International Journal of Agricultural Technology 2026 Vol. 22(6):2687-2702
DOI: https://doi.org/10.63369/ijat.2026.22.6.2687-2702
Available online http://www.ijat-aatsea.com
ISSN 2630-0192 (Online)

Non-destructive determination of the physicochemical properties of melon (*Cucumis melo* var. *Inodorus*) using Visible Near-Infrared Reflectance Spectroscopy and machine learning models

Khuriyati, N.*, Sukartiko, A. C., Ushada, M. and Kasih, A. N.

Department of Agro-Industrial Technology, Faculty of Agricultural Technology, Universitas Gadjah Mada, Yogyakarta, Indonesia.

Khuriyati, N., Sukartiko, A. C., Ushada, M. and Kasih, A. N. (2026). Non-destructive determination of the physicochemical properties of melon (*Cucumis melo* var. *Inodorus*) using Visible Near-Infrared Reflectance Spectroscopy and machine learning models. International Journal of Agricultural Technology 22(6):2687-2702.

Abstract Melon (*Cucumis melo* L.) is widely produced and consumed fruit globally. Traditionally, the quality of melon fruits is assessed through destructive testing, which can be both time-consuming and damaging to the fruit. To overcome this challenge, a study was undertaken to investigate non-destructive methods for evaluating the physicochemical quality of melons using Visible Near-infrared Reflectance (Vis-NIR) Spectroscopy. The study developed predictive models using Partial Least Square (PLS) and Random Forest (RF) techniques. The spectral data from 120 melons were non-destructively collected, while physicochemical properties, including fruit firmness, soluble solids content, and dry matter were determined through destructive testing. The data analysis was performed using Unscramber CAMO for PLS and Google Collab for RF. The RF analysis showed high reliability with R^2 test values of 0.8697, 0.9012, 0.9183, and 0.9142 for the raw spectrum, multiplicative scatter correction, 1st derivative Savitsky-Golay, and De-Trending data pre-treatment, respectively in the simultaneous prediction of all physicochemical properties. The results revealed that RF provided more high-accuracy predictions compared to PLS. This use of Vis-NIR and RF in melon quality prediction contributes to more sustainable agricultural practices by reducing food waste, minimizing chemical and energy use, and supporting supply chain efficiency. This research is offered an innovative approach to improve the sustainability of the agricultural sector through non-destructive technologies that are fast, accurate, and environmentally friendly.

Keywords: Melons, Physicochemical properties, Partial Least Square, Random Forest, Visible Near-Infrared Reflectance Spectroscopy

Introduction

Melon (*Cucumis melo* L.) is a horticultural commodity with high economic value and has become one of the main sources of income for many farmers in several countries, including Indonesia. This commodity is not only consumed directly by the community because of its sweet and refreshing taste but is also

* **Corresponding Author:** Khuriyati, N.; **Email:** nafis.khuriyati@ugm.ac.id

utilized as a raw material in the processed food industry. Melon is known as a good source of vitamins for human health, thus further strengthening its position as a valuable horticultural commodity. The relatively short harvest period of melons, accompanied by high selling prices, makes melons one of the leading commercial commodities in the horticultural market (Annisa and Gustia, 2018). Therefore, success in maintaining melon quality is an important factor in ensuring its competitiveness in local and international markets.

Traditionally, assessments of melon quality such as sweetness and firmness are based on subjective perceptions by producers or consumers. This is usually based on external indicators of the fruit, such as visual appearance, size, and aroma. However, this assessment method does not provide an accurate and consistent picture of the internal quality of the fruit. On the other hand, objective methods used to assess melon quality, such as measuring sugar content or ripeness, are generally destructive. The process requires sampling which damages the product making it unsalable. In addition, destructive methods tend to be time-consuming and require high operational costs. Thus, there is an urgent need for more efficient, accurate, and non-destructive quality evaluation methods to meet market demands and improve the efficiency of the melon supply chain.

One solution to address these challenges is the application of non-destructive testing (NDT) in assessing melon quality. Non-destructive testing (NDT) offers an approach that allows analysis of the internal quality of a product without altering or damaging the integrity or function of the material being tested. By using this method, the product can still be sold or consumed after going through the evaluation process. Some of the advantages offered by NDT methods include increased product reliability, reduced risk of damage or accidents during the testing process, and decreased production costs compared to destructive testing (Phey *et al.*, 2020). NDT is therefore an attractive approach to adopt in the agricultural industry, including in melon quality assessment.

