Agricultural products quality determination by means of near infrared spectroscopy

A A Munawar1*, Z Zulfahrizal1, R Hayati2, Syahrul1

1Department of Agricultural Engineering, Universitas Syiah Kuala, Jalan Tgk. Hasan Krueng Kalee 3, Banda Aceh 23111, Indonesia
2Department of Agro-technology, Universitas Syiah Kuala, Jalan Tgk. Hasan Krueng Kalee 3, Banda Aceh 23111, Indonesia

*Email: aamunawar@unsyiah.ac.id

Abstract. Cocoa is one of main agricultural products cultivated in many tropical countries and processed onto several derivative products. To determine cocoa beans qualities, laboratory procedures based on solvent extractions were mainly used, however most of them are destructive and may cause environmental pollutions. The main purpose of this present study is to employ near infrared spectroscopy (NIRS) for rapid and non-destructive assessment of cocoa beans in form of fat content. Near infrared spectral data of cocoa bean samples were measured as diffuse reflectance in wavelength range from 1000 to 2500 nm. Reference fat contents were measured using standard laboratory methods. Prediction models were developed using principal component regression with raw and baseline corrected spectra data. The results showed that fat contents of cocoa beans can be predicted and determined with maximum correlation coefficient ($r$) of 0.89 and ratio prediction to deviation (RPD) index of 2.87 for raw spectra and $r$ of 0.91, RPD of 3.18 for baseline spectra correction. It may conclude that NIRS was feasible to be applied as a rapid and non-destructive method for cocoa bean quality assessment.

1. Introduction

Cocoa beans as raw material for making chocolate are one of the strategic plantation export commodities to generate large foreign exchange for Indonesia. However, Indonesian cocoa bean products are generally not fermented, so the price is low in the market. Cocoa quality control such as moisture content and fat content has not been carried out intensively [1,2]. Quality assurance of cocoa beans through the development of fast and accurate quality estimation methods is the key word to increase the competitiveness of Indonesian cocoa beans exports at the world level [3,4].

Postharvest handling is the key to the success of improving the quality of Indonesian cocoa beans. The next thing that must be considered is the quality control of cocoa beans considering that the consistency of the quality of Indonesian agricultural products in general is still low. Fast and precise quality determination methods are needed to produce high quality standard cocoa commodities required by consumer countries [5]. Indonesia standard or known as SNI stipulates quality standards for cocoa beans physically, such as water content, contamination with insects, foreign objects and various aromas that can damage the distinctive aroma of cocoa. In particular, the quality of cocoa is determined by the yield of fat, aroma and taste, because these components usually determine the sensation of enjoying chocolate [2].
Determination and estimation of cocoa quality attributes are usually measured through laboratory tests (destructively), in which cocoa beans are crushed and extracts are extracted which are then analyzed using standard chemical methods common in laboratories [6,7]. In fact, this chemical method takes a long time and is expensive, so it is not suitable to be applied in industries that require a very fast and non-destructive method to analyze cocoa quality [8–10].

Fast and efficient detection of food quality can be realized through the development of Near Infrared Reflectance Spectroscopy (NIRS) technology. NIRS has become one of the most promising non-destructive methods and can be used for analysis in various fields, including in agriculture [11–14]. The advantages are simple preparation of the sample, fast detection process, and environmental friendliness as no chemicals are used. More importantly, NIRS has the potential ability to define multiple quality parameters simultaneously [15–18]. Through the development of computer science and chemometrics, the application capabilities of the NIRS technique have become more popular and have attracted the attention of researchers in the food sector. Components with a concentration percentage of 0.1% can be detected and evaluated using NIRS. Therefore, the main purpose of this present study is to employ near infrared spectroscopy (NIRS) for rapid and non-destructive assessment of cocoa beans in form of fat content.

2. Materials and methods

2.1. Cocoa bean samples
The material used in this study are cocoa beans as a bulk obtained from several cocoa plantations in Indonesia. Intact cocoa beans were delivered to Banda Aceh and stored in room temperature 25°C for two days to equilibrate.

