Vol. 14, No. 1 (2025): 202 - 214 http://dx.doi.org/10.23960/jtep-l.v14i1.202-2014

JURNAL TEKNIK PERTANIAN LAMPUNG

ISSN 2302-559X (print) / 2549-0818 (online)
Journal homepage : https://jurnal.fp.unila.ac.id/index.php/JTP

Chemical Content Evaluation of Peaberry Robusta Green Bean Using FT-NIRS Method

Bagus Setyawan[1,✉], Yuvita Lira Vesti Arista[2], Arfiati Ulfa Utami[1]

[1] Department of Agricultural Product Technology, Faculty of Agriculture, PGRI University of Banyuwangi, East Java, INDONESIA.

[2] Department of Food Technology, Faculty of Industrial Engineering and Technology, Kalimantan Institute of Technology, INDONESIA.

Article History:

Received : 17 October 2024
Revised : 21 November 2024
Accepted : 27 November 2024

Keywords:

Absorbance,
Caffeine,
Lipid,
Moisture Content,
Pre-Treatment.

Corresponding Author:
✉ bagusetyawan89@gmail.com
(Bagus Setyawan)

ABSTRACT

Mount Ijen is a prominent region for peaberry Robusta coffee beans, which has entered international markets. Accurate real-time estimation of its chemical components is crucial for export activities. This study evaluated moisture content, lipid, and caffeine in Robusta peaberry coffee beans from Ijen using FT-NIRS (Fourier Transform – Near Infrared Spectroscopy). A total of 50 samples were scanned in triplicate, generating 150 spectral data points. The data were optimized for wavelength selection and pre-treated using Standard Normal Variate Transformation (SNV), Second Derivative (dg2), Multiplicative Scatter Correction (MSC), and normalization. Results showed that FT-NIRS proved effective for rapid and accurate estimation of these components. The best calibration model used Kubelka-Munk transformation with dg2 pre-treatment in the 1000-2500 nm wavelength range. Optimal Partial Least Squares (PLS) factors were PLS 4 for lipid content (R^2 = 0.98, SEP = 0.013%, SEC = 0.012%, CV = 0.81, RPD = 2.03, consistency = 95.21%), PLS 5 for moisture content (R^2 = 0.94, SEP = 0.014%, SEC = 0.014%, CV = 0.80, RPD = 4.88, consistency = 101.02%), and PLS 5 for caffeine content (R^2 = 0.94, SEP = 0.014%, SEC = 0.014%, CV = 0.80, RPD = 4.88, consistency = 101.02%).

1. INTRODUCTION

The slope of Mount Ijen is one of the major coffee-producing areas in Banyuwangi Regency, known for their significant productivity. The most widely cultivated coffee variety in this region is peaberry Robusta, commonly known as peaberry coffee. Peaberry coffee is produced from single-seeded coffee beans that have a whole, round shape (Wulandari & Agustin, 2022). Siregar *et al*. (2020) stated that coffee is a commodity that dominates the Indonesian export market, and currently Indonesia ranks fourth as the largest coffee exporter in the world after Brazil, Vietnam and Columbia. Data from the Ministry of Agriculture in 2020, coffee exports of Indonesia amounted to 375,671 ton (Widaningsih, 2022). Based on data from the Central Statistics Agency of East Java, Banyuwangi Regency is ranked second largest for coffee production with a total of 12,504 ton in 2022.

Quality testing for food products is often done conventionally (destructively) where it damages the test sample, takes a long time and produces chemical residues (Zhu *et al*., 2021a). Arista *et al*. (2022), stated that there is an alternative in testing the quality of food products to be exported, namely by using non-destructive methods to shorten the time and reduce residues, one of which is by using FT-NIRS (Fourier Transform-Near Infrared Spectroscopy). The method offers several advantages, including rapid analysis, non-destructive testing, chemical-free procedures, zero waste production, suitability for various sample conditions, and the ability to analyze multiple parameters (Zhang *et al*, 2014). The working principle of NIRS is to absorb molecules in the wavelength range of 780-2500 nm (Kljusurić *et al*., 2019) where the NIRS spectra are created due to absorption of hydrogen bonds in the material including OH, CH, SH, NH combined with basic vibrations in the infrared area (Abbas *et al*., 2020; Shan *et al*., 2014).

Near Infrared Spectroscopy (NIRS) is a fast, non-destructive detection method that has high accuracy in estimating the chemical content of food ingredients (Gao *et al*., 2021). NIR has long been used in the industrial sector to predict the quality content of raw materials or finished products and the authenticity (variety) of agricultural products (Fernández Pierna *et al*., 2018). NIRS combined with chemometrics has been proven to be used in measuring the chemical content of coffee including caffeine, lipids, carbohydrates, phenolics, caffeine, trigonelline, theobromine and theophylline in coffee beans in real time (Ayu *et al*., 2024). The chemical content of coffee beans not only affects the quality but also affects the taste of the coffee produced (Rindang *et al*., 2022; Sahfitri *et al*., 2020). Zhu *et al*. (2021), stated that the lipid, caffeine and water content in coffee beans are important indicators that determine the final quality of the product.

Estimation of chemical content closely related to the quality and taste of green beans, especially water, caffeine and lipid content in real time is necessary to support export activities. Until now the chemical content of peaberry coffee still uses conventional testing based on laboratory and sensory data. It takes a lot of time and money even though the demand for Ijen Robusta peaberry coffee continues to increase every year. Another alternative is needed to test the quality of peaberry coffee with a more effective and efficient method. One of the methods is FT-NIRS, which offers several advantages, including rapid analysis, non-destructive testing, chemical-free procedures, zero waste production, suitability for various sample conditions, and the ability to analyze multiple parameters (Zhang *et al*., 2014). This study aims to estimate the chemical content of Robusta peaberry coffee from Ijen Slope, Banyuwangi, using FT-NIRS, focusing on moisture, lipid, and caffeine content. The goal is to identify the optimal wavelength, the best spectral data transformation, and pretreatment methods to produce a model with high prediction accuracy. This approach is expected to enhance cost and time efficiency in the export activities of Robusta peaberry green beans.

2. MATERIALS AND METHODS

2.1. Tools and Materials

The tool used in non-destructive determination of chemical content is FT-NIRS Flex N-500 (BUCHI Labortechnic AG, Switzerland) in the wavelength range of 1,000-2,500 nm, while the equipment used in destructive testing includes analytical balance (OHAUS), oven (ISUZU), clamp, weighing bottle, desiccator, soxhlet, petri dish (pyrex), fat flask (Iwaki), aluminum cup, condenser, desiccator (DURAN), 100 mesh sieve, LCMS (SHIMADZU), centrifuge.

