NON-DESTRUCTIVE PREDICTION OF OIL AND FREE FATTY ACID OF OIL PALM FRUITLETS USING NEAR-INFRARED SPECTROSCOPY AND HYBRID CALIBRATION METHOD

/

PREDIKSI KANDUNGAN MINYAK DAN ASAM LEMAK BEBAS BUAH SAWIT SECARA NON-DESTRUKTIF MENGGUNAKAN NEAR-INFRARED SPECTROSCOPY DAN METODE KALIBRASI HIBRID

I Wayan BUDIASTRA*[1,2], Sutrisno MARJAN[1], Nissa ADIARIFIA[1], Inna NOVIANTY[3], Yunisa Tri SUCI[1]
[1]IPB University, Department of Mechanical and Biosystem Engineering / Indonesia;
[2] IPB University, Centre for Research on Engineering Application in Tropical Agriculture / Indonesia;
[3] IPB University, Study Program of Computer Engineering / Indonesia;
Tel: +622518623026; E-mail: wbudiastra@apps.ipb.ac.id
DOI: https://doi.org/10.35633/inmateh-73-39

Keywords: *free fatty acid, hybrid calibration, NIR, oil content, oil palm fruitlets*

ABSTRACT

The study used near-infrared reflectance spectroscopy and hybrid calibration methods to predict oil and free fatty acid content of oil palm fruitlets non-destructively. The reflectance and chemical content of oil palm fruitlets were measured and the calibration between near infrared spectra and chemical content was performed using hybrid calibration methods. The developed hybrid calibration model was validated to predict oil and free fatty acid of other oil palm fruitlets. The best models to predict oil and free fatty acid of oil palm fruitlets respectively were the hybrid model of Partial Least Square-Artificial Neural Network with 25 factor components (R^2 = 0.96; SEP= 2.21%, RPD = 4.79) and 19 factor components (R^2 = 0.96; SEP= 0.25%, RPD = 4.24) using Savitzky-Golay first derivative spectra pre-treatment.

ABSTRAK

Penelitian ini menggunakan spektroskopi inframerah dekat dan metode kalibrasi hibrid untuk memprediksi kandungan minyak dan asam lemak bebas buah sawit secara nondestruktif. Reflektan dan kandungan kimia buah sawit diukur dan kalibrasi antara spektra inframerah dekat dan kandungan kimia dilakukan menggunakan metode kalibrasi hibrid. Model kalibrasi hibrid yang dibangun divalidasi untuk memprediksi kandungan minyak dan asam lemak bebas buah sawit lainnya. Model terbaik untuk memprediksi masing-masing kandungan minyak dan asam lemak bebas buah sawit adalah model hibrid Partial Least Square-Artificial Neural Network dengan input 25 komponen faktor (R^2 = 0,96; SEP = 2,21%, RPD = 4,79) dan 19 komponen faktor (R^2 = 0,96, SEP = 0,25%, RPD = 4,24) menggunakan spectra pre-treatment turunan pertama Savitzky-Golay.

INTRODUCTION

Palm oil is one of the most important commodities in the world for vegetable oil production *(Murphy et al., 2021*). Oil and free fatty acid (FFA) content are the important criteria for quality of oil palm fruitlets (OPF) or crude palm oil (CPO). According to quality standard for CPO in Indonesian Standard of SNI 01-2901-2006, the maximum allowable FFA content in CPO is 5% (*BSN, 2006*). When the FFA content exceeds this limit, it can cause rancidity and damage to the oil, reducing its shelf life (*Akbar et al., 2022*).

Presently, determination of oil and free fatty acid content of OPF is carried out destructively in the chemical laboratory that require a long time and complicated procedures. Several non-destructive technologies for assessing the oil and free fatty acid of OPF have been explored, including those based on electrical properties (*Sinambela et al., 2020; Mellyana et al., 2024*), image analysis (*Makky, 2016*), and near-infrared spectroscopy (NIRS) (*Makky and Soni, 2014; Iqbal, 2015; Novianty et al., 2020, Novianty et al., 2022*). The advantages of implementing NIR technology in agriculture include rapid response times, competitive costs, and being non-destructive (*Blanco and Villarroya, 2002*). *Makky and Soni, (2014),* predicted the FFA content of OPF using artificial neural network (ANN) calibration methods with a VIS-NIR series spectrometer.

---

[1] *I Wayan BUDIASTRA, Assoc.Prof. Dr.Agr.Eng; Sutrisno MARJAN, Prof.Dr.Agr.Eng; Nissa ADIARIFIA, MS Agr.Eng; Inna NOVIANTY, Dr.Agric. Eng; Yunisa Tri SUCI, PhD.Stud.Agr. Eng*

Iqbal et al., (2018), developed a predictive model for the FFA content of OPF using absorbance in the wavelengths of 700-2500 nm from NIRFlex N-500 spectrometer and Partial Least Square Regression (PLSR) calibration, but the result was not satisfied. The better results of NIR method in predicting oil content of OPF were obtained using absorbance in the wavelengths of 1000-1500 nm from NIRFlex N-500 spectrometer and using Partial Least Square (PLS) calibration (*Novianty et al., 2020*) and empirical mode decomposition calibration method *(Novianty et al., 2022*).

Recently, a hybrid method combining Partial Least Square-Artificial Neural Network (PLS-ANN) is studied as an alternative prediction model in many fields, including near infrared spectroscopy. *Idrus and Kim, (2019)*, successfully applied PLS-ANN in predicting blood haemoglobin using near-infrared spectrum. PLS-ANN showed a better quality compared to PLS in terms of coefficient of determination and mean square error value. A similar result is also shown by *Yu et al., (2018),* in the prediction of consumer liking scores of ready-to-drink green tea beverages and *Song et al., (2014),* for inversion of inland water chlorophyll-a using PLS-ANN. Therefore, this research aims to assess hybrid calibration method and NIR spectroscopy to predict oil and FFA content in OPF non-destructively.

