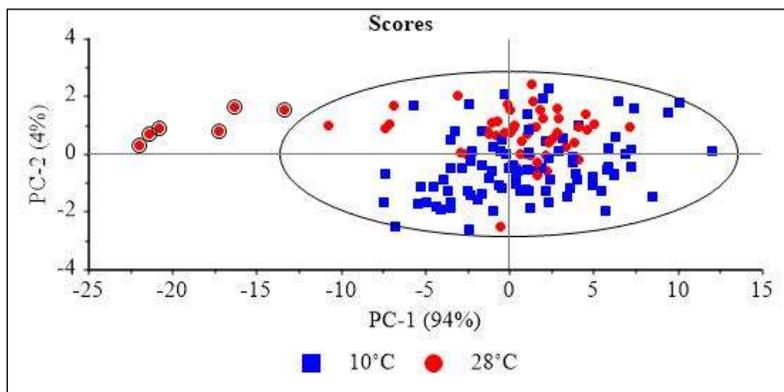
Figure 2. Projection of the PCA.

Table 1. Descriptive statistics of the SSC.

Parameter	Set Calibration	Set Prediction
Number of Samples	98	48
Mean	4.18	4.26
Median	4.13	4.22
Maximal	5.67	5.30
Minimal	2.87	3.20
Standard Deviation	0.59	0.53

3.3 Pretreatment

This study applied several pretreatment methods, including Standard Normal Variate (SNV) to minimize light scattering and inconsistencies caused by surface or size variations of tomatoes, gap-segment derivative to correct baseline drift and enhance spectral resolution, and a combination of SNV and Gap-Segment derivative to integrate both advantages for optimizing spectral information related to quality parameters (Rinnan et al., 2009). The original and pretreated spectra obtained using these methods are shown in Figure 3(a-c).
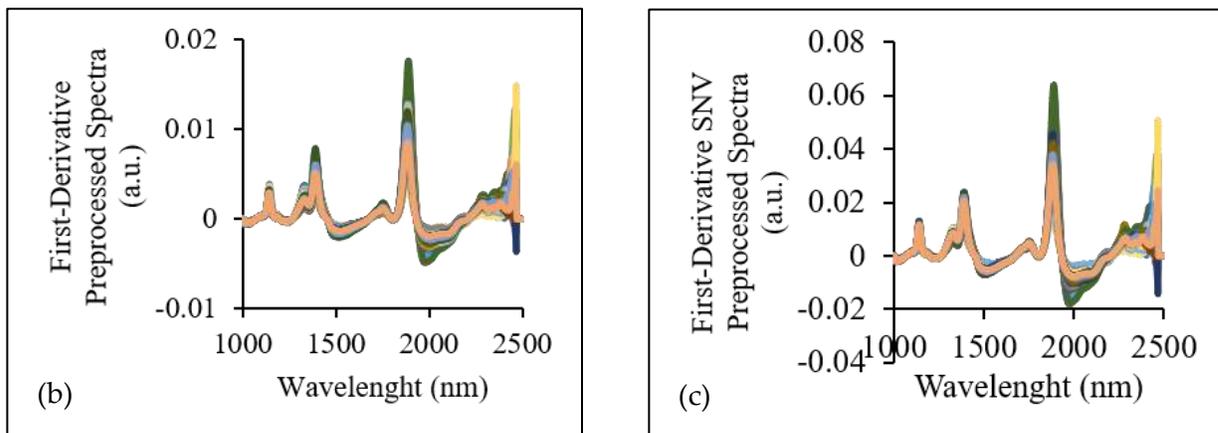

(a)
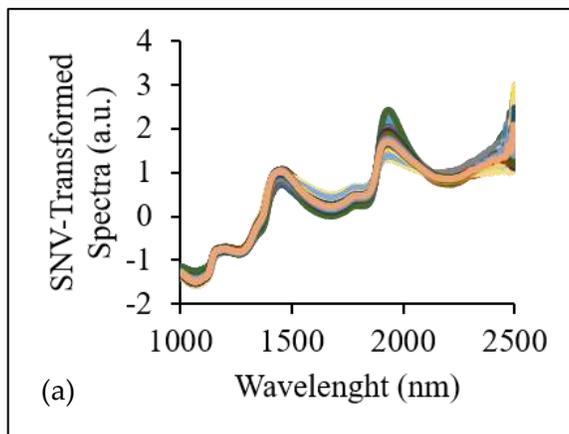
Continue

Figure 3. Characteristics of the NIR spectra using pretreatment (a.) SNV (b.) GS9 1^{st}D (c.) SNV + GS9 1^{st}D.