The materials used in the study consisted of green bean Robusta peaberry coffee obtained from coffee farmers in the Ijen Slope area, Licin District, Banyuwangi Regency, while the materials used in the destructive chemical content analysis consisted of filter paper, 95% neutral ethanol, hexane, cotton, standard caffeine solution , formic acid, acetonitrile, 95% methanol, deionized water.

2.2. Methods

Research on the water, lipid and caffeine content of Ijen Slope Robusta Peaberry Green Beans was carried out destructively and non-destructively (using FT-NIRS). The initial stage carried out was:

2.2.1. Sample Preparations

Robusta peaberry green beans from Mount Ijen slope was first refined using a grinder. Conversion of coffee beans into powder was carried out using the chemometric method to reduce cavities that affecting absorbance. The coffee powder was classified by sieving using 100 mesh or 150 µm sieve (Purningsih *et al*., 2018). The sample was placed on a petri dish to measure absorbance with FT-NIRS. Sample of coffee beans and powder were presented in Figure 1.

2.2.2. Chemical Scanning with FT-NIRS

A total of 50 samples of Ijen Robusta peaberry green beans powder were placed on a petri dish, then a reflectance scan was performed with FT-NIRS (BUCHI Labortechnic AG Switzerland) at a wavelength of 1000-2500 nm. Scanning on the test sample was carried out with 150 samples with a data retrieval speed of 3 scans/s, the operating temperature of the tool in the range of 22-25°C with an electrical power requirement of 100-230 VAC ±10%, 50/60 Hz, 350 watts.

Figure 1. Robusta peaberry coffee from Mount Ijen slope: (a) Green beans, and (b) Ground coffee

2.2.3. Lipid Testing (AOAC, 2012)

Measurement of lipid content was carried out using the Soxhlet method. The initial stage of determining lipid content is to oven the fat flask at a temperature of 105°C for 30 minutes and then cool it in a desiccator. The cooled fat flask is then weighed (X). The next stage is to wrap ± 2 g of sample with filter paper and insert it into a sleeve and then close it with cotton. The next stage is to insert it into the extractor and then fill it with hexane solution. Next, extraction is carried out for a period of 5 hours and then oven at a temperature of 105°C for 20-30 minutes. The sample is then cooled in a desiccator and weighed (Y). Lipid content can be determined using the following equation:

$$\text{L (\% wb)} = \frac{Y-X}{W} \times 100\% \quad (1)$$

$$\text{L (\% db)} = \frac{KL\ (\%bb)}{(100-KA)\%bb} \times 100\% \quad (2)$$

where W is dry weight sample (g), KL is fat content, and KA is water content

2.2.4. Water Content Measurement (AOAC, 2012)

Water content of Ijen Robusta peaberry green beans carried out using the gravimetric method. The initial stage is to dry the weighing bottle at a temperature of 100-105°C for 30 minutes and continued with cooling in a desiccator for 30 minutes and weighing (J). The next stage is to insert 2 g of sample into the weighing bottle and weigh it (K). The sample is then oven-dried for 24 hours at a temperature of 100-105°C and cooled in a desiccator and continued with weighing (L). The water content (KA) of green beans was determined using the following equation:

$$\text{KA (\%)} = \frac{K-L}{K-J} \times 100\% \quad (3)$$

where J is the weight of empty cup, K is the weight of cup + sample before drying, and L is the weight of cup + sample after drying

2.2.5. Caffeine Measurement

Measurement of caffeine content of Ijen Robusta peaberry green beans from Banyuwangi was carried out using LCMS. The initial stage carried out was to prepare a standard caffeine solution of 2 ppm and continued with the introduction into the LCMS column. During the testing process, the C-18 water column in the LCMS was conditioned at a temperature of 40°C, where the LCMS mobile phase would be divided into 2 phases, namely phase A which is a mobile phase of 40% H_2O and 0.1% formic acid, while phase B is a mobile phase of 60% acetonitrile with a flow rate of 0.2 mL/min. Peak/valley caffeine concentration can be obtained at a certain time. The next stage is to weigh 0.2 g of

coffee and continue to enter it into a centrifuge tube. Extraction was carried out by adding 7 mL of 95% methanol and continued with centrifugation at a rotation speed of 20 rpm for 30 minutes. The supernatant obtained from the centrifugation process was separated, and the remaining residue was resuspended in the solvent. Furthermore, 0.25 of the filtrate was diluted with deionized water 5 times from the initial concentration. The sample solution was then diluted again with 15 mL of reagent which aims to remove/reduce the polymer. The sample solution and reagent were then centrifuged at a speed of 12000 rpm for 5 minutes. The supernatant obtained was then filtered and put into a chromatography bottle and continued with LCMS analysis so that the peak/valley of caffeine content would be obtained. The caffeine content of Ijen Robusta peaberry green beans was determined as the following:

$$Kf = \frac{100}{100-H} \tag{4}$$

$$I = Kf \times \frac{100}{1000} \times \frac{100}{Bs} \times Kf \tag{5}$$

2.2.6. FT-NIRS Spectra Data Transformation and Pretreatment

Spectral data generated from scanning as many as 150 scan data of Ijen Robusta peaberry green beans. Likewise with the destructive lipid and water chemistry data used for calibration and validation of 150 sample scan spectral data. Transformation and pretreatment of spectral data were carried out using the unscrumbler X trial version (CAMO) software. The reflectance spectra generated from measurements with FT-NIRS were then transformed into absorbance (Log 1/R) and the Kubelka-Munk equation (K/S) (Arista *et al*., 2022). Reflectance transformation into Log 1/R absorbance can be done using the following equation:

$$\mathrm{Y} = log\,\frac{1}{R} \tag{6}$$

where *Y* is the absorbance, and *R* is reflectance. Meanwhile, the transformation of the spectra into absorbance using the Kubelka-Munk equation (K/S) is as follows:

$$\frac{K}{S} = \frac{\{1-R\}^2}{2R} \tag{7}$$

where *S* is the light scattering coefficient, and *K* is the light absorption coefficient

The data from the spectral transformation are then grouped into 2 types, namely 2/3 of the data is separated for calibration and 1/3 of the data is separated for validation. The data used for calibration and validation is pre-treated with spectral data first (Kumar *et al*., 2016). This aims to reduce the scattering of solid sample radiation measured based on the reflectance of the spectral line which can affect the phenomenon (Arista *et al*., 2022). The data pretreatment used in estimating the lipid content and water content of Ijen Robusta peaberry green beans is Standard Normal Variable Transformation (SNV), Second Derivative (dg2), Multiplicative Scatter Correction (MSC), and normalization.

