JTEP Jurnal Keteknikan Pertanian,
Vol. 13 No. 4, p 513-528, 2025
P-ISSN 2407-0475 E-ISSN 2338-8439

Available online:
http://journal.ipb.ac.id/index.php/jtep
DOI: 10.19028/jtep.013.4.313-528

Portable NIR Spectroscopy for Cocoa Bean Re-fermentation Analysis: A Rapid and Reliable Technique

Sutrisno S Mardjan[1,2], Musimura K Mitaray[1], Samsudin[3*], Yogi Purna Raharjo[4], Jonni Firdaus[5]

[1]Department of Mechanical and Biosystem Engineering, Faculty of Agricultural Engineering and Technology, IPB University, Lingkar Akademik Street, IPB Dramaga Campus, Bogor, West Java 16002, Indonesia.
[2]Center for Research on Engineering Application in Tropical Agriculture, IPB University, Lingkar Akademik Street, IPB Dramaga Campus, Babakan, Bogor, West Java 16680, Indonesia.
[3]Research Center for Estate Crops, BRIN–National Research and Innovation Agency Indonesia, Soekarno Research and Technology Area, Jakarta – Bogor Street 46 KM, Cibinong, West Java, Indonesia, 16911.
[4]Research Center for Agroindustry, BRIN–National Research and Innovation Agency Indonesia, Soekarno Research and Technology Area, Jakarta – Bogor Street 46 KM, Cibinong, West Java, Indonesia, 16911.
[5]BRIN–National Research and Innovation Agency Indonesia, Jakarta, Indonesia, BJ Habibie Building, M. H. Thamrin Street, Kb. Sirih, Menteng, Central Jakarta, Indonesia 10340.
*Corresponding author, email: sams015@brin.go.id

Article Info

Submitted: 4 August 2025
Revised: 29 September 2025
Accepted: 14 November 2025
Available online: 4 Desember 2025
Published: December 2025

Keywords:
Cocoa beans, fermentation, NIR spectrometer, PLSR, ANN, chemometrics

How to cite:
Mardjan S.S., Mitaray, M. K., Samsudin., Raharjo, Y. P., Firdaus, J. (2025). Portable NIR Spectroscopy for Cocoa Bean Re-fermentation Analysis: A Rapid and Reliable Technique. Jurnal Keteknikan Pertanian, 13(4): 513-528. https://doi.org/10.19028/jtep.013.4.513-528.

Abstract

Portable near-infrared (NIR) spectrometers may examine samples directly on-site, speeding up data gathering. However, the NIR spectrometer has a limited wavelength, ranging from 740 to 1,070 nm, whereas previous studies used a longer wavelength. The research aims to determine the fermentation index, pH, and moisture content of fermented dry cocoa beans using a portable Near-Infrared (NIR) spectrometer. The NIR spectra were preprocessed using several methods, including the Savitzky-Golay first and second derivatives (SG1 and SG2), Linear Baseline Correction (LBC), Multiplicative Scatter Correction (MSC), comparative Partial Least Square Regression (PLSR) modeling, and Artificial Neural Network (ANN). To simulate different fermentation levels, unfermented cacao beans underwent pretreatment fermentation for durations of 0, 24, 48, and 72 hours. The prediction outcomes of ANN models, when applied to dried fermented cocoa beans with data preprocessing, offered better results in comparison to PLSR models, with strong correlation, lowest RMSEC, and highest residual predictive value. The most effective method for predicting fermentation index was ANN combined with LBC preprocessing, while optimal pH models were applied using the SG2 method. The effective moisture content models were developed using MSC preprocessing. The analytical approach of portable NIR spectroscopy produced rapid and accurate results to determine the quality of ground-dried cacao bean fermentation.

Doi: https://doi.org/10.19028/jtep.013.4.513-528

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

JTEP Jurnal Keteknikan Pertanian,
Vol. 13 No. 4, p 513-528, 2025
P-ISSN 2407-0475 E-ISSN 2338-8439
Available online:
http://journal.ipb.ac.id/index.php/jtep
DOI: 10.19028/jtep.013.4.313-528

1. Introduction

The objective of cocoa fermentation is to activate biochemical activities in the beans, resulting in the synthesis of aroma precursors, increased flavor and color, reduced bitterness, and improved physical appearance of cocoa (Ministry of Agriculture Republic of Indonesia, 2021). Unfermented or improperly fermented cocoa beans usually have an astringent flavor and bitter taste, making them unsuitable for chocolate production (Sabahannur et al., 2016). Quality parameters such as fermentation index, pH, free fatty acids, polyphenols, and water content directly influence the fermentation process (Dzelagha et al., 2020). Bean color changes occur throughout the fermentation process, induced by a decrease in anthocyanin concentration, and are detected by measuring the fermentation index in cocoa beans (Raharja et al., 2023). The pH of cocoa beans also affects flavor development and fermentation levels (Sunoj et al., 2016). Cocoa beans with a moisture content exceeding 7.5% may reduce the yield of the product and have a risk of bacterial and fungal infestation, including contamination by mycotoxigenic fungi such as aflatoxins.

Cocoa farmers in Indonesia commonly manufacture dried cocoa beans without fermentation because of long fermentation timeframes and limited quantities in recurring harvests. Unfermented cocoa beans are relatively inexpensive, and when processed by the industry, they must be blended with fermented cocoa to produce high-quality chocolate paste. Another alternative for fermenting cocoa beans is to re-ferment dried beans (Priambodo et al., 2022) by rehydrating dry cocoa beans from farmers and re-fermenting them by adding a starter or enzyme (Hernani et al., 2019; Raharja et al., 2023).

