Non-Destructive Evaluation of Moisture Content in Single Soybean Seed Using Vis-NIR Spectroscopy

Muhammad Fahri Reza Pahlawan[1] Devi Alicia Surya Saputri[1] Rudiati Evi Masithoh[1,*]

[1]Department of Agricultural and Biosystems Engineering, Faculty of Agricultural Technology, Universitas Gadjah Mada, Jl. Flora No. 1 Bulaksumur, Yogyakarta 55281, Indonesia
[*]Corresponding author. Email: evi@ugm.ac.id

ABSTRACT

In this study, moisture content of soybean was evaluated using visible near infrared (Vis-NIR) spectroscopy. Soybean were dried at 60˚C for up to 10 hours to get moisture variation. A total of 200 soybean were used in this study which made a total of 600 reflectance spectra scanned with Vis-NIR spectrometer at 400-1000 nm. All samples were randomly divided into calibration set (2/3 samples) and prediction set (1/3 samples). Partial least square regression (PLSR) was used for developing calibration model for determining moisture content of soybean seed. Original and several pre-processed spectra such as area normalization, standard normal variate (SNV), multiple scatter correction (MSC), Savitzky-Golay smoothing, and Savitzky-Golay derivative were used in PLSR. The best PLSR model was obtained using 2nd Savitzky-Golay derivative with determination coefficient of calibration (R^2C) of 0.93 and root mean square error of calibration (RMSEC) of 0.004%. The PLSR model was then applied to prediction data set which resulted in determination coefficient of prediction (R^2P) of 0.82 and root mean square error of prediction (RMSEP) of 0.006%. The result showed the potency of Vis-NIR spectroscopy to predict moisture content in soybean seed.

Keywords: *moisture content, soybean seed, spectroscopy, Vis-NIR*

1. INTRODUCTION

Soybean is usually processed as food such as tofu, tempeh, and soymilk, used as livestock fee, or processed into biodiesel [1]. Soybean has many health benefits since it contains 8 essential aminos acids and other useful nutrients [2]. Being a popular food product causes soybean to become one of the most important products in the food world trade. Soybean is usually dried to a certain moisture content to extend shelf life and reduce weight making it is easy to transport. Moisture content can affect the physical properties of soybean and also affect seed germination; therefore, information of soybean moisture content is important.

Commonly, moisture content is determined using thermogravimetry method or moisture tester. Those methods require samples destruction prior to analysis which affect future usefulness of samples and require long-time analysis. Hence, a simple, fast, accurate, and non-destructive method is needed to measure soybean moisture content. Visible-near infrared (Vis-NIR) spectroscopy is a non-destructive method that can measure moisture content in agricultural products. Vis-NIR spectroscopy is able to measure moisture content in guava [3], cacao bean [4], tea leaf [5], and sugar beet [6]. Measurement using Vis-NIRs is based on colour pigment in visible region and C-H-O-N molecules interaction with light in near infrared region [7].

Spectroscopy is usually combined with chemometrics analysis to relate spectra variable to desired chemical composition of samples [8]. Common chemometrics analysis method is partial least square regression (PLSR). PLSR is a multivariable analysis method with 2 stages of analysis. First, dimension reduction of spectra variables into new variables which can be up to 20 variables called PLS factors. When creating a new variable, PLS will also take into account the desired variable [9]. The second stage is relate the new variable with desired variable using linear regression [9]. PLS factor with the highest R^2 and lowest RMSE will be selected.

A. D. Saputro et al. (Eds.): ICOSEAT 2022, ABSR 26, pp. 396–400, 2023.
https://doi.org/10.2991/978-94-6463-086-2_52

The modular type Vis-NIR spectrometer is a relatively low-cost spectrometer compared to other IR spectrometers. This modular type Vis-NIR spectroscopy can be used to evaluate fruit parameters in-situ quickly with minimal preparation. Hence, in this study, calibration model to predict moisture content of soybean will be created based on reflectance spectra at visible and near infrared region.

2. METHODOLOGY

2.1 Sample

A total of 220 soybean sample were used in this study. Soybean samples were collected from local market in Yogyakarta, Indonesia. Initial moisture content of soybean was ±12.30% (%wb). To obtain moisture content diversity, sample was dried at 60˚C for up to 10 hours. Every 1 hour, 10 bean samples were removed from the dryer and the spectra were measured.

