ASIAN JOURNAL OF AGRICULTURE
Volume 10, Number 1, June 2026
Pages: 9

E-ISSN: 2580-4537
https://doi.org/10.13057/asianjagric/g100139

Effect of sampling depth on °Brix prediction in melon (*Cucumis melo*) using Visible Near Infrared Spectroscopy (Vis-NIRS)

RAYZRAN LAKSAMANA WIRAWAN, NAFIS KHURIYATI♥, AGUNG PUTRA PAMUNGKAS, MUHAMMAD IQBAL RAMADANI

Department of Agroindustrial Technology, Faculty of Agricultural Technology, Universitas Gadjah Mada. Jl. Flora 1, Bulaksumur, Sleman 55281, Yogyakarta, Indonesia. Tel.: +62-274-551219, Fax: +62-274-551219, ♥email: nafis.khuriyati@ugm.ac.id

Manuscript received: 11 September 2025. Revision accepted: 19 March 2026.

Abstract. *Wirawan RL, Khuriyati N, Pamungkas AP, Ramadani MI. 2026. Effect of sampling depth on °Brix prediction in melon (*Cucumis melo*) using Visible Near Infrared Spectroscopy (Vis-NIRS). Asian J Agric 10 (1): g100139. https://doi.org/10.13057/asianjagric/g100139.* The accuracy of non-destructive melon sweetness (°Brix) prediction using Visible Near Infrared Spectroscopy (Vis-NIRS) depends strongly on the quality and representativeness of destructive reference measurements. However, standardized reference sampling depth for °Brix determination has not been clearly established. This study systematically evaluated the effect of reference sampling depth on the performance of Vis-NIRS-based prediction models in melon (*Cucumis melo*). Spectral data were collected from six surface points per fruit, while reference °Brix values were obtained from two different tissue depths to represent differences in internal chemical composition. Full flesh depth sampling represented the entire edible mesocarp from the outer flesh to the fruit center, while partial flesh depth sampling represented only the outer mesocarp close to the rind. Artificial neural network models were developed to compare predictive performance. The model calibrated with full-depth reference values achieved higher prediction accuracy (92.88%) and stronger correlation (R^2=0.75) than the model calibrated with partial-depth references (86.66%; R^2=0.50), with the difference statistically confirmed by Welch's t-test ($p<0.001$). These results demonstrate that aligning destructive reference measurements with fruit internal heterogeneity is essential for improving the reliability of non-destructive sweetness assessment. The findings provide practical guidance for developing more robust non-destructive quality evaluation protocols of melon and other horticultural crops with spatially heterogeneous internal structures.

Keywords: ºBrix prediction, Artificial Neural Network, fruit quality, melon, non-destructive testing

INTRODUCTION

Melon (*Cucumis melo*) is a widely cultivated horticultural crop valued for its distinctive flavor, high sweetness, and nutritional quality. The sweetness and nutrient content of this commodity are enhanced by phytochemicals, such as antioxidants and sugars (Chen et al. 2023). Organoleptic properties, including flavor, aroma, and texture, strongly influence fruit quality. Furthermore, sweetness plays a key role in determining market value and consumer preference. This factor is typically measured using Total Soluble Solids (TSS), expressed in degrees Brix (°Brix). The °Brix value reflects the concentration of soluble sugars such as sucrose, glucose, and fructose, which are the primary contributors to perceived sweetness in fruit. °Brix value is commonly measured with a refractometer due to simplicity and strong correlation with perceived sweetness (Magwaza and Opara 2015). However, this method is destructive, requiring slicing and juice extraction, which limits the efficiency for large-scale postharvest quality control.

Visible Near Infrared Spectroscopy (Vis-NIRS) is a promising method used for detecting physicochemical properties by measuring the interaction of light with biological tissues, without damaging the fruit. This method operates in the 400-1000 nm spectral range and captures absorbance patterns related to sugars, water, and tissue structure (Rady et al. 2021). It offers advantages of speed, non-invasiveness, and the ability to assess multiple quality parameters, making it highly suitable for precision agriculture and postharvest evaluation (Wieme et al. 2022). The effectiveness has been reported in pomegranate (Khodabakhshian et al. 2017), strawberry (Shen et al. 2018), and cherry (Shao et al. 2019). However, these studies primarily focused on predictive performance using conventional reference sampling approaches and did not explicitly evaluate how differences in tissue sampling depth influence calibration reliability. Vis-NIRS combined with machine learning algorithms such as Artificial Neural Networks (ANNs) is used to develop predictive models that correlate spectral data with reference parameters, including sweetness, where the accuracy of the models depends heavily on the quality and consistency of reference data derived through destructive testing. In this context, the interaction between light and fruit tissue in the Vis-NIRS region is governed by both chemical composition and optical properties of the tissue, such as absorption and scattering, as well as interactions with cellular structures, which together influence light propagation and penetration depth (Rodríguez-Ortega et al. 2023).

Several studies have developed non-destructive °Brix prediction models; however, most rely on reference values obtained from juice extraction, and do not account for the fruit's internal spatial heterogeneity (Guthrie et al. 2006;

Hadiwijaya et al. 2020). In melon, soluble solids are unevenly distributed, with higher sugar concentrations in the inner tissues than near the rind (Peiris et al. 1999), while the effective Vis-NIRS sampling depth is biased toward near-surface tissues, even though longer NIR wavelengths may penetrate deeper depending on optical properties and measurement configuration (Escribano et al. 2017; Walsh et al. 2020). This mismatch between the dominant optical sensing region and the reference sampling depth can weaken model calibration and reduce the representativeness of predicted °Brix values. Such discrepancies between optical sensing depth and destructive reference sampling may introduce systematic bias during model calibration, leading to reduced prediction accuracy and model generalizability.

