Non-Destructive Method for Predicting Sapodilla Fruit Quality Using Near Infrared Spectroscopy

Kusumiyati¹, S Mubarok¹, W Sutari¹, Farida¹, J S Hamdni¹, Y Hadiwijaya² and I E Putri²

¹ Lecturer of Agronomy Studies Program, Agriculture Faculty, Padjadjaran University, Jalan Raya Bandung Sumedang KM.21, Hegarmanah, Jatinangor, Kabupaten Sumedang 45363, Indonesia
² Alumni of Agronomy Studies Program, Agriculture Faculty, Padjadjaran University, Jalan Raya Bandung Sumedang KM.21, Hegarmanah, Jatinangor, Kabupaten Sumedang 45363, Indonesia

E-mail: kusumiyati@unpad.ac.id

Abstract. The non-destructive method was used to predict the moisture content, firmness, total dissolved solids, and color values (including L*, a*, b*, h° and C*) of sapodilla fruit using near infrared spectroscopy with wavelength of 312 - 1050 nm. Irradiated and measured sapodilla samples were analyzed using partial least squares (PLS) algorithm of multivariate calibration technique to build the calibration models. Validation analysis was performed to build validation models and verify the calibration models. The accuracy of the validation models were defined by the ratio performance deviation (RPD) and root mean squares error of prediction (RMSEP) values on each parameters. The RPD and RMSEP values of moisture content, firmness, total dissolved solids, L*, a*, b*, h° and C* were 1.35 (0.01), 3.82 (2.08), 2.04 (1.70), 3.32 (1.30), 3.20 (0.82), 3.41 (1.38), 3.64 (2.03) and 3.11 (1.32), respectively. The results showed that non-destructive method using near infrared spectroscopy was able predict the quality of sapodilla fruit with high accuracy and low error.

1. Introduction

Sapodilla is fruit originating from Central America. It grows wild in the forests of Central America and Mexico. During colonialism, Spanish introduced sapodilla to the Philippines. Nowadays sapodilla spreads to other countries including Indonesia where sapodilla growing commercially [1]. Quality of fruits could be describe by the value of color, firmness, and sugar content [2]. Generally, most researchers assess the quality of fruit by doing destructive analysis. This method damages the fruit, hence the fruit can not be sold afterwards. Destructive analysis also requires some chemical materials and consuming a lot times and costs [3].

Near infrared spectroscopy is new non-destructive method to predict fruit quality. It is a spectroscopic method based on atomic vibrations of the molecules. The infrared spectrum is generated by passing infrared radiation on the sample and determining the absorbed radiation. The advantages of using near infrared spectroscopy that this method is non-destructive, and also many sample forms can be investigated such as fluids, solvents, pastes, powders, fibers, and gases can be identified by appropriate sample selection techniques [4]. In addition, near infrared spectroscopy technique is environmentally friendly.

The near infrared spectroscopy technique was first used in agriculture to measure the moisture content of grain [5]. Nowadays, this technique has been done to determine the total soluble solids, firmness, acidity and dry matter of mangos [6, 7]. In another study near infrared spectroscopy technique was used to predict soluble solids content, moisture content, firmness and color of tomato fruits on-tree.
and after-harvesting [8, 9]. Other research reported that this technique was performed to predict the quality of bitter gourd [10]. Storage is one of the important things in post harvest handling of fruits, because the quality of the fruit may change during storage. The factors that affecting fruit quality changes are storage duration and storage temperature. This research aimed to predict the quality of sapodilla fruit, which stored for 0, 5 and 10 days at room temperature.

2. Materials and Method

2.1. Raw Materials
Sapodilla fruit for this study were obtained from Situraja, Sumedang, Indonesia. Each sapodilla sample was numbered before the analysis, then stored for 0, 5 and 10 days at room temperature (25°C). A total of 300 fruit samples were split into 2 groups, first group used as calibration set and second group used as validation set.

2.2. Spectra Data Acquisition
Sapodilla fruit samples were irradiated using near infrared spectrometer (NirVana AG410) with wavelength of 312-1050 nm. Each sample was irradiated 3 points on the front and 3 points on the back vertically starting from the top to the bottom of the fruit. After the spectra data were obtained, multivariate calibration of chemometric was performed to extract the information from near infrared spectrum data [11]. Then the spectra of data along with reference data were used to build the calibration models using partial least squares (PLS) algorithm.