2.2 Spectra measurement

Reflectance spectra were measured using visible near-infrared spectrometer (Flame-T-VIS-NIR Ocean Optics, 350-1000 nm) with tungsten halogen lamps (360-2400 nm, HL-2000-HP-FHSA Ocean Optics) and reflectance probes (QR400-7 VIS-NIR Ocean Optics). Before reflectance spectra measurement, instrument calibration procedures were done by measuring the spectra of white reflectance standard followed by background spectra. Instrument calibrations were done every 10 samples. Soybean samples were measured from below with 5 mm distance between probe and samples (Figure 1). Each sample was measured in triplicate.

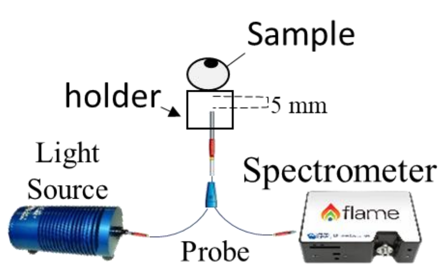

Figure 1 Spectra acquisition set-up

2.3 Moisture content analysis

Moisture content measurement were done based on thermogravimetric method. After spectra measurement, 10 were weighed to get $w_{before\ oven}$. Sample were then placed in an aluminium cup and dried in 105˚C using laboratory oven for 24 h. After dried, sample were weighed again to get $w_{after\ oven}$. Wet basis moisture content was measured using Equation 1.

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

[Equation 1]

2.4 Chemometrics analysis

Reflectance spectra obtained were compiled in MS. Excel and imported to Unscrambler®X software (CAMO, Oslo, Norway) for multivariate analysis. Reflectance spectra covered wavelength range of 345 – 1033 nm. However, due to the visible noises appeared at wavelength <400 nm and >1000 nm, only wavelength range of 400-1000 nm were used in analysis. In this study, multivariate analysis used was partial least square analysis (PLSR) used to relate 3182 variables of spectra to moisture content variable. A good model was identified by a high value of R^2 (maximum of 1) and a low value of RMSE. To improve model performance, several pre-processing methods were applied to original spectra such as standard normal variate (SNV), multiple scatter correction (MSC), area normalization (AN), 1st and 2nd order of Savitzky-Golay derivatives (SGD), and Savitzky-Golay smoothing (SGS). Model with the best performance of each pre-processing method was selected.

3. RESULTS AND DISCUSSION

3.1 Soybean moisture content

Table 1 showed the statistical parameter of soybean used in this study. From a total of 220 soybean samples, the moisture content had an average of 6.89%, standard deviation of 1.39%, moisture range of 4.61%, and ratio of standard deviation to range (std/range) of 30.10%. Based on [10], large variation of data was indicated by std/range > 20% which produced a robust calibration model.

Soybean was dried up to 10 hours to vary the moisture content of samples. Spectra data of soybean after drying were shown in Figure 2. Spectra of soybean showed that reflectance spectra tended to be lower as the drying time increased. However, all spectra showed similar pattern. Low reflectance at the beginning of the wavelength (400-500 nm) were influenced by carotenoids such as lutein content in soybeans [7]. Reflectance valley at around 670 nm was correlated to chlorophyll content [11]. At near infrared region (700-1000 nm), reflectance spectra were high and appeared flat due to low absorbance of OH and CH molecules of water and carbohydrate.

3.2. Partial least square regression of soybean ’s moisture content

Partial least square regression (PLSR) was used to relate multivariable data of spectra with desired variable (moisture content). To increase the model performances, several pre-processing methods were applied to the original spectra. The performances of PLSR model for the prediction of soybean moisture content was shown in Table 2. Using original spectra, R^2C obtained was 0.486 with RMSEC of 0.010. Table 2showed that pre-processing method such as SNV, MSC, AN, SGD1, and SGD2 increased model performance, but SGS reduced model performances. The best model performance was obtained using 2nd derivative of Savitzky-Golay (SGD2). The best model obtained has R^2C of 0.933 and RMSEC of 0.004. Prediction performance of the best model has R^2P of 0.823 and RMSEP of 0.006.