Spectroscopy, particularly Visible-Near Infrared Reflectance Spectroscopy (Vis-NIR), is one of the most promising NDT techniques for this application. This method has several advantages, including its ability to minimize the use of chemicals, ease of operation, and does not require an elaborate or complex testing environment (Lim *et al.*, 2015). In practice, Vis-NIR can measure the physicochemical properties of a product quickly, consistently, and accurately without the need to damage the sample. Mishra *et al.* (2017) showed that Vis-NIR spectroscopy, when combined with chemometric methods such as Partial Least Square (PLS) regression, can predict the pork red color parameter (a*) with a coefficient of determination (R^2) value of 0.86, indicating good predictive ability. In addition, the PLS method has also been successfully used to predict capsaicin concentration in chili peppers with a high accuracy of R^2 of 0.953

(Candan and Inan, 2014). These results show the great potential of the combination of Vis-NIR and PLS methods in assessing various quality parameters of agricultural products.

Besides PLS, machine learning algorithms such as Random Forest (RF) have also been used in NDT applications. RF methods offer advantages in terms of flexibility and the ability to handle high-dimensional data. In a study by Merzlyak *et al.* (2003), RF was successfully used to predict Total Dissolved Solids, pH, and total iron in river water with accuracy rates of 97.74%, 98.66%, and 98.65%, respectively. This success shows that RF has great potential to be applied in melon quality analysis, especially in predicting physicochemical parameters accurately and efficiently. The integration of NDT methods such as Vis-NIR spectroscopy with machine learning, such as PLS and RF, not only offers advantages in terms of accuracy and efficiency but also contributes significantly to sustainability in agricultural practices. By reducing the need for destructive testing that often results in product disposal, these approaches can help minimize waste as well as improve resource use efficiency (Nicolaï *et al.*, 2007). In addition, the use of this method allows for a reduction in chemical use and energy consumption during the quality evaluation process, thus supporting more environmentally friendly and sustainable agricultural practices (Corrêdo *et al.*, 2021).

Therefore, the research is sought to develop a prediction model that can be used to evaluate the physicochemical properties of melons to assess their quality non-destructively.

Materials and methods

Materials

A total of 120 Honey Globe melons (*Cucumis melo* var. *Inodorus*) were used in this research. They were obtained from the Greenhouse Field Research Center, Universitas Gadjah Mada, Kulon Progo Regency, Special Region of Yogyakarta, Indonesia. The melons used had three levels of maturity, namely unripe, ripe, and overripe. The unripe melons were harvested at 45 and 51 days after transplanting (DAT), ripe ones at 55 and 60 DAT, while the overripe fruits were harvested at 65 DAT, with a harvesting interval of 5 to 6 days.

Spectra acquisition

A self-assembled Vis-NIR instrument was used to record the reflectance spectrum of the sample, which consisted of a Vis-NIR Miniature Spectrometer (Flame-T-VIS-NIR Ocean Optics, 350-1000 nm), a tungsten halogen light source

(HL-2000-HP-FHSA Ocean Optics, 360-2400 nm) and a reflection probe (QR400-7-VIS–NIR Ocean Optics). The schematic is shown in Figure 1. The OceanView 1.67 software was used to perform spectral data acquisition, with configured parameters including an integrase period of 100 ms, an average of 100 scans, and one boxcar width. Before data collection, the spectrometer was calibrated using white and dark references. A ceramic diffuse reflectance standard (WS-1, Ocean Optics, USA) was used for the white reference, while a visible light source blockage was used for the dark reference. The spectrometer was calibrated every five samples to ensure instrument reliability and accurate results under optimal conditions. For each melon sample, three spectra at λ of 300–1000 nm were analyzed, and the average was calculated (Figure 2).

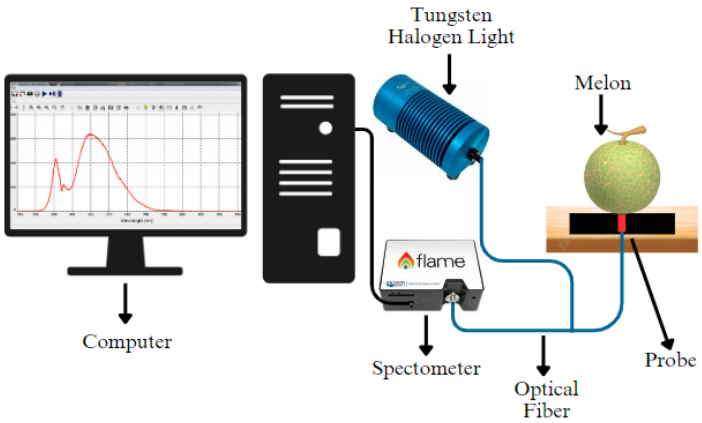


Figure 1. A schematic of the Vis-NIR spectroscopy

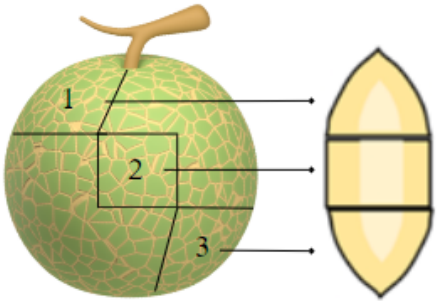


Figure 2. Melon parts of which spectra were acquired

Reference measurements

Physicochemical parameters such as firmness, soluble solid content (SSC), and dry matter (DM) were ascertained using destructive testing equipment. The Fujiwara KM-5 fruit hardness tester was utilized to measure the firmness of

