Food Research 8 (Suppl. 2) : 9 - 14 (2024)
Journal homepage: https://www.myfoodresearch.com

FOOD RESEARCH

RESEARCH PAPER

Non-destructive determination of water content in fruits using Vis-NIR spectroscopy

Amanah, H.Z., Pratiwi, E.Z.D., Rahmi, D.N., Pahlawan, M.F.R. and *Masithoh, R.E.

Department of Agricultural and Biosystems Engineering, Faculty of Agricultural Technology, Universitas Gadjah Mada, Jl. Flora No. 1 Bulaksumur, Yogyakarta 55281, Indonesia

Article history:
Received: 5 March 2023
Received in revised form: 27 September 2023
Accepted: 20 January 2024
Available Online: 30 March 2024

Keywords:
Fruit,
Partial least squares regression,
Vis-NIR spectroscopy,
Water content

DOI:
https://doi.org/10.26656/fr.2017.8(S2).14

Abstract

Visible near-infrared (Vis-NIR) spectroscopy is less accurate for detecting -OH molecules, therefore, the use of Vis-NIR spectroscopy is challenging for samples containing high water content, such as fruits. However, as Vis-NIR spectroscopy is a portable and economical instrument, it can be used in the field by small-scale farmers. This study aimed to evaluate Vis-NIR spectra to measure water content in fruits. Four fruits were used in this study, including dragon fruit, guava, sapodilla and banana (100 pieces each). All fruits were randomly divided into a calibration set (two-thirds of the samples) and a prediction set (one-third of the samples). Water content was predicted using partial least square regression (PLSR) analysis. The PLSR calibration model had a coefficient of determination (R^2) of 0.29 for dragon fruit, 0.63 for guava, 0.62 for sapodilla and 0.80 for banana. The prediction model had an R^2 of 0.11 for dragon fruit, 0.63 for guava, 0.52 for sapodilla and 0.75 for banana. These results show that Vis-NIR spectroscopy has the potential to predict water content in relatively low water-content fruits, such as bananas.

1. Introduction

Fruit is a perishable product in which quality easily deteriorates over time. One of the quality parameters of fruit is water content. Changes in the amount of water in fruit affect the physical and mechanical parameters of the fruit (Ehiem *et al*., 2016; Eboibi and Uguru, 2018). In addition, the high water content in fruit triggers microbial growth. Therefore, information on fruit water content is important because it affects consumer acceptance and determines the handling and storage of fruit.

Water content can be determined using the thermogravimetric method in which a sample is heated in an oven and the water content of the sample is determined from the loss of sample weight. However, this method requires an extended time and destroys the sample. Near-infrared (NIR) spectroscopy has been used to determine water content in whole wheat flour (Manley *et al*., 2002), processed cheese (Blazquez *et al*., 2004) and vegetable seeds (Szulc *et al*., 2020). Water is comprised of OH molecules, which dominate intensive absorption in the near-infrared region at 1,440-1,930 nm (Bogomolov *et al*., 2018). Water absorbs the majority of the photon energy and causes light scattering (Xie *et al*., 2008). The dominance of the water absorption band can interfere with the measurement of small concentrations of other products; therefore, developing a calibration model for water content using NIR spectroscopy usually produces good results. However, the price of NIR spectroscopy is relatively expensive, which hinders its use in small-scale industries.

However, spectroscopy working in the short infrared region (350-1,000 nm) or visible near-infrared (Vis-NIR) region is available at a relatively inexpensive price. Vis-NIR spectroscopy has been used to identify quality parameters, such as color (Magwaza *et al*., 2014), acidity, pH (Wati *et al*., 2021; Priambodo *et al*., 2022), and Brix (Pahlawan *et al*., 2021). The authors have also studied the effectiveness of Vis-NIR spectroscopy in qualitative analyses, such as classifications of cocoa beans (Saputro *et al*., 2022), soybean seeds, and flour (Pahlawan *et al*., 2022) or detecting soybean seed viability (Saputri *et al*., 2022). Unlike NIR spectroscopy, the absorption of the -OH molecule is weak in the Vis-NIR region (350-1,000 nm), as shown by its small and broad absorption spectrum (Manley *et al*., 2007). Thus, it is challenging to employ Vis-NIR spectroscopy to predict water content in fruits in which water is present in large quantities. Therefore, in this study, the effectiveness of Vis-NIR spectroscopy in determining