MATERIALS AND METHODS

Sample Preparation and NIR Measurement

About 396 samples of OPF (*Elaeis guineensis* Jacq. Var. tenera) were harvested from Cikabayan Farm, IPB University. The samples consisted of 10 groups of maturity age (number months from full bloom) namely 3 months, 4 months, 4 months 1 week, 4 months 2 weeks, 4 months 3 weeks, 5 months, 5 months 1 week, 5 months 2 weeks, 5 months 3 weeks, and 6 months. The identification of sample age was established by considering the point at which OPF undergo a significant change in oil content, which typically occurs within 4 to 6 months *(Novianty et al., 2020, Akbar et al., 2022*). NIR spectrum data acquisition was carried out using the NIRFlex N-500 spectrometer (BUCHI Labortechnic AG, Switzerland) equipped with fibre optic solid by shooting the gun of the fibre optic solid to the sample in the wavelength range of 1000-1500 nm with a 0.4 nm interval, a scan speed of 3 scans/s and the ambient temperature around 22 – 25 °C. Before NIR measurement, the spectrometer was calibrated by using a standard reference material with known characteristics. The selection of 1000-1500 nm wavelength was based on the previous research results (*Novianty et al., 2120, Novianty et al., 2022*) that showed more accurate prediction in determining chemical content of OPF. The light energy reflected by the OPF samples was captured by the detector as reflectance data or spectrum data, which was then recorded by the instrument.

Chemical Content Measurement

Oil content of OPF samples was determined using Soxhlet extraction method (*AOAC, 2000*) as described in *Novianty et al., (2020)*. The determination of free fatty acid content in the OPF is using titration method according to AOAC 940.28 procedures (*AOAC, 2003*; *Irfan et al., 2023*).

NIR Spectra Pre-treatments

Reflectance spectra was transformed to absorbance and some NIR spectra pre-treatments were carried out in order to reduce noise, baseline shift, and enhance peak or variation of spectra namely the first derivative Savitzky-Golay pre-treatment (SG1), normalization (N01), and standard normal variate (SNV). The spectra pre-treatments improved the accuracy of prediction by NIR spectroscopy (*Budiastra et al., 2020*).

The Development of Calibration Model

A hybrid calibration method that combined multivariate analysis and artificial neural network (PLS-ANN, PCA-ANN) and a single multivariate calibration of PCR and PLSR were applied to develop calibration between spectra data and chemical content of OPF. PCR is suitable for large datasets (NIR spectra) (*Perera et al., 2021*) and produces principal components (PC) as representations of new uncorrelated variables resulting from the transformation *(Costantini et al., 2024*). Meanwhile, PLS extracts predictors for the analysis of dependent and independent variables *(Liu et al., 2022*) and generates factor components (FC) as representations of new data related to the original data (NIR spectra). ANN is a tool for modelling non-linear statistical data where complex relationships between various inputs (NIR spectra) and outputs (chemical data) are modelled for pattern recognition (*Hamdani et al., 2023*).

PCR and PLSR calibration were carried out to obtain some numbers of PC and FC and then correlated with chemical data using software the Unscrambler X. Meanwhile, hybrid models are a combination of multivariate analysis of PCA or PLS and Artificial Neural Network model. In hybrid model (PCA-ANN and PLS-ANN), some numbers of PC extracted from PCA and FC from PLS were input to ANN model (Figure 1).

Hidden layer that was used in this study consisted of 1 (one) layer with 2 neuron per layer. The output layer represented the oil or FFA content that was predicted by the hybrid model. Hybrid model utilized a feed-forward neural network trained by a back propagation algorithm (multi-layer perceptron) using software Rapid Miner.

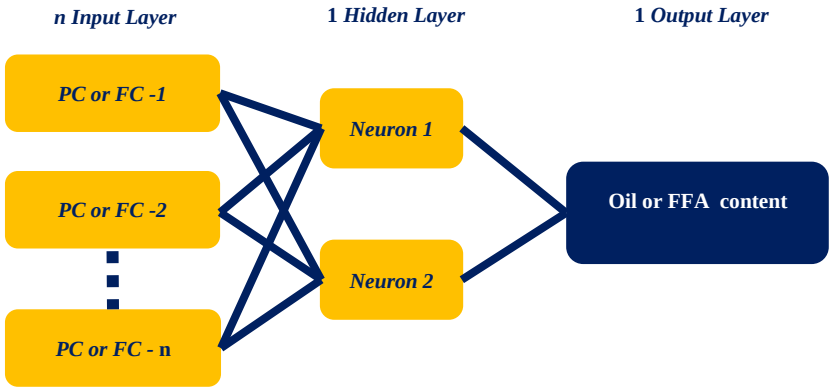

Fig. 1 - Hybrid calibration model (PCA-ANN, PLS-ANN)

Model Evaluation

The statistical parameters used to evaluate calibration and prediction models include the coefficient of determination (R^2), standard error of calibration (SEC), standard error of prediction (SEP), coefficient of variation (CV), residual predictive deviation (RPD), and consistency. A good model's validity can be assessed based on the following statistical parameters, such as R^2 value ≈ 1, SEP ≈ SEC≈0, RPD ≥ 2, consistency values within the range of 80-110%, and CV < 10% (*Munawar et al., 2019*; *Palou et al., 2023*).