Table 1. Moisture content (MC) of soybean

Average	Std	Min	Max	Range	Std/Range
6.89%	1.39%	4.93%	9.54%	4.61%	30.10%

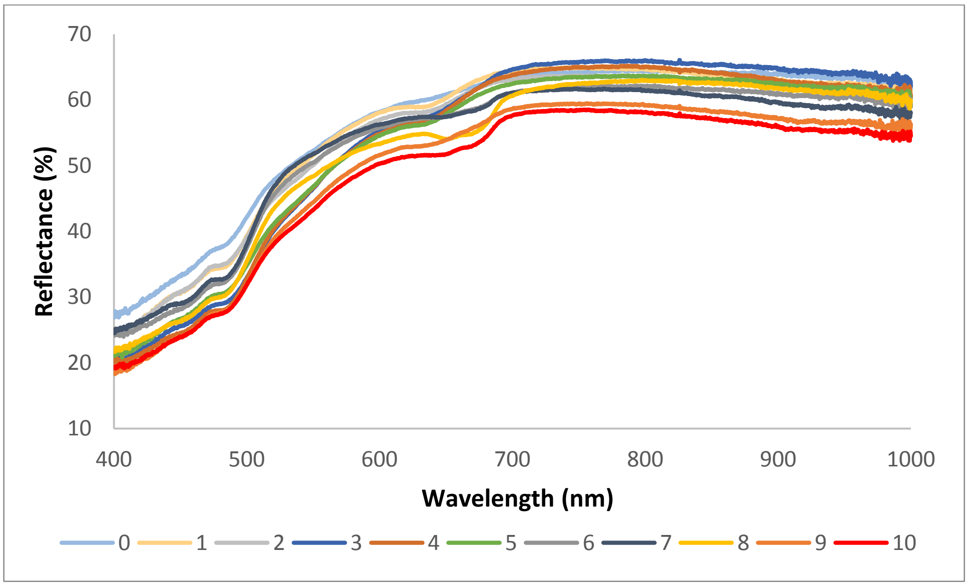

Figure 2 Visible near-infrared spectra of soybean after drying (The legends show variation of drying time)

Table 2. Performance of PLSR model for soybean 's moisture content

Pre-process	Calibration		Cross Validation		Prediction	
spectra	R^2C	RMSEC	R^2CV	RMSECV	R^2P	RMSEP
Original	0.486	0.010	0.514	0.010	0.501	0.009
SNV	0.707	0.008	0.611	0.009	0.693	0.007
MSC	0.728	0.008	0.629	0.009	0.650	0.008
AN	0.670	0.009	0.553	0.010	0.675	0.007
SGS	0.367	0.012	0.226	0.013	0.453	0.010
SGD1	0.876	0.005	0.769	0.007	0.799	0.006
SGD2	0.933	0.004	0.825	0.006	0.823	0.006

Note: R^2C= regression coefficient of calibration, R^2CV= regression coefficient of cross validation, R^2P= regression coefficient of prediction, RMSEC= root mean square error of calibration, RMSEC= root mean square error of cross validation, RMSEP= root mean square error of prediction, SNV= standard normal variate, MSC= multiple scatter correction, AN= area normalization, SGS= Savitzky-Golaysmoothing, SGD1= Savitzky-Golay 1st derivative, SGD2= Savitzky-Golay 2nd derivative.

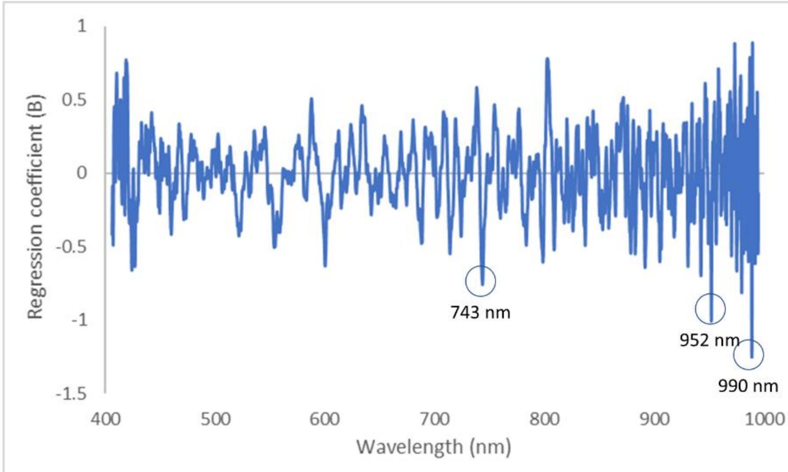

Figure 3 Regression coefficients (B) of PLSR model

Figure 3 was regression coefficients (B) obtained from the best calibration model to predict moisture content in soybean. Not only the spectra were noisy, but also the wavelengths responsible to the calibration model were not easily identified. Based on [12], wavelengths that might contribute to water absorption were at 760 nm and 970 nm. In this study, several wavelengths that could be noticed near water absorption wavelength were at 743 nm, 952 nm, and 990 nm.