3.4 PLSR Analysis

A predictive model was developed using the Partial Least Squares (PLS) method within the NIR wavelength range of 1000–2500 nm, covering key functional groups (O–H, C–H, and N–H) related to water, sugars, and amino acids (Munawar et al., 2022). Cross-validation was applied to assess model stability and determine the optimal number of latent variables to avoid overfitting or underfitting (Andasuryani et al., 2014).

As shown in Table 2, the model achieved determination coefficients (R^2 calibration and R^2 prediction) above 0.70, indicating a strong correlation between the spectral data and tomato quality parameters. This confirms that the model provides good interpretability and predictive performance. (Martín, 2022). The prediction of the Total Soluble Solids (SSC) content in tomatoes using NIR spectroscopy is illustrated in Figure 4(a–d).

Table 2. Prediction model performance of tomato SSC during storage.

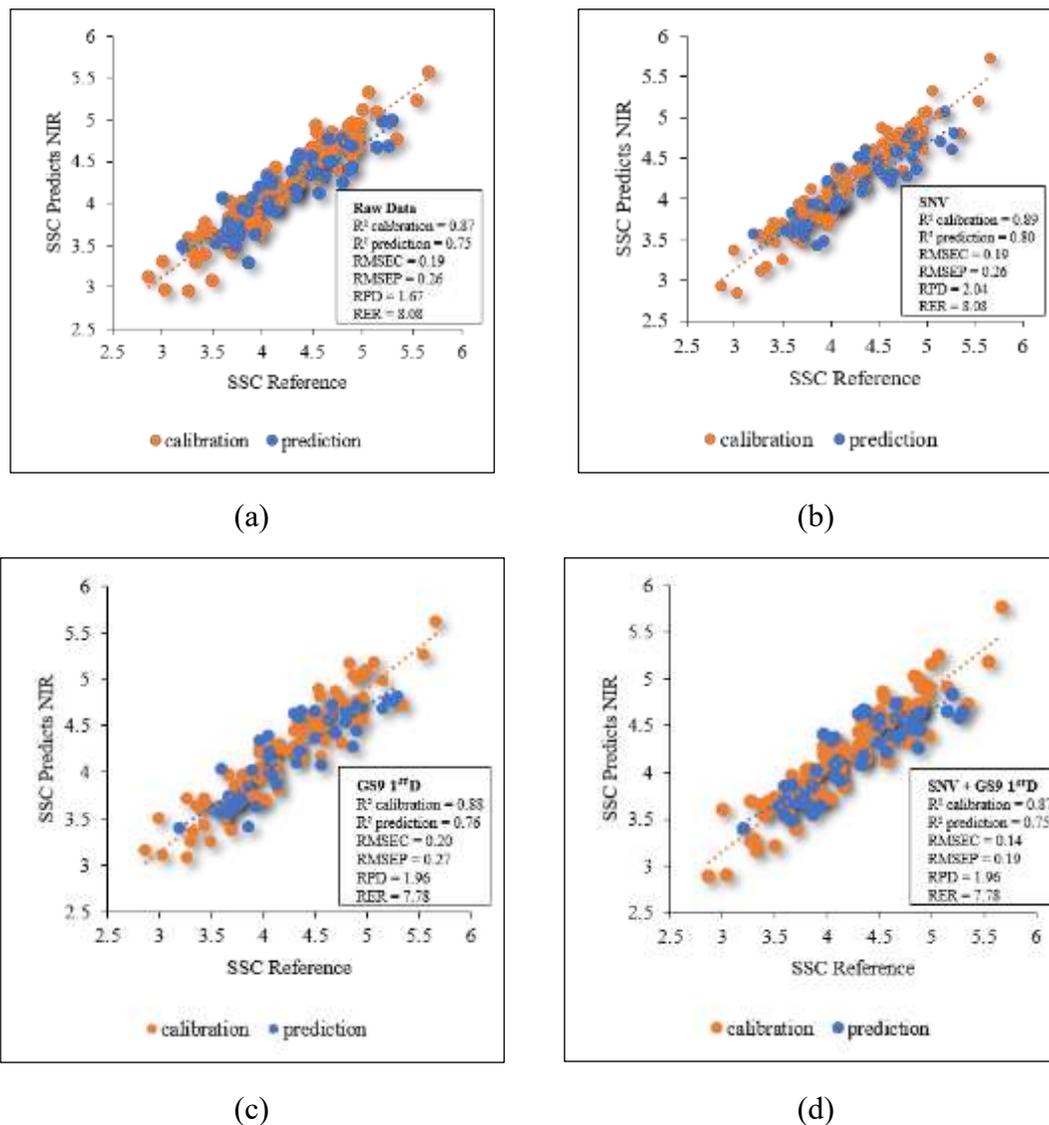| Pretreatment	LV	Calibration		Prediction		RPD	RER
		R^2	RMSEC (°Brix)	R^2	RMSEP (°Brix)		
Original	11	0.89	0.19	0.78	0.26	1.67	8.08
SNV	10	0.89	0.19	0.80	0.26	2.04	8.08
GS9 1^{st}D	9	0.87	0.20	0.76	0.27	1.96	7.78
SNV + GS9 1^{st}D	8	0.87	0.21	0.75	0.27	1.96	7.78