RESULTS

NIR spectra of oil palm fruitlets

Fig. 2(A) illustrates the original absorbance spectra obtained from NIRS measurements on oil palm fruitlets. There are some variations in absorbance spectra of ten maturities of OPF, indicating variation in its chemical content. The spectra peak at wavelengths of 1200 and 1450 nm indicates the water absorption. The water content of OPF is high so it is identified visually from spectra. The oil and FFA content absorption peaks (have C-O and O-H bonds) could not be identified, but it was estimated in the wavelength ranges of 1200-1300 nm and 1408-1470 nm (*Iqbal, 2015; Lengkey et al., 2013; Sudarno et al., 2017*). So, it is necessary to find out the methods to identify the absorption peak of oil and FAA in the NIR region.

One of the most powerful of spectra pre-treatment that could show more absorption peaks than original spectra is the first derivative Savitzky-Golay pre-treatment (SG1). The results of the spectrum refinement process with the first derivative Savitzky-Golay pre-treatment (SG1) can be seen in Fig. 2(B). The oil palm fruitlets spectra after correction with SG1 form a more slender pattern with clearer peaks and valleys compared to the original spectra, especially for FFA content. This is because the first derivative functions as a separator for chemical components that experience overlapping (*Bou-Orm et al., 2020*).

The result of another spectra pre-treatment, namely normalization (N01), can be observed in Fig. 2(C). Normalization pre-treatment forms a pattern that is similar to the original spectra but with a narrower range between spectra values. Normalization pre-treatment serves to reduce the range of reflectance values into the range of 0-1, thereby minimizing the effect of sample particle size differences (*Lengkey et al., 2013*). However, normalization could not exhibit especially oil and FFA absorption peaks, similar to original spectra.

Meanwhile, Fig. 2(D) represents the result of standard normal variate (SNV) pre-treatment. It can be observed from Fig. 2(D), SNV pre-treatment reduces the interference of noise waves, resulting in a smoother and denser spectrum (*Nurhasanah et al., 2019*). Similar to normalization spectra pre-treatment, the SNV also could not show the absorption peaks of oil and FFA.

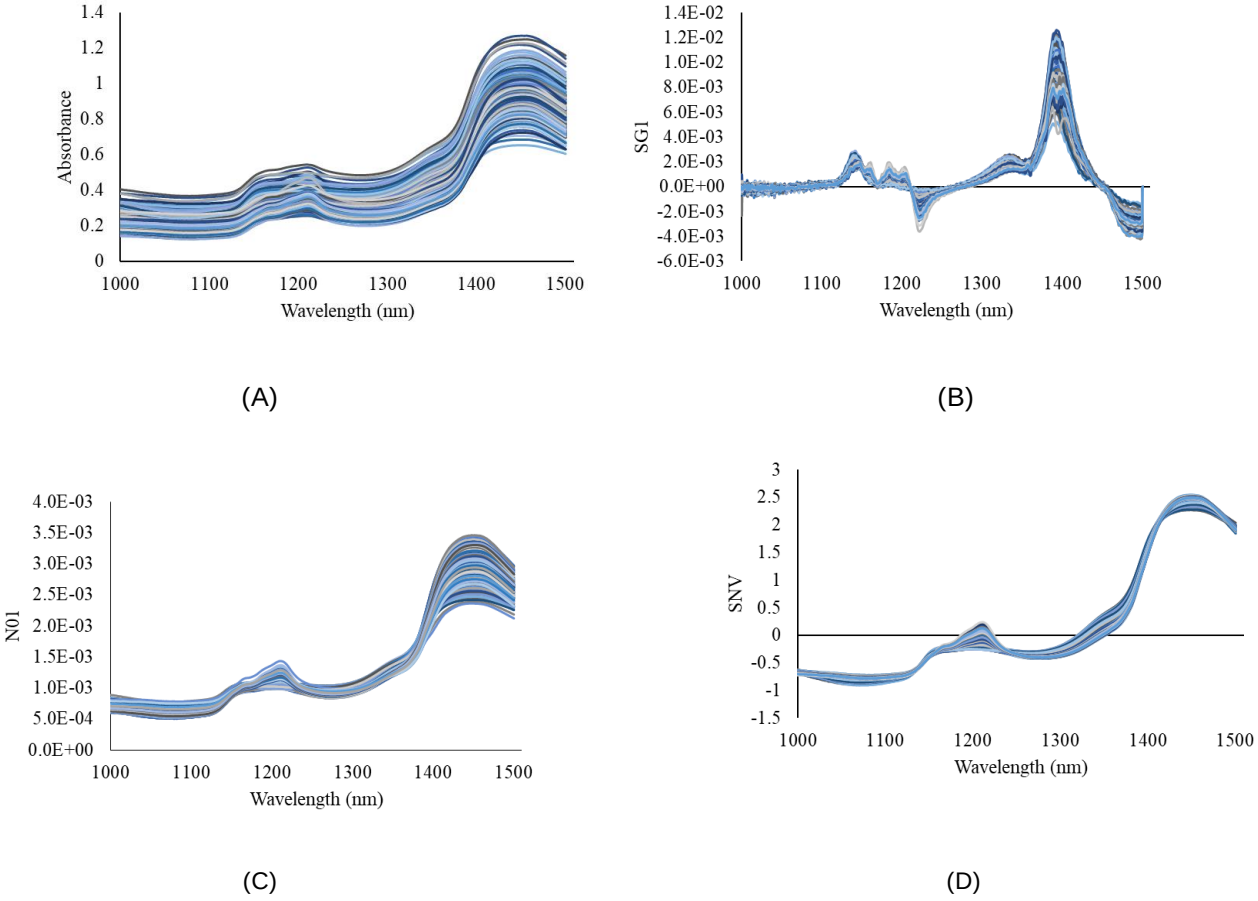

Fig. 2 - NIR Spectra of OPF :(A) -*Original* ;(B) -SG1 ;(C) - N01 ; (D) -SNV

Oil and Free Fatty Acid (FFA) Content in Oil palm fruitlets

The oil content and FFA of samples of 10 maturity age are shown in Table 1. The oil content of OPF ranges from 0.64% to 45.65%. As FFA content ranged from 1.08% to 5.51%. The oil and FFA content of OPF increased from unripe OPF (3 months age) to ripe OPF (6 months age). The oil content of OPF is significantly higher than FFA content. The variation of oil and FFA content of the samples is good enough to develop a good calibration, indicated by their standard deviation.