2.2.7. Spectral Data Evaluation (Calibration and Validation)

The lipid, caffeine and water content of Ijen Robusta peaberry green beans can be seen from the value of R^2 (coefficient of determination), SEC (standard error of calibration), SEP (standard error of validation set), RPD (ratio of performance to deviation) and CV (coefficient of variation) (Mouazen *et al*., 2005). SEC describes the level of accuracy of the resulting calibration equation; a low value indicates minimal error in predicting calibration data. On the other hand, SEP measures the level of model inaccuracy during the validation process, with a smaller SEP value indicating better model accuracy. Consistency of SEC and SEP values in the range of 80-110% indicates that the model is not overfitting. CV indicates the magnitude of the error relative to the average chemical test results of the material. Meanwhile, RPD measures the ratio between the standard deviation of the reference data and the SEC and SEP values. RPD values in the range of 1.5-1.9 indicate that the prediction is still rough and requires improvement in calibration. RPD values between 2.0 -2.5 indicate an adequate calibration model, while values above 2.5 reflect very good prediction model accuracy. Mendes *et al*. (2022),states that a good prediction model has an R^2 value in the range of values above 0.88, SEC (standard error of calibration) and SEP (standard error of validation set) approaching 0, RPD more than 2.5.

3. RESULTS AND DISCUSSION

3.1. Chemical Content of Ijen Robusta Peaberry Coffee

The FT-NIRS spectra data processing requires destructive chemical data used as a reference. Destructive testing for moisture content of Ijen Robusta peaberry coffee was carried out using the oven method and lipid testing refers to the soxhlet method. The chemical content of peaberry green beans is presented in Table 1. The average water content of green lanang coffee meets the quality standards set by BSN in SNI 01-2907 (BSN, 2008) where the general quality requirements for green beans contain a maximum water content of 12.5% and lipid content in the range of 12-14% and caffeine in the range of 2.2-2.5%. While the caffeine content in Ijen Slope Robusta Peaberry Green Beans is in the range of 2.46-3.59%. The standard deviation (SD) value obtained in estimating the lipid content and water content of Ijen Slope Robusta Peaberry Green Beans varies greatly. Standard deviation is a value used to assess the diversity of data, where a higher SD value will affect the accuracy of FT-NIRS predictions.

Table 1 Chemical content of Ijen Robusta Peaberry Green Beans

Chemical components	Level (%)	Average (%)	SD (Standard Deviation) (%)
Water	4.92-6.45	5.97	0.40
Fat	15.8-17.5	16.5	0.51
Caffeine	2.46-3.59	3.06	0.42

3.2. Wavelength Optimization with Absorbance Spectra (Log/R)

Zhou *et al*. (2024), stated that near infrared can be used to determine the chemical content of organic materials because the molecular bonds of organic materials are very sensitive to near infrared waves. Organic materials generally consist of O–H, N–H, C–H, and S–H atoms bound covalently and electrovalently to form molecules. When these molecules are irradiated with an external energy source, the molecules experience changes in potential energy (Budiastra *et al*, 2020). The results of scanning Ijen Robusta peaberry coffee with FT-NIRS are in the form of spectral data in the wavelength range of 1,000-2,500 nm. The average absorbance spectra of Ijen Robusta peaberry coffee has peaks and valleys which are presented in Figure 2. The peak of water absorption is at wavelengths of 1450, 1940 and 2200 nm (Gauglitz & Vo-Dinh, 2006; Foley, 2003; Miller & Miller, 2010), lipids at wavelengths of 1,730 and 2,300 nm (Beauchemin, 2009; Aernouts & Baerdemaeker, 2009) and caffeine at wavelengths of 1,380, 2,100 and 2,150 nm (Miller & Miller, 2010; Kuk & Jolly, 2010). Wang (2019) stated that the optimum absorption of CH, C-O and OH is in the combination band absorption region which is in the wavelength range of 2100-2500 nm.

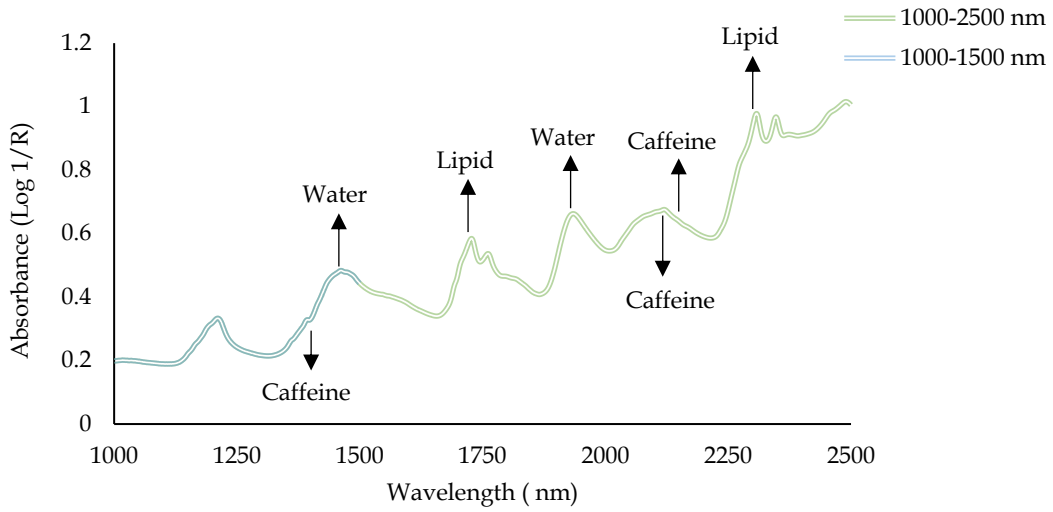

Figure 2. Scanning results of Robusta peaberry green beans from Mount Ijen slope in the range of 1000-2500 nm

The results of scanning lipid, water and caffeine content produced various absorption peaks at wavelengths of 2300 nm (lipids), 1940 nm (water) and 2100 nm (caffeine). Purningsih *et al.* (2018) stated that the higher the concentration of chemical content in a food ingredient, the greater the absorbance value. This is in line with the statement Zhu *et al.* (2021) that the higher the concentration of chemical content in food ingredients, the absorbance value measured also tends to increase. This shows a direct relationship between the concentration of chemical components in the sample and the absorbance value measured in the spectroscopic technique

The scanning results of Ijen Robusta peaberry coffee (Table 2) show that the wavelength range of 1000-2500 nm has more optimum lipid, water and caffeine absorption compared to the wavelength range of 1,000-1,500 nm. This is because the optimum lipid, water and caffeine content in the combination band region is in the wavelength range of 2000-2500 nm as evidenced in Figure 2. The overall model built has relatively good prediction accuracy as evidenced by statistical evaluations covering the values of r, SEC, SEP, CV, RPD and consistency that are in accordance with the provisions. Arista *et al.* (2022) states that a prediction model with good accuracy has high coefficient of determination.