4. CONCLUSION

In near infrared region, the absoption of water molecules was weak thus the use of Vis-NIR spectroscopy to detect moisture content was challenging. The study proved that reflectance spectra at visible and near infrared region could be used to predict moisture content in soybean seed. The best calibration PLSR model was obtained using 2nd derivative of Savitzky-Golay spectra with R^2C of 0.933 and RMSEC of 0.004%, while prediction performance had R^2P of 0.823 and RMSEP of 0.006%.

REFERENCES

[1] A. Barrett, Long-Term World Soybean Outlook. US Soy US Soy. 2019.

[2] A. Kakade, S. Khodke, S. Jadhav, et al., Effect of Moisture Content on Physical Properties of Soybean. Int. J. Curr. Microbiol. Appl. Sci. vol. 8, 2019, pp. 1770–1782, doi: 10.20546/ijcmas.2019.804.206.

[3] Kusumiyati, Y. Hadiwijaya, I. E. Putri, et al., Water content prediction of "crystal" guava using visible-near infrared spectroscopy and chemometrics approach. IOP Conf. Ser. Earth Environ. Sci. vol. 393, 2019, pp. 0–5, doi: 10.1088/1755-1315/393/1/012099.

[4] D. C. Priambodo, D. Saputro, M. F. R. Pahlawan, et al., Determination of Acid Level (pH) and Moisture Content of Cocoa Beans at Various Fermentation Level Using Visible Near-Infrared (Vis-NIR) Spectroscopy. IOP Conf. Ser. Earth Environ. Sci. 985: in IOP Conference Series: Earth and Environmental Science, 2022. doi: 10.1088/1755-1315/985/1/012045.

[5] Z. Huang, A. Sanaeifar, Y. Tian, et al., Improved generalization of spectral models associated with Vis-NIR spectroscopy for determining the moisture content of different tea leaves. J. Food Eng. vol. 293, 2021, 110374, doi: 10.1016/j.jfoodeng.2020.110374.

[6] L. Pan, R. Lu, Q. Zhu, et al., Measurement of moisture, soluble solids, sucrose content and mechanical properties in sugar beet using portable visible and near-infrared spectroscopy. Postharvest Biol. Technol. vol. 102, 2015, pp. 42–50, doi: 10.1016/j.postharvbio.2015.02.005.

[7] M. F. R. Pahlawan, B. M. A. Murti, and R. E. Masithoh, The potency of Vis/NIR spectroscopy for classification of soybean based of colour. IOP Conf. Ser. Earth Environ. Sci. 2022.

[8] R. E. Masithoh, S. Lohumi, H. Z. Amanah, et al. 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, 2020.

[9] G. M. Vasques, S. Grunwald, and J. O. Sickman. Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. Geoderma vol. 146,

2008, pp. 14–25, doi: 10.1016/j.geoderma.2008.04.007.

[10] B. Savenije, G. H. Geesink, J. G. P. Van Der Palen, et al., Prediction of pork quality using visible/near-infrared reflectance spectroscopy. Meat Sci. vol. 73, 2006, pp. 181–184, doi: 10.1016/j.meatsci.2005.11.006.

[11] V. Cortés, C. Ortiz, N. Aleixos, et al., A new internal quality index for mango and its prediction by external visible and near-infrared reflection spectroscopy. Postharvest Biol. Technol. vol. 118, 2006, pp. 148–158, doi: 10.1016/j.postharvbio.2016.04.011.

[12] Q. Zhang, Q. Li, and G. Zhang. (2012) Rapid determination of leaf water content using VIS/NIR spectroscopy analysis with wavelength selection. Spectrosc. (New York) vol. 27, 2012, pp. 93–105, doi: 10.1155/2012/276795.