2.3. Reference Data Measurement
The water content of sapodilla is calculated by calculating fresh wet weight compared with dry fruit weight. The sapodilla fruits were sliced then dried in the oven with temperature of 80°C. Dried samples were calculated every day until the weight were constant. The water content value was expressed by percentage (%) of water content. Firmness of sapodilla was calculated using AD-4932A-50N tension gauge. The firmness value of sapodilla fruit was expressed with Newton (N). TDS (total dissolved solids) were measured using PR-101α refractometer (Atago, Model 41325, Japan). Sapodilla fruits were extracted for the analysis. The obtained data was expressed as %Brix. The reference data of color was measured using chromameter (Konica Minota CR-400). The obtained data were the values of L*, a* and b*. L* shows the color brightness of the sapodilla fruit, the value of 0 means black while 100 means white. a* is a gradation of green to red, negative a* indicates green while positive a* denotes red. b* is a gradation of blue to yellow, negative value of b* defines blue while positive b* describes yellow. Hue (hθ) is color appearance parameter or color dimension, it describes whether the color is red, blue, yellow, green, brown, purple, orange, and other colors. Chroma (C*) is the color intensity, it explains the strength and weakness of the color. The formula of hθ and C* are described in the following equations :

\[ h^\theta = tg^{-1}\left(\frac{b^*}{a^*}\right) \]  
\[ C^* = \sqrt{a^{*2} + b^{*2}} \]

2.4. Chemometrics Analysis
Chemometrics is a science of extracting data using mathematics and statistics by analyzing chemical data. Chemometrics in spectroscopic analysis consists of three aspects, pre-processing of spectral data, model building and transfer model. One of the most important chemometric analysis is to construct a calibration model through a pattern recognition technique to identify the major similarities and patterns of data. Furthermore, this model is used to predict unknown samples [12]. Nowadays chemometrics is used to analyze multivariate data that involving multiple variables and each variable can be mutually correlated.
2.5. Partial Least Squares

Partial least squares is one of the most common used multivariate calibration technique in chemometrics due to the quality of calibration model and the ease of application. PLS is able to analyze a large number of data, having high level of collinearity, a lot number of X variables and some Y response variables [13]. This is dimensional data reduction method to find the most relevant factors in predicting and interpreting data. PLS reduces predictor data that is not relevant to the data diversity. In addition, this method can also be applied for classification purposes. PLS analysis was done using second derivative absorbance (d2a) spectra data of each quality parameter to build the calibration models.

2.6. Model Accuracy

Root mean square error (RMSE) is a method to evaluate the forecasting technique that is used to assess the accuracy level of forecasting models. RMSE measures the difference of values predicted by a model and the actual measured values. Root mean squares error of calibration (RMSEC) is RMSE for calibration model, on the other hand the root mean squares error of prediction (RMSEP) is the RMSE for validation model. RMSE informs the average inexactness that can be expected for predictions of impending samples. The equations of RMSEC and RMSEP are defined in the following explanation [14]:

\[ \text{RMSEC} = \sqrt{\frac{1}{n_c} \sum_{i=1}^{n_c} (\hat{y}_i - y_i)^2} \]  \hspace{1cm} (3)

\[ \text{RMSEP} = \sqrt{\frac{1}{n_p} \sum_{i=1}^{n_p} (\hat{y}_i - y_i)^2} \]  \hspace{1cm} (4)

where:
\( \hat{y}_i \) : near infrared prediction,
\( y_i \) : actual data,
\( n_c \) : amount of data in the calibration set,
\( n_p \) : amount of data in the validation set.

RPD (Ratio performance deviation) is determined by dividing the standard deviation of reference data in validation set with root mean squares error of prediction. RPD value ranged from 1.5 to 2 describes the model is able to distinguish low from high values of the reference data; value ranged from 2 to 2.5 defines the decent quantitative model prediction; RPD value above 2.5 indicates the excellent model prediction [15, 4]. RPD is calculated using the following formula:

\[ \text{RPD} = \frac{\text{SD}}{\text{RMSEP}} \]  \hspace{1cm} (5)

3. Result and Discussion

The spectra data were preprocessed by second derivative pretreatment to examine the peaks of sapodilla fruit samples. The second derivative (d2a) spectra were used to reduce the ratio of noise. The d2a spectra showed the peaks at around 360, 640 and 900 nm (Figure. 1).
Figure 1. Typical spectra data (d2a) of sapodilla in wavelength of 312 – 1050 nm.

