Development of Partial Least Square (PLS) Prediction Model to Measure the Ripeness of Oil Palm Fresh Fruit Bunch (FFB) by Using NIR Spectroscopy

Zaqlul Iqbal1, Sam Herodian2, Slamet Widodo2

1 Department of Agricultural Engineering, Universitas Brawijaya, Malang, ZIP 65145, Indonesia
2 Department of Agricultural and Biosystem Engineering, Bogor Agricultural University, Bogor, ZIP 23111, Indonesia.
E-mail: zaqluliqbal@ub.ac.id

Abstract. In order to develop a model for predicting the oil palm Fresh Fruit Bunch (FFB) ripeness, a rapid and non-destructive method such as NIR spectroscopy is utilized. This method has shown its capability to determine the quality of some crops by predicting their internal chemical contents. The objective of the research is to investigate the feasibility of NIR spectroscopy to predict water and oil content in FFB by developing a calibration model. Sixty samples of FFB were scanned by using NIRFlex N-500 spectrometer ranging from 1000 to 2500nm. Water and oil content of samples were measured after scanned. To develop a calibration model, Partial Least Square (PLS) Regression and pre-processing were conducted using Unscrambler X 10.3. The results showed that PLS performs well to establish a calibration model to predict water content using MSC pre-processing with r², factor, RSMECV, and RPD are 0.93, 3, 5.24, and 2, respectively. On the other hand, PLS could not be used well for establishing oil content calibration model because the result did not meet statistic parameters. For laboratory measurement, the model could predict water content of FFB; but it was limited to samples taken from the same variety and plantation. However, NIR Spectroscopy proposed a promising method to detect the ripeness of oil palm FFB.

1. Introduction
The yield of palm oil extraction is influenced by the quality of oil palm Fresh Fruit Bunch (FFB). Before production process, the quality of oil palm FFB can be evaluated by grading: classifying the ripeness of FFB on plantation and examining harvested FFB on palm oil factory. The first step is an important process because it determines the quality and yield of crude oil production. In order to manage this step, ripeness-fraction is used as a standard to predict ripening stages of FFB on field. This method uses human vision according to fruit colours. Table 1 shows the type of fraction employed in Indonesia [1]. Fraction 2 and 3 are preferred because they contain high oil content and acceptable Free Fatty Acid (FFA).

Nowadays, fraction level is commonly opted because it is easy to use; workers just use their vision to examine the redness level of FFB subjectively to predict the ripe ones. However, in a practical manner, the classification of FFB fraction depends on each worker decision leading to non-standardized results creating bias to other workers. Moreover, it requires lengthy experience to understand the exact level of fraction. Consequently, oftentimes, workers harvest unripe or over ripe FFB for oil production. It causes a decrease in yield and quality of crude oil production because of its low oil content or high FFA. To solve this problem, an objective and quantitative approach must be conducted to replace the recent...
method. NIR Spectroscopy has shown its capability to determine the quality of some fruits by assessing their internal attributes [2-8]. As a rapid and non-destructive method, NIR Spectroscopy can also predict an internal chemical content during ripening [9-11] which is very useful to determine the ripening stages. Moreover, previous research has been conducted to establish water content prediction model of FFB to perceive the ripening stages [12]. However, even the research gives a good coefficient of determination ($r^2$), it can be improved by recalculating the model for predicting purpose.

### Table 1. Fraction of Oil Palm Fresh Fruit Bunch (FFB)

| Fraction | Exocarp Colour     | Oil Yield (%) | Free Fatty Acid (%) | Ripeness Level          |
|----------|--------------------|---------------|---------------------|-------------------------|
| F00      | Black              | -             | -                   | Unripe Extremely        |
| F0       | Black Reddish      | 16            | 1.6                 | Unripe                  |
| F1       | Red Blackish       | 21.4          | 1.7                 | Unripe Slightly         |
| F2       | Yellow Reddish     | **22.1**      | **1.8**             | Ripe 1                 |
| F3       | Red Yellowish      | **22.2**      | **2.1**             | Ripe 2                 |
| F4       | Red Yellowish      | 22.2          | 2.6                 | Over ripe 1             |
| F5       | Red Yellowish      | 21.9          | 3.8                 | Over ripe 2             |