Table 2. Accuracy of prediction of lipid, water and caffeine content of green bean Robusta at two wavelength ranges using PLS.

Long Wave	Parameter	PLS Factor	R^2	SEC (%)	SEP (%)	CV (%)	RPD	Consistency (%)
1000-1500 nm	Water	9	0.78	0.073	0.070	1.31	1.65	79.07
	Lipid	8	0.80	0.064	0.059	1.51	1.79	80.65
	Caffein	10	0.75	0.069	0.063	1.48	1.62	78.23
1000-2500 nm	Water	8	0.82	0.065	0.063	1.21	1.74	80.28
	Lipidk	7	0.85	0.053	0.052	1.45	1.82	82.01
	Caffeine	8	0.80	0.057	0.054	1.15	1.81	82.03

3.3. Characterization of Spectral Data in 2 Types of Transformation and Four Types of Data Pretreatment

Spectral data from green bean scanning Ijen Robusta peaberry has not been fully able to conduct information mining because there is still noise. Peiqiang & Shi (2017) stated that the spectra data obtained from the scanning results not only contain chemical data information of the material but also contain noise and overlapping spectra. Therefore, it is necessary to have spectra data pretreatment. Spectra data pretreatment used in estimating lipid, water and caffeine content in Ijen Slope Robusta Peaberry Green Beans includes using MSC (Multiple Scatter Correction), Normalization, SNV (Standard Normal Variate) and dg2 (second derivative) (Mehari *et al*, 2015). The results of the pretreatment of green bean scanning spectra data Ijen Slope Robusta Peaberry Green Beans is presented in Figure 3.

3.4. Evaluation of NIRS Prediction on Various Types of Spectral Data Transformation

The PLS (Partial Least Square) method is combined with FT-NIRS to build a model with high prediction accuracy. In building a model with PLS, both chemical data and spectral data are centered, so that the data extraction obtained can be more optimal. Chen *et al.* (2024) stated that PLS is the method that is most often combined with infrared spectroscopy in building a predictive model because it can reduce noise, and has strong correlation and covariance. Research conducted by Deng *et al*. (2024), shows that the use of PLS combined with FT-NIRS can be used in real-time adenosine detection in porcini mushrooms/boletes.

The spectra that have been analyzed for the optimum wavelength range, then the type of transformation analysis and also the pretreatment of the spectral data are carried out. Li *et al*. (2023), stated that particle size, distribution and shape of the sample can affect the scattering of the spectra, therefore it is necessary to pretreat the spectral data. This is also in line with the statement of Mishra *et al*. (2021), that it is necessary to make efforts to pretreat the spectral data for scattering which can affect the accuracy of the prediction. In the study, 2 types of spectral data transformations were used, including using the Kubleka Munk (K/S) and absorbance (Log 1/R) equations, and 4 types of data pretreatment were used, including using MSC (Multiple Scatter Correction), Normalization, SNV (Standard Normal Variate) and dg2 (second derivative). The accuracy of the transformation prediction and also the pretreatment of the spectral data are presented in Tables 3, 4, and 5.