Table 1

The oil and FFA content of oil palm fruitlets

Chemical content	Calibration/Training [n-277]				Validation/Testing [n=119]			
	Mean	Standard Deviation	Minimum	Maximum	Mean	Standard Deviation	Minimum	Maximum
Oil [%]	21.79	9.60	0.64	45.65	21.64	10.58	0.67	45.56
FFA [%]	3.28	1.11	1.08	5.51	3.41	1.07	1.09	5.25

Calibration and Prediction Results for Oil Content

Table 2 shows the results of calibration and validation using four kinds of calibration methods at different spectra pre-treatments. PLSR calibration results are better than PCR ones. These results are as predicted since PLSR include the chemical content in the transformation of the spectra into the factor components which is used in calibration with chemical content of OPF. Meanwhile, PCR used spectra data only in transformation of the spectra to the principle components without involving chemical content of OPF.

Table 2

The calibration and validation results for predicting oil content in oil palm fruitlets

Pre-Treatment	Method	Input	Calibration [n=277]		Validation [n=119]			Consistency [%]
			R^2	SEC[%]	SEP [%]	CV [%]	RPD	
Original	PCR	22 PC	0.82	4.10	4.52	20.89	2.34	90.79
		23 PC	0.83	3.99	4.24	19.61	2.49	93.93
		24 PC	0.82	4.06	4.25	19.63	2.49	95.61
	PLSR	12 FC	0.84	3.83	3.82	20.41	2.77	86.67
		14 FC	0.88	3.27	3.80	17.57	2.78	85.95
		15 FC	0.90	3.01	3.82	17.66	2.77	78.74
	PCA-ANN	22 PC	0.88	3.30	4.23	19.55	2.50	78.03
		23 PC	0.90	3.05	3.75	17.32	2.82	81.22
		24 PC	0.91	2.89	3.91	18.07	2.71	73.87
	PLS-ANN	20 FC	0.94	2.38	2.66	12.28	3.98	89.60
		21 FC	0.94	2.32	2.66	12.18	3.98	88.05
		22 FC	0.94	2.36	2.59	11.99	4.08	91.14
SG1	PCR	23 PC	0.82	4.11	4.66	21.54	2.27	88.25
		24 PC	0.82	4.08	4.62	21.36	2.29	88.30
		25 PC	0.82	4.07	4.62	21.36	2.29	88.10
	PLSR	7 FC	0.87	3.52	4.21	19.45	2.51	83.72
		8 FC	0.89	3.24	4.01	18.51	2.64	80.97
		9 FC	0.90	3.00	3.85	17.79	2.75	78.02
	PCA-ANN	15 PC	0.87	3.43	4.05	18.70	2.61	84.68
		16 PC	0.88	3.32	4.01	18.53	2.64	82.87
		17 PC	0.87	3.49	4.37	20.18	2.42	79.92
	PLS-ANN	16 FC	0.94	2.28	2.54	11.73	4.17	89.80
		24 FC	0.96	1.88	2.22	10.24	4.77	84.92
		25 FC	**0.96**	**1.86**	**2.21**	**10.21**	**4.79**	**84.33**
N01	PCR	19 PC	0.82	4.06	4.41	20.37	2.40	92.01
		20 PC	0.83	3.99	4.37	20.17	2.42	91.49
		25 PC	0.83	3.92	4.36	20.13	2.43	89.98
	PLSR	12 FC	0.87	3.47	4.17	19.29	2.53	83.08
		13 FC	0.88	3.47	3.93	18.18	2.69	83.98
		14 FC	0.90	3.08	3.90	18.02	2.71	78.90
	PCA-ANN	19 PC	0.89	3.24	4.44	20.52	2.38	72.95
		20 PC	0.86	3.55	4.27	19.71	2.48	83.15
		21 PC	0.89	3.20	4.28	19.78	2.47	74.71
	PLS-ANN	22 FC	0.94	2.35	2.66	12.27	3.98	88.50
		23 FC	0.95	2.35	2.35	10.87	4.50	92.19
		25 FC	0.95	2.09	2.42	11.20	4.37	86.21
SNV	PCR	23 PC	0.82	4.12	4.43	20.45	2.39	93.07
		24 PC	0.82	4.06	4.43	20.45	2.39	91.84
		25 PC	0.82	4.04	4.39	20.27	2.41	92.19
	PLSR	11 FC	0.85	3.76	4.55	21.04	2.32	82.47
		14 FC	0.89	3.18	3.96	18.31	2.67	80.23
		15 FC	0.90	3.09	4.01	18.52	2.64	77.08
	PCA-ANN	20 PC	0.88	3.29	4.27	19.72	2.48	77.03
		23 PC	0.89	3.22	4.13	19.06	2.56	78.04
		24 PC	0.90	3.04	4.23	19.56	2.50	71.76
	PLS-ANN	24 FC	0.94	2.37	2.53	11.68	4.19	93.80
		25 FC	0.95	2.10	2.55	11.80	4.14	82.35
		26 FC	0.95	2.14	2.49	11.51	4.25	85.99