*Corresponding author.
Email: *evi@ugm.ac.id*

eISSN: 2550-2166
Published by Rynnye Lyan Resources

fruit water content was determined using several fruits with various peel thicknesses. This study also employed chemometrics analysis, which exports spectral information, and relates the spectra with a desired variable (Masithoh *et al*., 2020). Partial least squares regression (PLSR) analysis based on the original and various pre-processed spectra was used to evaluate the calibration and prediction models.

2. Materials and methods

2.1 Sample and spectra measurement

Dragon fruit, sapodilla, guava and banana (100 pieces of each) were used as fresh samples. The reflectance spectra of each sample were measured using a modular type Vis-NIR spectrometer (Flame-T-VIS-NIR Ocean Optics, Dunedin, USA; 350-1,000 nm with a resolution of 0.22 nm), equipped with a tungsten halogen lamp (360-2,400 nm, HL-2000-HP-FHSA Ocean Optics) and a reflectance probe (QR400-7 VIS-NIR Ocean Optics). Spectra acquisition was set as explained by (Masithoh *et al*., 2021). The spectra were acquired using OceanView 1.6.7 software with 100 integration times, 100 scans to average, and 1 boxcar width.

2.2 Water content analysis

Water content was measured on the fruit flesh without the fruit rind. Therefore, the dragon fruit and banana peels were removed before the water content measurements. Water content was determined using the thermogravimetric method. The samples (~3 g) were dried at 105°C for 24 hrs in an oven ($w_{before\ oven}$). After 24 hrs, the dried samples were weighed ($w_{after\ oven}$), and water content was calculated using Equation 1. Each sample was analyzed in triplicate. Descriptive statistics, such as mean, minimum, maximum range, standard deviation (SD), and SD/range, were calculated.

$$MC\ (\%wb) = \frac{w_{before\ oven} - w_{after\ oven}}{w_{before\ oven}} \times 100\%$$

2.3 Chemometric analysis

The acquired spectra were compiled in Excel and imported into Unscrambler® X (10.5.1 version software, CAMO, Oslo, Norway). Three spectra were used from each sample without being averaged. Spectra <400 nm and >1,000 nm were cut before the chemometrics analysis due to heavy noise. Spectra within the wavelength range of 400-1,000 nm were used to build the PLSR calibration model. Before the analysis, all data were randomly divided so that two-thirds of the data were used as the calibration set and one-third of the data were used as the prediction set. The calibration set was used to build the calibration model, and the prediction set was used to validate the model. Calibration models were built separately for each fruit using the spectra as the predictor (X) and water content as the reference variable (Y). The performance of the model was evaluated based on the coefficient of determination (R^2) and the root mean square error (RMSE). A good model has high R^2 and low RMSE values. Several pre-processing techniques were applied to the spectra, such as Savitsky-Golay's first and second derivatives (side points = 60, polynomial = 2), Savitsky-Golay's smoothing (side points = 60, polynomial = 2), normalization, standard normal variation (SNV), and multiple scatter correction.

3. Results and discussion

3.1 Water content of the fruits

Table 1 shows that the water contents of the dragon fruit, guava, sapodilla and banana were 85-91%, 81.3-90.5%, 70.2-80.0%, and 62.90-72.1%, respectively, for the calibration model, and 85.6-90.5%, 84.8-90.4%, 72.5 -78.2%, and 62.6-72.1%, respectively, for the prediction model. The range in water content for the prediction data set was in the range of the calibration set (interpolation prediction), which agreed with Chia *et al*. (2012) in which interpolated prediction performance was better than the extrapolated prediction. Table 1 also shows that the highest water content was detected in dragon fruit followed by guava, sapodilla and banana. Moreover, the broadest range in water content was found in sapodilla followed by banana and guava, while the narrowest range occurred in dragon fruit.