(a)

(b)

(c)

(d)

Figure 4. Comparison of tomato SSC for calibration and prediction data using different pretreatment methods.

The best model for predicting Soluble Solids Content (SSC) was obtained using the SNV pretreatment with 10 latent variables, yielding R^2 calibration = 0.89 and RMSEC = 0.19 in calibration, and R^2 prediction = 0.80 and RMSEP = 0.26 in prediction. The RPD (2.04) and RER (8.08) values indicated good reliability for quantitative prediction (Maurel et al., 1994; Barbin et al., 2015). Compared to the original spectrum and other pretreatments (GS9 1st derivative and SNV+GS9), SNV provided the most stable and accurate results with higher RPD and RER values. The choice of pretreatment strongly affects the model quality, as it enhances the relevant spectral information while

minimizing noise. Therefore, SNV was selected as the most effective pretreatment for SSC prediction in tomatoes.

3.5 Linier Discriminant Analysis (LDA)

3.5.1 Classification Based on Storage Temperature

LDA classification based on storage temperature showed that the NIR spectra of tomatoes could clearly distinguish between room-temperature and cold-stored samples. The spectral data pretreated with the SNV + GS9 1stD method produced a distinct separation pattern, indicating spectral differences influenced by the temperature effects on fruit quality. This distinction is associated with storage duration, where tomatoes at room temperature lasted only up to day 7, while those at cold storage remained stable until day 28 owing to slower respiration and quality degradation. The best separation results between 10 °C and 28 °C storage are shown in Figure 5.
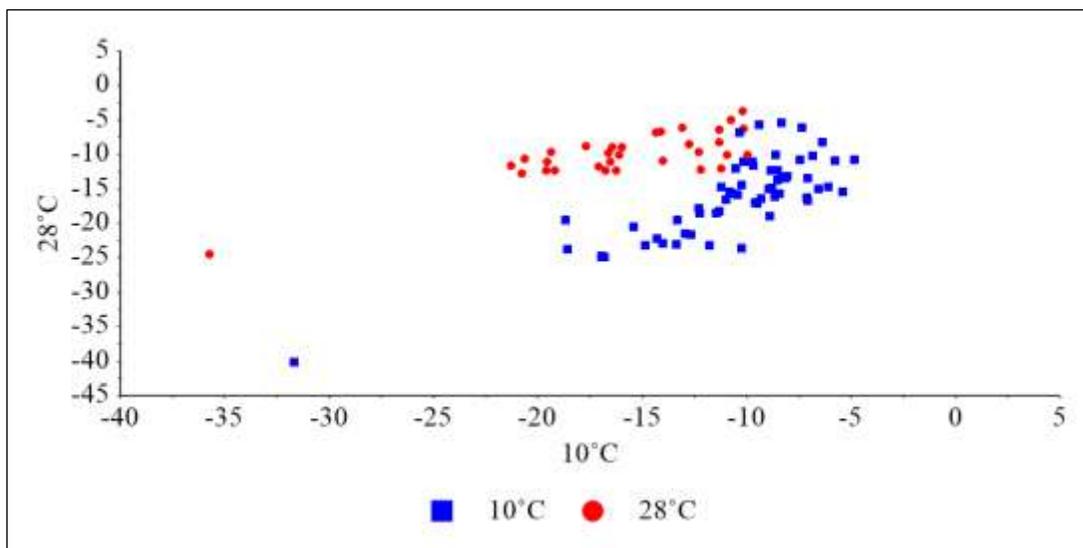