Notes: Blue Block = Best model for hybrid calibration method with and without pre-treatments ; Yellow Block = The Best Model

This result agrees with *Budiastra et al., 2020* that compared PLS and PCR calibration method in predicting trigonelline and chlorogenic acid of intact coffee bean and *Kurniasari et al., 2018* in predicting soluble solid of persimmon. Similarly, PLS-ANN calibration method is much better than PCA-ANN in any pre-treatments. This result confirms that PLS calibration method is always better than PCR. PLS-ANN calibration method is better than PLSR and PCA-ANN calibration method is better than PCR. It means that hybrid model gives the best performance in predicting oil content of OPF. This result agrees with the results of previous research which examined hybrid methods (PLS-ANN) in predicting blood haemoglobin *(Idrus and Kim, 2019*) and inversion of inland water chlorophyll-a (*Song et al., 2014*). The hybrid model is better than single PLSR or PCR since it could compensate the non-linear relationships between NIR spectra and chemical content of OPF.

Commonly, spectra pre-treatment improves calibration and validation performance compared to the original spectra (without pre-treatment). However, in this study, this only happened with the PLS-ANN method. Spectra pre-treatments combined with the PLS-ANN method provided higher calibration and validation performance compared to the original spectra as indicated by lower errors (SEC, SEP), and higher RPD values (Table 2). This also shown that the PLS-ANN method requires a larger number of components (FC) to develop better calibration and validation results.

The best model to predict oil content of OPF is PLS-ANN with 25 Factor Component (FC) input using SG1 pre-treatment (R^2 = 0.96; SEC= 1.86%, SEP= 2.21%, CV= 10.21%, RPD = 4.79). This result agrees with *Lengkey et al., (2013),* that also found that the best pre-treatment for oil prediction of oil content of Jatropha Curcas L was using first derivative Savitzky-Golay. The first derivative of Savitzky-Golay pre-treatment was effective in improving accuracy (*Bou-Orm et al., 2020*) because the percentage of moisture content in OPF in this study was high enough in unripe OPF. This result is slightly better than *Novianty et al., (2022),* that used EMD-ANN calibration method. The result is much better than *Novianty et al., 2020* that used similar instrument and PLS calibration in predicting oil content of OPF. But the result is not better than *Makky and Soni, (2014),* (R^2=0.98) and *Sudarno et al., (2017),* (R^2=0.95) that used different NIR instruments and wavelength region. However, the PLS-ANN is a promising hybrid method to improve accuracy of NIR spectroscopy in determining oil content of OPF.

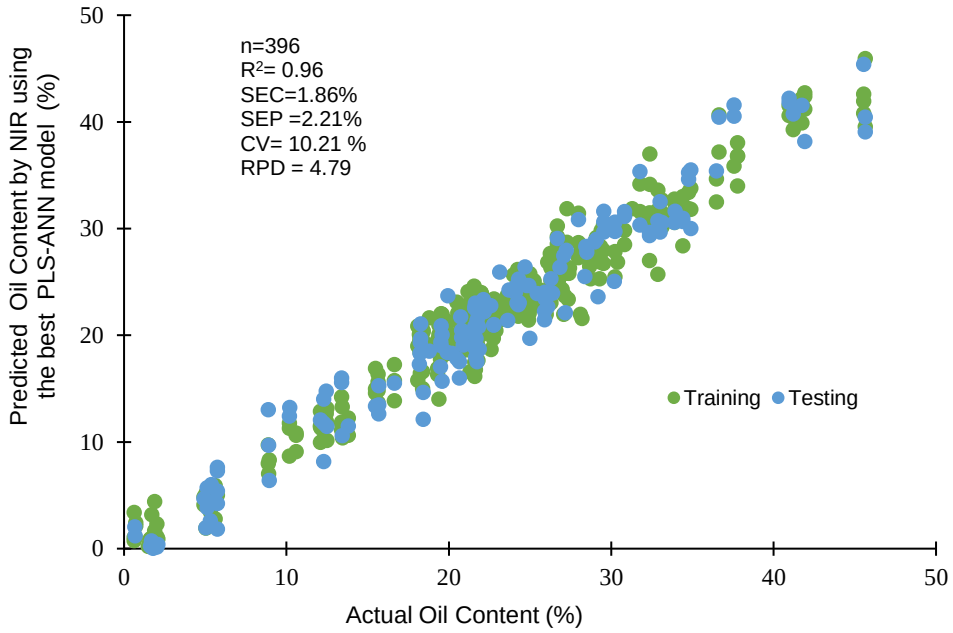

Fig. 3 - Plot of predicted oil content by NIR using the best PLS-ANN model vs actual oil content

Fig. 3 shows the plot of predicted oil content by NIR using the best PLS-ANN model and the chemically tested one of oil content. The high coefficient of determination (R^2) value of 0.96, a low SEC value of 1.86%, a low SEP value of 2.21%, a low CV value of 10.21%, and high RPD value of 4.79 indicate that this model is excellent for predicting oil content in OPF non-destructively.

Calibration and Prediction Results for FFA Content

The calibration results of the model using the original absorbance spectra could not produce satisfactory models for predicting FFA content, except for PLS (17 FC) and hybrid model (Table 3).