Although incorporating spatially diverse spectral inputs has been reported to improve model robustness (Suhandy and Yulia 2017; Arendse et al. 2018), the combined consideration of multi-point spectral acquisition and depth-aligned reference sampling remains limited, including in Hadiwijaya et al. (2020) and Kasih et al. (2024). In contrast to previous studies that primarily emphasize spectral modeling strategies, the present study systematically evaluates the influence of reference sampling depth on non-destructive °Brix prediction performance. Furthermore, most previous studies have focused primarily on optimizing spectral preprocessing techniques and model architectures, with limited attention to the physiological representativeness of the reference data, which represents a critical research gap in aligning optical sensing depth with destructive reference sampling.

By comparing partial and full-flesh-depth sampling in combination with six-point spectral acquisition on the melon surface, this study examines how reference representativeness affects ANN-based model accuracy and reliability. The findings provide methodological guidance for improving the robustness and practical applicability of handheld Vis–NIRS systems for non-destructive sweetness assessment in melon. The results of this study are also expected to support the development of more reliable non-destructive quality evaluation systems for horticultural commodities with heterogeneous internal structures, particularly in industrial grading and precision agriculture applications.

MATERIALS AND METHODS

Fruit material

A total of 75 Honey Globe melons (*C. melo* var. *inodorus*) were harvested from the Fablab, Universitas Gadjah Mada Greenhouse, Kulon Progo, Special Region of Yogyakarta, Indonesia. Fruits were sampled at 71-76 days after pollination, corresponding to commercial ripeness.

Spectral measurements

Spectral measurements were performed using a handheld Vis-NIR spectrophotometer (F-750 Produce Quality Meter, Felix Instruments, USA) over the wavelength range 429-1044 nm. Sample preparation consisted of marking a portion of the fruit for spectral acquisition, with six points spread across the top, middle, and bottom. This number of acquisition points was selected to adequately represent internal heterogeneity, considering the approximately spherical fruit geometry and the limited penetration depth of Vis-NIRS, which primarily captures information from superficial to sub-surface tissue layers. Each spectral acquisition corresponded to a destructively measured reference value. Spectral acquisition points are shown in Figure 1.

TSS measurements

Total Soluble Solids (TSS) were measured destructively using a digital refractometer after spectral data acquisition. All measurements were performed on the same fruits used for Vis-NIRS analysis.

The extraction procedure was adapted from the juice-based reference approach described by Guthrie et al. (2006). However, unlike the original study, which used a single homogenized reference strategy for calibration, the present study systematically compared two sampling depths to evaluate the effect of reference-spectral depth alignment on prediction performance. Two reference sampling strategies were therefore applied: (i) Partial reference sampling, representing the outer mesocarp region near the rind (approximately 1-1.5 cm depth from the exocarp). (ii) Full reference sampling, representing the entire edible mesocarp from the rind to the inner flesh near the seed cavity. The anatomical regions of the melon fruit and the distinction between partial and full reference sampling depths are illustrated in Figure 2.

This comparison was motivated by the known radial gradient of soluble solids in melon, in which sugar concentration increases toward the inner mesocarp (Mori et al. 2019). Such internal heterogeneity may affect optical absorption and scattering during spectral acquisition (McGlone and Kawano 1998), potentially reducing calibration reliability when the depth of reference sampling does not reflect the effective sensing depth. For each sampling strategy, pulp tissue was homogenized to obtain juice for refractometric analysis. All measurements were conducted in duplicate, and the mean value was used as the reference target for model development.

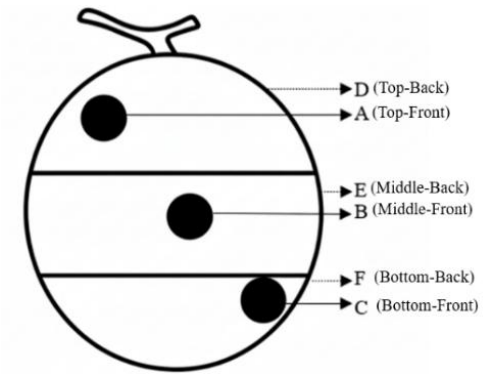

Figure 1. Distribution of six spectra acquisition points across the top, middle, and bottom of the melon

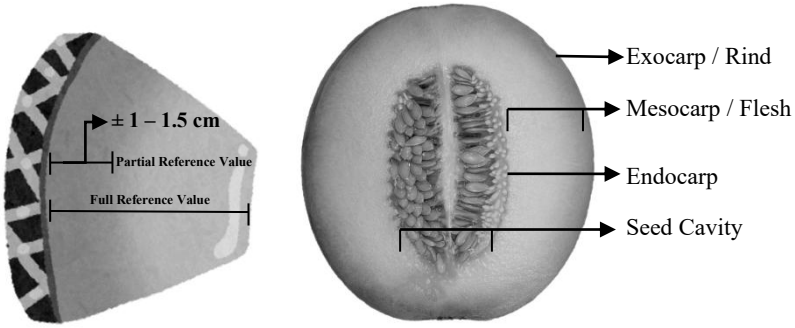

Figure 2. Morphological traits of melon fruit assessed for TSS measurement

ANN training

The dataset obtained from Vis-NIRS spectral acquisition and TSS reference measurement was used to develop the ANN model. A total of 75 melon fruits were used, with each fruit measured at 6 points, yielding 450 paired spectral and TSS data for model development. All data preprocessing and ANN training were performed using MATLAB R2023a (MathWorks Inc., USA). Absorbance values between 429 and 1044 nm, measured at 3 nm intervals, were used as input variables, resulting in 206 input nodes. The output variable corresponded to the TSS value measured destructively at each respective spectral acquisition point.