2.2. Near infrared spectrum
Spectra data of cocoa bean samples were collected and acquired in form of absorbance or diffuse reflectance spectrum. Data were recorded using PSD NIRS i16 (USK) as shown in Figure 1 with 4x optical gain and co-added of 64 scans per acquisition in wavelength range from 1000 to 2500 nm [19–21]. Spectral data of cocoa bean samples were saved as spa and csv file formats for further data analysis.

Figure 1. spectra data acquisition of cocoa bean sample using PSD NIRS i16.
2.3. Cocoa beans fat content prediction
Cocoa quality parameter in form of fat content was predicted by means of spectral data with calibration models approach. Prediction models were developed using principal component regression with raw and baseline corrected spectra data and validated using n-fold cross validation [22,23].

2.4. Prediction accuracy and robustness
The accuracy and robustness of the prediction models were observed by looking several statistical parameters, namely the correlation coefficient (r), root mean square error (RMSE) in calibration or validation, and the ratio between standard deviation and RMSEP or known as RPD. A good model should have high r and RPD value, with low RMSE value that is smaller than the standard deviation (SD) of the reference data [15,24,25].

2.5. Relevant wavelengths
From selected models, relevant and optimum wavelengths for cocoa beans quality parameter can be observed by determining wavelength range from the highest peak and lowest valley in regression coefficient plot or loading plot [10,26].

3. Results and discussion

3.1. near infrared spectrum of cocoa beans
The spectra feature of intact cocoa beans samples which corresponds to chemical information of inner quality of the samples is shown in Figure 2.

![Near Infrared Spectrum of Cocoa Beans](image)

**Figure 2.** Near infrared spectrum of intact cocoa bean samples.

The raw spectrum of intact cocoa beans represents the presence of procyanidin and ammonia, which are markers of fermentation level. The peak formed in the spectrum indicating the presence of procyanidin and ammonia was clearly visible. In previous studies, it was found that the wavelength ranges thought to play a role in providing information for the presence of procyanidins were 1400-1480 nm, 1900-2000 nm and 2060-2160 nm respectively. Cocoa beans fermentation is an important thing in the cocoa bean processing.

The ability of NIRS to differentiate fermentation is an important problem that requires a solution. The spectra correction method used in this present study to enhance the prediction performance is baseline shift correction (BSC). It refers to the previous results of the first study which showed that this
BSC approach had good performance for agricultural product. The spectrum of corrected spectral data by means of BSC method is presented in Figure 3.

![Enhanced near infrared spectrum of intact cocoa bean samples using BSC method](image1)

**Figure 3.** Enhanced near infrared spectrum of intact cocoa bean samples using BSC method.

The baseline shift correction algorithm seeks to eliminate multiplicative interference on scattering, particle size and changes in beam distance. This method is able to improve the effect of multiplicative and additive scatter [24,27]. It is often used as a pretreatment because it focuses on the differences between observations rather than the absolute values of the data. The BSC method ensures that the resulting data or model can be interpreted in variations around the data mean. Moreover, spectral data were subjected onto principal component analysis to observe data distribution for similarities as shown in Figure 4.

![Spectral data of cocoa samples projected on principal component analysis](image2)

**Figure 4.** Spectral data of cocoa samples projected on principal component analysis.

The NIRS method works on the principle that every biological object has certain optical and electromagnetic characteristics that can be analyzed into information about the content of the chemical
object. Several industries use this method to predict the protein, fat and carbohydrate content of agricultural products, as well as to analyze the degree of bruising and damage to fruit. Each biologic material has different optical characteristics and the shape of the electromagnetic spectrum where the shape of this spectrum will characterize the chemical content of the material. This phenomenon has prompted many scientists to research the possibility of applying this method to predict the quality of an organic material such as fruits, flour, and herbal leaves that will be used as herbal ingredients.