Table 3

The calibration and validation results for predicting FFA content in oil palm fruitlets

Pre-Treatment	Method	Input	Calibration [n = 277]		Validation [n = 119]			Consistency [%]
			R^2	SEC [%]	SEP [%]	CV [%]	RPD	
Original	PCR	10PC	0.54	0.75	0.74	21.74	1.45	101.82
		13PC	0.60	0.71	0.67	19.53	1.62	106.18
		20PC	0.77	0.53	0.61	17.79	1.77	87.92
	PLSR	10FC	0.77	0.54	0.61	17.82	1.77	88.60
		12FC	0.83	0.46	0.61	16.97	1.77	80.19
		15FC	0.90	0.36	0.51	14.90	2.12	70.47
	PCA-ANN	10PC	0.56	0.78	0.78	22.93	1.38	99.35
		13PC	0.62	0.70	0.71	20.98	1.50	97.94
		17PC	0.83	0.46	0.53	15.58	2.03	86.98
	PLS-ANN	10FC	0.77	0.54	0.53	15.52	2.03	102.09
		13FC	0.85	0.43	0.53	14.03	2.03	89.32
		15FC	0.88	0.38	0.43	12.69	2.49	88.63
SG1	PCR	10PC	0.75	0.56	0.59	17.23	1.83	95.20
		13PC	0.78	0.52	0.54	16.00	1.97	95.36
		25PC	0.81	0.48	0.52	15.29	2.06	93.07
	PLSR	6FC	0.83	0.46	0.54	15.81	2.00	84.99
		16FC	0.97	0.19	0.62	18.22	1.73	30.56
		19FC	0.98	0.15	0.64	18.90	1.67	23.56
	PCA-ANN	10PC	0.58	0.73	0.80	23.51	1.34	91.52
		13PC	0.73	0.59	0.62	18.26	1.73	95.00
		24PC	0.87	0.40	0.48	14.00	2.25	83.45
	PLS-ANN	12FC	0.93	0.31	0.32	9.52	3.32	94.44
		16FC	0.95	0.26	0.28	8.29	3.81	90.58
		19FC	**0.96**	**0.23**	**0.25**	**7.45**	**4.24**	**90.36**
N01	PCR	10PC	0.53	0.77	0.75	21,96	1.44	102.55
		13PC	0.68	0.63	0.66	19.24	1.64	96.47
		25PC	0.80	0.49	0.52	15.33	2.06	94.36
	PLSR	13FC	0.87	0.41	0.51	14.91	2.12	80.69
		25FC	0.97	0.51	0.65	19.14	1.65	31.01
		29FC	0.98	0.16	0.68	20.08	1.57	23.68
	PCA-ANN	10PC	0.45	0.83	0.84	24.61	1.28	98.55
		13PC	0.71	0.63	0.64	18.76	1.68	99.06
		15PC	0.81	0.48	0.57	16.87	1.87	84.34
	PLS-ANN	22FC	0.93	0.29	0.34	9.94	3.17	84.56
		25FC	0.95	0.34	0.30	8.67	3.64	88.05
		29FC	0.96	0.22	0.27	7.93	3.98	82.25
SNV	PCR	10PC	0.57	0.73	0.77	22.58	1.40	94.76
		13PC	0.67	0.65	0.69	20.18	1.56	93.98
		25PC	0.81	0.49	0.53	15.63	2.02	92.00
	PLSR	13FC	0.86	0.41	0.51	14.94	2.11	80.77
		24FC	0.96	0.22	0.61	18.05	1.75	35.36
		26FC	0.97	0.19	0.63	18.40	1.72	30.55
	PCA-ANN	10PC	0.44	0.83	0.89	26.05	1.21	93.97
		13PC	0.68	0.66	0.70	20.45	1.54	95.39
		17PC	0.84	0.44	0.54	15.79	2.00	82.39
	PLS-ANN	23FC	0.94	0.28	0.32	9.44	3.34	87.67
		24FC	0.93	0.29	0.32	9.54	3.31	90.30
		26FC	0.94	0.27	0.30	8.89	3.55	87.57

Notes: Blue Block = Best model for hybrid calibration method with and without pre-treatments; Yellow Block = The Best Model

Similar to oil content, the PLSR is more accurate than PCR. The PLS-ANN is more accurate than PCA-ANN in any pre-treatment spectra. Spectra pre-treatments also increase calibration and prediction performance in FFA content. It is caused by the ability of spectra pre-treatment to reduce noise, overlapping absorption peaks (SG1) and enhance spectra differences (N01 or SNV) in relatively small amount of FFA content of oil palm fruitlets.

Table 3 also shows that the first derivative Savitzky-Golay (SG1) pre-treatment is the best pre-treatment for all calibration models. This result agrees with *Lengkey et al., 2013* that also found that the best pre-treatment for prediction of FFA content of Jatropha Curcas L was using first derivative Savitzky-Golay.

The PCR needs more numbers of PC in order to make a good calibration and prediction in any spectra pre-treatment. Conversely, The PLS needs less numbers of FC to develop a good calibration and prediction.

The best model for predicting FFA content is the PLS-ANN model with a pre-treatment first derivative Savitzky-Golay (SG1) of 19 FC (Table 3). This result agrees with *Lengkey et al., (2013)*, that also found that the first derivative of Savitzky-Golay pre-treatment was the best pre-treatment for predicting FFA content of Jatropha Curcas using PLS. The result is better than *Iqbal et al., (2018),* (R^2=0.24 and RPD=1.27) that used the similar instrument, which might be due to the ability of PLS-ANN to compensate a non-linearities relationship between NIR spectra and FAA content of OPF. The accuracy is lower than *Makky and Soni (2014)* due to the different instrument and wavelength region used (400-1000 nm).

Fig.4 illustrates the plot of predicted FFA content by NIR using the best model (PLS-ANN using 19 FC with SG1) and the chemically tested one of FFA content. The high coefficient of determination (R^2) value of 0.96, a low SEC value of 0.23%, a low SEP value of 0.25%, a low CV value of 7.45%, and high RPD value of 4.24 indicate that this model is excellent in predicting FFA content of OPF.