ANN models were trained using the Levenberg-Marquardt backpropagation algorithm (the trainlm function) in the MATLAB Neural Network Toolbox. The network architecture, including the number of hidden layers and neurons, activation functions (purelin, logsig, and tansig), learning rate, and number of training epochs, was determined through a structured trial-and-error approach. In this process, several architectures were tested by varying the number of hidden neurons (from 5 to 40) and hidden layers (one to two layers). Different combinations of activation functions were applied. The learning rate was explored over 0.1-0.9, and the number of epochs ranged from 1000 to 9000 iterations. The optimal structure was selected based on model performance, emphasizing the lowest Mean Squared Error (MSE) and the highest coefficient of determination (R^2) across training, validation, and testing datasets. Initial weights and biases were randomly assigned and iteratively updated during training.

Although systematic hyperparameter optimization techniques such as grid search or Bayesian optimization can provide automated architecture tuning, these approaches were not implemented in the present study. The primary objective of this research was to evaluate the influence of reference sampling depth on prediction performance rather than to perform exhaustive hyperparameter optimization. Given the moderate dataset size (n=450) and the relatively simple network structure of a single hidden layer in the final model, a structured empirical exploration was deemed sufficient to yield a stable and representative architecture. Moreover, applying systematic optimization procedures would substantially increase computational complexity without necessarily providing meaningful additional insight into the core experimental variable. Therefore, the selected approach ensured methodological consistency while maintaining focus on the comparative evaluation of reference sampling strategies.

The trial-and-error procedure for optimizing ANN architecture has been successfully applied in various studies related to fruit quality prediction using spectral data (Amoriello et al. 2022). Network weights and biases were randomly initialized and iteratively updated during training to minimize prediction error. Throughout the training process, optimal weight and bias parameters were obtained for each neuron and layer configuration (Darvishi et al. 2017).

Data analysis

Error analysis and model reliability ANN prediction

ANN prediction model error was analyzed by calculating MSE to determine the smallest error, i.e., the lowest average. This parameter is a general measure of the discrepancy between the model's predicted and observed values. The smaller the MSE, the closer the observed value is to the predicted value. ANN training produces an R^2 value that indicates the model's reliability and its ability to predict melon sweetness level. Furthermore, R^2 is the square of the correlation between the observed and predicted value "y". The model formed is better when R^2 is close to 1.

Accuracy level analysis and application of Welch's t-test

The accuracy of the two models was compared by analyzing observed (ºBrix) against predicted TSS values. The difference and the percentage of the accuracy levels were calculated to obtain an average. This test aimed to determine the accuracy of the development of melon ºBrix prediction models. The formula for calculating accuracy was as follows (Deepthi et al. 2017).

$$\%\ Error = \frac{|predicted\ value - measured\ value|}{measured\ value}\ x\ 100\%$$

$$\%\ Accuracy = 100\% - \%error$$

An independent-samples Welch's t-test was used to compare the mean prediction errors between the ANN models trained with full- and partial-reference values. The

two datasets were treated as independent because each model was calibrated using different reference targets representing distinct tissue depths. Preliminary analysis indicated unequal variances between groups; therefore, Welch's t-test was selected as it does not assume homogeneity of variance and provides a more robust comparison under heteroscedastic conditions (Slaughter and Abbott 2004). In addition, the two reference datasets exhibited unequal variances, with the partial reference group displaying greater variability due to the inherently non-uniform distribution of soluble solids in outer tissue layers. Because the assumption of equal variances required by the classical Student's t-test and one-way ANOVA was not met, Welch's t-test was selected as the most appropriate method. Welch's t-test is specifically recommended for comparing two independent groups with heteroscedasticity, as it provides more accurate Type I error control and greater robustness than the standard t-test (Ruxton 2006; Delacre et al. 2017). Therefore, Welch's t-test offered the most statistically sound approach for evaluating how different reference sampling depths influence ANN prediction performance in this study. The t-test statistic was calculated using the following formula:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Where, $\bar{x}_1$ and $\bar{x}_2$ are the means, s_1 and s_2 are the standard deviations, and n_1 and n_2 are the sample sizes of the full and partial reference values, respectively. The analysis was conducted to test the hypothesis that the depth of tissue used in TSS measurement affected the resulting values. A statistically significant difference was inferred when the p-value was less than 0.05.

RESULTS AND DISCUSSION

Spectral acquisition

Each spectral acquisition point was measured three times to minimize instrumental and positioning errors, and the repeated measurements were averaged to obtain a representative spectrum for each sampling point. A total of 450 spectra were collected from 75 melons, with six acquisition points per fruit. Figure 3 presents the averaged reflectance spectra derived from all measurements. The spectral curves exhibited consistent patterns across samples, indicating stable acquisition conditions and reproducible signal quality. A prominent absorbance feature was observed around 680 nm, while additional absorbance regions were detected near 480 nm and 978 nm. Moderate variation among samples was also observed within the wavelength range of 740-840 nm. Although the overall spectral shapes were similar, differences in absorbance intensity were observed among samples with varying TSS values. These variations suggest that the selected visible and near-infrared wavelength range contains information relevant to internal fruit characteristics and may contribute to TSS prediction modeling.