*IOPRI*

To process NIR spectroscopy data, chemometrics has been executed to develop a calibration model [13, 14]. Chemometrics is a data processing stage which utilizes a statistical method using spectral data to develop a qualitative model or a combination of internal quality data such as water content, oil content, fatty acid, and soluble solid content to build a quantitative model. The objective of this research is to develop a calibration model based on NIR Spectroscopy approach using chemometrics for predicting water content and oil content of oil palm FFB.

### 2. Materials and Methods

#### 2.1. Sample Preparation

Sixty samples of oil palm FFB (Sungai Pancur Clone Variety) were obtained from Cikasungka PTPN 8 Oil Palm Plantation Field, Bogor, Indonesia. Sixty samples were derived from 15 FFBs of each ripening stage (4, 5, 6, and 7 month); in addition, basically, the stages were chosen based on ripeness-fraction method (Table 1). In order to achieve better understanding to ripening stage, monthly examination stage were employed instead of fraction. The ripening estimation was defined by workers after a small fruit start growing on the oil palm tree. Several group of fruitlets were then detached from FFB, stored in cool-box, then transported to Post-Harvest and Food Processing Laboratory (TPPHP Laboratory), IPB, Bogor. Afterwards, fruitlets was prepared and cleaned. From one FFB, two groups were divided for water and oil content measurement. It needed six fruitlet for water test and nine fruitlet for oil content measurement. Water content was measured in TPPHP laboratory and oil content measurement was conducted in BB-Pascapanen Laboratory, Ministry of Agriculture, Bogor, Indonesia.

![Figure 1](image-url)  
**Figure 1.** (A) An oil palm Fresh Fruit Bunch (FFB), (B) group of fruitlet (C) oil palm fruitlets.
2.2. NIR Spectral Data Acquisition
Spectral acquisition was conducted in TPPHP Laboratory. Reflectance of NIR Spectral data were obtained from all single detached fruits using fibre optic solid NIRFlex N-500 scanning from 1000 to 2500 nm. The data were utilized to develop water content and oil content calibration model.

2.3. Water and Oil Content Measurement
Water content was measured by using oven drying method [15]. Five grams of fruit was placed into a dish (a, gram) and was dried using oven at 105°C. After 21 hours the sample was stored into a desiccator until reaching room temperature then weighed (b, gram). The following step was calculating water content by using Equation 1.

\[ m = \frac{a-b}{a} \times 100\% \]  

(1)

After acquiring the spectral data, samples were delivered to BB-Pascapanen Laboratory, Bogor. Soxhlet method, then, was utilized to measure oil content of the samples [16]. One to two grams of sample (W, gram) was placed in paper and dried using oven at 80°C for an hour. After carrying out this drying phase, the samples were extracted using hexane for 7 hours in flask (W₁, gram), dried at 105°C, cooled down and weighted. The process was repeated until reaching stable weight measurement (W₂, gram). The oil content was determined by using Equation 2.

\[ \% \text{ oil content} = \frac{W - W_1}{W_2} \times 100\% \]  

(2)

2.4. Statistics Analysis
Chemometric was performed using Unscrambler X 10.3. Several pre-processing techniques were conducted to improve the spectral signal: Derivative Savitzky-Golay (DG1), Standard Normal Variance (SNV), and Multiplicative Scattering Correction (MSC). Partial Least Square (PLS) Regression was employed to develop a calibration model for both water and oil content. In order to build a robust calibration model, a high Determination Coefficient (r²) might be achieved. Moreover, to avoid over-fitting model, the optimum Factor and low Root Mean Square Error of Cross-Validation (RSMCV) were defined. In addition, Ratio of Prediction to Deviation (RPD) was calculated as a ratio of Standard Deviation (SD) and RMSECV [17].