Reference data were obtained from the destructive analysis. 6 reference data and 6 spectra data were collected from each samples. The amount of fruit used in this study was 300 fruit samples. Data selection was then performed for calibration and validation purposes. Validation set used different data from the calibration set. The data used in calibration set were 600 reference data and 600 spectra data, as well as the validation set. Table 1 shows the reference data used in the calibration and validation set.

Firmness value average of calibration set with destructive method yielded 22.68, 13.23 and 12.21, total dissolved solids generated 24.93, 21.94 and 20.38, L* value measurement obtained 57.77, 54.63 and 52.28, b* value generated 27.04, 22.69 and 21.26, C* obtained 27.33, 23.66 and 21.73 for H0, H5, and H10 treatments, respectively (Table 1). These results showed the trend of impairment along with the length of storage duration as illustrated in Fig. 2. The decline trend also occurred in validation set. This revealed that the length of storage duration had effect in decreasing the quality of sapodilla on firmness, TDS, L*, b* and C* parameters. This is affirmed by [16] that reported there were firmness
and total dissolved solids loss of mature dutch eggplant fruit in storage treatment of 5 and 10 days at room temperature.

3.1. Calibration models
The results of PLS analysis for calibration set showed the $R^2$ and RMSEC values were 0.57 (0.01), 0.93 (2.02), 0.82 (1.47), 0.91 (1.32), 0.93 (0.72), 0.92 (1.35), 0.93 (2.00), and 0.93 (2.00) for the moisture content, firmness, total dissolved solids, $L^*$, $a^*$, $b^*$, $h^\circ$ and $C^*$ values, respectively (Table 2). The highest value of $R^2$ were shown in $a^*$ and firmness (Figure 2). High $R^2$ value indicates the calibration models are sufficiently good and feasible to be analyzed in validation set. In accordance with [17] statement, which states that the value of $R$ calibration as 0.71 ($R^2=0.5$) or higher indicates good value for calibration screening and also to be proceeded for validation analysis.

Table 1. Reference data of quality parameters of the calibration and validation set.

| Quality Parameter | Calibration (n=600) | Validation (n=600) |
|------------------|---------------------|--------------------|
|                  | Range   | Mean   | SD     | Range   | Mean   | SD   |
| MC               |         |        |        |         |        |      |
| H0               | 0.66-0.75 | 0.70   | 0.01   | 0.66-0.75 | 0.70 | 0.01 |
| H5               | 0.63-0.78 | 0.71   | 0.02   | 0.63-0.78 | 0.71 | 0.02 |
| H10              | 0.65-0.75 | 0.71   | 0.02   | 0.65-0.75 | 0.71 | 0.01 |
| Firmness         |         |        |        |         |        |      |
| H0               | 18.03-27.81 | 22.68 | 1.93   | 15.59-27.81 | 22.42 | 1.99 |
| H5               | 0.62-25.57 | 13.23 | 8.30   | 0.62-25.57 | 14.06 | 8.14 |
| H10              | 0.23-25.60 | 12.21 | 8.42   | 0.23-26.13 | 12.45 | 8.56 |
| TDS              |         |        |        |         |        |      |
| H0               | 20.4-32.20 | 24.93 | 2.21   | 20.5-30.60 | 24.92 | 2.19 |
| H5               | 15.9-29.40 | 21.94 | 3.02   | 12.3-28.80 | 21.93 | 3.00 |
| H10              | 11.7-29.20 | 20.38 | 3.77   | 11.6-28.00 | 20.36 | 3.86 |
| $L^*$             |         |        |        |         |        |      |
| H0               | 47.63-65.62 | 57.77 | 2.49   | 50.22-64.18 | 57.59 | 2.30 |
| H5               | 45.30-63.00 | 54.63 | 3.63   | 42.45-63.47 | 54.75 | 4.07 |
| H10              | 33.65-61.33 | 52.28 | 5.27   | 37.30-60.48 | 52.74 | 4.57 |
| $a^*$             |         |        |        |         |        |      |
| H0               | (-1.55)-11.81 | 3.94 | 2.84   | (-1.34)-11.06 | 4.27 | 2.39 |
| H5               | (-3.01)-12.72 | 6.55 | 2.50   | (-1.17)-11.34 | 6.30 | 2.62 |
| H10              | (-3.72)-10.55 | 5.10 | 2.37   | (-2.25)-10.34 | 4.92 | 2.48 |
| $b^*$             |         |        |        |         |        |      |
| H0               | 19.98-33.69 | 27.04 | 2.51   | 21.46-32.37 | 26.91 | 2.15 |
| H5               | 11.44-31.31 | 22.69 | 4.48   | 12.01-31.22 | 22.92 | 4.34 |
| H10              | 10.29-30.38 | 21.26 | 5.03   | 10.22-32.70 | 21.56 | 5.30 |
| $h^\circ$         |         |        |        |         |        |      |
| H0               | 66.70-89.79 | 80.43 | 4.93   | 62.92-89.36 | 80.69 | 4.71 |
| H5               | 60.01-88.00 | 73.64 | 6.75   | 60.63-89.47 | 73.40 | 6.69 |
| H10              | 47.95-88.89 | 73.90 | 8.72   | 56.30-88.82 | 74.89 | 8.13 |
| $C^*$             |         |        |        |         |        |      |
| H0               | 21.94-33.37 | 27.33 | 2.13   | 21.31-32.26 | 27.57 | 2.12 |
| H5               | 13.71-31.43 | 23.66 | 4.04   | 14.56-31.27 | 23.69 | 3.68 |
| H10              | 11.65-30.74 | 21.73 | 4.79   | 12.83-30.53 | 22.40 | 4.32 |