3.2. Quality prediction based on NIRS models
The main part in near infrared application is to develop prediction models used to determine inner quality parameters on intact agricultural object. This can be completed by regressing spectral data obtained from NIRS instrument and reference chemical data quality attributes measured by means of standard laboratory methods. In this study, near infrared spectral data were used to determine fat content on intact cocoa bean samples. At first, prediction model was established using raw untreated spectra data to determine fat content. The prediction performance of the model is presented in Figure 5. The prediction accuracy of the model shows that the model can predict fat content with determination coefficient 0.758, correlation coefficient 0.89 and ratio prediction to deviation index is 2.87, which categorized as good model performance.

![Figure 5. Prediction performance of raw spectrum for fat content using PCR regression.](image)

NIRS is in the wavelength of 780 – 2500 nm (12,500 – 4,000 cm\(^{-1}\)) and contains more complex information structures due to the combination pattern of the bonds. The recording of the NIRS electromagnetic wave region is the response of the molecular bonds of O-H, C-H, C-O and N-H. This bond causes a change in vibration energy when irradiated by the NIRS frequency, namely stretching and bending vibrations. Each form or response that occurs from this electromagnetic radiation carries a photon energy of different magnitude. Photons, as defined by [28] is the lowest energy radiation contained in electromagnetic radiation.

some of these photons cause electron transfer, while others cause molecular vibrations because biological materials contain molecular bonds between atoms. The molecular vibrations that occur cause the molecular bands to move up and down or pull and stretch at certain frequencies and wavelengths (Batten 1998). This event causes the shape of the spectrum that is different for each biological material.
An important thing that must be considered in choosing an NIR measurement set-up is the penetration of NIR radiation that can enter the material network. This penetration will usually decrease with increasing depth of the material to be penetrated. Quella and Criollo [4,29] found the penetration depth of apples that can penetrate up to 4 mm at a wavelength of 900 – 1900 nm. The depth of penetration will differ significantly based on the thickness of the material. Variations in the size and temperature of the sample particles affect the dispersion of NIR radiation as it passes through the material. Large particles cannot spread NIR radiation as much as absorbed radiation, so the higher the absorbance value and the absorbed wavelength will also be larger and stronger. The prediction performance was improved when the models is constructed using BSC corrected spectra data as shown in Figure 6.

![Figure 6](image.png)

**Figure 6.** Prediction performance of BSC spectrum for fat content using PCR regression.

Organic matter will only reflect about 4% of the light it receives from a source through the outer surface as regular reflection and the remaining 96% will enter the product which then undergoes absorption, body reflectance, scattering and forwarding, light or transmittance.

In principle, the NIRS instrument consists of a light source, a beam splitter system, a sample detector, a beam detector and a data processing analysis system. For light sources, cheap Halogen Tungsten lamps are usually used or expensive lamps can also be used. The beam splitter system is useful for translating multi-coloured rays into a single beam such as filter beams, interferometers and gratings. The sample detector is adapted to the shape of the sample such as liquid or solid. Computers are used for data acquisition, communication control analysis and numerical analysis on the spectrometer system. The parameters of the NIR spectrometer are considered to obtain the optimum performance of the instrument. The selection of the wavelength region, solution, scanning speed, number, mode, and sampling interval will affect the accuracy and repeatability of measurements.

4. Conclusions

Cocoa is one of main agricultural products cultivated in many tropical countries and processed onto several derivative products. The main purpose of this present study is to employ near infrared spectroscopy (NIRS) in wavelength range from 1000 to 2500 nm for rapid and non-destructive assessment of cocoa beans in form of fat content. The results showed that fat contents of cocoa beans can be predicted and determined with correlation coefficient (r) of 0.89 and ratio prediction to deviation (RPD) index of 2.87 for raw spectra and r of 0.92, RPD of 3.18 for baseline spectra correction. It may
conclude that NIRS was feasible to be applied as a rapid and non-destructive method for cocoa bean quality assessment.