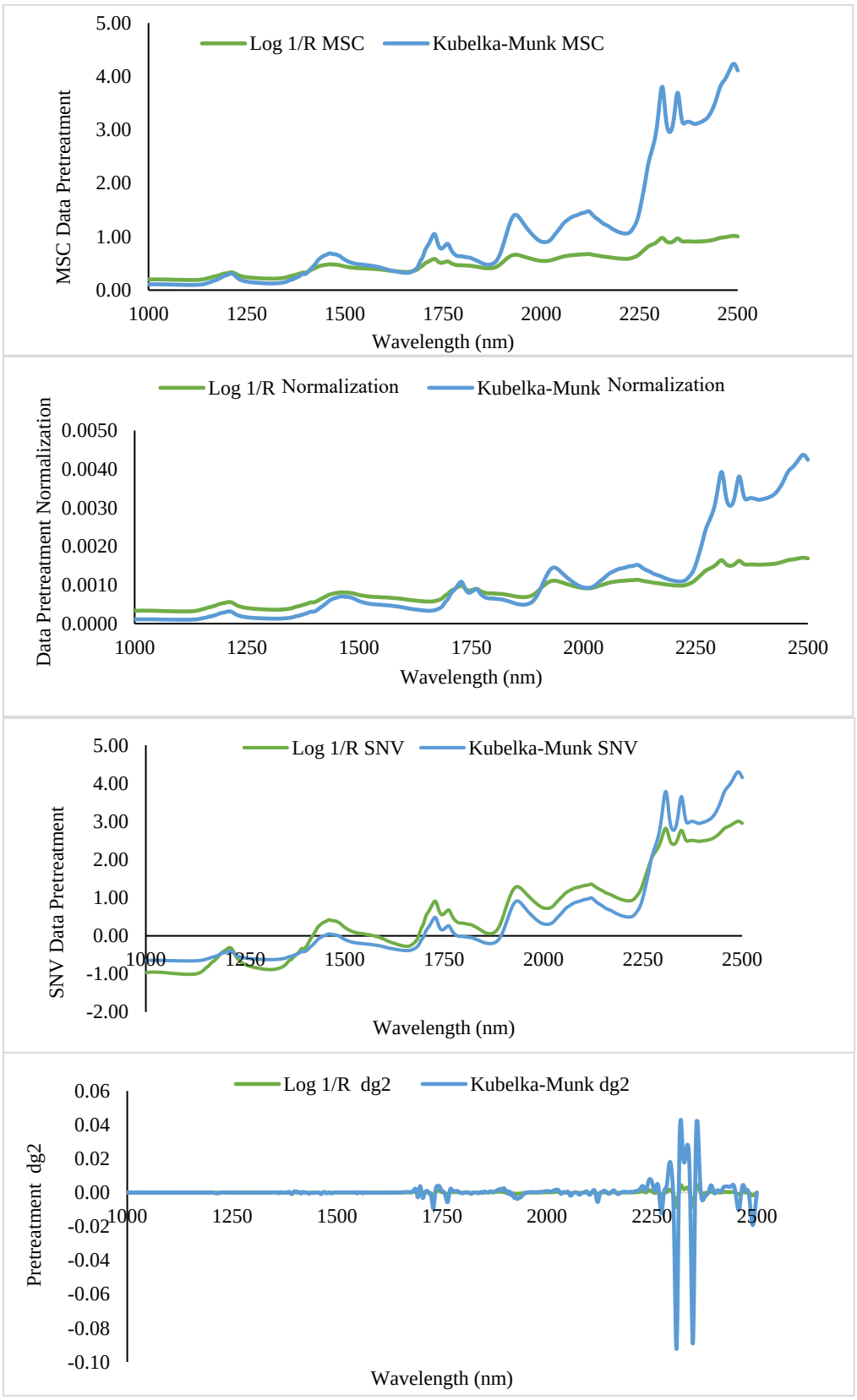

Figure 3. Pretreatment results of spectral data (a) MSC, (b) SNV, (c) Normalization, (d) dg2

Table 3. Calibration and validation results for lipid of Ijen Robusta peaberry coffee with 2 types of data transformation and 4 types of data pretreatment using the PLS method.

Transfor-Still	Pre-Treatment Data	Factor PLS	R^2	SEC (%)	SEP (%)	CV (%)	RPD	Consistency (%)
Absorban	Original	7	0.85	0.053	0.052	1.45	1.82	82.01
	MSC	6	0.89	0.045	0.040	1.14	2.04	91.78
	Normalisasi	6	0.87	0.032	0.030	1.02	1.97	87.04
	SNV	5	0.90	0.021	0.019	0.93	3.83	96.87
	dg2	5	0.91	0.019	0.017	0.82	4.45	99.01
Kubelka-Monk	Original	6	0.88	0.045	0.041	1.11	1.96	88.63
	MSc	6	0.92	0.039	0.035	0.94	2.25	95.54
	Normal thing	5	0.89	0.025	0.020	1.03	1.99	92.68
	SNV	5	0.94	0.016	0.015	0.85	4.87	101.56
	dg2	4	0.98	0.013	0.012	0.81	4.96	103.92

Note: r: koefisien korelasi; SEC: standard error calibration; SEP: standard error prediction; CV: coefficient of variation; RPD: ratio of performance to deviation.

Table 4. Calibration and validation results of water content of Ijen Robusta peaberry green bean with 2 types of data transformation and 4 types of data pretreatment using the PLS method.

Transformation	Pre-Treatment Data	Factor PLS	R^2	SEC (%)	SEP (%)	CV (%)	RPD	Consistency (%)
Absorbance	Original	8	0.82	0.065	0.062	1.21	1.74	80.28
	MSC	8	0.85	0.058	0.055	1.18	1.92	90.52
	Normalization	7	0.82	0.041	0.042	1.13	2.05	86.24
	SNV	7	0.87	0.030	0.029	1.09	3.09	95.02
	dg2	7	0.89	0.024	0.020	0.97	3.76	98.21
Kubelka-Monk	Original	7	0.84	0.051	0.050	1.15	2.05	86.18
	MSc	7	0.87	0.044	0.043	1.03	2.47	93.04
	Normalization	6	0.85	0.026	0.026	1.01	2.07	91.39
	SNV	6	0.91	0.015	0.017	0.95	4.81	98.43
	dg2	5	0.94	0.011	0.014	0.80	4.88	101.02

Lipid estimation results in Ijen Slope Robusta peaberry green bean shows that the transformation of spectral data using the Kubelka-Munk (K/S) equation produces a prediction model with high prediction accuracy compared to using the absorbance transformation (Log 1/R) both without and using pretreatment data. The best prediction model was obtained from the use of the Kubelka-Munk (K/S) transformation with second derivative (dg2) pretreatment data on the PLS factor 4, with an R^2 value = 0.98, SEP = 0.013%, SEC = 0.012%, CV = 0.81, RPD = 4.96 and consistency = 103.92%. The use of the Kubelka-Munk (K/S) spectral data transformation was able to reduce scattering in the sample, this is because the sample used in the study was refined with a particle size of 100 mesh. Similar results were also shown from the research results of Purningsih *et al*. (2018), where the use of Kubelka Munk data transformation and dg2 data pretreatment produced the best prediction model in estimating trigonelline, caffeine and caffeine content.

Mohseni *et al*. (2018) stated that Kubelka-Munk (K/S) is generally used in building prediction models on samples that have small particle sizes with high scattering values. Research conducted Otsuka (2004), by comparing the particle size of phenacetin powder at particle sizes of 41-710 µm using Kubelka-Munk data transformation showed that the smaller particle size of phenacetin powder produced a higher scattering value. Further analysis with data transformation using the Kubelka-Munk equation showed that phenacetin powder with a particle size of 41 µm had a higher prediction accuracy than phenacetin with a particle size of 710 µm. In addition, the transmittance value of phenacetin powder with a particle size of 41 µm was also better than that of 710 µm.

The water content estimation in Ijen Robusta peaberry bean showed that the use of Kubelka-Munk transformation (K/S) with pretreatment second derivative (dg2) data on PLS factor 5, with R^2 value = 0.94, SEC = 0.011%, SEP =

0.014%, CV = 0.80, RPD = 4.88 and consistency = 101.02% is the best predictive model. Prediction models with high accuracy must meet the provisions of statistical evaluation values including R^2 approaching 1, SEP and SEC approaching 0, and consistency in the range of 80-110% (Purningsih *et al*., 2018; Pomeranstev & Rukhin, 2008).

Table 5. Calibration and validation results of Ijen Robusta peaberry green bean with 2 types of data transformation and 4 types of data pretreatment using the PLS method.