Several methods are widely employed to determine the quality parameters of cocoa fermentation based on liquid and solvent extraction, which are often laborious, time-consuming, destructive, and complicated (Wang, 2019). The near-infrared reflectance spectroscopy (NIRS) method has been proven to be a powerful nondestructive method for carrying out rapid and robust analysis (Feng et al., 2022). These spectrometers can detect vibrational patterns in the functional groups of chemical elements in cocoa beans, providing specific data used to build mathematical models (Wang, 2019). NIR spectroscopy, when combined with mathematical models, has shown promising results in food quality assessment. NIR spectra are used to gather wide-ranging information from samples; however, they frequently have overlapping signals and noise. The use of data preprocessing techniques and chemometric methodologies is crucial for effectively extracting relevant information and preventing unwanted interference with the data (Feng et al., 2022). The processing of spectral data can be categorized into two primary classifications: spectral derivatives and scatter-correlation approaches. Spectral derivatives, such as the Savitzky-Golay (SG) polynomial derivative filters and the Norris Williams algorithm, have been employed to eliminate baseline variations and spectral peak overlap, resulting in enhanced spectral resolution and improved signal-to-noise ratio. The scatter-correlation methods frequently used in research include MSC, de-trending (DT), and Standard Normal Variate

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

Available online:
http://journal.ipb.ac.id/index.php/jtep
DOI: 10.19028/jtep.013.4.313-528

(SNV). These methods are used to reduce the effects of additive factors, such as dispersion, particle size, and signal strength variations (Moghaddam et al., 2022).

The use of an NIR spectrometer alone does not provide a direct assessment of the sample composition. To ensure greater precision, it is necessary to first subject the sample to analysis using chromatography and mass spectroscopy techniques. The collected data were subsequently integrated into a calibration and classification model for a specific sample. Calibration models are constructed by establishing a relationship between the chemical composition of a substance and the absorption, reflectance, or transmittance of near-infrared radiation at specific wavelengths (Wang, 2019). Subsequently, the model was tested for robustness and reliability using reference data and predictability measures. A more stable model is indicated by a higher correlation coefficient (R) or coefficient of determination (R2).

NIR applications have been used to identify, classify, characterize, authenticate, and differentiate cocoa beans from different varieties and countries of origin. The quality of cocoa beans is determined by their variety, geographic zone, and differences in fermentation and drying methods. This allows us to characterize and differentiate cocoa beans from various cocoa-growing regions in Ghana using NIR spectroscopy (Basile et al., 2022). The MC+ SVM model is the best tool for the optimal classification of different geographic origins (El Orche et al., 2021). The SVM model can accurately distinguish between fermented and unfermented cocoa beans (Teye et al., 2016). The quality of cocoa beans has also been quantitatively measured using NIR spectroscopy in terms of water content, fat content, caffeine, polyphenols, and volatile and non-volatile compounds. The total fat content can be determined using the MSC technique and coupled SI-PLS-SVMR, and the biochemical content, fermentation duration, and pH can be quantified well with NIR using the PCA-PLS model (Osborne, 2006; Teye et al., 2020). The equipment used in this research was a traditional NIR instrument, which has a higher resolution and precession.

Portable near-infrared (NIR) spectrometers offer several advantages over standard laboratory NIR spectrometers. One notable advantage is their ability to assess samples immediately at the location, which accelerates data creation. Additionally, these spectrometers enable noninvasive testing and are cost-effective. However, the wavelength range of the portable NIR spectrophotometer used was only approximately 700-1100 nm, whereas several studies that used NIR as an identification process used the NIR wavelength range of 800-2498 nm (Basile et al., 2022) and 3600-12500 nm (Grgić et al., 2022). The aim of this research was to determine the fermentation index, pH, and water content of dry fermented cocoa beans using a Portable NIR Spectrometer with a wavelength range of 740 nm – 1,070 nm. The results of this study will encourage the use of portable NIR spectroscopy as a quality detector for fermented cocoa.

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

Available online:
http://journal.ipb.ac.id/index.php/jtep
DOI: 10.19028/jtep.013.4.313-528

2. Material and Methods

2.1 Cocoa Beans Sample

Selected dried cocoa beans (forastero cocoa) were obtained from local farmers in Sukabumi Regency, West Java, Indonesia, and re-fermented using the Raharja and Kadow method, which was modified with cellulase enzymes (0.04 g/ml) and acidic liquid from glacial acetic acid solution (150 mmol/L) (Raharja et al., 2023; Kadow et al., 2015). The process involved varying fermentation durations of 0, 24, 48, and 72 h, with approximately 300 g of sample collected in each phase. The beans were first rehydrated in water for eight hours, then soaked in a glacial acetic acid solution in a tightly sealed plastic container, and air was added. The plastic container was then immersed in a water bath filled with warm water at 45 °C for 12 h to facilitate incubation. The fluid was replenished after 12 h and incubated with the enzyme solution the following day. Following enzyme incubation, the cocoa beans were subjected to a temperature treatment for 12 h without adding water. After fermentation, the beans were dried in an oven at 55°C for 12 h to achieve a moisture content below 5%. The beans were ground, packaged in zip-lock bags, labeled, and stored in a dark room before the studies.

2.2 Fermentation Index

The fermentation index (FI) was calculated using a modified method described by Gourieva and Tserevitinov in 1979 (Gourieva & Tserevitinov, 1979). The process involved adding 0.5 g of ground cocoa beans to a dark bottle, adding 50 mL of 97:3 v/v methanol: HCl, agitating for 2 min, refrigerating for 20 h, filtering using a Whatman No. 1 filter, and measuring the absorbance at wavelengths of 460 and 530 nm using UV/Vis spectrophotometry (Biochrom Ultrospec 7500, United Kingdom). The FI is calculated using Equation 1.

$$Fermentation\ index\ (FI) = \frac{Absorbance\ at\ 460\ nm}{Absorbance\ at\ \ 530\ nm} \tag{1}$$

2.3 The pH and Moisture Content

The pH level of ground cocoa beans was assessed using a modified method from the AOAC 1995 publication (AOAC 1995). The process involved incorporating ground cocoa beans into hot water at a temperature of 60 ± 5 °C. The mixture was stirred until homogeneous, the extract was filtered, and the pH was determined using a digital pH meter (Hanna Instrument HI5522) calibrated with buffer solutions at pH values of 4.01, 7.01, and 10.01, respectively. The moisture content of cocoa powder was measured using the SNI 3747:2013 standard (National Standardization Agency of Indonesia 2013).

2.4 NIRS Data Acquisition of Ground Cocoa Beans and Dana Preprocessing

This research used the SCiO™ (Consumer Physics Inc., Tel Aviv, Israel) Portable NIR Spectrometer molecular sensor to measure the NIRS spectrum in the wavelength range of 740 nm to 1,070 nm. Radiation referred to as near-infrared has a range between 780 and 2500 nm but portable NIR has a

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

limited wavelength value between 700 -1100nm or 1100-2500nm. The selected wavelengths have an electromagnetic spectrum close to the visible region and correspond to the overtones and combination bands of the fundamental molecular vibrations of the observed C–H, N–H, O–H, and S–H functional groups. Prior to data collection, the operational principle of the Portable NIR Spectrometer involved the utilization of infrared light by directing it to the sample.