melon flesh. This device can measure up to 5 kg with a precision of ± 50 g, providing accurate information about the texture of the fruit. The Pocket refractometer "ATAGO" model PAL 15S was employed to measure the SSC expressed in Brix degrees. Melon samples were dehydrated using a hot air oven at 75 ℃ for 48 h until a constant weight (Subedi and Walsh, 2020). The DM content was calculated by applying Equation [1]:

$$\frac{W1}{W2} X\ 100 \quad (1)$$

W1 refers to the dry weight, and W2 is the wet weight of the sample.

Chemometric analysis

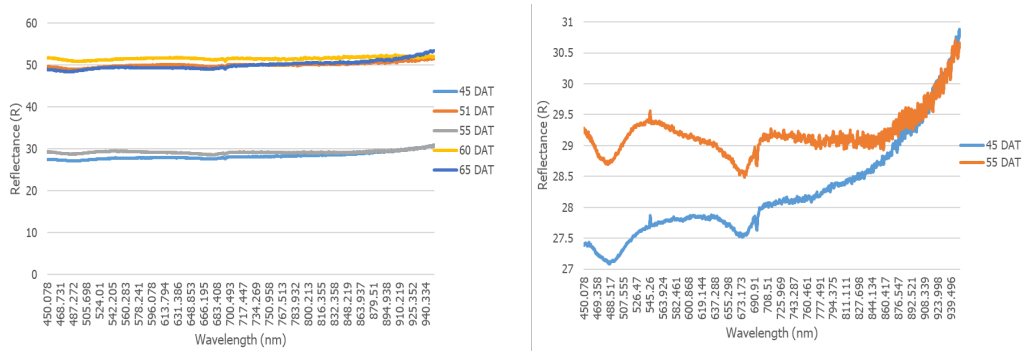
All measured spectra were compiled in MS Excel and subsequently imported into Unscrambler® X software (CAMO, Oslo, Norway) for further analysis. The spectrum utilized for chemometric analysis was 450-950 nm which provides more detailed information for further analysis. Pre-treatment steps were applied to the spectra to reduce noise. The PLS method was utilized to predict the physicochemical properties of the samples and to perform full cross-validation of the calibration model. A high R^2 value (maximum 1) and low error values indicate good model performance (Cortés *et al.*, 2019). The model with a Ratio of Prediction to Deviation (RPD) value of ≥2.5 was considered effective for screening applications, while an RPD value of 1 indicated inaccurate predictions (Xie *et al.*, 2008). In addition to PLS, RF was applied to predict the physicochemical properties. RF was applied to all spectral data (input) and physicochemical properties (output), with each factor corresponding to potential characteristics associated with varying wavelengths of Vis–NIR data. The dataset was randomly split, with 80% used for training and 20% for testing. To minimize noise caused by initial deviations and various scattering effects in the spectra, pre-treatment techniques such as Multiple Scatter Correction, Normalization, Standard Normal Variate, Savitsky-Golay's Smoothing, and De-trending were applied (Gordillo-Delgado *et al.*, 2017).

Results

Vis-NIR spectra of melons

In this study, the spectral data of each sample was collected three times and averaged. The 120 samples of raw spectral reflectance data are shown in Figure 3(a). These spectra provide clear information when plotted against fruit age. The spectrum appears flat with no peaks or valleys across the tested wavelength range. In addition, the raw spectra of 120 samples at harvest ages 45 and 55 DAT

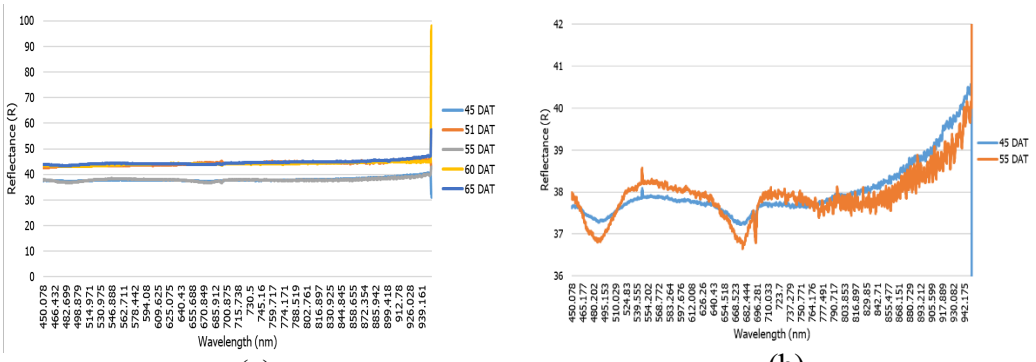
(Day After Transplanting) with the same reflectance interval results are shown in Figure 3(b). These spectra showed the peaks and valleys along the wavelengths tested.