Figure 5. Linear Discriminant Analysis (LDA) of tomatoes based on storage temperature using SNV + GS9 1stD pretreatment spectra.

The classification of tomatoes based on storage temperature using the LDA method demonstrated excellent performance. The best model with SNV + GS9 1stD pretreatment achieved 92.55% accuracy in the calibration data, confirming the strong ability of the NIR spectrum to distinguish tomatoes stored at room and cold temperatures. The actual classification rates were 91.67% (10 °C) and 93.75% (28 °C), with low misclassification. In the prediction data, the accuracy slightly decreased to 89.13%; however, the model remained robust and reliable in differentiating the storage conditions. The classification values based on the storage temperature are presented in Table 3.

Table 3. Classification results of tomatoes based on the storage temperature.

	Actual	Predicted		Average Classification
		10°C	28°C	
Calibration	10°C	**93.34%**	6.67%	**92.55%**
	28°C	6.25%	**93.75%**	
Prediction	10°C	**89.65%**	10.35%	**89.13%**
	28°C	11.76%	**88.23%**	

The classification results were consistent with the independent t-test on SSC values of tomatoes stored at 28 °C (room) and 10 °C (cold), showing a p-value < 0.05, indicating a significant difference between the two storage conditions. The analysis used SSC data from days 0 and 7 to ensure a valid comparison. These findings confirm that storage temperature affects SSC changes during storage, aligning with Pathare and Al-Dairi (2021), who reported that cold storage better preserves SSC than room temperature. The average SSC comparison between the two temperatures is shown in Figure 6.
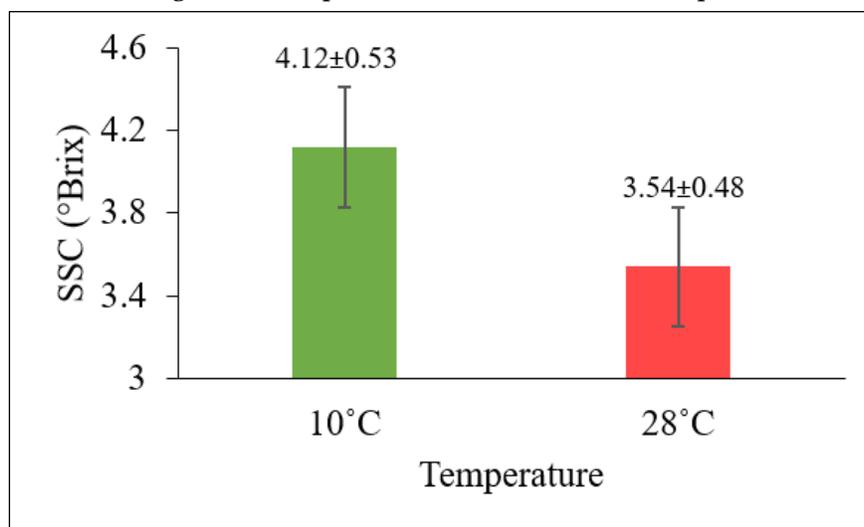

Figure 6. Comparison of average SSC values of tomatoes stored at different temperatures.