Data preprocessing was performed using The Unscrambler X 10.4 (CAMO, Norway) using methods such as Savitzky-Golay first and second derivatives (SG1 and SG2), LBC, SNV, and MSC. The primary objective was to remove inherent physical phenomena and enhance multivariate regression (PLSR and ANN modeling) and classification models or data exploration analysis (Rinnan et al., 2009). The selection of the optimal spectrum preprocessing method is mostly determined through experimentation and evaluation based on the highest number of validation criteria.

2.5 Partial Least Squares Regression Model

Partial least squares regression (PLSR) is a statistical technique used for multivariate data analysis. It is employed to estimate the chemical content using linear combinations of independent variables (Burns & Ciurczak, 2007). It presupposes a robust connection between the variables and allocates equal weights to the response variables. Determining the ideal number of PLSR elements is essential when constructing an NIR calibration model to avoid overfitting, which occurs when the root mean square error of prediction (RMSEP) is greater than the root mean square error of calibration (RMSEC). Conversely, underfitting is indicated by the RMSEC value being higher than the RMSEP value. The calibration created using Partial Least Squares Regression (PLSR) was observed using The Unscrambler X 10.4 software (CAMO, Norway).

2.6 Artificial Neural Network Architecture

Artificial Neural Networks (ANN) have a strong pattern recognition capacity, enabling them to adapt their problem-solving abilities to complex systems with multiple variables. The ANN model is a parallel distributed architecture with three layers: an input, hidden, and output layer. The success of achieving targets can be influenced by the network architecture because different challenges may require different architectural approaches (Sharabiani et al., 2023). The process of constructing a trained ANN includes generating data, designing the architecture, conducting training and testing procedures, and evaluating outcomes. This process leads to the identification and selection of an ideally trained ANN model (Grgić et al., 2022). The regression study was performed using MATLAB's 'net fitting' tool, and the Levenberg-Marquardt algorithm was used for optimization. The classification model was optimized using a feedforward backpropagation neural network as the training strategy. The ANN architecture comprised three layers: the first covering wavelengths from 740 nm to 1070 nm with a total of 331 nodes, the second using an optimal number of hidden nodes, and the third with a single node referring to the fermentation index, pH, and water content.

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

JTEP Jurnal Keteknikan Pertanian,
Vol. 13 No. 4, p 513-528, 2025
P-ISSN 2407-0475 E-ISSN 2338-8439
Available online:
http://journal.ipb.ac.id/index.php/jtep
DOI: 10.19028/jtep.013.4.313-528

2.7 Validation

The calibration model for predicting the accuracy of the ground cocoa bean content was validated using parameters such as the correlation coefficient (r), root mean square error of calibration (RMSEC), root mean square error of prediction (RMSEP), and residual predictive deviation (RPD). The correlation coefficient measures the model's ability to represent the dependent variables, with a value close to 1 indicating higher accuracy. The correlation is calculated using Equation 2.

$$r = \frac{\sum(Y-\hat{Y})(Y_{NIRS}-\hat{Y}_{NIRS})}{\sqrt{\sum(Y-\hat{Y})^2\sum(Y-\hat{Y}_{NIRS})^2}} \quad (2)$$

Where: Y= reference value of dried cocoa beans content (fermentation index, pH, moisture content), Y_NIRS= NIRS prediction content of fermentation index, pH, and water content, Yˆ= average reference value of dried cocoa beans content (fermentation index, pH, moisture content), Yˆ_NIRS= average NIRS prediction content of fermentation index, pH, and water content.

Root mean square error calibration (RMSEC) and root mean square error prediction (RMSEP) refer to the difference between the estimated near-infrared spectroscopy (NIRS) outcomes and the observed test outcomes. A lower standard error is indicative of a more desirable model outcome. The standard error is calculated using Equations 3 and 4.

$$RMSEC = \sqrt{\frac{\sum(Y_{calibration}-Y_{NIRS})^2}{n}} \quad (3)$$

$$RMSEP = \sqrt{\frac{\sum\left(Y_{prediction}-Y_{NIRS}\right)^2}{n}} \quad (4)$$

Where: RMSEC= root mean square error calibration, RMSEP= root mean square error prediction, Y= reference value of dried cocoa beans (fermentation index, pH, moisture content), Y_NIRS= NIRS prediction content of fermentation index, pH, and water content, N= number of samples.

The ratio of prediction to deviation (RPD) is a measure of a model's ability to predict the chemical components. RPD values between 2.5 and 3 indicate excellent predictive capabilities, whereas values between 2 and 2.5 indicate good performance. Between 1.5 and 2.0, the high and low values can be distinguished, and the model indicates sufficient performance and is deemed suitable for utilization; below 1.5, calibrations are not usable (Hernández-Hernández et al., 2022). The RPD equation is given by Equation 5.

$$RPD = \frac{Standard\ deviation\ of\ reference\ values}{RMSEP} \quad (5)$$

2.8 Statistical Analysis

The collected data on fermentation index, pH, and moisture content were analyzed using one-way analysis of variance (ANOVA). If a significant effect was observed, further analysis was conducted

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

Available online:
http://journal.ipb.ac.id/index.php/jtep
DOI: 10.19028/jtep.013.4.313-528

using Duncan's advanced test at a significance threshold of 5% in the IBM SPSS Statistics 25 software. In the present study, ANOVA and Duncan's post hoc test were employed to examine the impact of fermentation time on the fermentation index, pH, and moisture content.

3. Results and Discussion

3.1 Chemical Analysis of Grournd Cocoa Beans Quality

The reference data for developing a prediction model included the fermentation index, pH, and water content, which were correlated with the NIR spectrum data obtained from tests conducted using the SCiO Portable NIR Spectrometer. The results of the fermentation index, pH, and water content of cocoa beans obtained by chemical analysis are presented in Table 1 and show the influence of fermentation index, pH, and water content based on fermentation time, with a significance level of 5%.

Table 1. Cocoa bean quality parameters in chemical analysis.