Transformation	Pre-Treatment Data	Factor PLS	R^2	SEC (%)	SEP (%)	CV (%)	RPD	Consistency (%)
Absorbance	Original	8	0.80	0.057	0.055	1.15	1.81	82.03
	MSC	8	0.83	0.053	0.050	1.11	1.90	90.31
	Normalisasi	8	0.81	0.044	0.041	1.09	2.02	87.36
	SNV	7	0.84	0.037	0.036	1.04	3.15	96.28
	dg2	6	0.86	0.031	0.029	0.92	3.54	97.01
Kubelka- Monk	Original	7	0.82	0.045	0.043	1.09	1.97	87.02
	MSc	7	0.85	0.032	0.031	1.01	2.04	94.11
	Normalization	6	0.84	0.028	0.027	0.98	1.92	92.06
	SNV	6	0.89	0.023	0.022	0.93	4.17	96.23
	dg2	6	0.91	0.019	0.016	0.84	4.35	99.82

The Kubelka-Munk (K/S) transformation can reduce scattering in peaberry coffee powder. Gowen *et al*. (2007) stated that samples with high scattering, such as powder or powder, require model adjustments to account for the significant impact of scattering. The Kubelka-Munk equation helps reduce the impact of excessive scattering and improves the accuracy of component concentration predictions using reflectance data (Kokhanovski, 2007). Meanwhile, the use of dg2 (second derivative) pretreatment data is able to separate overlapping spectra, especially in materials with very complex chemical content, including materials that contain a lot of water, protein, fat, carbohydrates, etcetera (Cen & He, 2007).

The results of the estimation of caffeine content in Ijen Robusta peaberry green bean showed that the use of Kubelka-Munk transformation (K/S) with second derivative (dg2) pretreatment data on PLS factor 6, with R^2 value = 0.91, SEP = 0.019%, SEC = 0.016%, CV = 0.84, RPD = 4.35 and consistency = 99.82% is the best predictive model. The overall prediction model including water, lipid and caffeine content showed that the use of Kubelka-Munk data transformation (K/S) and the use of dg2 (second derivative) pretreatment data produced a prediction model with high accuracy. Dai *et al*. (2018) states that Kubelka-Munk (K/S) is an equation that connects the absorption coefficient k (depending on concentration) and the scattering coefficient s (depending on particle size), therefore the transformation of spectral data with the Kubelka-Munk (K/S) equation is very suitable for application to samples with particle sizes in the range 90–180µm.

The prediction model of water, fat and caffeine also shows that the use of dg2 pretreatment data produces a predictive model with high accuracy. The dg2 can produce new peaks in materials or samples that have quite complex chemical content (water, fat, protein, caffeine etc.) and are diverse (Chen *et al*., 2014). Gowen *et al*. (2007) stated that dg2 is able to reduce noise and baseline in spectral data, thus clarifying the peaks or valleys of spectral data that affect prediction accuracy.

The selection of PLS factors is also an important factor in determining the accuracy of predictions where the higher the selection of PLS factors makes the prediction model overfitting , which will result in a greater SEP (standard error of validation set) value than the SEC (standard error of calibration) value, thereby reducing the accuracy of predictions (Mitchell & weller, 2012). The best prediction model is presented in the prediction plot In this plot, the values predicted by the model are compared with the actual values measured. This provides an idea of how well the model can predict into the actual model. The prediction plot for estimating the water, fat and caffeine content of Ijen Robusta peaberry green bean is best shown in Figures 4, 5, and 6.

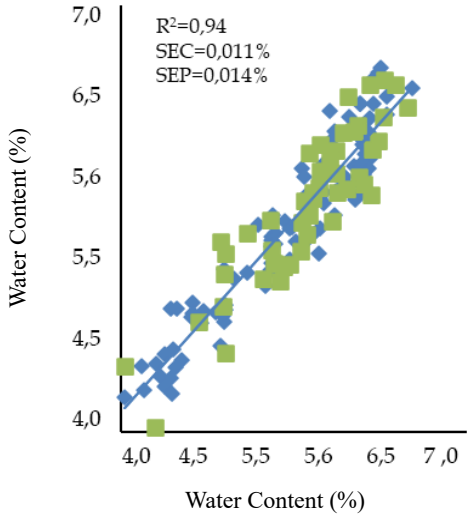

Figure 4. Plot of prediction of the best water content of Ijen Robusta peaberry green bean (Green = calibration; Blue = validation)

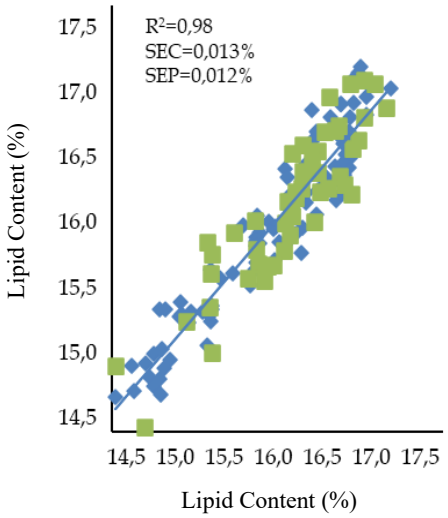

Figure 5. Plot of best predicted lipid levels of Ijen Robusta peaberry green bean (Green = calibration; Blue = validation)

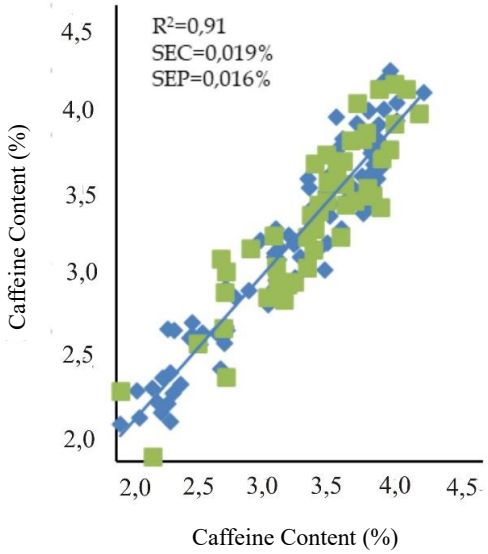

Figure 6. Plot of the best caffeine content prediction for Ijen Robusta peaberry green bean (green = calibration; blue = validation)

4. CONCLUSION

FT-NIRS can be used as a method to estimate the water, fat and caffeine content of Ijen Robusta peaberry green bean quickly and accurately. The best calibration for predicting the chemical content of Ijen Robusta peaberry green bean is using Kubelka-Munk data transformation with dg2 data pretreatment at a wavelength of 1000-2500 nm with a PLS factor of 4 for lipid content (R^2 = 0.98, SEP = 0.013%, SEC = 0.012%, CV = 0.81, RPD = 2.03 and consistency = 95.21%) PLS factor 5 for water content (R^2 = 0.94, SEP = 0.014%, SEC = 0.014%, CV = 0.80, RPD = 4.88 and consistency = 101.02%) and PLS factor 5 for caffeine content (R^2 = 0.94, SEP = 0.014%, SEC = 0.014%, CV = 0.80, RPD = 4.88 and consistency = 101.02%).