The highest mean water content was detected in dragon fruit followed by guava, sapodilla and banana. Dragon fruit had the lowest SD. The highest SD was observed for bananas followed by sapodilla. A low SD indicates a small variation in the water content that is close to the mean value. A broader data range and a larger SD are required to obtain a good model, although it results in a large RMSE (Kuang and Mouazen, 2011). The SD range (16-29%) obtained for the calibration set indicates that the variation in the data will produce a good calibration model (Savenije *et al*., 2006).

3.2 Reflectance spectra of the fruits

Figure 1 shows the reflectance spectra of all fruits. Each fruit had different characteristics, particularly in the visible light region, which is correlated with pigments. Figure 1 also shows the absorbance profiles of chlorophyll in all fruits at 680 nm. The highest absorbance was detected in sapodilla, followed by guava, banana and dragon fruit. However, the NIR spectra showed a similar high reflectance trend, indicating low absorbance of the 700-1,000 nm spectra by water (Zhang *et al*., 2012), which is typical of Vis/NIR spectroscopy

https://doi.org/10.26656/fr.2017.8(S2).14

Table 1. Statistical analysis of water content in samples.

Sample	Mean	SD	Max	Min	Range	SD/Range
Calibration set						
Dragon Fruit	88.70%	1.20%	91.00%	85.00%	6.00%	19.20%
Guava	87.20%	1.50%	90.50%	81.30%	9.20%	16.50%
Sapodilla	74.80%	2.00%	80.00%	70.20%	9.80%	20.20%
Banana	68.00%	2.70%	72.10%	62.90%	9.20%	29.10%
Prediction set						
Dragon Fruit	88.60%	1.10%	90.50%	85.60%	4.90%	22.45%
Guava	87.50%	1.30%	90.40%	84.80%	5.60%	23.21%
Sapodilla	75.20%	1.20%	78.20%	72.50%	5.70%	21.05%
Banana	68.20%	2.20%	72.10%	62.60%	9.50%	23.16%

SD: standard deviation, Max: maximum, Min: minimum.

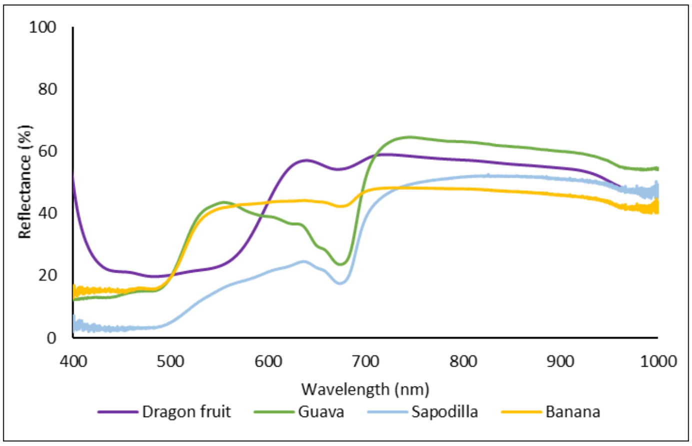

Figure 1. Reflectance spectra of samples.

(Manley *et al*., 2007).

Dragon fruit had a broad reflectance valley at 420-580 nm, which correlated with betacyanin content (Sánchez *et al*., 2006) and created the red-purple colour of the dragon fruit. Guava is characterized by a low reflectance valley indicating high absorbance at about 680 nm, which is correlated with chlorophyll and creates the green colour of the guava fruit. Sapodilla had the lowest reflectance in the visible wavelength region (400-500 nm) due to the dark brown colour of the sapodilla fruit. Banana had a low reflectance indicating high absorbance of chlorophyll at 680 nm and causing the green colour of the banana (Masithoh *et al*., 2021).