Duration of Fermentation (Hour)	Samples	Fermentation Index	pH	Water Content (wb%)
0	11	1.11 ± 0.14^{a}	6.40 ± 0.04^{c}	4.68 ± 0.04^{a}
24	11	1.28 ± 0.06^{b}	5.68 ± 0.06^{b}	4.29 ± 0.06^{a}
48	11	1.37 ± 0.06^{b}	5.31 ± 0.05^{a}	4.63 ± 0.05^{a}
72	11	1.41 ± 0.15^{c}	5.30 ± 0.11^{a}	4.44 ± 0.11^{a}

3.2 NIR Spectrum Characteristics

The study used a SCiO portable NIR spectrometer to analyze the reflectance data from 740 nm to 1,070 nm. The reflectance data were transformed into absorbance data using log (1/R). The transformation of reflectance into absorbance was performed to determine whether the composition of a material has a linear relationship with the absorbance spectrum. The absorbance spectrum data in Figure 1 show peaks within the wavelength range of 920–950 nm.

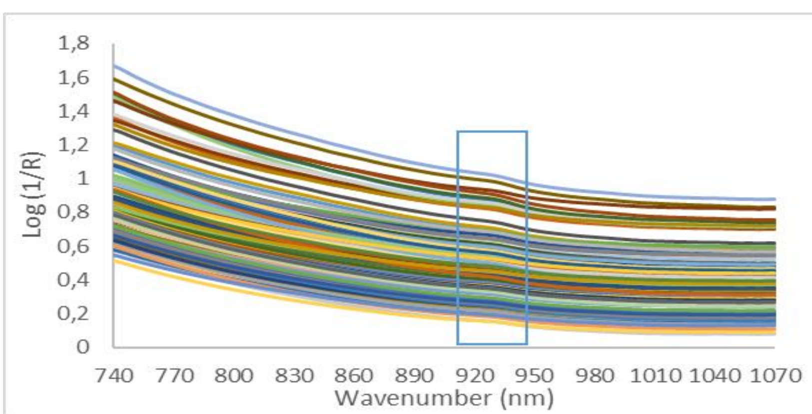


Figure 1. Portable NIR spectra absorbance obtained for the cocoa beans.

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

The presence of peaks and troughs in the spectrum indicates the absorption of atomic bonds, including O-H, N-H, and C-H bonds. Materials with a higher chemical content and thickness absorb a considerable amount of light. The amount of light reflected or absorbed can indicate the quality of the object obtained using mathematical models (Osborne, 2006, Pasquini, 2003, Osborne et al., 1993). The absorbance values of all samples were consistent but varied due to factors such as orientation, surface properties, and biochemical composition. Figure 2 shows the absorbance regression coefficients for (a) the fermentation index, (b) pH, and (c) water content.

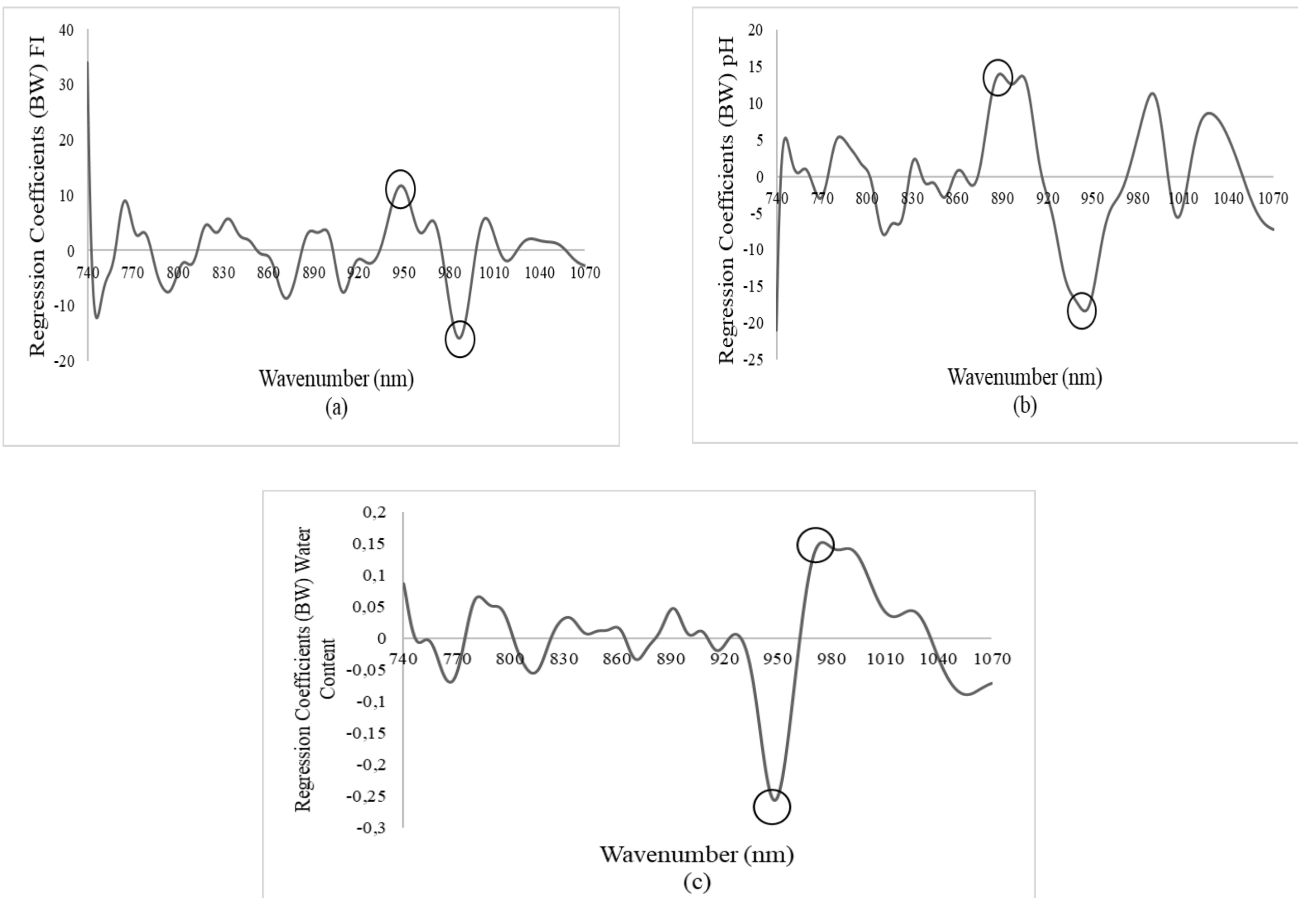


Figure 2. Absorbance regression coefficient (BW) (a) FI, (b) pH, (c) water content