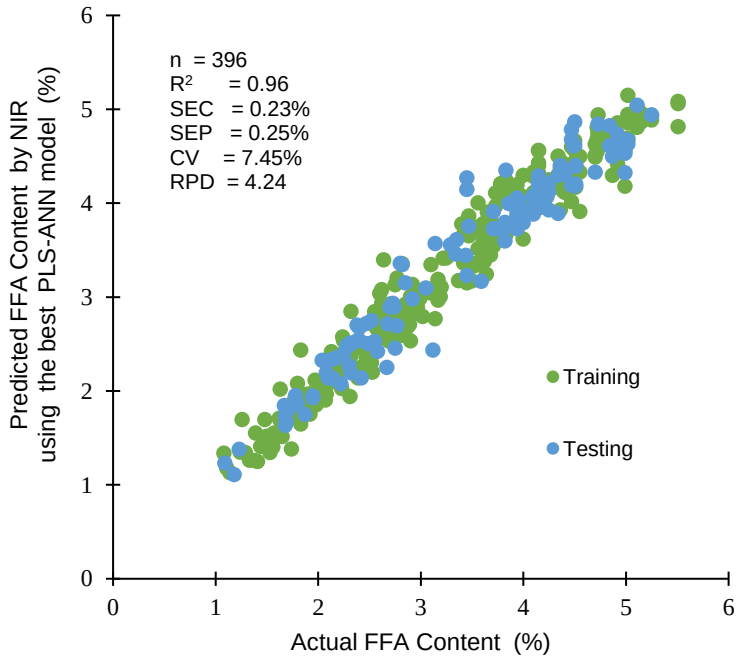

Fig. 4 - Plot of predicted FFA content by NIR using the best PLS-ANN model vs actual FFA content

CONCLUSIONS

The hybrid method of PLS-ANN increased accuracy of NIR spectroscopy in predicting oil and FFA content of OPF non-destructively. The best model to predict oil content of OPF was PLS-ANN with 25 Factor Component (FC) input and using Savitzky-Golay first derivative pre-treatment (R^2 = 0.96 ; SEC=1.86%, SEP=2.21%, CV=10.21%, RPD = 4.79). For FFA content, the best model was also PLS-ANN with 19 FC input using Savitzky-Golay first derivative pre-treatment (R^2 = 0.96; SEC=0.23%, SEP=0.25%, CV=7.45%, RPD = 4.24). The PLS-ANN model developed can be used as a rapid and non-destructive method for determining oil and FFA content in oil palm fruitlets.

ACKNOWLEDGEMENTS

We would like to thank the Directorate of Research, Technology and Community Services, Ministry of Education, Culture, Research and Technology, for cooperation and support through the research grant in Fiscal Year 2023 (contract number: 18901/IT3.D10/PT.01.02/M/T/2023).

REFERENCES

[1] Akbar, A.R., Legowo M.A.C., and Rustiani K. (2022). Determination of Harvesting Time Based on the Variation of Fruitlet Abortion Days in Smallholder Oil Palm Plantations: A Case Study in Tanah Laut Regency, South Kalimantan. *Agritech*, vol. 24, no. 2, pp.167-173, Indonesia;

[2] AOAC. 2000. *Official Method of Analysis of The Association of Official Analytical of Chemist*. 17th edition. Gaithersburg, MD, USA: The Association of Official Analytical Chemist, Inc;

[3] AOAC. 2003. *AOAC 940.28 procedures Fatty Acids (Free) in Crude and Refined Oils. Titration Method*. The Association of Official Analytical Chemist, Inc;

[4] Blanco, M., I. Villarroya. (2002). NIR spectroscopy: a rapid-response analytical tool. *TrAC*, vol. 21, no. 4, pp.240-250;

[5] [BSN] National Standardization Agency. 2006. *SNI 01-2901-2006*. Jakarta: BSN;

[6] Budiastra I. W., Sutrisno, S. Widyoutomo, and P.C. Ayu. (2020). Determination of trigonelline and chlorogenic acid (CGA) concentration in intact coffee beans by NIR spectroscopy. *Agricultural Engineering International: CIGR Journal,* vol. 22, no. 1, pp.163-168;

[7] Bou-Orm, N., AlRomaithi, A. A., Elrmeithi, M., Ali, F. M., Nazzal, Y., Howari, F. M., & Al Aydaroos, F. (2020). Advantages of first-derivative reflectance spectroscopy in the VNIR-SWIR for the quantification of olivine and hematite. *Planetary and Space Science*, *188*;

[8] Costantini, E., Lang, K. M., Sijtsma, K., & Reeskens, T. (2024). Solving the many-variables problem in MICE with principal component regression. *Behavior Research Methods*, vol.56, no.3, pp.1715–1737;

[9] Hamdani, A., N. A. Ganai, and J. Bashir. (2023). Artificial neural networks for data mining in animal sciences. *Bulletin of the National Research Centre*, vol. 47, no.68, pp.1-8;

[10] Idrus, M.N.E and S.C. Kim. (2019). Artificial neural network and partial least square in predicting blood hemoglobin using near-infrared spectrum. *Indonesian Journal of Electrical Engineering and Computer Science*, vol.16, no. 2, pp. 701-708;

[11] Iqbal, Z. (2015). *Prediction of Oil Palm Fresh Fruit Bunch (FFB) Chemical Content Using NIR Spectroscopy*. M.S. thesis. Agricultural and Biosystem Engineering Department. IPB University, Indonesia.