3.3 Partial least squares regression to determine water content

Table 2 shows the performance of the PLSR model developed using the original and several spectral pre-processing methods. The PLSR model yielded coefficients of determination of calibration (R^2C) of 0.07 -0.3 for dragon fruit, 0.27-0.63 for guava, 0.38-0.62 for sapodilla and 0.43-0.8 for banana. The root mean square of errors of calibration (RMSEC) was 0.97-1.11% for dragon fruit, 0.89-1.23% for guava, 1.22-1.56% for sapodilla and 1.18-2.02% for banana. The best PLSR model to quantify water content in dragon fruit was

Table 2. PLSR model performance.

Fruits		ORI	SNV	MSC	AN	SGS	SGD1	SGD2
Dragon Fruit	R^2C	0.17	0.18	0.07	0.16	0.21	0.20	0.30
	RMSEC	1.04	1.05	1.11	1.06	1.03	1.04	0.97
	R^2CV	0.08	0.08	0.05	0.07	0.11	0.12	0.18
	RMSECV	1.10	1.12	1.13	1.13	1.10	1.10	1.06
Guava	R^2C	0.30	0.42	0.45	0.43	0.27	0.53	0.63
	RMSEC	1.23	1.10	1.08	0.98	1.16	0.98	0.89
	R^2CV	0.23	0.34	0.37	0.35	0.19	0.47	0.55
	RMSECV	1.30	1.18	1.16	1.05	1.22	1.04	1.00
Sapodilla	R^2C	0.54	0.62	0.62	0.38	0.44	0.46	0.53
	RMSEC	1.34	1.22	1.22	1.56	1.48	1.46	1.36
	R^2CV	0.40	0.50	0.51	0.31	0.33	0.35	0.38
	RMSECV	1.54	1.40	1.39	1.65	1.62	1.60	1.56
Banana	R^2C	0.70	0.70	0.43	0.75	0.68	0.72	0.80
	RMSEC	1.47	1.45	2.02	1.34	1.52	1.42	1.18
	R^2CV	0.66	0.65	0.29	0.69	0.64	0.68	0.74
	RMSECV	1.57	1.59	2.26	1.48	1.61	1.53	1.36

ORI: original spectra, SNV: standard normal variate, MSC: multiple scatter correction, AN: area normalization, SGS: Savitsky-Golay's smoothing, SGD1: Savitsky-Golay's 1st derivative, SGD2: Savitsky-Golay's 2nd derivative

https://doi.org/10.26656/fr.2017.8(S2).14

obtained from the SGD2 spectra with an R²C value of 0.30 and RMSEC of 0.97%; the R²C value for guava obtained from the SGD2 spectra was 0.63, that obtained for sapodilla from the SNV spectra was 0.62 with an RMSEC of 1.22%, an R²C of 0.80, and RMSEC of 1.18% was obtained for banana from the SGD2 spectra.

The regression coefficient is used to understand the relationship between the wavelength and a predicted variable (water content). Strong correlations between wavelength and water content are shown as peaks or valleys, depending on whether the correlation is positive or negative. The regression coefficient can also be used to select important variables (Amanah *et al*., 2021). The regression coefficient values of the best model for each fruit are shown in Figure 2. Dragon fruit had more distinct peaks at 400-600 nm than the other fruits, which are highly correlated with anthocyanin. Guava had peaks at 650-700 nm, which correlate with chlorophyll. Presumably chlorophyll content in guava is higher than in sapodilla, banana and dragon fruit. The regression coefficients at 700-950 nm were small and relatively noisy due to the limitations of the instrument in the near-infrared region. The regression coefficients were also very noisy at 950-1,000 nm where the absorption of water molecules occurs. Dragon fruit had a more distinct peak at about 970 nm compared to the other fruits probably because of the high water content of dragon fruit (Table 1). From the regression coefficient profiles (Figure 2), specific variables that contribute to water content cannot be clearly defined. However, since the PLSR models were developed using the whole wavelength, the water content resulted from the pigments in the visible region and C-H and O-H vibrations in the shortwave near-infrared region of 900-1,000 nm (Magwaza and Opara, 2015).