3.3 Calibration Model

3.3.1 Fermentation Index Calibration Models

The results of calibration using PLSR and ANN afforded the best prediction, and data pre-processing using LBC at a PLSR factor of 13 and ANN node at 3 produced a better fermentation index model than the other pre-processing methods as shown in Table 2. Data pre-processing using LBC with PLC produced a calibration correlation coefficient (r) of 0.821 and a validation of 0.769, indicating

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

JTEP Jurnal Keteknikan Pertanian,
Vol. 13 No. 4, p 513-528, 2025
P-ISSN 2407-0475 E-ISSN 2338-8439

Available online:
http://journal.ipb.ac.id/index.php/jtep
DOI: 10.19028/jtep.013.4.313-528

that it is good. The RMSEC value was 0.097, the RMSEP was 0.769, and the RPD value was 1.562. The ANN model significantly improved, offering better predictions than the PLSR models.

Table 2. Calibration and validation performances for the prediction of the Fermentation Index.

Preprocessing method	Factor/ Node	Calibration		Validation		RPD
		r	RMSEC	r	RMSEP	
SG1	PLS (7)	0.787	0.104	0.768	0.100	1.592
SG2	PLS (4)	0.715	0.118	0.752	0.092	1.722
LBC	**PLS (13)**	**0.821**	**0.097**	**0.769**	**0.102**	**1.562**
SNV	PLS (10)	0.803	0.101	0.718	0.101	1.570
MSC	PLS (13)	0.830	0.094	0.713	0.104	1.528
SG1	ANN (4)	0.881	0.084	0.824	0.093	1.576
SG2	ANN (3)	0.896	0.070	0.818	0.086	1.664
LBC	**ANN (3)**	**0.931**	**0.059**	**0.900**	**0.069**	**2.191**
SNV	ANN (3)	0.919	0.068	0.867	0.075	1.951
MSC	ANN (3)	0.845	0.080	0.932	0.080	2.286

Table 2 shows that all data preprocessing methods using ANN were significant, providing improved values for fermentation index parameters using LBC data preprocessing. The model obtained using the ANN method can estimate the fermentation index based on the parameters obtained using NIR spectroscopy. Figure 3 shows a graph of the NIR prediction results with reference. The ANN models show plots that are mostly close to a linear line, but the distribution of sample data for the fermentation index is not evenly distributed.

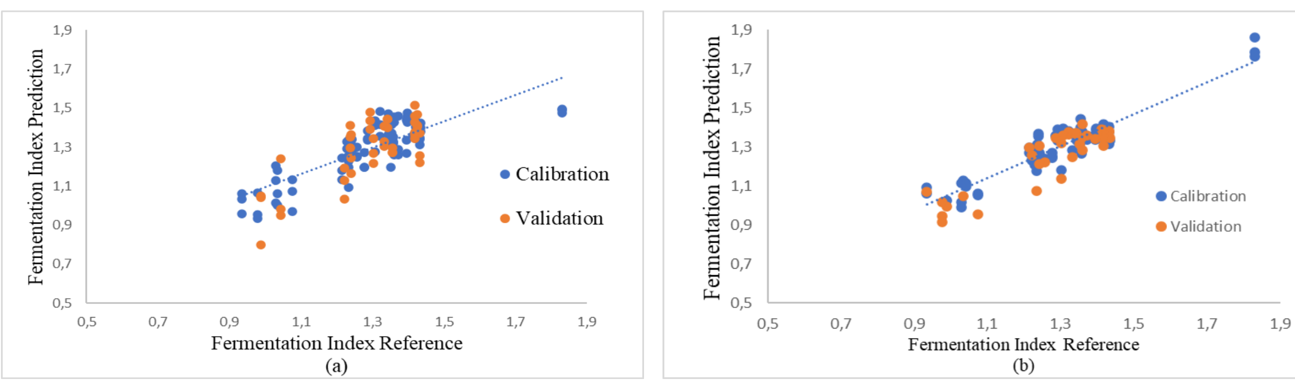


Figure 3. NIR the fermentation index prediction and reference (a) PLSR model, (b) ANN model (NIR serves as the prediction and reference for the fermentation index. (a) PLSR model, (b) ANN model.)

3.3.2 The pH and Water Content Calibration Models

Calibration and validation analyses using data pre-processing using SG1, SG2, LBC, SNV, and MSC were also carried out for the pH and water content parameters. In Table 3, the calibration and validation results of the pH and water content parameters are presented. Figure 4 shows the NIR

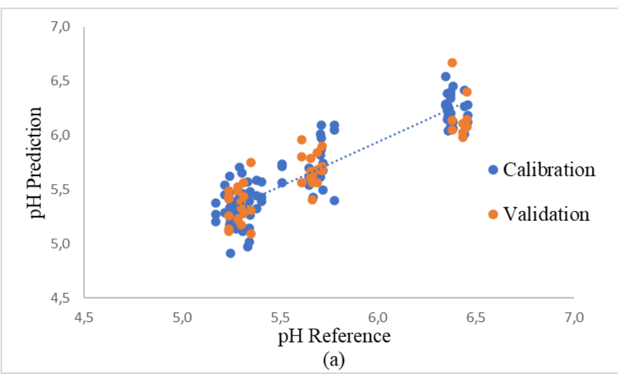
Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

JTEP Jurnal Keteknikan Pertanian,
Vol. 13 No. 4, p 513-528, 2025
P-ISSN 2407-0475 E-ISSN 2338-8439
Available online:
http://journal.ipb.ac.id/index.php/jtep
DOI: 10.19028/jtep.013.4.313-528

graph, the pH prediction and reference, which compared the ANN and PLSR methods. The results of the ANN method were mostly close to a linear line, but the distribution of the sample pH data was not even. In contrast to the pH distribution results, the water content values tended to be more evenly distributed, and the ANN method results were closer to a linear line than the PLSR method, as shown in Figure 5.

Table 3. Calibration and validation performances for the prediction of pH and water content.