3.5.2 Classification Based on Ripeness Level

The best classification model was obtained using GS9 1stD pretreatment, which enhanced the spectral information by emphasizing small changes and reducing baseline effects. Although GS9 1stD improved spectral stability, the classification results still showed overlaps between classes, particularly between the breaker and pink ripeness stages. This is because these stages differ by only one level in the USDA classification, resulting in minor biochemical changes, such as sugar accumulation and chlorophyll degradation. Consequently, the SSC values and NIR spectra showed no significant distinction, making LDA unable to clearly separate the two stages. Figure 7 presents the best spectral separation between the breaker and pink tomatoes.
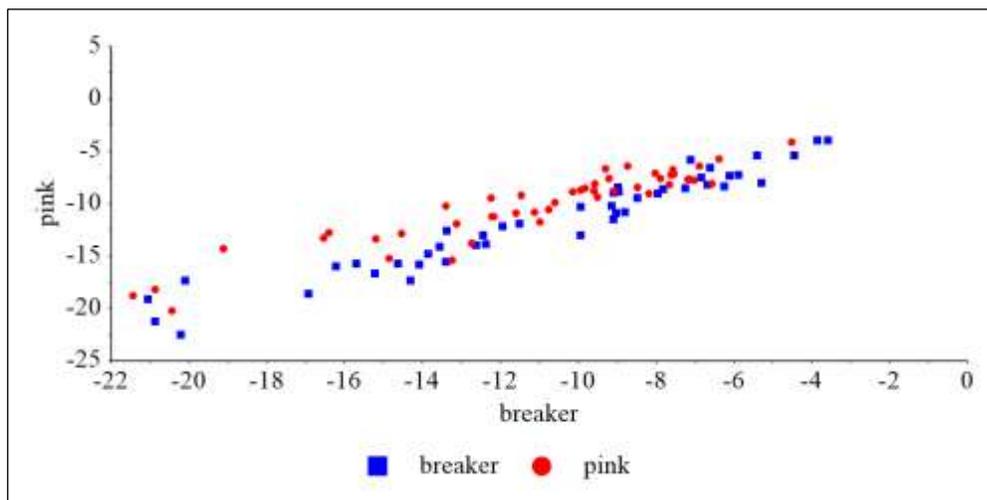
Figure 7. Linear Discriminant Analysis (LDA) of tomatoes based on ripeness level using GS9 1^stD pretreatment spectra.

Table 4 presents the classification results of the tomato ripeness levels (breaker and pink) using the LDA method. The calibration model achieved an accuracy of 80.43%, with true classification rates of 82.23% for breaker and 78.72% for pink, although the misclassification rates ranged between 17 and 21%. During the prediction stage, the accuracy decreased to 65.21%, mainly because of the reduced recognition in the breaker class. This indicates a spectral overlap between closely related ripeness stages. Overall, the LDA model shows potential but remains limited in distinguishing similar ripeness levels in tomatoes, suggesting the need for further model development.

Table 4. Results of tomato classification based on ripeness level.

	Actual	Predicted		Average Classification
		Breaker	Pink	
Calibration	Breaker	**82.23%**	17.77%	**80.43%**
	Pink	21.28%	**78.72%**	
Prediction	Breaker	**59.09%**	40.01%	**65.21%**
	Pink	29.17%	**70.83%**	

The classification results based on tomato ripeness levels were consistent with the t-test results for SSC values at the two ripeness stages, namely breaker and pink. Based on the independent t-test analysis, a p-value > 0.05 was obtained, indicating that there was no significant difference in SSC values between breaker and pink tomatoes. The data used in this analysis were SSC values obtained on the same observation day to ensure uniform storage and handling conditions. These results suggest that the transition from the breaker to the pink stage does not cause a significant change in the SSC content. Thus, the early ripening process is not accompanied by a marked increase in sugar

concentration or other soluble compounds. A comparison of SSC values between breaker and pink tomatoes is shown in Figure 8.
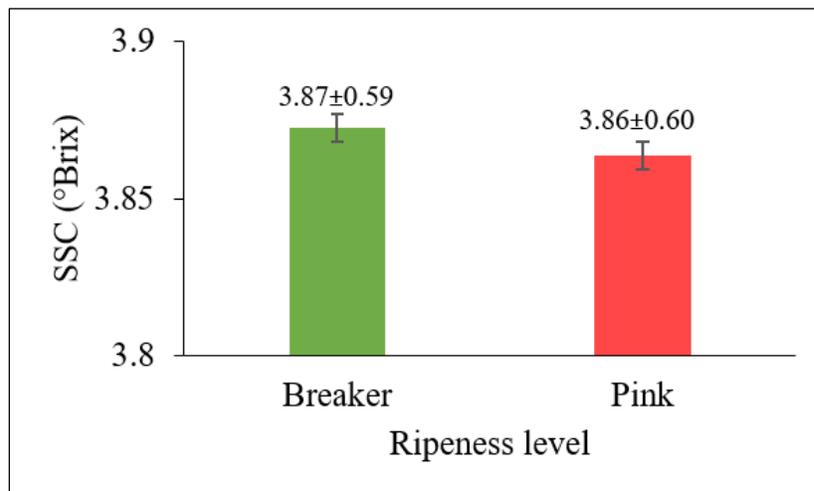