Figure 3 shows the scatterplot between each fruit's actual and predicted water content. Bananas performed the best compared to the other fruits with an R² of prediction (R2P) of 0.75 and an RMSE of prediction (RMSEP) of 1.11%, followed by guava and sapodilla with R²P of 0.63 and RMSEP of 0.77%, and R²P of 0.52 and RMSEP of 0.89%, respectively. Dragon fruit had the worst PLSR model with an R² of 0.11 and RMSEP of 1.05% possibly due to its high water content and thick peel.

4. Conclusion

The absorption of water in the Vis and NIR region (350–1,000 nm) is very weak which makes it a challenge to determine the water content of products. However, as water is a primary quality parameter and the price of Vis-NIR spectroscopy is quite low, the use of Vis-NIR spectroscopy to determine water content is feasible. This study demonstrated the effectiveness of Vis-NIR spectroscopy combined with PLSR to predict water content in fruits. The best PLSR model was developed for banana, which has low water content, while the worst model was for dragon fruit, which has high water content. Nevertheless, Vis-NIR spectroscopy can be used

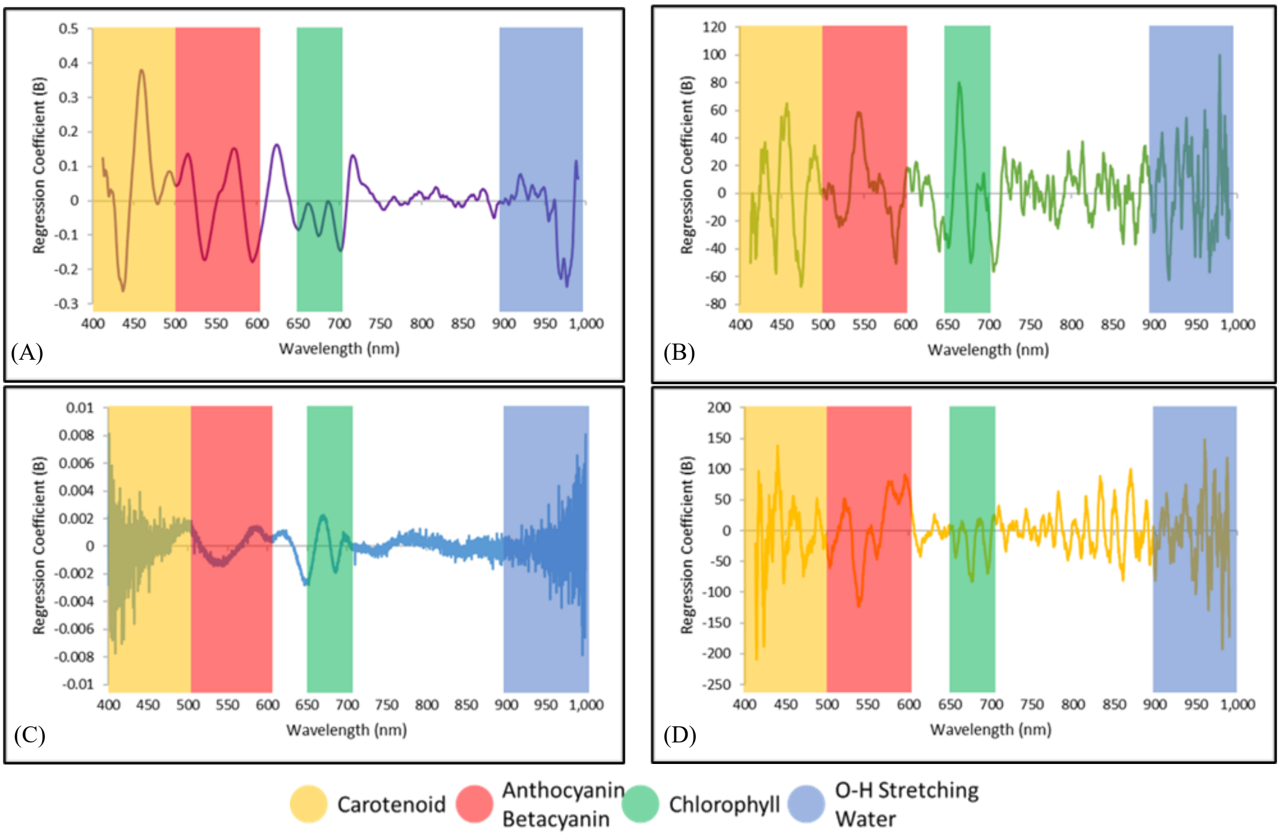

Figure 2. Regression coefficient from the best model for each fruit: (A) dragon fruit, (B) Guava, (C) Sapodilla, (D) Banana.

https://doi.org/10.26656/fr.2017.8(S2).14

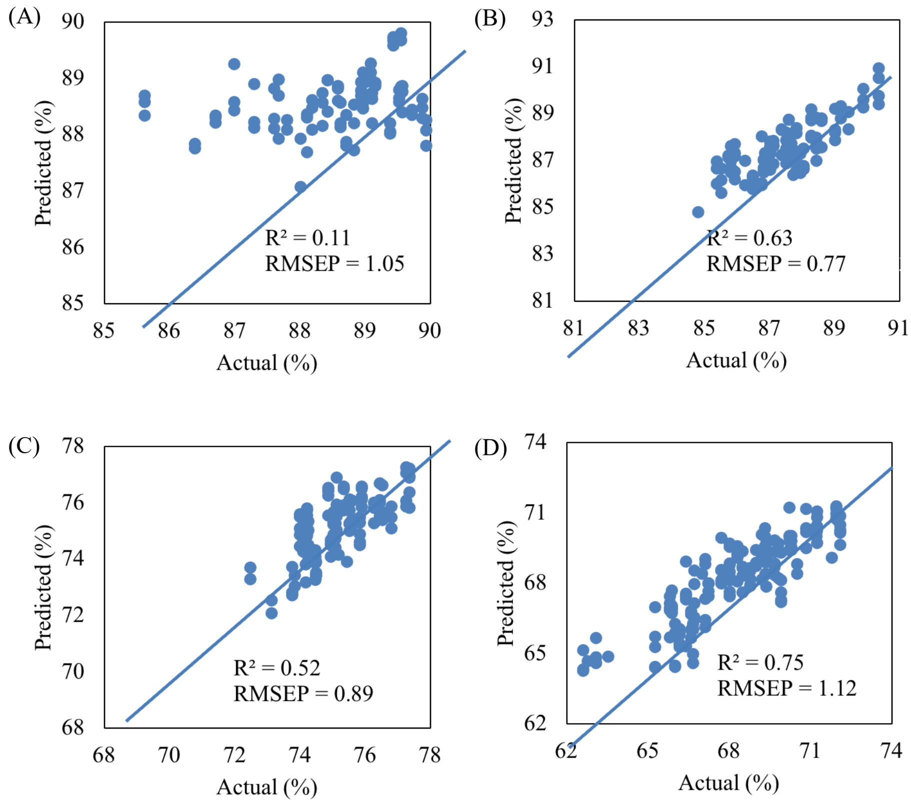

Figure 3. Scatter plots between actual water content and Vis-NIR predicted water content: (a) dragon fruit, (b) guava, (c) sapodilla and (d) banana.

to determine water content in fruits with low water content.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

Data used in this paper are parts of Evia Zunita Dwi Pratiwi's Graduate (Master) thesis and Diah Nur Rahmi's Undergraduate research reports from the Department of Agricultural and Biosystems Engineering, Faculty of Agricultural Technology, Universitas Gadjah Mada.