3. Results and Discussion
3.1. Chemical Content of oil Palm FFB
Based on the chemical content measurement of oil palm FFB, different FFB ripeness showed different performance of its chemical content. As the ripening stage increases, the water content decreases while the oil content slightly increases. From the measurement, the water content of FFB showed various values, ranging from 22.05% to 86.15%. Meanwhile, the oil content was varied from 1.25% to 53.91%. Table 2 illustrates a water content decline followed by oil content increment during ripening. From the results of the previous research, water content experiences derivation over the month [18] whereas the oil content is synthesized from the first month of ripeness stage until the fourth. After four-month duration, the oil content will be only produced in the mesocarp of the fruit [19].

The best condition to harvest FFB is approximately 27% of water content [20]; on this condition, the oil content reaches the optimum concentration with low FFA. Therefore, water content can be used as an indicator to determine a ripe FFB. Moreover, the analysis of simple regression and correlation shows that oil content had a high and positive correlation with its water content (r² = 0.732, Figure 2). The oil content becomes one of primary indicators to determine the optimum ripeness of oil palm fruit because it was the main composition to get crude oil.
As presented by Table 2, both water and oil content have high variation data on every month of ripening. The results prove that workers could not accurately predict a ripe FFB. Hence, a qualitative measurement will be needed to get a low deviation, high precision and data accuracy. After acquiring information related to the chemical contents, they will be combined with spectral data to establish a calibration model as a basic qualitative measurement.

3.2. NIR Spectral of Oil Palm FFB

The absorbance of oil palm spectral from the interval of wavelength from 1000 to 2500 nm is presented in Figure 3. Although only a little chemical information is defined by that spectral, several dominant peaks on the absorbance recorded from 1400-1500 nm and 1900-2000 nm. The wavelength of 1400-1500 nm corresponds to CH2, CH, and ROH indicating oil content; the spectral ranging from 1900 to 2000 nm corresponds to H2O indicating the presence of water [21]. In order to gain a good calibration model, the raw data might be fixed by using pre-processing technique. Overlapping signal in interval wavelength of 1150-1250 nm and 1667-1800 nm can be fixed by using First Derivative Savitzky Golay (DG1); a multiplicative Scattered Correction (MSC) or Standard Normal Variate (SNV). Pre-processing technique needs to be conducted to remove noise from 2000 to 2500 nm.
Figure 3. Absorbance spectrum of oil palm FFB

Full spectrum of spectral reflectance was used for establishing calibration model. Unlike the previous research which employs all data for developing calibration model [12], sixty samples of data in this research were divided into two groups: 40 (2/3 of 60 samples) sets of data for developing calibration model and 20 (1/3 of 60 samples) set of data for testing the model. To gain a robust calibration model in a small number of data, cross validation method was utilized to validate developed model. According to Figure 3, the calibration model without any pre-processing technique results r2 of 0.908, RMSECV of 6.35, RPD of 1.73 and factor of 4. When it was modified by using some pre-processing techniques, the model offers a better prediction ability (Table 3). The standard of error (RMSE) must below 1 in order to establish a good model [22][23][24]; however, when developing a model from high deviation data, an error value of approximately 4 could also explain a good result [17]. The calibration model of oil content (Figure 4) resulted r2 was 0.031, RMSECV was 11.51, RPD was 1.05 and factor was 1. Unlike water content model, when spectral data were modified by some pre-processing, the result does not show any significant difference. Hence, this model could not be used properly to predict oil content.

Because of limitation in obtaining a new sample for testing the model, 20 set of data of total sample from both water and oil content model development were used as a new set of samples. From the statistic parameter (Figure 4), the model could predict water content well. However, the prediction shows a limitation; this model could only predict samples taken from the same plantation as the samples built for model, and it is only used in laboratory scale.
Table 3. Model from several pre-processing techniques