Figure 8. Comparison of the average SSC values of tomatoes at the breaker and pink ripeness levels.

4. Conclusion

The results showed that applying pretreatment to the NIR spectral data significantly improved the model quality. The best SSC prediction model, developed using PLS with SNV pretreatment (R^2 calibration = 0.89; R^2 prediction = 0.80; RPD = 2.04; RER = 8.08), demonstrated good quantitative prediction performance. SNV provided the best results among all pretreatments. Tomato classification based on storage temperature using LDA achieved high accuracy of 92.55% (calibration) and 89.13% (prediction), indicating effective differentiation between room and cold storage. In contrast, classification by ripeness showed lower accuracy (calibration: 80.43% (calibration) and 65.21% (prediction) owing to spectral overlap between the closely related breaker and pink stages.

5. References

Amr, A., & Raie, W. (2022). Tomato components and quality parameters. a review. Jordan Journal of Agricultural Sciences, 18(3), 199–220. https://doi.org/10.35516/jjas.v18i3.444

Andasuryani, Purwanto Aris, Y., Wayan Budiastra, I., & Syamsu, K. (2014). Prediksi kandungan katekin gambir (*Uncaria gambir* Roxb.) dengan spektroskopi NIR. Teknologi Industri Pertanian, 24(1), 43–52.

Bapary, M. S., Islam, M. N., Kumer, N., Tahery, M. H., Noman, M. A. Al, & Mohammed Mohi-Ud-Din. (2024). Postharvest physicochemical and nutritional properties of tomato fruit at different maturity stages affected by physical impact. Applied Food Research, 100636(4), 1–11.

Barbin, D. F., Kaminishikawahara, C. M., Soares, A. L., Mizubuti, I. Y., Grespan, M., Shimokomaki, M., & Hirooka, E. Y. (2015). Prediction of chicken quality attributes by near infrared spectroscopy. Food Chemistry, 168, 554–560. https://doi.org/10.1016/j.foodchem.2014.07.101

Beć, K. B., & Huck, C. W. (2019). Breakthrough potential in near-infrared spectroscopy: spectra simulation. A Review of Recent Developments. Frontiers in Chemistry, 7(FEB), 1–22. https://doi.org/10.3389/fchem.2019.00048

Cao, X., Li, L., Liu, Y., Zhou, X., Qiu, R., Dai, S., Wang, F., & Zong, X. (2025). NIRS combined with chemometrics for nondestructive quality monitoring of components during soymilk boiling process. Food Science and Technology, 229, 1–10. https://doi.org/10.1016/j.lwt.2025.118247

Collins, E. J., Bowyer, C., Tsouza, A., & Chopra, M. (2022). Tomatoes: an extensive review of the associated health impacts of tomatoes and factors that can affect their cultivation. Biology, 11(2), 1-44. https://doi.org/10.3390/biology11020239

Cortés, V., Blasco, J., Aleixos, N., Cubero, S., & Talens, P. (2019). Monitoring strategies for quality control of agricultural products using visible and near-infrared spectroscopy: A Review. Trends in Food Science & Technology, 85(1), 138–148. https://doi.org/10.1016/j.tifs.2019.01.015

Evi, T., & Rachbini, W. (2022). Partial Least Square (Teori dan Praktek) (1st ed.). Sukoharjo: Tahta Media Group.