References

Amanah, H.Z., Wakholi, C., Perez, M., Faqeerzada, M. A., Tunny, S.S., Masithoh, R.E., Choung, M.G., Kim, K.H., Lee, W.H. and Cho, B.K. (2021). Near-infrared hyperspectral imaging (NIR-HSI) for nondestructive prediction of anthocyanins content in black rice seeds. *Applied Sciences*, 11(11), 2-14. https://doi.org/10.3390/app11114841

Blazquez, C., Downey, G., Donnell, C.O. and Callaghan, O. (2004). Prediction of moisture, fat and inorganic salts in processed cheese by near infrared reflectance spectroscopy and multivariate data analysis. *Journal of Near Infrared Spectroscopy*, 12, 149-157. https://doi.org/10.1255/jnirs.420

Bogomolov, A., Mannhardt, J. and Heinzerling, O. (2018). Accuracy Improvement of In-line Near-Infrared Spectroscopic Moisture Monitoring in a Fluidized Bed Drying Process. *Frontiers in Chemistry*, 6, e00388. https://doi.org/10.3389/fchem.2018.00388

Chia, K.S., Abdul Rahim, H. and Abdul Rahim, R. (2012). Prediction of soluble solids content of pineapple via non-invasive low cost visible and shortwave near infrared spectroscopy and artificial neural network. *Biosystems Engineering*, 113(2), 158-165. https://doi.org/10.1016/j.biosystemseng.2012.07.003

Eboibi, O. and Uguru, H. (2018). Effect of Moisture Content on the Mechanical Properties of Cucumber Fruit. *International Journal of Scientific and Engineering Research*, 9(4), 671-678.

Ehiem, J.C., Ndirika, V.I.O. and Onwuka, U.N. (2016). Effect of moisture content on some physical properties of *Canarium schweinfurthii* Engl. fruits. *Research in Agricultural Engineering*, 62(4), 162-169. https://doi.org/10.17221/11/2015-RAE

Kuang, B. and Mouazen, A.M. (2011). Calibration of visible and near infrared spectroscopy for soil analysis at the field scale on three European farms. *European Journal of Soil Science*, 62(4), 629-636. https://doi.org/10.1111/j.1365-2389.2011.01358.x

Magwaza, L.S. and Opara, U.L. (2015). Analytical methods for determination of sugars and sweetness of horticultural products-A review. *Scientia Horticulturae*, 184, 179-192. https://doi.org/10.1016/j.scienta.2015.01.001

Magwaza, L.S., Opara, U.L., Cronje, P.J.R., Landahl, S., Nieuwoudt, H.H., Mouazen, A.M., Nicolaï, B.M. and Terry, L.A. (2014). Assessment of rind quality of “Nules Clementine” mandarin fruit during postharvest storage: 2. Robust Vis/NIRS PLS models for prediction of physico-chemical attributes. *Scientia Horticulturae*, 165, 421-432. https://doi.org/10.1016/j.scienta.2013.09.050

Manley, M., Joubert, E., Myburgh, L. and Kidd, M. (2007). Prediction of soluble solids content and post-storage internal quality of Bulida apricots using near infrared spectroscopy. *Journal of Near Infrared Spectroscopy,* 15, 179-188. https://doi.org/10.1255/jnirs.725

Manley, M., Zyl, L.V. and Osborne, B.G. (2002). Using Fourier transform near infrared spectroscopy in determining kernel hardness, protein and moisture content of whole wheat flour. *Journal of Near Infrared Spectroscopy*, 10, 71-76. https://doi.org/10.1255/jnirs.323

Masithoh, R.E., Lohumi, S., Amanah, H.Z., Yoon, W.S. and Cho, B.K. (2020). Development of multi-product calibration models of various root and tuber powders by Fourier Transform Near Infra-red (FT-NIR) spectroscopy for the quantification of polysaccharide contents. *Heliyon,* 6(10), e05099. https://doi.org/10.1016/j.heliyon.2020.e05099