Parameters	Preprocessing method	Factor / Node	Calibration		Validation		RPD
			r	RMSEC	R	RMSEP	
pH	SG2	PLS (6)	0.908	0.1932	0.871	0.2124	2.129
	SG2	ANN (4)	0.971	0.1210	0.962	0.1250	3.668
Water Content	MSC	PLS (15)	0.847	0.2930	0.734	0.3530	1.534
	MSC	ANN (3)	0.930	0.2100	0.885	0.2270	2.060



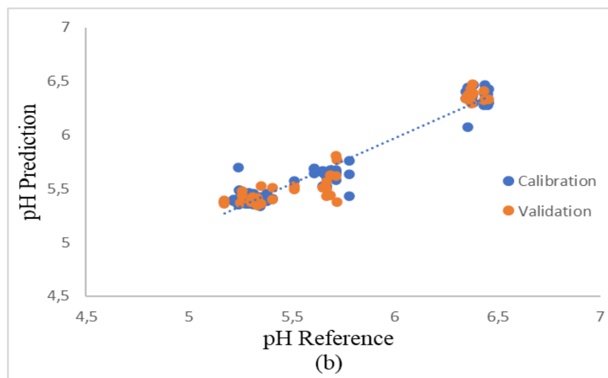
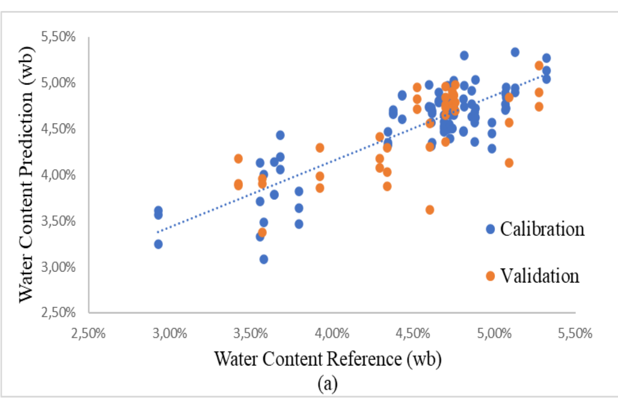


Figure 4. NIR the pH prediction and reference (a) PLSR model, (b) ANN model



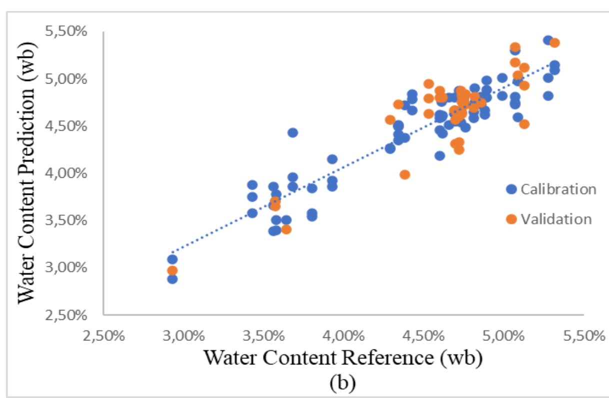


Figure 5. NIR water content prediction and reference (a) PLSR model, (b) ANN model

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

Available online:
http://journal.ipb.ac.id/index.php/jtep
DOI: 10.19028/jtep.013.4.313-528

3.3.3 Discussions

The fermentation index (IF) measures the color change in cocoa bean cotyledons due to a decrease in anthocyanin content during fermentation (Raharja et al., 2023). The methanolic acid extract, which absorbs at 460 and 530 nm, can be used to assess fermentation processes (Balcázar-Zumaeta et al., 2023). The (IF) average value of 1.29 aligns with previous research, indicating that cocoa beans with a fermentation index value exceeding 1 are considered completely fermented, and those below 1 indicate incomplete fermentation. Furthermore, the study found that cocoa beans with a fermentation index value over 1.6 were classified as over-fermented (Romero-Cortes et al., 2023). The results showed a positive correlation between the duration of dried bean fermentation and the fermentation index value. Re-fermented dry cocoa beans have a pH range of 5.2–5.49, indicating that the beans have been well fermented under conditions comparable to the traditional procedure, and a pH range of 5.5–5.7 is achieved if the beans have not been entirely fermented with a low fermentation index. However, in traditional fermentation techniques, cocoa beans with a pH range of 4.75–5.19 are frequently fully fermented (Afoakwa et al., 2008). The moisture content of cocoa powder ranged from 2.92% to 5.32%, with an average of 4.51%. Fermented cocoa beans with a fermentation period of 48-72 hours have a substantial fermentation index and reduced pH level, qualifying them as fermented cocoa beans. NIRS requires reference data in the form of chemical content information obtained through chemical analysis.

The regression coefficient graph (Figure 2) provides evidence of a strong correlation between the absorption spectrum and fermentation index, pH, and moisture content. The wavelength of 948 nm exhibited a strong correlation with the fermentation index, as depicted in Figure 2(a), owing to its significantly elevated absorption value. The absorption at 888 nm in Figure 2(b) has the highest value, indicating a strong correlation between this wavelength and the pH parameter. The correlation between the wavelength range of 975 nm-980 nm and water content, as depicted in Figure 2(c), is strong (high value of coefficient regression/BW) due to the significant absorption characteristics exhibited by this range. Similar research results have been presented by Priambodo et al., (2022).

The second derivative preprocessing method (SG2) was proven to be more suitable for the pH model than the other four preprocessing methods were. The ANN model provided the best results, with a calibration correlation coefficient (r) of 0.971 and a validation value of 0.962, indicating a high correlation calibration model. Meanwhile, the RPD value is 2.191, which shows very good prediction accuracy because it is between the values 2.5 and 3 and greater than 3 (Hernández-Hernández et al., 2022). In the water content analysis, the preprocessing technique using MSC showed a strong correlation ($r > 0.7$). The MSC data processing technique produced the largest RPD values in the ANN and PLSR models, proving its efficacy in significantly increasing the RPD value for predicting water content.

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

Available online:
http://journal.ipb.ac.id/index.php/jtep
DOI: 10.19028/jtep.013.4.313-528

The calibration and validation models for estimating the fermentation index, pH, and water content using PLSR and ANN methods with a pre-processing method for NIR spectra data (SG1, SG2, LBC, SNV, and MSC) were found to be effective in predicting these parameters. The resulting model had an RPD value between 1.5 and 2.0, making it possible to observe the difference between high and low values (Andasuryani et al., 2013). Based on the obtained r and RMSE values for training, testing, and validation, ANN modeling outperformed PLSR modeling in terms of the core between experimental and model-predicted data. This can be explained by the nature of the model. The ANN model is attributable to its nonlinear mapping capabilities, which are lacking in PLSR models (Grgić et al., 2022, Panagou et al., 2011).