TSS measurement

Table 1 presents the descriptive statistics of TSS measured at two sampling depths, full reference and partial reference. The full reference values showed a higher mean TSS (12.62 °Brix) compared to the partial reference values (10.71 °Brix). The standard deviation was 2.38 for the full reference and 2.53 for the partial reference, indicating slightly greater variability in the partial sampling depth. These results demonstrate that sampling depth influences the measured TSS values and affects the representativeness of sweetness assessment.

Developing ºBrix value prediction model

In this research, 450 absorbance data points were split into training, validation, and testing data. Data were split according to general proportions used in the prediction model development, 70% (n=315), 15% (n=68), and 15% (n=67) for training, validation, and testing, respectively. This scheme aims to maintain the model's stability and ensure it can generalize to new data. Validation data is used to monitor the model's performance during training and to prevent overfitting, which occurs when the model memorizes the training data and loses the ability to generalize to unseen information. These data allow the training process to be halted once model performance declines, enabling the final model to balance accuracy and generalization (Kamilaris and Prenafeta-Boldú 2018). ºBrix value prediction model was developed using full and partial reference values. Each prediction model was built with the same data split. Table 2 reports the distribution of data used in developing the melon fruit ºBrix value prediction model.

Selecting an optimal ANN architecture is an important step in developing a reliable prediction model. As stated by Díaz et al. (2017), several parameters need to be set in advance, including the network architecture and various hyperparameters. There are no established rules or guidelines for manually selecting parameters. The success of the training process largely depends on selecting appropriate values. Therefore, effective hyperparameter optimization methods are necessary to determine optimal values and automatically improve model performance. In this research, architecture exploration is conducted by comparing the performance of models built using both types of reference values. ANN architecture results indicate differences in network complexity based on reference value type.

The optimal ANN architecture for the full-reference model uses a 206-5-1 structure. Both ANN models were developed with 206 input nodes, corresponding to absorbance values measured across the 429-1044 nm spectral range. The optimal ANN architecture for the full-reference model was 206-5-1, while the partial-reference model required a more complex 206-15-1 structure. The hidden layer contained 5 or 15 neurons depending on the model, and the output layer consisted of a single neuron predicting the ºBrix value. The simpler architecture in the

full-reference model indicates that the reference dataset better reflected the fruit's internal sugar distribution, enabling the ANN to achieve good performance with fewer neurons. In contrast, the partial-reference model, which captured only localized tissue information, exhibited higher variability and therefore required a larger number of neurons to approximate the underlying nonlinear patterns. The hidden layer consists of 5 or 15 neurons, depending on the model, and the output layer consists of a single neuron that predicts the ºBrix value. The more complex architecture of the partial model reflects the need for the network to cope with high local data variation, which does not fully represent the overall fruit sugar content profile. To recognize nonlinear patterns in this partial-reference-value data, a large number of hidden-layer neurons is required. This difference suggests that the full-reference dataset was more representative of the actual soluble solids distribution in melon tissues, allowing accurate prediction with a simpler architecture. The partial-reference dataset, which captured only localized tissue information, exhibited greater variability and required a more complex network to approximate the underlying patterns. Detailed technical parameters of the ANN models, including activation functions, epochs, and learning rates, are provided in Table 3. At the same time, the corresponding ANN structure diagrams are presented in Supplementary Figures 4 and 5.

Table 1. Reference values of melon

Melon flesh depth	Number of data (n)	Total Soluble Solids (ºBrix)			
		Min	Max	Mean	Std. Deviation
Full Reference Values	450	3.70	18.35	12.62	2.38
Partial Reference Values	450	3.48	17.22	10.71	2.53

Table 2. Comparison data distribution of reference values

Melon flesh depth	Total Soluble Solid (ºBrix)														
	Train					Validation					Testing				
	Number of data	Min	Max	Mean	Std. dev.	Number of data	Min	Max	Mean	Std. dev.	Number of data	Min	Max	Mean	Std. dev.
Full reference value	315	3.71	18.35	12.61	2.35	68	3.94	18.25	12.66	2.39	67	6.40	17.00	12.62	2.21
Partial reference value	315	3.48	17.22	10.71	2.54	68	3.60	16.56	10.75	2.58	67	5.95	15.95	10.70	2.42

Table 3. Comparison data distribution statistics of reference values

Melon flesh depth	Total Soluble Solid (ºBrix)														
	Train					Validation					Testing				
	Number of data	Min	Max	Mean	Std. dev.	Number of data	Min	Max	Mean	Std. dev.	Number of data	Min	Max	Mean	Std. dev.
Full reference value	315	3.71	18.35	12.61	2.35	68	3.94	18.25	12.66	2.39	67	6.40	17.00	12.62	2.21
Partial reference value	315	3.48	17.22	10.71	2.54	68	3.60	16.56	10.75	2.58	67	5.95	15.95	10.70	2.42