Masithoh, R.E., Pahlawan, M.F.R. and Wati, R.K. (2021). Non-destructive determination of SSC and pH of banana using a modular Vis/NIR spectroscopy: comparison of Partial Least Square (PLS) and Principle Component Regression (PCR). *IOP Conference Series: Earth and Environmental Science*, 752(1), 012047. https://doi.org/10.1088/1755-1315/752/1/012047

Pahlawan, M.F.R., Murti, B.M.A. and Masithoh, R.E. (2022). The potency of Vis/NIR spectroscopy for classification of soybean based of colour. *IOP Conference Series: Earth and Environmental Science,* 1018(1), 012015. https://doi.org/10.1088/1755-1315/1018/1/012015

Pahlawan, M.F.R., Wati, R.K. and Masithoh, R.E. (2021). Development of a low-cost modular VIS/NIR spectroscopy for predicting soluble solid content of banana. *IOP Conference Series: Earth and Environmental Science*, 644(1), 012047. https://doi.org/10.1088/1755-1315/644/1/012047

Priambodo, D.C., Saputro, D., Pahlawan, M.F.R., Saputro, A.D. and Masithoh, R.E. (2022). Determination of Acid Level (pH) and Moisture Content of Cocoa Beans at Various Fermentation Level Using Visible Near-Infrared (Vis-NIR) Spectroscopy. *IOP Conference Series: Earth and Environmental Science*, 985(1), 012045. https://doi.org/10.1088/1755-1315/985/1/012045

Sánchez, F.D., López, E.M.S., Kerstupp, S.F., Ibarra, R.V. and Scheinvar, L. (2006). Colorant extraction from red prickly pear (*Opuntia lasiacantha*) for food application. *Electronic Journal of Environmental, Agricultural and Food Chemistry*, 5(2), 1330-1337.

Saputri, D.A.S., Pahlawan, M.F.R., Masithoh, R.E. and Murti, B.M.A. (2022). Vis/NIR spectroscopy for non-destructive method in detecting soybean seeds viability. *IOP Conference Series: Earth and Environmental Science,* 1038(1), 012043. https://doi.org/10.1088/1755-1315/1038/1/012043

Saputro, D., Priambodo, D.C., Pahlawan, M.F.R. and Masithoh, R.E. (2022). Classification of Cocoa Beans Based on Fermentation Level Using PLS-DA Combined with Visible Near-Infrared (VIS-NIR) Spectroscopy. In 2nd International Conference on Smart and Innovative Agriculture, p. 100-106. USA: Atlantis Press. https://doi.org/10.2991/absr.k.220305.015

Savenije, B., Geesink, G.H., Van Der Palen, J.G.P., Hemke, G., Hopkins, D. and Ouali, A. (2006). Prediction of pork quality using visible/near-infrared reflectance spectroscopy. *Meat Science*, 73(1), 181-184. https://doi.org/10.1016/j.meatsci.2005.11.006

Szulc, J., Gozdecka, G. and Poćwiardowski, W. (2020). The application of NIR spectroscopy in moisture determining of vegetable seeds. *Czech Journal of Food Sciences*, 38(2), 131-136. https://doi.org/10.17221/57/2019-CJFS

Wati, R.K., Pahlawan, M.F.R. and Masithoh, R.E. (2021). Development of calibration model for pH content of intact tomatoes using a low-cost Vis/NIR spectroscopy. *IOP Conference Series: Earth and Environmental Science*, 686(1), 012049. https://doi.org/10.1088/1755-1315/686/1/012049

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(2), 111-115. https://doi.org/10.1007/s11694-008-9031-7

Zhang, Q., Li, Q. and Zhang, G. (2012). Rapid determination of leaf water content using VIS/NIR spectroscopy analysis with wavelength selection. *Spectroscopy*, 27(2), 93-105. https://doi.org/10.1155/2012/276795