[12] Iqbal, Z., S. Herodian, S. Widodo. (2018). Non-Destructive Evaluation Of Free Fatty Acid (FFA) Of Oil Palm Fresh Fruit Bunch (FFB) Using NIR Spectroscopy. *Jurnal Teknik Pertanian Lampung,* vol. 7, no.2, pp. 80-87, Indonesia;

[13] Irfan, I., Aji, I. F., & Yunita, D. (2023). Application of Near Infrared Spectroscopy (NIRS) in the Measurement of Oil Loss and Its Accuracy Compared To Soxhlet Method. *Jurnal Agroindustri*, vol. 13, no. 2, pp. 107–120, Indonesia;

[14] Kurniasari I, Y.A. Purwanto, I.W Budiastra, Ridwani S. (2017). Prediction of Tannin and NonSoluble Solid Content of Persimmon (Diospyros kaki L.) Using NIR Spectroscopy. *Jurnal Teknik Pertanian,* vol. 5 , no. 3, pp. 245-252;

[15] Lengkey, L.C.E., Sari, D.P., Budiastra, I.W., Seminar, K.B., Purwoko, B.S. (2013). Prediction of chemical contents of Jatropha curcas L. Seeds IP-3P by near infrared reflectance method. *Acta Horticulturae* , 1011, pp.335–342;

[16] Liu, C., X. Zhang, T. T. Nguyen, J. Liu, T. Wu, E. Lee, and X. M. Tu. (2022). Partial least squares regression and principal component analysis: similarity and differences between two popular variable reduction approaches. *Gen Psychiatr*, 35, pp.1-5;

[17] Makky, M, and P. Soni. (2014). In situ quality assessment of intact oil palm fresh fruit bunches using rapid portable non-contact and non-destructive approach. *Journal of Food Engineering,* 120, pp.248-259;

[18] Makky, M. (2016). A Portable Low-cost Non-destructive Ripeness Inspection for Oil Palm FFB. *Agriculture and Agricultural Science Procedia,* 9, pp. 230-240;

[19] Mellyana V., I W. Budiastra, Irmansyah, Y.A. Purwanto. (2024). Electrical Properties for Non-destructive Determination of Free Fatty Acid and Moisture Content in Oil Palm Fruit. *International Journal on Advanced Science Engineering Information Technology*, vol.14, no.2, pp. 641-649.

[20] Munawar, A. A., Kusumiyati, Hafidh, Hayati, R., & Wahyuni, D. (2019). The application of near infrared technology as a rapid and non-destructive method to determine vitamin C content of intact mango fruit. *INMATEH - Agricultural Engineering*, vol. 58, no. 2, pp.1–12;

[21] Murphy, D. J., Goggin, K., & Paterson, R. R. M. (2021). Oil palm in the 2020s and beyond: challenges and solutions. *CABI Agriculture and Bioscience*, vol. 2, no.1, pp. 1–22;

[22] Novianty, I. K. B. Seminar, Irzaman, I. W. Budiastra. (2020). Improving the accuracy of near-infrared (NIR) spectroscopy method to predict the oil content of oil palm fresh fruits. *IOP Conf. Series: Earth and Environmental Science,* 460, pp.1-8;

[23] Novianty, I., R. G. Baskoro, M. I. Nurulhaq, and M.A. Nanda. (2022). Empirical mode decomposition of near-infrared spectroscopy signals for predicting oil content in palm fruits. *Information Processing in Agriculture*, vol. 10, no. 3, pp.289-300;

[24] Nurhasanah, Siregar K, Zulfahrizal. (2019). Prediction of Rice Moisture Content Using NIRS with PLS Method and Pre-Treatment Standard Normal Variate, Derivative I, Savitzky Golay Smoothing. *Jurnal Ilmu Mahasiswa Pertanian,* vol. 4, no. 1, pp. 628-637, Indonesia;

[25] Palou A., P. Jimenez, J. Casals, I. Masalo. (2023). Evaluation of the Near Infrared Spectroscopy (NIRS) to predict chemical composition in *Ulva ohnoi*. *Journal of Applied Phycology* 35, pp. 2007-2015;

[26] Perera, K. D. C., G. K. Weragoda, R. Haputhanthri, S. K. Rodrigo. (2021). Study of concentration dependent curcumin interaction with serum biomolecules using ATR-FTIR spectroscopy combined with Principal Component Analysis (PCA) and Partial Least Square Regression (PLS-R). *Vibrational Spectroscopy*, 116, pp. 1-9;

[27] Sinambela, R., T. Mandang, I. D. M. Subrata, W. Hermawan. (2020). Application of an inductive sensor system for identifying ripeness and forecasting harvest time of oil palm. *Scientia Horticulturae,* 265, pp. 1-6;

[28] Song, K, Li L., Li S., Tedesco, L., Duan H., Li Z., Shi K., Du J., Zhao Y., Shao T. (2014). Using Partial Least Squares-Artificial Neural Network for Inversion of Inland Water Chlorophyll-a. *IEEE Transactions on Geoscience and Remote Sensing* vol. 52, no. 2, pp.1502-1517;

[29] Sudarno, Divo, D.S., Tauvik, R., Baiq, L. W., Fabrice, D., Yong, Y.Y., and Jean, P.C. (2017). Rapid determination of oil content in dried-ground oil palm mesocarp and kernel using near infrared spectroscopy. *JNIRS,* vol. 25, no. 5, pp. 338–347;

[30] Yu Peigen, Mei Yin Low, Weibiao Zhou. (2018). Development of a partial least squares-artificial neural network (PLS-ANN) hybrid model for the prediction of consumer liking scores of ready-to-drink green tea beverages. *Food Research International* 103, pp. 68-75.