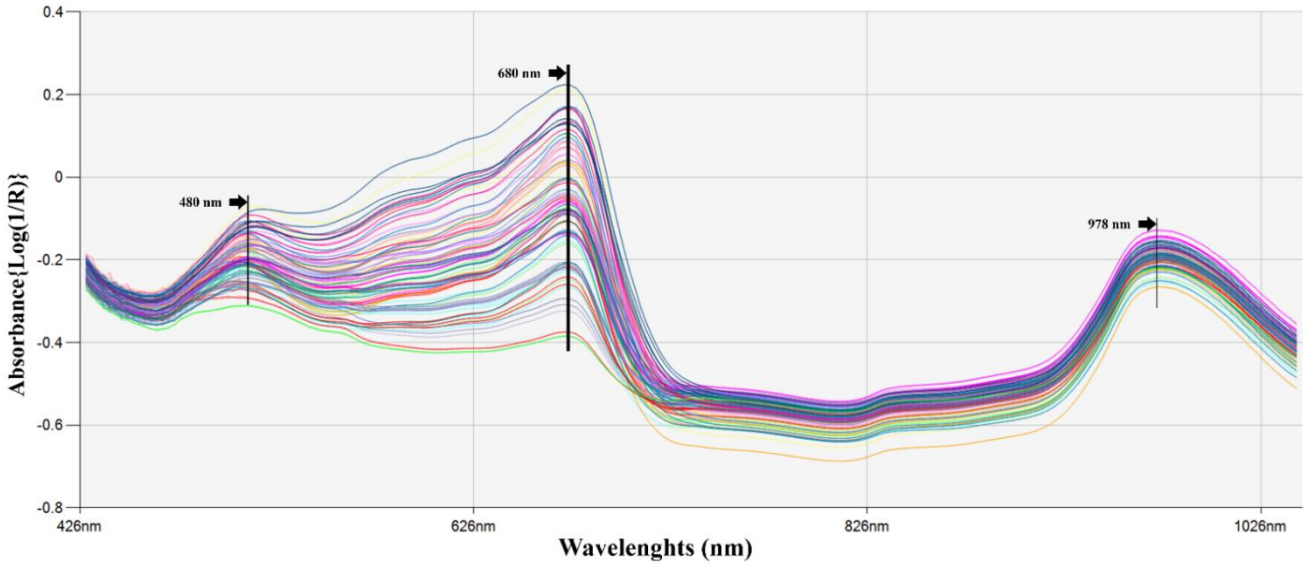

Figure 3. Averaged Vis-NIRS absorbance spectra of melon fruits (429-1044 nm) used for °Brix prediction modeling

Prediction model performance evaluation

The performance of the prediction models was evaluated using MSE, R^2, and average prediction accuracy, as summarized in Table 4. Overall, the model calibrated with full-reference values consistently outperformed the partial-reference model across the training, validation, and test datasets. It showed lower prediction error and stronger correlation between predicted and observed ºBrix values, indicating better generalization capability. The full-reference model achieved higher overall accuracy (92.88%) compared to the partial-reference model (86.66%), confirming that broader tissue representation improved predictive reliability. These results demonstrate a clear performance trend favoring full-depth reference sampling, while detailed numerical comparisons are provided in Table 3. The difference between the two models was statistically evaluated using Welch's t-test. The analysis yielded a t-value of 11.67 with a p-value <0.001, indicating a statistically significant difference in predictive performance between the two reference sampling strategies.

Discussion

The spectral acquisition results confirm that absorbance patterns in the 429-1044 nm range are closely associated with key biochemical constituents that influence melon sweetness. The dominant absorbance peak at approximately 680 nm corresponds to chlorophyll absorption, which is retained in rind-adjacent tissues and reflects physiological maturity (Zeb et al. 2021; Ustin and Jacquemoud 2020). Additional absorption features in the 740-760 nm and ~840 nm regions are associated with O-H and C-H molecular bonds characteristic of water and soluble carbohydrates, respectively (Walsh et al. 2020). Although multiple chemical constituents contribute to absorbance within this spectral region, this wavelength range is particularly relevant for °Brix prediction because it encompasses key absorption features associated with water, carbohydrates, and other tissue components that modulate light-tissue interactions. The main soluble carbohydrates in melon include glucose, sucrose, and fructose, whose concentrations directly determine fruit sweetness (Alves et al. 2019). The absorbance peak near 978 nm further indicates strong water-related absorption. The peak around 480 nm is associated with carotenoid pigments such as β-carotene and lutein, which influence light scattering within the mesocarp tissue (Ustin and Jacquemoud 2020). Collectively, these spectral features demonstrate that the Vis-NIRS wavelength range employed captures relevant information on water status, carbohydrate composition, and tissue structure, which are essential for °Brix prediction (Nicolaï et al. 2007; Zeb et al. 2021). The presence of overlapping absorption features explains the nonlinear relationship between spectral data and °Brix values, thereby justifying the application of artificial neural networks for model development (Kamilaris and Prenafeta-Boldú 2018).

Table 4. Comparison of prediction models based on MSE and R^2

Melon flesh depth	Mean Squared Error (MSE)				Coefficient of Determination (R^2)			
	Train	Val	Test	All	Train	Val	Test	All
Full Reference Values	0.77	2.81	2.81	1.39	0.86	0.53	0.48	0.75
Partial Reference Values	2.95	4.78	4.79	3.30	0.55	0.33	0.42	0.50