| Attribute     | Pre-processing | Factor | $r^2$ | Calibration        | Prediction       | RPD |
|---------------|----------------|--------|-------|--------------------|------------------|-----|
|               |                |        |       | RMSEC  | RMSECV | RMSEP |       |
| Water Content | None           | 3      | 0.908 | 5.61   | 6.35   | 8.12  | 1.73 |
|               | DG1            | 4      | 0.961 | 3.64   | 5.51   | 6.21  | 2.27 |
|               | SNV            | 3      | 0.939 | 4.58   | 5.42   | 7.02  | 2    |
|               | MSC            | 3      | 0.938 | 4.59   | 5.24   | 7.06  | 2    |
| Oil Content   | None           | 1      | 0.031 | 10.27  | 11.51  | 8.63  | 1.05 |
|               | DG1            | 1      | 0.051 | 10.16  | 10.67  | 8.13  | 1.12 |
|               | SNV            | 1      | 0.055 | 10.14  | 10.96  | 8.19  | 1.11 |
|               | MSC            | 1      | 0.055 | 10.14  | 10.71  | 8.19  | 1.11 |

4. Conclusion
Two calibration models were developed by applying PLS regression combined with pre-processing technique. Several results of pre-processing technique were investigated to obtain good calibration model. To predict water content, the calibration model using MSC pre-processing technique proposes better accuracy and more stable calibration model with $r^2$, factor, RMSECV, and RPD are 0.93, 3, 5.24, and 2 respectively. On the other hand, PLS could not be conducted well for establishing oil content calibration model. In conclusion, NIR spectroscopy had showed a promising method to predict the ripeness of oil palm FFB based on water content. The value of water prediction can also be converted into oil content using equation derived from water and oil correlation. However, the model could only predict samples taken from the same plantation.

References

[1] A. U. Lubis, Kelapa Sawit (Elaeis guineensis Jacq.) di Indonesia. Medan: Indonesian Oil Palm Research Institut (IOPRI), 1992.
[2] S. Srivichien, A. Terdwongworakul, and S. Teerachaichayut, “Quantitative Prediction of Nitrate Level in Intact Pineapple Using Vis – NIRS,” J. Food Eng., vol. 150, pp. 29–34, 2015.
[3] T. S. Rizvi, F. Mabood, L Ali, M Al-Broumi, H. K. M. Al-Rabani, J. Hussain, F. Jabeen, S. Manzoor, and A. Al-Harrasi, “Application of NIR Spectroscopy Coupled with PLS Regression for Quantification of Total Polyphenol Contents from the Fruit and Aerial Parts of Citrullus colocynthis,” Phytochem. Anal., vol 29, no. 01, pp. 16-22, 2017.
[4] K. Ncama, U. L. Opara, S. Z. Tesfay, O. A. Fawole, and L. S. Magwaza, “Application of Vis/NIR Spectroscopy for Predicting Sweetness and Flavour Parameters of ‘Valencia’ Orange (Citrus sinensis) and ‘Star Ruby’ Grapefruit (Citrus x paradisi Macfad),” J. Food Eng., vol. 193, pp. 86–94, 2017.
[5] P. Maniwara, K. Nakano, D. Boonyakiat, S. Ohashi, M. Hiroi, and T. Tohyama, “The Use of Visible and Near Infrared Spectroscopy for Evaluating Passion Fruit Postharvest Quality,” J. Food Eng., vol. 143, pp. 33–43, 2014.
[6] J. Wang, J. Wang, Z. Chen, and D. Han, “Development of Multi-Cultivar Models for Predicting the Soluble Solid Content and Firmness of European Pear (Pyrus communis L.) Using Portable Vis–NIR Spectroscopy,” Postharvest Biol. Technol., vol. 129, pp. 143–151, 2017.
[7] M. Li, R. R. Pullanagari, T. Pranamornkith, I. J. Yule, and A. R. East, “Quantitative Prediction of Post Storage ‘Hayward’ Kiwifruit Attributes Using at Harvest Vis-NIR Spectroscopy,” J. Food Eng., vol. 202, pp. 46–55, 2017.
[8] C. Camps and C. Gilli, “Prediction of Local and Global Tomato Texture and Quality by FT-NIR Spectroscopy and Chemometric,” Eur. J. Hortic. Sci., vol. 82, no. 3, pp. 126–133, 2017.
[9] S. N. Jha, K. Narsaiah, P. Jaiswal, R. Bhardwaj, M. Gupta, R. Kumar, and R. Sharma, “Nondestructive Prediction of Maturity of Mango Using Near Infrared Spectroscopy,” J. Food Eng., vol. 124, pp. 152–157, 2014.