(a) (b)

Figure 3. Raw spectra of melons at various ages (a) and at a specific age (b)

The spectra after pre-treatment with MSC, 1st derivative Savitsky-Golay, and De-trending are presented in Figures 4 to 6. It showed the spectra after pre-treatment using MSC (Figure 4). The change in spectral data after pre-treatment applying the 1st derivative Savitzky-Golay is shown in Figure 5, which was significantly compared to the raw spectrum. The results of the spectra after pre-treatment using De-trending are illustrated in Figure 6. This finding is indicated a marked variation from the raw spectral data, displaying them in the form of valleys and peaks, although not as noticeable. However, did not seem to display the peaks and valleys at the Vis-NIR wavelengths as observed in Figures 4(a) and 5(a).



(a) (b)

Figure 4. Spectra using Pre-Treatment Multiplicative Scatter Correction (MSC) at various ages (a) and at a specific age (b) of melons

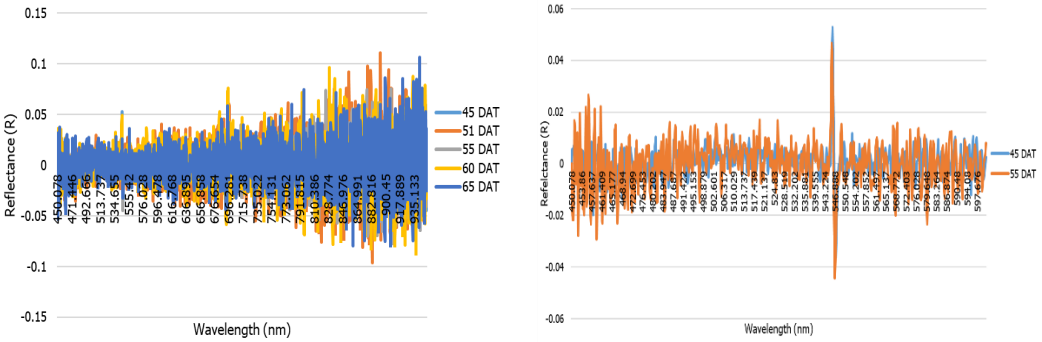


Figure 5. Spectra using Pre-Treatment 1st derivative Savitsky-Golay at various ages (a) and at a specific age (b) of melons

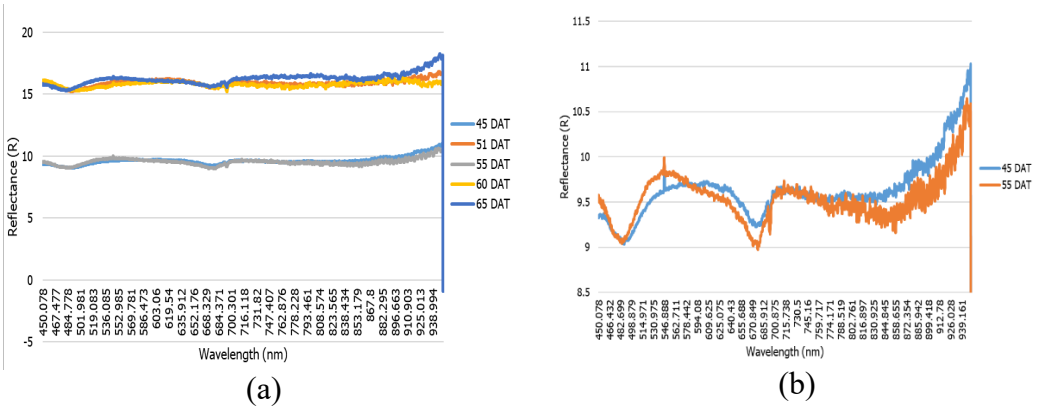


Figure 6. Spectra using Pre-Treatment De-trending at various ages (a) and at a specific age (b) of melons

The spectra plotted based on fruit age are shown in Figure 3 to 6. Fruits aged 60 and 65 days after transplanting (DAT) showed higher reflectance values compared to fruits at the younger ages of 45, 51, and 55 DAT in all treatments.

Physicochemical properties of melons at various growing stages

The firmness of melons declined significantly during the growth. At the same time, the soluble solids content (SSC) continued to increase, which means that the accumulation of sugar in melons was high during growth. In addition, the dry matter content (DM) also increased significantly during the growth period (Table 1).

Table 1. Physicochemical properties of fruit samples at various ages

Fruit age (DAT)	Number of samples	Firmness (kgf)	SSC (°Brix)	Dry matter (%)
45	24	1.26±0.07	13.34±1.21	9.04±2.31
51	24	1.11±0.09	14.02±2.03	8.69±1.02
55	24	0.95±0.03	16.62±0.97	10.91±1.87
60	24	0.69±0.06	17.51±1.82	11.29±0.98
65	24	0.53±0.03	17.11±0.88	11.16±0.92

Partial least square prediction model

The PLS model was developed to predict the physicochemical properties of melons. In this study, four types of pre-treatments were applied to the spectra of melon samples. For each pre-treatment, model performance was evaluated based on a high coefficient of determination (R^2), low standard error of calibration (SEC) and standard error of prediction (SEP), as well as high RPD values, as shown in Table 3. Statistical description of the samples for calibration and validation are presented in Table 2.