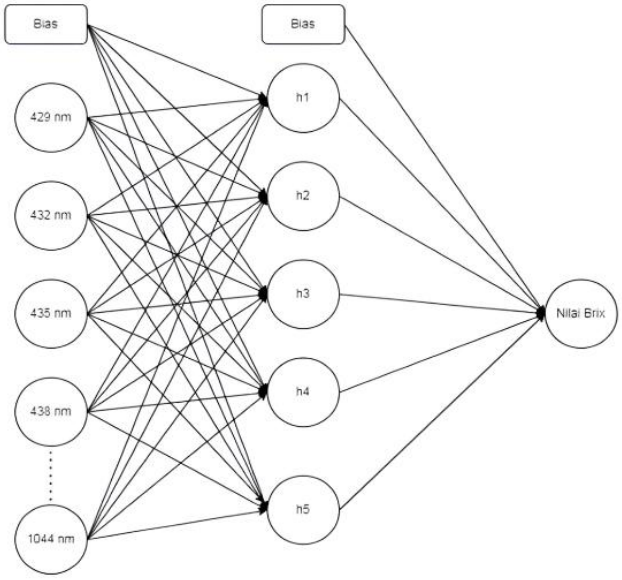

Figure 4. ANN structure for full reference values

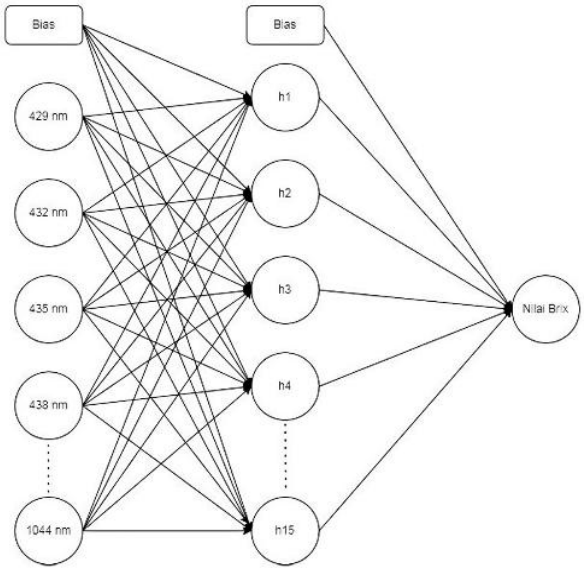

Figure 5. ANN structure for partial reference values

Variability in TSS measurements highlights the importance of reference sampling depth. Previous studies have consistently reported a pronounced radial gradient in melon sugar content, with higher soluble solids toward the inner mesocarp, making localized or shallow sampling less representative of overall sweetness (Peiris et al. 1999; Liu et al. 2019; Lan et al. 2021). Such internal heterogeneity has been identified as a major source of error in destructive and non-destructive quality assessment when reference measurements rely on localized sampling (García-Villela et al. 2020). In agreement with these findings, full-reference measurements integrating the entire edible mesocarp yielded higher mean TSS values than partial-reference measurements obtained from outer tissues alone. The greater variability observed in partial reference values indicates increased sensitivity to tissue heterogeneity and reduced representativeness (Peiris et al. 1999; Mori et al. 2019). Partial reference values tend to underestimate internal °Brix levels due to the radial sugar gradient and limited sampling depth.

The influence of reference sampling depth was clearly reflected in the structure and performance of the ANN models. Although both models were trained on identical spectral datasets, the ANN calibrated with full-reference values converged to a simpler architecture (206-5-1) than the model trained with partial reference values (206-15-1). This difference indicates that the full reference dataset provided a more consistent and representative target variable, enabling the ANN to learn underlying nonlinear relationships with fewer hidden neurons (Díaz et al. 2017; Barbedo 2018). In contrast, higher variability and weaker correspondence between spectral inputs and partial reference targets required increased model complexity, elevating the risk of overfitting and reducing interpretability (Barbedo 2018).

Performance metrics further confirmed the superiority of the full-reference approach, with lower mean squared error and higher coefficients of determination across the training, validation, and test datasets. These findings are consistent with previous reports indicating that mismatches between optical sensing depth and reference sampling depth degrade model performance (Guthrie et al. 2006; Hadiwijaya et al. 2020). Because Vis-NIRS penetrates several millimeters to approximately one centimeter into melon tissue, reference values that integrate deeper mesocarp regions provide a closer physiological match to the tissue volume contributing to the spectral signal (Nicolaï et al. 2007; Walsh et al. 2020). Partial reference values, which primarily reflect outer tissue composition, weaken the correspondence between spectral data and reference measurements.

Statistical analysis confirmed that the reference sampling depth significantly influenced model performance. Welch's t-test was applied due to unequal variances and independence between datasets (Ruxton 2006; Delacre et al. 2017), demonstrating a statistically significant difference between prediction errors. Effect size analysis using Cohen's d indicated a small to moderate practical advantage of the full reference model (Kim 2015; Wilcox 2016; Rachman et al. 2017; Suratmi et al. 2022).

In conclusion, this study demonstrates that destructive reference sampling depth plays a decisive role in determining the robustness of non-destructive °Brix prediction models based on Vis-NIRS and artificial neural networks. Although both models were developed using identical spectral datasets, differences in reference sampling strategy substantially influenced calibration consistency, model complexity, and predictive performance. Integrating full mesocarp reference values provided a more representative target variable, facilitating more efficient nonlinear learning and improved alignment between spectral responses and chemical composition. These findings highlight that calibration robustness in optical quality assessment systems depends not only on the

model architecture but also on the physiological compatibility between optical penetration depth and reference sampling design. Aligning these components reduces variability arising from internal compositional gradients and enhances predictive reliability. Consequently, this work provides a methodological framework for improving non-destructive prediction of sweetness in horticultural commodities that exhibit internal heterogeneity.

Several limitations should be considered. The study focused on a single cultivar grown under controlled environmental conditions, which may restrict generalizability across genotypes and field environments. Although the dataset was sufficient for model development, larger multi-season and multi-location datasets would strengthen external validity. Furthermore, inherent biological heterogeneity and optical penetration constraints may limit maximum achievable prediction accuracy. Future research should therefore extend validation across multiple cultivars, environmental conditions, and production systems. From an applied perspective, incorporating representative full-depth reference sampling during calibration may enhance the robustness of industrial quality grading systems. At the same time, sensor development strategies should account for tissue penetration characteristics to ensure physiological compatibility between optical signals and chemical reference values.