[10] J. A. Cayuela and J. F. García, “Sorting Olive Oil Based on Alpha-Tocopherol and Total Tocopherol Content Using Near Infrared Spectroscopy (NIRS) Analysis” J. Food Eng., vol. 202, pp. 79-88, 2017.

[11] C. Watanawan, T. Wasusri, V. Sraung, C. Wongs-Aree, and S. Kanlayanarat, “Near Infrared Spectroscopic Evaluation of Fruit Maturity and Quality of Export Thai Mango (Mangifera indica L. var. Namdokmai),” Int. Food Res. J., vol. 21, no. 3, pp. 1073–1078, 2014.

[12] Z. Iqbal, S. Herodian, and S. Widodo, “Prediction of Oil Palm Fresh Fruit Bunch (FFB) Water Content and Total Carotene Using NIR Spectroscopy,” J. Keteknikan Pertanian., vol. 2, no. 2, pp. 111–116, 2014.

[13] S. A. Mireei and M. Sadeghi, “Detecting Bunch Withering Disorder in Date Fruit by Bear Infrared Spectroscopy,” J. Food Eng., vol. 114, pp. 397–403, 2013.

[14] A. Pannico, R. E. Schouten, B. Basile, R. Romano, E. J. Woltering, and C. Cirillo, “Non-Destructive Detection of Flawed Hazelnut Kernels and Lipid Oxidation Assessment Using NIR Spectroscopy,” J. Food Eng., vol. 160, pp. 42–48, 2015.

[15] Badan Standarisasi Nasional (BSN), Cara Uji Makanan dan Minuman, SNI 01-2891-1992. Indonesia: BSN, 1992.

[16] Badan Standarisasi Nasional (BSN), Cara Uji Minyak dan Lemak, SNI 01-3555-1998. Indonesia: BSN, 1998.

[17] R. Karoui, A. M. Mouazen, E. Dufour, L. Pillonel, E. Schaller, J. D. Baerdemaeker, and J. O. Bosset, “Chemical Characterisation of European Emmental Cheeses by Near Infrared Spectroscopy Using Chemometric Tools,” Intl. Dairy. J., vol. 16, pp. 1211–1217, 2006.

[18] A. Keshvadi, J. B. Endan, H. Harun, D. Ahmad, and F. Saleena, “The Reflection of Moisture Content on Palm Oil Development During The Ripening Process of Fresh Fruits,” J. Food Agri. Env., vol. 10, no. 1, pp. 203-209, 2012.

[19] A. S, Razali, M. A.  Halim, and S. Z. Abidin, “A Review on Crop Plant Production and Ripeness Forecasting,” Intl. J. Agri. Crop., vol. 4, no. 2, pp. 54-63, 2012.

[20] S. Ketaren, Minyak dan Lemak Pangan. Jakarta: UI-Press, 1986.

[21] B. G. Osborne, Practical NIR Spectroscopy with Application in Food and Beverage Analysis. Singapore: Longman Singapore Publishers (Pte) Ltd, 1993.

[22] G. Fan, J. Zha, R. Du, and L. Gao, “Determination of Soluble Solids and Firmness of Apples by Vis / NIR Transmittance,” J. Food Eng., vol. 93, pp. 416–420, 2009.

[23] S. Saranwong, J. Somsrivichai, and S. Kawano, “Prediction of Ripe-Stage Eating Quality of Mango Fruit from Its Harvest Quality Measured Nondestructively by Near Infrared Spectroscopy,” Postharvest Biol. Tech., vol. 31, pp. 137–145, 2004.

[24] A. Moghimi, M. H. Aghkhani, A. Sazgarnia, and M. Sarmad, “Vis / NIR Spectroscopy and Chemometrics for The Prediction of Soluble Solids Content and Acidity (pH) of Kiwifruit,” Biosyst. Eng., vol. 106, pp. 295–302, 2010.