ACKNOWLEDGMENTS

Thanks are due to the Directorate of Research, Technology, and Community Service, Directorate General of Higher Education, Research and Technology, Ministry of Education, Culture, Research and Technology for providing research grant funds entitled Estimation of Chemical Content of Green Bean Peaberry Robusta Ijen Banyuwangi Using FT-NIRS as part of the Beginner Lecturer Research scheme for the 2024 fiscal year so that this research can be carried out properly.

REFERENCES

Abbas, O., Pissard, A., & Baeten, V. (2020). Near-infrared, mid-infrared, and Raman spectroscopy. *Chemical Analysis of Food: Techniques and Applications.* 2nd edition: 77–134. https://doi.org/10.1016/B978-0-12-813266-1.00003-6

Aernouts, B., & De Baerdemaeker, J. (2009). Application of near infrared spectroscopy for the characterization of fat content in dairy products. *Journal of Near Infrared Spectroscopy*, ***17***(4), 285-295.

AOAC (Association of Official Analytical Chemistry). (2012). *Official Method of Analysis*. Association of Official Analytical Chemists. 19th edition, Washington DC., 121-130.

Arista, Y.L.V.., Budiastra, I.W., & Hasbullah, R. (2022). Evaluation of chemichal content of desiccated coconut using FT-NIR spectroscopy. *IOP Conference Series: Earth and Environmental Science*, ***1024***(1). https://doi.org/10.1088/1755-1315/1024/1/012020

Ayu, P.C., Rindang, A., & Sebayang, N.U.W. (2024). NIR spectroscopy application to determine the carbohydrate and lipid content in green bean coffee. *IOP Conference Series: Earth and Environmental Science*, ***1302***(1), 012114. http://dx.doi.org/10.1088/1755-1315/1302/1/012114

Beauchemin, K.A., & McKinnon, J.J. (2009). *Infrared spectroscopy for the analysis of animal feed*. CRC Press.

BSN (Badan Standarisasi Nasional). (2008). *SNI 01-2907-2008: Syarat Mutu Biji Kopi*. Badan Standarisasi Nasional, Jakarta.

Budiastra, W., Sutrisno, M., Widyotomo, S., & Ayu, P.C. (2020). Determination of trigonelline and chlorogenic acid (CGA) concentration in intact coffee beans by NIR spectroscopy. *CIGR Journal*, ***22***(1), 162-168.

Cen, H., & He, Y. (2007). Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends in Food Science & Technology*, ***18***, 72-83. https://doi.org/10.1016/j.tifs.2006.09.003

Chen, H., Tan, C., & Lin, Z. (2024). Geographical origin identification of ginseng using near-infrared spectroscopy coupled with subspace-based ensemble classifiers. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, ***304***. https://doi.org/10.1016/j.saa.2023.123315

Dai, S., Pan, X., Ma, L., Huang, X., Du, C., Qiao, Y., & Wu, Z. (2018). Discovery of the linear region of near infrared diffuse reflectance spectra using the Kubelka-Munk theory. *Frontiers in Chemistry*, ***6***(May), Article 154. https://doi.org/10.3389/fchem.2018.00154

Deng, G., Li, J., Liu, H., & Wang, Y. (2024). A fast method for predicting adenosine content in porcini mushrooms using Fourier transform near-infrared spectroscopy combined with regression model. *LWT*, ***201***. https://doi.org/10.1016/j.lwt.2024.116243

Fernández Pierna, J.A., Manley, M., Dardenne, P., Downey, G., & Baeten, V. (2018). Spectroscopic technique: Fourier Transform (FT) Near-Infrared Spectroscopy (NIR) and Microscopy (NIRM). *Modern Techniques for Food Authentication*, 103–138. https://doi.org/10.1016/b978-0-12-814264-6.00004-9

Foley, M. (2003). *Application of near infrared spectroscopy to food quality and safety*. CRC Press.

Gauglitz, G., & Vo-Dinh, T. (2006). *Handbook of spectroscopic data of molecular beverages*. Elsevier.

Gao, B., Xu, X., Han, L., & Liu, X. (2021). A novel near infrared spectroscopy analytical strategy for meat and bone meal species discrimination based on the insight of fraction composition complexity. *Food Chemistry*, ***344***. https://doi.org/10.1016/j.foodchem.2020.128645

Gowen, A.A., & O'Donnell, C.P. (2007). The use of Kubelka-Munk theory in the analysis of high scattering materials. *Food Chemistry*, ***105***(3), 1218-1226. https://doi.org/10.1016/j.foodchem.2007.01.025

Kljusurić, J.G., Jurina, T., Valinger, D., Benkovi, M., & Tušek, A.J. (2019). NIR spectroscopy and management of bioactive components, antioxidant activity, and macronutrients in fruits. *Fruit Crops: Diagnosis and Management of Nutrient Constraints*, 95–109. https://doi.org/10.1016/B978-0-12-818732-6.00008-3

Kokhanovsky, A. (2007). Physical interpretation and accuracy of the Kubelka–Munk theory. *Journal of Physics D: Applied Physics*, ***40***, 2210-2216. https://doi.org/10.1088/0022-3727/40/7/053

Kuk, M.L., & Jolly, C.A. (2010). Quantitative analysis of caffeine using near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, ***18***(3), 181-188.

Kumar, S., Lahlali, R., Liu, X., & Karunakaran, C. (2016). Infrared spectroscopy combined with imaging: A new developing analytical tool in health and plant science. *Applied Spectroscopy Reviews*, ***51***(6), 466-483. https://doi.org/10.1080/05704928.2016.115780

Li, Y., Logan, N., Quinn, B., Hong, Y., Birse, N., Zhu, H., Haughey, S., Elliott, C. T., & Wu, D. (2023). Fingerprinting black tea: When spectroscopy meets machine learning—a novel workflow for geographical origin identification. *Food Chemistry*, ***438***, Article 138029. https://doi.org/10.1016/j.foodchem.2023.138029

Mendes, G. de A., de Oliveira, M.A.L., Rodarte, M.P., de Carvalho dos Anjos, V., & Bell, M.J.V. (2022). Origin geographical classification of green coffee beans (*Coffea arabica* L.) produced in different regions of the Minas Gerais state by FT-MIR and chemometric. *Current Research in Food Science*, ***5***, 298–305. https://doi.org/10.1016/j.crfs.2022.01.017