ACKNOWLEDGEMENTS

The authors are grateful to the Universitas Gadjah Mada, Indonesia, for Research Grant Number 5286/UN1.P1/PT.01.03/2024 under the Final Project Recognition Program. This funding was crucial to completing this research. Furthermore, the authors are grateful to the Greenhouse Fablab Jogja Field Research Center (FRC) Universitas Gadjah Mada, for providing access to its facilities.

REFERENCES

Alves LC, Llerena JPP, Mazzafera P, Vicentini R. 2019. Diel oscillations in cell wall components and soluble sugars as a response to short-day in sugarcane (*Saccharum* sp.). BMC Plant Biol 19: 215. https://doi.org/10.1186/s12870-019-1837-4.

Amoriello T, Ciccoritti R, Ferrante P. 2022. Prediction of strawberries' quality parameters using artificial neural networks. Agronomy 12 (4): 963. https://doi.org/10.3390/agronomy12040963.

Arendse E, Fawole OA, Magwaza LS, Opara UL. 2018. Non-destructive prediction of internal and external quality attributes of fruit with thick rind: A review. J Food Eng 217: 11-23. https://doi.org/10.1016/j.jfoodeng.2017.08.009.

Barbedo JGA. 2018. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. Comput Electron Agric 153: 46-53. https://doi.org/10.1016/j.compag.2018.08.013.

Chen H, Cheng J, Huang Y, Kong Q, Bie Z. 2023. Comparative analysis of sugar, acid, and volatile compounds in CPPU-treated and honeybee-pollinated melon fruits during different developmental stages. Food Chem 410: 134072. https://doi.org/10.1016/j.foodchem.2022.134072.

Darvishi E, Khotanlou H, Khoubi J, Giahi O, Mahdavi N. 2017. Prediction effects of personal, psychosocial, and occupational risk factors on low back pain severity using artificial neural networks approach in industrial workers. J Manipulative Physiol Ther 40: 486-493. https://doi.org/10.1016/j.jmpt.2017.03.012.

Deepthi S, Rajesh B, Vyshnavi N, Moni K. 2017. Automatic key term extraction from research article using hybrid approach. Intl J Comput Appl 166 (6): 17-21. https://doi.org/10.5120/ijca2017914039.

Delacre M, Lakens D, Leys C. 2017. Why psychologists should by default use Welch's t-test instead of Student's t-test. Intl Rev Soc Psychol 30: 92-101. https://doi.org/10.5334/irsp.82.

Díaz GI, Fokoue-Nkoutche A, Nannicini G, Samulowitz H. 2017. An effective algorithm for hyperparameter optimization of neural networks. IBM J Res Dev 61: 9:1-9:11. https://doi.org/10.1147/JRD.2017.2709578.

Escribano S, Biasi WV, Lerud R, Slaughter DC, Mitcham EJ. 2017. Non destructive prediction of soluble solids and dry matter content using NIR spectroscopy and its relationship with sensory quality in sweet cherries. Postharvest Biol Technol 128: 112-120. https://doi.org/10.1016/j.postharvbio.2017.01.016.

García-Villela KM, Preciado-Rangel P, Sifuentes-Ibarra E, Salas-Pérez L, Núñez-Ramírez F, González-Fuentes JA. 2020. Ecological nutrient solutions on yield and quality of melon fruits. Terra Latinoam 38: 39-44. https://doi.org/10.28940/terra.v38i1.527.

Guthrie JA, Liebenberg CJ, Walsh KB. 2006. NIR model development and robustness in prediction of melon fruit total soluble solids. Aust J Agric Res 57: 411-418. https://doi.org/10.1071/AR05123.

Hadiwijaya Y, Kusumiyati K, Munawar AA. 2020. Prediksi total padatan terlarut buah melon golden menggunakan Vis-SWNIRS dan analisis multivariat. Jurnal Penelitian Saintek 25: 103-114. https://doi.org/10.21831/jps.v25i2.34487. [Indonesian]

Kamilaris A, Prenafeta-Boldú FX. 2018. Deep learning in agriculture: A survey. Comput Electron Agric 147: 70-90. https://doi.org/10.1016/j.compag.2018.02.016.

Kasih AN, Khuriyati N, Falah MAF. 2024. Prediksi fisikokimia melon (*Cucumis melo* L.) secara non-destruktif dengan impuls akustik dan jaringan saraf tiruan. Agrointek: Jurnal Teknologi Industri Pertanian 18 (3): 733-741. https://doi.org/10.21107/agrointek.v18i3.21746. [Indonesian]

Khodabakhshian R, Emadi B, Khojastehpour M, Golzarian MR, Sazgarnia A. 2017. Non-destructive evaluation of maturity and quality parameters of pomegranate fruit by visible/near infrared spectroscopy. Intl J Food Prop 20: 41-52. https://doi.org/10.1080/10942912.2015.1126725.

Kim TK. 2015. T test as a parametric statistic. Kor J Anesthesiol 68: 540-546. https://doi.org/10.4097/kjae.2015.68.6.540.

Lan W, Jaillais B, Renard CMGC, Leca A. 2021. A method using near infrared hyperspectral imaging to highlight the internal quality of apple fruit slices. Postharvest Biol Technol 175: 111497. https://doi.org/10.1016/j.postharvbio.2021.111497.

Liu W, Qi S, Lu J, Han D. 2019. Non-destructive measurement of soluble solids content of three melon cultivars using portable visible/near infrared spectroscopy. Biosyst Eng 188: 31-39. https://doi.org/10.1016/j.biosystemseng.2019.10.003.