References

[1] Kamal M, Munawar A A and Sulaiman M I 2021 Comparison of principal component and partial least square regression method in NIRS data analysis for cocoa bean quality assessment IOP Conf. Ser. Earth Environ. Sci. 667

[2] Hayati R, Zulfarihazal Z and Munawar A A 2021 Robust prediction performance of inner quality attributes in intact cocoa beans using near infrared spectroscopy and multivariate analysis Heliyon 7 e06286

[3] Teye E, Anyidoho E, Agbemafle R, Sam-Amoah L K and Elliott C 2020 Cocoa bean and cocoa bean products quality evaluation by NIRS spectroscopy and chemometrics: A review Infrared Phys. Technol. 104

[4] Criollo R, Bayona O, Ochoa D, Cevallos-Cevallos J and Liao W 2020 Improving the detection of cocoa bean fermentation-related changes using image fusion Hyperspectral Remote Sensing (Elsevier) pp 343–56

[5] Cruz-Tirado J P, Fernández Pierna J A, Rogez H, Barbin D F and Baeten V 2020 Authentication of cocoa (Theobroma cacao) bean hybrids by NIR-hyperspectral imaging and chemometrics Food Control 118

[6] Zareef M, Mehedi Hassan M, Arslan M, Ahmad W, Ali S, Ouyang Q, Li H, Wu X and Chen Q 2020 Rapid prediction of caffeine in tea based on surface-enhanced Raman spectroscopy coupled multivariate calibration Microchem. J. 159

[7] Álvarez C, Pérez E, Clos E, Lares M, Assemat S, Boulanger R and Davrieux F 2012 The Use of near Infrared Spectroscopy to Determine the Fat, Caffeine, Theobromine and (−)-Epicatechin Contents in Unfermented and Sun-Dried Beans of Criollo Cocoa J. Near Infrared Spectrosc. 20 307–15

[8] Munawar A A, Kusumiyati and Wahyuni D 2019 Near infrared spectroscopic data for rapid and simultaneous prediction of quality attributes in intact mango fruits Data Br. 27 104789

[9] Munawar A A, von Hörsten D, Wegener J K, Pawelzik E and Mörlein D 2016 Rapid and non-destructive prediction of mango quality attributes using Fourier transform near infrared spectroscopy and chemometrics Food Control 52 208–15

[10] Subedi P P and Walsh K B 2011 Assessment of sugar and starch in intact banana and mango fruit by SWNIR spectroscopy Postharvest Biol. Technol. 62 238–45

[11] Nordey T, Joas J, Davrieux F, Chillet M and Léchaudel M 2017 Robust NIRS models for non-destructive prediction of mango internal quality Sci. Hortic. (Amsterdam). 216 51–7

[12] Comino F, Aranda V, García-Ruiz R, Ayora-Cañada M J and Domínguez-Vidal A 2018 Infrared spectroscopy as a tool for the assessment of soil biological quality in agricultural soils under contrasting management practices Ecol. Indic. 87 117–26

[13] Yusmanizar, Setiasih I S, Nurjanah S, Muhaemin M, Nurhadi B, Rosniawaty S and Munawar A A 2019 Fast and Non-Destructive Prediction of Moisture Content and Cholegenic Acid of Intact Coffee Beans Using Near Infrared Reflectance Spectroscopy IOP Conference Series: Materials Science and Engineering vol 506 (IOP Publishing Ltd.) p 012033

[14] Munawar A A, Kusumiyati, Hafidh, Hayati R and 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 - Agric. Eng. 58 1–12

[15] Pasquini C 2018 Near infrared spectroscopy: A mature analytical technique with new perspectives – A review Anal. Chim. Acta 1026 8–36

[16] Santos U J dos, Dematé J A de M, Menezes R S C, Dotto A C, Guimarães C C B, Alves B J R, Primo D C and Sampaio E V de S B 2020 Predicting carbon and nitrogen by visible near-infrared (Vis-NIR) and mid-infrared (MIR) spectroscopy in soils of Northeast Brazil Geoderma Reg. 23
Acknowledgement

We sincerely thank the directorate of research and community services LPPM Universitas Syiah Kuala for funding this study through PHI scheme 2021.