Table 2. Statistical descriptive of the calibration and validation sample sets

Physicochemical properties	Subset	Number of samples	Mean	Std dev	Min	Max
Firmness (Kgf)	Calibration	80	0.93	0.29	0.32	1.43
	Validation	40	0.91	0.33	0.26	1.42
SSC (°Brix)	Calibration	80	15.41	2.86	8.28	21.88
	Validation	40	15.70	3.26	8.28	21.88
Dry matter (%)	Calibration	80	10.06	1.99	4.87	15.77
	Validation	40	10.06	2.04	4.87	13.94

The R^2 values across all pretreatments for firmness ranged from 0.0070 to 0.8015; for SSC, from 0.0162 to 0.531; and for DM contents, from 0.108 to 0.9001. The models for predicting the physicochemical properties of melons using PLS showed have low performance, indicated by the very low R^2 value for all pretreatments, except the prediction model of firmness employing MSC spectra (Table 3). Additionally, datasets not included in calibration showed lower R^2 values (R^2 validation) for firmness, SSC, and DM content across all pre-treatment.

Table 3. Performance model for predicting the physicochemical properties of melons using Partial Least Square

Physical properties	Pre-treatment	Calibration		Validation	
		R^2	RMSE	R^2	RMSE
Firmness	Raw spectra	0.7321	0.1519	0.6014	0.1876
	MSC	0.7321	0.1519	0.6014	0.1876
	1st derivative Savitsky Golay	0.0079	0.2923	NA	0.3044
	De-Trending	0.8015	0.1307	0.6901	0.1654
SSC	Raw spectra	0.3311	2.3182	0.1964	2.5732
	MSC	0.0162	2.8115	NA	3.053
	1st derivative Savitsky Golay	0.4376	2.1256	0.1712	2.6131
	De-Trending	0.5319	1.9390	0.1890	2.5849
Dry matter	Raw spectra	0.4599	1.4380	0.2295	1.739
	MSC	0.0108	1.9466	NA	2.0816
	1st derivative Savitsky Golay	0.9001	0.6184	0.3723	1.5702
	De-Trending	0.7045	1.0630	0.2945	1.6640

Random forest prediction model

Prediction of physicochemical properties in melon fruits was carried out using RF in Google Collab. Vis-NIR data was utilized as the input and output data were in the form of physicochemical properties such as firmness, SSC, and DM. After applying several pretreatments, the prediction performance was evaluated employing the R^2 accuracy value, which indicated a high value close to 1. The best result was achieved with Savitsky-Golay's pre-treatment, which yielded an R^2 value of 0.9183 or 91.83%. The prediction model performance for each pre-treatment is summarized in Table 4.

Table 4. Performance model for predicting the physicochemical properties of melons using the Random Forest prediction model

Physical properties	Pre-treatment	R^2	RMSE
Firmness, SSC, Dry matter	Raw Spectra	0.8697	0.2208
	MSC	0.9012	0.1924
	1st derivative Savitsky Golay	0.9183	0.1574
	De-Trending	0.9142	0.1741

Discussion

This research will utilize data obtained through the Vis-NIR spectroscopy method, which is then analyzed using PLS and RF techniques. By using a combination of these methods, it is expected to be possible to analyze the internal quality of melons quickly, accurately, and efficiently without damaging the samples. In addition to offering an innovative approach, this research also seeks to provide practical and applicable solutions to address challenges in the agricultural industry, particularly in ensuring the quality and competitiveness of melons in the global market. This approach is expected to replace the subjective assessment methods that have been widely used and provide significant benefits in supporting the sustainability and efficiency of the agricultural industry in the future.

Vis-NIR spectra of melons

When a sample is scanned with light, it absorbs and reflects light energy, with the resulting reflections captured by a spectrometer and visualized on a computer. This process generates reflectance spectrum data within the visible near-infrared and short λ ranges of 450-950 nm. Multiple spectral data are needed for each sample and the average value is calculated to minimize random errors (Guo *et al.*, 2016). They appeared flat with no peaks or valleys throughout the wavelength range tested. This was due to the wide range of reflectance intervals (20%–60%) generated from each sample, as well as differences in the internal structure or chemical content, surface characteristics, light scattering, and instrument-related factors. The curvature of the fruit surface impacts how light is absorbed and transported within the fruit, while the thickness of the fruit pulp layer affects the diffuse reflection pattern and overall inspection efficiency. Inspection efficiency decreases when the distance between the light source and the detector is too small, whereas a greater distance necessitates higher light intensity and increased detector sensitivy (Ding *et al.*, 2016).