Magwaza LS, Opara UL. 2015. Analytical methods for determination of sugars and sweetness of horticultural products: A review. Sci Hortic 184: 179-192. https://doi.org/10.1016/j.scienta.2015.01.001.

McGlone VA, Kawano S. 1998. Firmness, dry-matter and soluble-solids assessment of postharvest kiwifruit by NIR spectroscopy. Postharvest Biol Technol 13: 131-141. https://doi.org/10.1016/S0925-5214(98)00007-6.

Mori K, Beauvoit BP, Biais B, Chabane M, Allwood JW, Deborde C, Maucourt M, Goodacre R, Cabasson C, Moing A, Rolin D, Gibon Y. 2019. Central metabolism is tuned to the availability of oxygen in developing melon fruit. Front Plant Sci 10: 594. https://doi.org/10.3389/fpls.2019.00594.

Nicolaï BM, Beullens K, Bobelyn E, Peirs A, Saeys W, Theron KI, Lammertyn J. 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. Postharvest Biol Technol 46: 99-118. https://doi.org/10.1016/j.postharvbio.2007.06.024.

Peiris KHS, Dull GG, Leffler RG, Kays SJ. 1999. Spatial variability of soluble solids or dry-matter content within individual fruits, bulbs, or tubers: Implications for the development and use of NIR spectrometric techniques. HortScience 34: 114-118. https://doi.org/10.21273/HORTSCI.34.1.114.

Rachman A, Irawan S, Suastika I. 2017. Soil quality indicators of reclaimed mine soils. J Sumber Daya Lahan 11: 1-10. https://doi.org/10.2018/jsdl.v11i1.8185.

Rady A, Guyer DE, Watson N. 2021. Near-infrared spectroscopy and hyperspectral imaging for sugar content evaluation in potatoes over multiple growing seasons. Food Anal Methods 14: 581-595. https://doi.org/10.1007/s12161-020-01886-1.

Rodríguez-Ortega A, Aleixos N, Blasco J, Albert F, Munera S. 2023. Study of light penetration depth of a Vis-NIR hyperspectral imaging system for the assessment of fruit quality. A case study in persimmon fruit. J Food Eng 358 (2023): 111673. https://doi.org/10.1016/j.jfoodeng.2023.111673.

Ruxton GD. 2006. The unequal variance t-test is an underused alternative to the Student's t-test and the Mann-Whitney U test. Behav Ecol 17 (4): 688-690. https://doi.org/10.1093/beheco/ark016.

Shao Y, Xuan G, Hu Z, Gao Z, Liu L. 2019. Determination of the bruise degree for cherry using Vis-NIR reflection spectroscopy coupled with multivariate analysis. PLoS One 14: e0222633. https://doi.org/10.1371/journal.pone.0222633.

Shen F, Zhang B, Cao C, Jiang X. 2018. Online discrimination of storage shelf-life and prediction of post-harvest quality for strawberry fruit by visible and near infrared spectroscopy. J Food Process Eng 41: e12866. https://doi.org/10.1111/jfpe.12866.

Slaughter DC, Abbott JA. 2004. Analysis of fruits and vegetables. In: Roberts CA, Workman J Jr, Reeves JB III (eds.). Agronomy Monographs. Volume 44. American Society of Agronomy, Madison, WI. https://doi.org/10.2134/agronmonogr44.c14.

Suhandy D, Yulia M. 2017. The use of partial least squares regression and spectral data in UV-visible region for quantification of adulteration in Indonesian palm civet coffee. Intl J Food Sci 2017: 6274178. https://doi.org/10.1155/2017/6274178.

Suratmi S, Chotimah HENC, Syahid A. 2022. The effects of KNO_3 fertilizer and growth regulator of mung bean sprout extract on the growth, sweetness increase and yield of melon (*Cucumis melo* L.). AgriPeat 23: 29-35. https://doi.org/10.36873/agp.v23i1.4454.

Ustin SL, Jacquemoud S. 2020. How the optical properties of leaves modify the absorption and scattering of energy and enhance leaf functionality. In: Cavender-Bares J, Gamon JA, Townsend PA (eds.). Remote Sensing of Plant Biodiversity. Springer, Cham. https://doi.org/10.1007/978-3-030-33157-3_14.

Walsh KB, Blasco J, Zude-Sasse M, Sun X. 2020. Visible-NIR 'point' spectroscopy in postharvest fruit and vegetable assessment: The science behind three decades of commercial use. Postharvest Biol Technol 168: 111246. https://doi.org/10.1016/j.postharvbio.2020.111246.

Wieme J, Mollazade K, Malounas I, Zude-Sasse M, Zhao M, Gowen A, Argyropoulos D, Fountas S, Van Beek J. 2022. Application of hyperspectral imaging systems and artificial intelligence for quality assessment of fruit, vegetables and mushrooms: A review. Biosyst Eng 222: 156-176. https://doi.org/10.1016/j.biosystemseng.2022.07.013.

Wilcox RR. 2016. Introduction to Robust Estimation and Hypothesis Testing. 4th Edition. Academic Press, Cambridge. https://doi.org/10.1016/B978-0-12-804733-0.00001-9.

Zeb A, Qureshi WS, Ghafoor A, Malik A, Imran M, Iqbal J, Alanazi E. 2021. Is this melon sweet? A quantitative classification for near-infrared spectroscopy. Infrared Phys Technol 114: 103601. https://doi.org/10.1016/j.infrared.2021.103645.