The root mean square errors (RMSE) of the chosen models were small, and the coefficients of determination for all four preprocessing methods were higher than 0.84. ANN and PLSR, supported by preprocessing, consistently produced predictions with good performances. ANN, with the help of SG1, SG2, LBC, SNV, and MSC, consistently predicted the fermentation index, pH, and water content of dried fermented cocoa beans with better results than PLSR. Data preprocessing techniques, specifically SG1, SG2, and LBC, were applied to handle the baseline effect in the spectral data resulting from noise or other sources of interference. These techniques involve smoothing and differentiation procedures to enhance the spectra of the SG1 and SG2 methods. In the case of LBC, the baseline variations in spectral measurements and chemical content are reduced to improve the signal-to-noise ratio and obtain a more accurate signal (Kusumaningrum et al., 2018, Meenu et al., 2016). Scatter-correlation methods, such as standard normal variance (SNV) and multiplicative scatter correction (MSC), were used to eliminate variations in scattering from light sources in spectral data, alterations in light conditions, and variations in the sizes of scanned objects. The primary objective of these methods is to eliminate multiplicative interference and separate the influences of physical light scattering and chemical light absorbance (Kusumaningrum et al., 2018, Ramadhan et al., 2016).

Research on predicting fermentation index and pH using FT-NIR spectroscopy on cocoa beans treated with long natural fermentation and long post-harvest fruit storage has also shown promising results. The best fermentation index was estimated using the PLSR method with the first derivative pre-processing method (R2 = 0.88, SEC = 0.06, and RPD = 2.47). Moreover, the best pH value obtained using the PLSR method with first derivative data pre-processing produced a value of R2 = 0.76, SEC = 0.026, and RPD = 2.05 (Sunoj et al., 2016). NIR spectroscopy, which mainly refers to the chemical composition of the sample, only records absorbance changes due to molecular interactions at various wavenumbers. Therefore, NIR spectroscopy cannot directly measure pH, as pH meters require processing. However, the model performance can be improved by increasing the number of samples and narrowing the range of values. Based on this research, fermentation time is one of the factors that influences the predicted results.

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

Available online:
http://journal.ipb.ac.id/index.php/jtep
DOI: 10.19028/jtep.013.4.313-528

4. Conclusion

The results of this study demonstrate that the Portable NIR Spectrometer SCiO, in the wavelength range of 740–1,070 nm, has the potential to reliably determine the fermentation index, pH, and water content of dried fermented cocoa beans. The optimal duration for fermenting cocoa beans to achieve better quality in terms of fermentation index and pH ranged from 48 to 72 h. Cocoa beans with a fermentation index ranging from 1.3 to 1.4 and a relatively low pH level of approximately 5.3 can be classified as fully fermented cocoa beans. The prediction outcomes of ANN models, when applied to dried fermented cocoa beans with data preprocessing, offered better results than those of PLSR models. The optimal pH was determined using the ANN method with preprocessing techniques. The most accurate water content determination was achieved using the ANN method with MSC preprocessing. The application of NIR spectra data pre-processing techniques enhances the model's performance.

5. References

Ministry of Agriculture Republic of Indonesia (2021) Outlook Kakao 2020, Jakarta (ID), Pusdatin Pertanian.

Sabahannur, S., Nirwana, N., & Subaedah, S. (2016). Kajian mutu biji kakao petani di Kabupaten Luwu Timur, Soppeng dan Bulukumba. Jurnal Industri Hasil Perkebunan, 59–66. https://doi.org /10.33104 /jihp.v11i2.3412

Dzelagha, B. F., Ngwa, N. M., & Nde Bup, D. (2020). A review of cocoa drying technologies and the effect on bean quality parameters. International Journal of Food Science, 2020, 1–11. https://doi.org/10.1155/2020/8830127

Raharja, S., Rahardjo, Y. P., & Syamsu, K. (2023). Aroma precursor enhancing in dried cocoa beans fermentation using enzyme and heat addition. AIMS Agric Food, 8(2), 674-686. https://doi.org/ 10.3934/agrfood.2023037

Sunoj, S., Igathinathane, C., & Visvanathan, R. (2016). Nondestructive determination of cocoa bean quality using FT-NIR spectroscopy. Comput Electron Agric 124: 234–242. https://doi.org/ 10.1016/j.compag.2016.04.012

Priambodo, D. C., Saputro, D., Pahlawan, M. F. R., Saputro, A. D., & Masithoh, R. E. (2022). Determination of acid level (ph) and moisture content of cocoa beans at various fermentation level using visible near-infrared (vis-NIR) spectroscopy. IOP Conference Series: Earth and Environmental Science, 985(1), 012045. https://doi.org/10.1088/1755-1315/985/1/012045

Hernani, Hidayat, T., & Mulyawanti, I. (2019). The usage of dried starter for re-fermentation of unfermented cocoa beans. IOP Conference Series: Earth and Environmental Science, 012061. https://doi.org/10.1088/1755-1315/309/1/012061

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

Available online:
http://journal.ipb.ac.id/index.php/jtep
DOI: 10.19028/jtep.013.4.313-528

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

Feng, H., Zhang, Z., & Gao, X. (2022). Rapid quality assessment of Succus Bambusae oral liquid based on near infrared spectroscopy and chemometrics. Industrial Crops and Products, 189: 115862. https://doi.org/10.1016/j.indcrop.2022.115862

Moghaddam, H. N., Tamiji, Z., & Akbari Lakeh, M. (2022). Multivariate analysis of food fraud: A review of NIR based instruments in tandem with chemometrics. Journal of Food Composition and Analysis, 107: 104343. https://doi.org/10.1016/j.jfca.2021.104343

Basile, T., Marsico, A. D., & Perniola, R. (2022). Use of artificial neural networks and nir spectroscopy for non-destructive grape texture prediction. Foods, 11(3), 281. https://doi.org/ 10.3390/ foods11030281

El Orche, A., Mamad, A., Elhamdaoui, O., Cheikh, A., El Karbane, M., & Bouatia, M. (2021). Comparison of machine learning classification methods for determining the geographical origin of raw milk using vibrational spectroscopy. Journal of Spectroscopy, 2021, 1–9. https://doi.org/ 10.1155/2021/5845422

Teye, E., Uhomoibhi, J., & Wang, H. (2016). Nondestructive authentication of cocoa bean cultivars by FT-NIR spectroscopy and multivariate techniques. Focus on Sciences, 2(3), 1–10. https://doi.org/ 10.21859/focsci-020347

Osborne, B. G. (2006). Near‐infrared spectroscopy in food analysis. . In encyclopedia of analytical chemistry: applica-tions, theory and instrumentation. John Wiley & Sons, Ltd.