Huang, Y., Li, Z., Bian, Z., Jin, H., Zheng, G., Hu, D., Sun, Y., Fan, C., Xie, W., & Fang, H. (2025). Overview of deep learning and nondestructive detection technology for quality assessment of tomatoes. Foods, 14(2), 1–33. https://doi.org/10.3390/foods14020286

Martín, J. F. G. (2022). Potential of near-infrared spectroscopy for the determination of olive oil quality. Sensors, 22(8), 1–26. https://doi.org/10.3390/s22082831

Maurel, V. B., Vigneau, J. L., & Sévila, F. (1994). Infrared and near-infrared technology for the food industry and agricultural uses: on-line applications. Food Control, 5(1), 21–27. https://doi.org/10.1016/0956-7135(94)90129-5

Munawar, A., Zulfahrizal, Z., & Syahrul. (2022). Agricultural products quality determination by means of near infrared spectroscopy. IOP Conference Series: Earth and Environmental Science, 1–8. https://doi.org/10.1088/1755-1315/951/1/012112

Najjar, K., & Abu-khalaf, N. (2021). Nondestructive quality measurement for three varieties of tomato using vis / nir spectroscopy. Sustainability, 13(19), 97-108.

Pathare, P. B., & Al-Dairi, M. (2021). Bruise damage and quality changes in impact-bruised, stored tomatoes. Horticulturae, 7(5), 1-20. https://doi.org/10.3390/horticulturae7050113

Polinar, Y. Q., Yaptenco, K. F., Peralta, E. K., & Agravante, J. U. (2019). Near-infrared spectroscopy for nondestructive prediction of maturity and eating quality of 'carabao' mango (*Mangifera indica* L.) fruit. Agricultural Engineering International: CIGR Journal, 21(1), 209–219.

Priyankara, G. G. D. S., Karunarathne, C. L. S. M., Sarananda, K. H., & Ariyaratne, M. (2017). Effect of maturity stage on ripening and quality characters of four tomato (*Solanum lycopersicum* L.) Varieties of Sri Lanka. Tropical Agricultural Research, 28(4), 496-502. https://doi.org/10.4038/tar.v28i4.8249

Quinet, M., Angosto, T., Yuste-Lisbona, F. J., Blanchard-Gros, R., Bigot, S., Martinez, J. P., & Lutts, S. (2019). Tomato fruit development and metabolism. Frontiers in Plant Science, 10(November), 1–23. https://doi.org/10.3389/fpls.2019.01554

Ribeiro, J. P. O., Medeiros, A. D. de, Caliari, I. P., Trancoso, A. C. R., Miranda, R. M. de, Freitas, F. C. L. de, Silva, L. J. da, & Dias, D. C. F. dos S. (2021). FT-NIR and linear discriminant analysis to classify chickpea seeds produced with harvest aid chemicals. Food Chemistry, 342(October 2020), 1-12. https://doi.org/10.1016/j.foodchem.2020.128324

Rinnan, Å., Berg, F. van den, & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. TrAC - Trends in Analytical Chemistry, 28(10), 1201–1222. https://doi.org/10.1016/j.trac.2009.07.007

Silalahi, D. D., Reaño, C. E., Lansigan, F. P., Panopio, R. G., & Bantayan, N. C. (2016). Linear discriminant analysis vs. genetic algorithm neural network with principal component analysis for hyperdimensional data analysis: A Study on Ripeness Grading of Pil Palm (*Elaeis guineensis* Jacq.) Fresh Fruit. Philippine Statistician, 65(2), 107–120.

Sun, D., Cruz, J., Alcalà, M., Romero del Castillo, R., Sans, S., & Casals, J. (2021). Near infrared spectroscopy determination of chemical and sensory properties in tomato. Journal of Near Infrared Spectroscopy, 29(5), 289–300. https://doi.org/10.1177/09670335211018759

Thole, V., Vain, P., & Martin, C. (2021). Effect of elevated temperature on tomato postharvest properties. Plants, 10, 2–18.

Yarce, C. J., & Rojas, G. (2012). Near infrared spectroscopy for the analysis of macro and micro nutrients in sugarcane leaves. Sugar Industry, 11(137), 707–710.

Zeaiter, M., & Rutledge, D. (2009). Comprehensive chemometrics chemical and biochemical data analysis (Vol. 3). Amsterdam: Elsevier Science. https://doi.org/10.1016/B978-044452701-1.00074-0