Measurement results from a sample spectrum can be influenced by various physical and chemical factors. Pre-treatment minimizes noise or interference in the spectrum, thereby enhancing the accuracy of the analysis. Each pre-treatment method has its own set of advantages and disadvantages and hence, is essential to identify the most suitable one. MSC is a normalization technique commonly utilized in spectroscopic methods to correct scale variations and shifts in a series of spectra (Windig *et al.*, 2008). It has significance in correcting scattered light as it can eliminate spectral variations caused by non-uniform particle sizes within the same sample set (Chen *et al.*, 2013). In addition, MSC can also reduce the number of latent variables needed for analysis compared to those that do not

involve pre-processing (Khuriyati *et al.*, 2023). Savitzky-Golay is a method designed to reduce noise in spectral data without sacrificing the relevant information while reducing the impact of noise on the signal (Candan and Inan, 2014). The SG derivative approach involves spectral smoothing steps before the derivation process, which aims to balance the signal-to-noise ratio and determine the number of adjacent variables to be utilized in the calculation (Bruning *et al.*, 2019). The first derivative of the SG method is often applied to remove baselines that may appear in the spectra (Zaid *et al.*, 2020). De-trending is a method that aims to eliminate the baseline in spectral data under the assumption that it can be modeled as a lower-order polynomial. This approach involves fitting the spectrum into a polynomial of a predetermined degree and then replacing the spectrum with the matched residuals (Roger *et al.*, 2022).

The fruits at 60 and 65 DAT had a higher reflectance value than those at 45, 51, and 55 DAT for all pretreatments, which may be influenced by physiological and biochemical changes in melon fruits. These include alterations in the fruit texture that affect cell wall structures as well as in SSC. Ripening of melon fruits causes softening that occurs due to cell wall degradation, which involves ethylene and cell wall modifying enzymes that play a role in modifying and remodeling pectin (Peng *et al.*, 2022). Underlying this process are alterations in tissue density and the attenuation of cells and cell structures, thereby impacting the dispersion of light. Compared to tougher fruits, softer fruits generally exhibit a wider scattering profile, which consequently increases reflectance (Lu and Peng, 2006). The SSC value increases with maturity, which reduces photon scattering, seen as an enhancement in inner luminescence due to increased light absorption by the fruit tissue (Romano *et al.*, 2011).

Physicochemical properties of melons at various growing stages

The firmness declined while SSC and DM increased during the growth of melons. The softening of cell walls during maturation reduces the firmness, which is attributed to changes in the polymer structure of the cell wall pectin (Khoshnam *et al.*, 2015). SSC is closely linked to carbohydrates, as the chemical composition of fruit carbohydrates is similar to sugars. Immature melons have low sweetness due to the starch (polysaccharide) content not being converted into saccharides. Enzymes such as glucoamylase, amylase, and phosphorylase break down fruit carbohydrates. As the respiration rate increases to produce sucrose, glucose, and fructose, the SSC levels rise with fruit age (Setiawati and Bafdal, 2020). The DM content of melons elevates throughout fruit maturation due to the accumulation of organic compounds. Utilizing samples from various growth

stages can provide a wide range of physical properties, which is beneficial for developing reliable prediction models.

Partial least square prediction model

PLS is a standard tool used for predicting compounds from complex and high-dimensional spectral data. In this study, the firmness, SSC, and DM values of calibration and validation sample sets were designed to be similar to obtain excellent prediction models. However, the RMSE values obtained during the validation process for firmness, SSC, and DM content were higher than those during calibration, indicating overfitting. This occurs when a statistical machine learning model becomes overly tailored to the training dataset, leading to a decline in performance on new unseen datasets. This phenomenon indicates that the models predicted values are excessively aligned with the observed values in the training dataset, or that the model fails to generalize to new data because it has memorized specific patterns from the training dataset (Montesinos López *et al.*, 2022).

Random forest prediction model

PLS is widely utilized multivariate analysis technique in chemometrics, especially in quantitative spectral analysis (Jena *et al.*, 2023). PLS regression utilizes a limited number of latent variables (LVs), which are linear combinations of the original ones. The objective of this method is to identify the LVs that capture the majority of the information in variable X, thereby facilitating accurate prediction of variable Y while reducing the dimensionality of the regression problem by using a minimum number of LVs (Nawar and Mouazen, 2017). On the other hand, RF is an ensemble learning technique that combines and averages the predictions from multiple decision trees (DTs) to improve predictive performance and control overfitting (Ghasemi and Tavakoli, 2013). This method evaluates each predictor individually as well as in interactions with other predictors, even when the number of variables exceeds the sample size (Breiman, 2001). In the context of rapid and non-destructive prediction using NIR spectroscopy, RF demonstrates potensial and is comparable to traditional chemometric methods (Vigneau *et al*., 2018). Additionally, the RF technique, when compared with PLS for the online prediction of total nitrogen (TN), exhibited excellent spatial similarity with TN measurements (Hwang *et al*., 2023), and can successfully predict them.