MC – Moisture content; TDS – Total dissolved solids; SD – Standard deviation
The calibration models that have been built need to be reassessed to verify the calibration accuracy, therefore validation analysis was required. Validation method used in this research was internal validation. The data were divided into 2 groups, calibration and validation set. The calibration models that have been built using calibration set was then applied to the validation set to build the validation models.

3.2. Validation models
Validation models that have high correlation coefficient value (closed to 1) indicate the high validity. The values of $R^2$ and RMSEP of validation set were 0.46 (0.01), 0.93 (2.08), 0.76 (1.70), 0.90 (1.30), 0.90 (0.82), 0.91 (1.38), 0.92 (2.03), and 0.89 (1.32) for moisture content, firmness, total dissolved solids, $L^*$, $a^*$, $b^*$, $h^0$ and $C^*$, respectively (Table 2). The obtained results showed the validation models didn’t differ much from the calibration model. This accords with [15], which explain a good model should have a high $R$ value in calibration and validation set, as well as having low RMSEC and RMSEP values. In addition, RMSEC and RMSEP are expected to have almost the same value. RPD value of moisture content (1.35) had the lowest accuracy compared to the other parameters, followed by TDS (2.04), $C^*$ (3.11), $a^*$ (3.20), $L^*$ (3.32), $b^*$ (3.41), $h^0$ (3.64) and firmness (3.82). The use of RPD to evaluate fruit color was also investigated in tomato with feasible values of $L^*$ (3.66), $a^*$ (4.76), $b^*$ (1.77), $h^0$ (2.50) and $C^*$ (2.06) [9]. RPD values alone can not be used as benchmark for calibration potential, [3] suggest to consider other statistical analysis results such as $R^2$ value of validation set, RMSEP, and the number of factors used in the calibration set.

| Elements                      | Calibration (n=600) | Validation (n=600) | RPD  |
|-------------------------------|--------------------|--------------------|------|
|                               | $R^2$ | RMSEC | $R^2$ | RMSEP |       |
| Moisture Content              | 0.57  | 0.01  | 0.46  | 0.01  | 1.35  |
| Firmness                      | 0.93  | 2.02  | 0.93  | 2.08  | 3.82  |
| Total Dissolved Solids        | 0.82  | 1.47  | 0.76  | 1.70  | 2.04  |
| $L^*$                         | 0.91  | 1.32  | 0.90  | 1.30  | 3.32  |
| $a^*$                         | 0.93  | 0.72  | 0.90  | 0.82  | 3.20  |
| $b^*$                         | 0.92  | 1.35  | 0.91  | 1.38  | 3.41  |
| $h^0$                         | 0.93  | 2.00  | 0.92  | 2.03  | 3.64  |
| $C^*$                         | 0.89  | 1.40  | 0.89  | 1.32  | 3.11  |

$R^2$ – Coefficient of determination; RPD – Ratio performance deviation; RMSEC – Root mean standard error of calibration; RMSEP – Root mean square error of prediction

4. Conclusion
The results showed that non-destructive method using near infrared spectroscopy had high accuracy values to predict the firmness and color of the sapodilla fruit (including $L^*$, $a^*$, $b^*$, $h^0$, and $C^*$).
addition, good accuracy values were shown on moisture content and total dissolved solids parameters. Firmness, TDS, L*, b*, and C* parameters presented the trend of quality impairment along with the longer storage duration. Non-destructive method using near infrared spectroscopy technique was able to predict the quality of sapodilla which stored for 0, 5 and 10 days storage at room temperature.