Mehari, B., Redi-Abshiro, M., Chandravanshi, B.S., Atlabachew, M., Combrinck, S., & McCrindle, R. (2016). Simultaneous determination of alkaloids in green coffee beans from Ethiopia: Chemometric evaluation of geographical origin. *Food Analytical Methods*, ***9***(6), 1627–1637. https://doi.org/10.1007/s12161-015-0340-2

Mishra, P., Rutledge, D.N., Roger, J.M., Wali, K., & Khan, H.A. (2021). Chemometric pre-processing can negatively affect the performance of near-infrared spectroscopy models for fruit quality prediction. *Talanta*, ***229***, Article 122303. https://doi.org/10.1016/j.talanta.2021.122303

Mitchell,.R & Weller,.C. (2022). Evaluation of predictive performance of PLS regression models after being transferred from benchtop to handheld NIR spectrometers. *Biosystems Engineering*, ***218***(3-4), 245-255. https://doi.org/10.1016/j.biosystemseng.2022.04.014

Miller, J.C., & Miller, J.N. (2010). *Statistics and chemometrics for analytical chemistry*. Pearson Education.

Mohseni, M., Bastani, S., Allonas, X., Moradian, S., Ley, C., Croutxé-Barghorn, C., & Mohajerani, E. (2018). Effect of scattering nanoparticles on the curing behavior and conversion gradient of UV-curable turbid systems: two-flux Kubelka-Munk approach. *Progress in Organic Coatings*, ***115***, 65–73. https://doi.org/10.1016/j.porgcoat.2017.11.003

Mouazen, A.M., De Baerdemaeker, J., & Ramon, H. (2005). Towards development of on-line soil moisture content sensor using a fibre-type NIR spectrophotometer. *Soil & Tillage Research*, 80, 171-183. https:// doi.org/10.1016/j.still.2004.03.022

Otsuka, M. (2004). Comparative particle size determination of phenacetin bulk powder by using Kubelka-Munk theory and principal component regression analysis based on near-infrared spectroscopy. *Powder Technology*, ***141***(3), 244–250. https://doi.org/10.1016/j.powtec.2004.01.025

Peiqiang, Y & Shi, H. (2017). Comparison of grating-based near-infrared (NIR) and Fourier transform mid-infrared (ATR-FT/MIR) spectroscopy based on spectral preprocessing and wavelength selection for the determination of crude protein and moisture content in wheat. *Food Control*, ***82***, 57-65. https://doi.org/10.1016/j.foodcont.2017.06.015

Purningsih, V., Budiastra, I.W., Sutrisno, S., & Widyotomo, S. (2018). Accuracy of NIRS method for prediction of chemical content of Bondowoso green coffee powder using Kubelka-Munk model. *Jurnal Keteknikan Pertanian,* ***6***(3), 271–278.

Rindang, A., Ayu, P.C., & Sahfitri. (2022). NIR spectroscopy as an environmentally friendly method to determine the lipid and carbohydrate concentration in Lintong Arabica coffee bean. *IOP Conference Series: Earth and Environmental Science*, ***977***(1). https://doi.org/10.1088/1755-1315/977/1/012075

Sahfitri, Rindang, A., & Ayu, P.C. (2020). Prediction of water content in Lintong green bean coffee using FT-NIRS and PLS method. *IOP Conference Series: Earth and Environmental Science,* ***454***(1), 012047. https://doi.org/10.1088/1755-1315/454/1/012047

Shan, J., Suzuki, T., Suhandy, D., Ogawa, Y., & Kondo, N. (2014). Chlorogenic acid (CGA) determination in roasted coffee beans by Near Infrared (NIR) spectroscopy. *Engineering in Agriculture, Environment and Food*, ***7***(4), 139–142. https://doi.org/10.1016/j.eaef.2014.08.003

Siregar, S.D., Rindang, A., & Ayu, P.C. (2020). Principle Component Analysis (PCA) - Classification of Arabica green bean coffee of North Sumatera Using FT-NIRS. *IOP Conference Series: Earth and Environmental Science*, ***454***(1). https://doi.org/10.1088/1755-1315/454/1/012046

Wang, X. (2019). Near-infrared spectroscopy for food quality evaluation. *Evaluation Technologies for Food Quality*, 105–118. https://doi.org/10.1016/B978-0-12-814217-2.00007-X

Widaningsih, R. (2022). *Buku Outlook Komoditas Perkebunan Kopi*. Pusat Data dan Sistem Informasi Pertanian, Sekretariat Jenderal – Kementerian Pertanian, Jakarta: 83 pp.

Wulandari, S., & Agustin, Y. (2022). Biji kopi robusta peaberry green bean : skrining fitokimia, formulasi herbal lotion. *Journal of Pharmaceutical and Sciences* (*JPS*), ***5***(2), 355-363. https://doi.org/10.36490/journal-jps.com.v5i2.152

Zhang, X-L., Liu, L-F., Zhu, L-Y., Bai, Y-J., Mao, Q., Li, S-L., Chen, S-L., & Xu, H-X. (2014). A high performance liquid chromatography fingerprinting and ultra high performance liquid chromatography coupled with quadrupole time-of-flight mass spectrometry chemical profiling approach to rapidly find characteristic chemical markers for quality evaluation of dispensing granules, a case study on Chuanxiong Rhizoma. *Journal of Pharmaceutical and Biomedical Analysis*, ***88***, 391-400. https://doi.org/10.1016/j.jpba.2013.09.023

Zhang, X., Li, W., Yin, B., Chen, W., Kelly, D. P., Wang, X., Zheng, K., & Du, Y. (2013). Improvement of near-infrared spectroscopic (NIRS) analysis of caffeine in roasted Arabica coffee by variable selection method of stability competitive adaptive reweighted sampling (SCARS). *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, ***114***, 350–356. https://doi.org/10.1016/j.saa.2013.05.053

Zhu, M., Long, Y., Chen, Y., Huang, Y., Tang, L., Gan, B., Yu, Q., & Xie, J. (2021). Fast determination of lipid and protein content in green coffee beans from different origins using NIR spectroscopy and chemometrics. *Journal of Food Composition and Analysis*, ***102***. https://doi.org/10.1016/j.jfca.2021.104055

Zhou, X., Li, L., Zheng, J., Wu, J., Wen, L., Huang, M., Ao, F., Luo, W., Li, M., Wang, H., & Zong, X. (2024). Quantitative analysis of key components in Qingke beer brewing process by multispectral analysis combined with chemometrics. *Food Chemistry*, ***436***, 137739. https://doi.org/10.1016/j.foodchem.2023.137739