Non-destructive methods offer the potential to enhance sustainability in agriculture and food processing by minimizing waste, improving food safety, and

boosting efficiency in production and processing, leading to cost reductions. Nicolaï *et al*. (2007) highlights that NIR spectroscopy technology can significantly cut food waste by enabling the assessment of fruit and vegetable quality without causing damage. This research supports those findings, as the Vis-NIR spectroscopy technique was used to non-destructively predict the physicochemical properties of melons. By using this approach, damage to melons during quality evaluation is avoided, increasing the quantity of marketable or consumable fruit and reducing waste.

Moreover, the application of Vis-NIR spectroscopy allows for the rapid acquisition of spectral data from 120 samples, leading to significant time savings and reduced operational costs. Cen and He (2007) emphasized that NIR spectroscopy is a swift and efficient technique, utilizing less energy and resources compared to conventional destructive methods, which are more time-consuming.

Traditionally, fruit quality assessment relies on destructive methods that use chemicals. However, Corrêdo *et al*. (2021) has demonstrated that Vis-NIR spectroscopy is a rapid and effective technique for evaluating the concentration of chemical compounds. In this study, the quality of melon fruit was assessed using Vis-NIR spectroscopy combined with machine learning algorithms, specifically Random Forest (RF) and Partial Least Squares (PLS). The results indicated that RF was more effective in processing complex spectral data and delivering more accurate predictions of melon quality compared to PLS, achieving a higher R^2 value. These results suggest that integrating machine learning enhances the accuracy of melon quality predictions and highlights the significant potential of Vis-NIR spectroscopy to reduce chemical use in fruit assessment. Additionally, this method provides a more efficient and environmentally friendly alternative.

Acknowledgements

This research has been financed by Universitas Gadjah Mada under the grant of the *Peningkatan* Academic Excellence 2023 (contract no: 7725/UN1.P.II/Dit-Lit/PT.01.03/2023).

Conflict of interest

The authors declare no conflict of interest.

References

Annisa, P. and Gustia, H. (2018). Respon Pertumbuhan dan Produksi Tanaman Melon Terhadap Pemberian Pupuk Organik Cair Tithonia diversifolia. Prosiding SEMNASTAN, pp.104-114.

Breiman, L. (2001). Random Forests. Machine Learning 45:5-32.

Bruning, B., Liu, H., Brien, C., Berger, B., Lewis, M. and Garnett, T. (2019). The Development of Hyperspectral Distribution Maps to Predict the Content and Distribution of Nitrogen and Water in Wheat (*Triticum aestivum*). Frontiers in Plant Science, 10:1-16.

Candan, Ç. and Inan, H. (2014). A unified framework for derivation and implementation of Savitzky-Golay filters. Signal Processing, 104:203-211.

Cen, H. and He, Y. (2007). Theory and application of near infrared reflectance spectroscopy in determination of food quality. Trends in Food Science and Technology, 18:72-83.

Chen, H., Song, Q., Tang, G., Feng, Q. and Lin, L. (2013). The Combined Optimization of Savitzky-Golay Smoothing and Multiplicative Scatter Correction for FT-NIR PLS Models. ISRN Spectroscopy, 1-9.

Corrêdo, L. de P., Maldaner, L. F., Bazame, H. C. and Molin, J. P. (2021). Evaluation of minimum preparation sampling strategies for sugarcane quality prediction by vis-nir spectroscopy. Sensors, 21:1-23.

Cortés, V., Blasco, J., Aleixos, N., Cubero, S. and Talens, P. (2019). Monitoring strategies for quality control of agricultural products using visible and near-infrared spectroscopy: A review. Trends in Food Science and Technology, 85:138-148.

Ding, C., Shi, S., Chen, J., Wei, W. and Tan, Z. (2016). Influence of surface curvature on light-based nondestructive measurement of stone fruit. Computers and Electronics in Agriculture, 121:200-206.

Ghasemi, J. B. and Tavakoli, H. (2013). Application of random forest regression to spectral multivariate calibration. Analytical Methods, 5:1863-1871.

Gordillo-Delgado, F., Bedoya, A. and Marín, E. (2017). Study of the Pigments in Colombian Powdered Coffee Using Photoacoustic Spectroscopy. International Journal of Thermophysics, 38:1-8.

Guo, Z., Huang, W., Peng, Y., Chen, Q., Ouyang, Q. and Zhao, J. (2016). Color compensation and comparison of shortwave near infrared and long wave near infrared spectroscopy for determination of soluble solids content of “Fuji” apple. Postharvest Biology and Technology, 115:81-90.