Acknowledgement
This research was financially supported by LPPM of Padjadjaran University.

Conflicts of interest
He authors declare that there is no conflict of interest regarding the publication of this paper.

References
[1] Lakshminarayana S, Rivera M A M 1979 Proximate characteristics and composition of sapodilla fruits grown in Mexico Proceedings of the Florida State Horticultural Society 92: 303-305.
[2] Shewfelt R L 1999 What is quality? Postharvest Biology and Technology 15: 197-200.
[3] Ignat T, Schmilovitch Z, Fefoili J, Steiner B, Alkalai-Tuvia S 2012 Nondestructive measurement of ascorbic acid content in bell peppers by VIS–NIR and SWIR spectroscopy Postharvest Biology and Technology 74: 91–99.
[4] Stuart B 2004 Infrared spectroscopy : Fundamental and Applications Saunders College Pub. Philadelphia.
[5] Norris K H 1964 Design and development of a new moisture meter Journal of Agricultural Engineering 45: 370-372.
[6] Schmilovitch Z, Mizrah A, Hoffman A, Egozi H, Fuchs Y 2000 Determination of mango physiological indices by near-infrared spectrometry Postharvest Biology and Technology 19(3): 245–252.
[7] Nagle M, Mahayothee B, Rungpichayapichet P, Janjai S, Müller J 2010 Effect of irrigation on near-infrared (NIR) based prediction of mango maturity Scientia Horticulturae 125(4): 771–774.
[8] Kusumiyati, Akinaga T, Yonemori S, Kawasaki S, Tanabe T 2007 Evaluation of tomato quality on tree and after harvesting using portable NIR spectroscopy Journal Society of Agricultural Structures Japan 38(2): 117-126.
[9] Kusumiyati, Akinaga T, Tanaka M, Kawasaki S 2008a On tree and after harvesting evaluation of firmness, color, and lycopene content of tomato fruit using portable NIR spectroscopy Journal of Food Agriculture and Environment 6(2): 132-137.
[10] Kusumiyati, Akinaga T, Yonemori S, Shikanai T, Okamoto H, Tanabe T, Kawasaki S 2008b Internal quality sensor of fruit jagged surface: preliminary study of bitter gourd (Momordica charantia Linn) Acta Horticulturae 768: 391-398.
[11] Stchur P, Cleveland D, Zhou J, Michel R 2002 A review of recent applications of near infrared spectroscopy, and the characteristics of a novel PbS CCD array-based near infrared spectrometer Applied Spectroscopy Reviews 37: 383–428.
[12] Africia S M 2007 Kombinasi Metode Spektrofotometri dan Kalibrasi Multivariat untuk Penentuan Simultan Natrium Benzoat dan Kalsium Sorbat Bogor Indonesia Thesis Bogor Agricultural University.
[13] Kuno A, Matsuo M 2000 Nondestructive specification of solid mixtures by multivariate calibration of x-ray absorption near edge structure using artificial neural networks and partial least square Analytical Sciences 16: 597-602.
[14] Yahaya O K M, MatJafri M Z, Aziz A A, Omar A F 2015 Visible spectroscopy calibration transfer model in determining pH of Sala mangoes Journal of Instrumentation 10.
[15] Nicolaï B M, Beullens K, Bobelyn E, Peirs A, Saey S, Theron K I, Lammertyna J 2007 Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review Postharvest Biology and Technology 46: 99–118.
[16] Julianti E 2011 Pengaruh tingkat kematangan dan suhu penyimpanan terhadap mutu buah terong belanda (Cyphomandra betacea) Jurnal Hortikultura Indonesia 2(1): 14-20.
[17] Williams P C 1987 Variables affecting near-infrared reflectance spectroscopic analysis In: Williams P, Norris K (ed.) Near-infrared technology in the agricultural and food industries
American Association of Cereal Chemists Inc. Saint Paul pp. 143–167

[18] Viscarra R R V, McGlyn R N, McBratney A B 2006 Determining the composition of mineralorganic mixes using UV–vis–NIR diffuse reflectance spectroscopy Geoderma 146: 403-411.

[19] Williams P 2007 Application to agricultural and marine products In: Ozaki Y, McClure W F, Christy A A (Ed.) Near-Infrared spectroscopy in food science and technology John Wiley and Sons, Inc. Pub. New Jersey. pp. 165–218.

[20] Gómez A H, He Y, Pereira G 2006 Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using Vis/NIR-spectroscopy techniques Journal of Food Engineering 77(2): 313–319.