Teye, E., Anyidoho, E., Agbemafle, R., Sam-Amoah, L. K., & Elliott, C. (2020). Cocoa bean and cocoa bean products quality evaluation by NIR spectroscopy and chemometrics: A review. Infrared Physics & Technology, 104, 103127. https://doi.org/10.1016/j.infrared.2019.103127

Grgić, F., Jurina, T., Valinger, D., Gajdoš Kljusurić, J., Jurinjak Tušek, A., & Benković, M. (2022). Near-infrared spec-troscopy coupled with chemometrics and artificial neural network modeling for prediction of emulsion drop-let diameters. Micromachines, 13(11), 1876. https://doi.org/10.3390/ mi13111876

Kadow, D., Niemenak, N., & Rohn, S. (2015). Fermentation-like incubation of cocoa seeds (Theobroma cacao L.)–Reconstruction and guidance of the fermentation process. LWT-Food Sci Technol 62: 357–361. https://doi.org/10.1016/j.lwt.2015.01.015

Gourieva, K. B., Tserevitinov, O. B. (1979). Method of evaluating the degree of fermentation of cocoa beans. USSR Pat 646.

AOAC. (1995). Official Methods of Analysis of AOAC International (16th ed.) pH of acidified foods. Section 42.1.04, method 981.12, Gaithersburg (MD), AOAC International.

National Standardization Agency of Indonesia. (2013). Kakao Bubuk, SNI 3747-2013. 1–29.

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

Available online:
http://journal.ipb.ac.id/index.php/jtep
DOI: 10.19028/jtep.013.4.313-528

Rinnan, Å., Van Den Berg, F., Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. TrAC Trends Anal Chem 28: 1201–1222. https://doi.org/ 10.1016/ j.trac.2009.07.007

Burns, D. A., Ciurczak, E. W. (2007). Handbook of near-infrared analysis, CRC press. https://doi.org/10.1016/j.trac.2009.07.007

Sharabiani, V. R., Nazarlo, A. S., Taghinezhad, E., Veza, I., & Szumny, A. (2023). Prediction of winter wheat leaf chlorophyll content based on VIS/NIR spectroscopy using ANN and PLSR. Food Science Nutrition, 11: 2166–2175. https://doi.org/10.1002/fsn3.3071

Hernández-Hernández, C., Fernández-Cabanás, V. M., &Rodríguez-Gutiérrez, G., Fernández-Prior, Á., & Morales-Silero, A. (2022). Rapid screening of unground cocoa beans based on their content of bioactive compounds by NIR spectroscopy. Food Control 131: 108347. https://doi.org/10.1016/ j.foodcont.2021.108347

Pasquini, C. (2003). Near infrared spectroscopy: fundamentals, practical aspects and analytical applications. Journal of The Brazilian Chememical Society 14: 198–219. https://doi.org/10.1590/ S0103-50532003000200006

Osborne, B. G., Fearn, T., & Hindle, P. H. (1993). Practical NIR spectroscopy with applications in food and beverage analysis., Longman scientific and technical.

Balcázar-Zumaeta, C. R., Castro-Alayo, E. M., & Cayo-Colca, I. S. (2023). Metabolomics during the spontaneous fermentation in cocoa (Theobroma cacao L.): An exploraty review. Food Res Int 163: 112190. https://doi.org/10.1016/j.foodres.2022.112190

Romero-Cortes, T., Salgado-Cervantes, M. A., & García-Alamilla, P. (2013). Relationship between fermentation index and other biochemical changes evaluated during the fermentation of Mexican cocoa (Theobroma cacao) beans. J Sci Food Agric 93: 2596–2604. https://doi.org/10.1002/jsfa.6088

Afoakwa, E. O., Paterson, A., & Fowler, M. (2008). Flavor formation and character in cocoa and chocolate: A Critical Review. Critical Reviews in Food Science and Nutrition. 48: 840–857. https://doi.org/10.1080/10408390701719272

Andasuryani., Purwanto, Y. A., Budiastra, I. W. (2013). Non destructive and rapid analysis of catechin content in Gambir (Uncaria gambir Roxb.) using NIR spectroscopy. IJSER 4: 383–389. https://doi.org/10.18517/ijaseit.4.5.423

Panagou, E. Z., Mohareb, F. R., & Argyri, A. A. (2011). A comparison of artificial neural networks and partial least squares modelling for the rapid detection of the microbial spoilage of beef fillets based on Fourier transform infrared spectral fingerprints. Food Microbiol 28: 782–790. https://doi.org/ .1016/j.fm.2010.05.014

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0

Kusumaningrum, D., Lee, H., & Lohumi, S. (2018). Non‐destructive technique for determining the viability of soybean (Glycine max) seeds using FT‐NIR spectroscopy. J Sci Food Agric 98: 1734–1742. https://doi.org/10.1002/jsfa.8646

Meenu, M., Kamboj, U., & Sharma, A. (2016). Green method for determination of phenolic compounds in mung bean (Vigna radiata L.) based on near‐infrared spectroscopy and chemometrics. Int J Food Sci Technol 51: 2520–2527. https://doi.org/10.1111/ijfs.13232

Ramadhan, S., Munawar, A. A., & Nurba, D. (2016). Aplikasi NIRS dan principal component analysis (PCA) untuk mendeteksi daerah asal biji kopi arabika (Coffea arabica). J Ilm Mhs Pertan 1: 954–960. https://doi.org/10.17969/jimfp.v1i1.1182

Copyright © 2024. This is an open-access article distributed under the CC BY-SA 4.0 License (https://creativecommons.org/licenses/by-sa/4.0