Hwang, S. W., Chung, H., Lee, T., Kim, J., Kim, Y. J., Kim, J. C., Kwak, H. W., Choi, I. G. and Yeo, H. (2023). Feature importance measures from random forest regressor using near-infrared spectra for predicting carbonization characteristics of kraft lignin-derived hydrochar. Journal of Wood Science, 69:1-12.

Jena, P. K., Rahaman, S. M., Mohapatra, P. K. Das, Barik, D. P. and Patra, D. S. (2023). Surface water quality assessment by Random Forest. Water Practice and Technology, 18:201-214.

Khoshnam, F., Namjoo, M. and Golbakhshi, H. (2015). Acoustic testing for melon fruit ripeness evaluation during different stages of ripening. Agriculturae Conspectus Scientificus, 80:197-204.

Khuriyati, N., Sukartiko, A. C., Falah, M. A. F. and Alfiani, R. N. (2023). Determination of Total Carotene and Vitamin C in Chili Powder (*Capsicum annuum* L.) Non-destructively Using Near-Infrared Spectroscopy. International Journal on Advanced Science, Engineering and Information Technology, 13:69-76.

Lim, J., Kim, G., Mo, C. and Kim, M. S. (2015). Design and fabrication of a real-time measurement system for the capsaicinoid content of Korean red pepper (*Capsicum annuum* L.) powder by visible and near-infrared spectroscopy. Sensors (Switzerland), 15:27420-27435.

Lu, R. and Peng, Y. (2006). Hyperspectral scattering for assessing peach fruit firmness. Biosystems Engineering, 93:161-171.

Merzlyak, M. N., Solovchenko, A. E. and Gitelson, A. A. (2003). Reflectance spectral features and non-destructive estimation of chlorophyll, carotenoid and anthocyanin content in apple fruit. Postharvest Biology and Technology, 27:197-211.

Mishra, S. P., Sarkar, U., Taraphder, S., Datta, S. and Swain, D. P. (2017). Multivariate Statistical Data Analysis- Principal Component Analysis (PCA). International Journal of Livestock Research, 7:60-78.

Montesinos López, O. A., Montesinos López, A. and Crossa, J. (2022). Overfitting, Model Tuning, and Evaluating of Prediction Performance. In: Multivariate Statistical Machine Learning Methods for Genomic Prediction, Mexico, Springer, pp.109-138.

Nawar, S. and Mouazen, A. (2017). Comparison between random forest and partial least squares regression of on-line vis-NIR spectroscopy measurement of soil total nitrogen. Towards Sustainable Agriculture and Biosystems Engineering, 109-130.

Nicolaï, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I. and Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. Postharvest Biology and Technology, 46:99-118.

Peng, Z., Liu, G., Li, H., Wang, Y., Gao, H., Jemrić, T. and Fu, D. (2022). Molecular and Genetic Events Determining the Softening of Fleshy Fruits: A Comprehensive Review. International Journal of Molecular Sciences, 23.

Phey, O., Hashim, N. and Maringgal, B. (2020). Quality evaluation of mango using non-destructive approaches: A review. Journal of Agricultural and Food Engineering, 1:1-8.

Roger, J. M., Mallet, A. and Marini, F. (2022). Preprocessing NIR Spectra for Aquaphotomics. Molecules, 27:1-18.

Romano, G., Nagle, M., Argyropoulos, D. and Müller, J. (2011). Laser light backscattering to

monitor moisture content, soluble solid content and hardness of apple tissue during drying. Journal of Food Engineering, 104:657-662.

Setiawati, R. and Bafdal, N. (2020). Dampak Kualitas Air Tanah Terhadap Kualitas Melon (*Cucumis melo* L.). Agrotekma: Jurnal Agroteknologi Dan Ilmu Pertanian, 4:83-93.

Subedi, P. P. and Walsh, K. B. (2020). Assessment of avocado fruit dry matter content using portable near infrared spectroscopy: Method and instrumentation optimisation. Postharvest Biology and Technology, 161:111078.

Vigneau, E., Courcoux, P., Symoneaux, R., Guérin, L. and Villière, A. (2018). Random forests: A machine learning methodology to highlight the volatile organic compounds involved in olfactory perception. Food Quality and Preference, 68:135-145.

Windig, W., Shaver, J. and Bro, R. (2008). Loopy MSC: A simple way to improve multiplicative scatter correction. Applied Spectroscopy, 62:1153-1159.

Xie, L., Ying, Y., Lin, H., Zhou, Y. and Niu, X. (2008). Nondestructive determination of soluble solids content and pH in tomato juice using NIR transmittance spectroscopy. Sensing and Instrumentation for Food Quality and Safety, 2:111-115.

Zaid, A., Abu-Khalaf, N., Mudalal, S. and Petracci, M. (2020). Differentiation between normal and white striped Turkey breasts by visible/near infrared spectroscopy and multivariate data analysis. Food Science of Animal Resources, 40:96-105.

(Received: 31 January 2025, Revised: 26 February 2026, Accepted: 10 March 2026)