



Research article

Robust prediction performance of inner quality attributes in intact cocoa beans using near infrared spectroscopy and multivariate analysis

Rita Hayati^{a,**}, Zulfahrizal Zulfahrizal^b, Agus Arip Munawar^{b,*}^a Department of Agro-technology, Syiah Kuala University, Banda Aceh, Indonesia^b Department of Agricultural Engineering, Syiah Kuala University, Banda Aceh, Indonesia

ARTICLE INFO

Keywords:

NIRS
Cocoa
Prediction
Fat
Moisture

ABSTRACT

Fast and simultaneous determination of inner quality parameters, such as fat and moisture contents, need to be predicted in cocoa products processing. This study aimed to employ the near-infrared reflectance spectroscopy (NIRS) in predicting the quality mentioned above parameters in intact cocoa beans. Near-infrared spectral data, in a wavelength ranging from 1000 to 2500 nm, were acquired for a total of 110 bulk cocoa bean samples. Actual fat and moisture contents were measured with standard laboratory procedures using the Soxhlet and Gravimetry methods, respectively. Two regression approaches, namely principal component regression (PCR) and partial least square regression (PLSR), were used to develop the prediction models. Furthermore, four different spectra correction methods, namely multiple scatter correction (MSC), de-trending (DT), standard normal variate (SNV), and orthogonal signal correction (OSC), were employed to enhance prediction accuracy and robustness. The results showed that PLSR was better than PCR for both quality parameters prediction. Spectra corrections improved prediction accuracy and robustness, while OSC was the best correction method for fat and moisture content prediction. The maximum correlation of determination (R^2) and residual predictive deviation (RPD) index for fat content were 0.86 and 3.16, while for moisture content prediction, the R^2 coefficient and RPD index were 0.92 and 3.43, respectively. Therefore, NIRS combined with proper spectra correction method can be used to rapidly and simultaneously predict inner quality parameters of intact cocoa beans.

1. Introduction

Chocolate is food in a paste or solid-state made from either roasted or ground cocoa and fat combination. It is typically sweetened with additional sugar and other essential ingredients, made into bars, and eaten as confectionery. It was made from raw cocoa beans extracted from the cocoa tree pod and are either roasted, fermented, or grounded in other to realize the processed products. Chocolate plays a strategic role in the food industry because it is also directly consumed in other foodstuffs [1]. Cocoa beans are the most used raw substances in chocolate production, and it is extremely popular among people. Presently, there are two acknowledged types of cocoa beans, namely bulk cocoa (standard quality) and flavor cocoa (high quality). In the highly competitive market, chocolate industries and manufacturers need to ensure that they are supplied with high-quality cocoa beans. Fat and moisture content are the two primary inner quality parameters of cocoa associated with these product qualities [2].

Generally, several methods are widely employed to determine cocoa beans' inner quality parameters and other agricultural products. However, most of these methods are based on liquid and solvent extraction followed by other laboratory procedures, such as titration to determine the quantity of ascorbic acid and its separation using a centrifuge [3]. These methods are often laborious, time-consuming, destructive, and complicated [4]. Therefore, there is a need for rapid, robust, and non-destructive methods to analyze raw material has been one of the essential objectives of this sector and manufacturers over the past decades. The industry needs to be equipped with a proper and ideal rapid method that can monitor real-time cocoa processing steps, thereby facilitating the urgent and important decision-making process.

Over the past ten years, studies have reported the near-infrared reflectance spectroscopy (NIRS) as one of the most powerful non-destructive methods that have been widely proven to be used to carry out rapid and robust analysis in several fields, including agriculture. Numerous advantages are associated with the use of NIRS, such as simple

* Corresponding author.

** Corresponding author.

E-mail addresses: rita@unsyiah.ac.id (R. Hayati), aamunawar@unsyiah.ac.id (A.A. Munawar).

sample preparation, versatile, rapid, and environmentally friendly due to the absence of chemical materials [5]. In addition, it has the potential to predict and determine several inner quality parameters simultaneously [6, 7, 8].

Numerous studies and publications have reported that the NIRS method has been able to predict and determine inner quality parameters of food, dairy products, and other agricultural products [9, 10, 11], such as meat, animal feed [10, 12, 13], horticulture [14, 15, 16, 17], soil nutrients and properties [18, 19]. This technique was discovered to be rapidly applied, effective, and non-destructive in determining several quality attributes of raw intact organic materials and other derivative products. Based on these studies, it was argued that the NIRS method has the potentials and feasibility to determine inner quality parameters of organic materials. The proposed novelty includes a self-developed portable near infrared reflectance spectroscopy, which covered near-infrared regions. This device namely PSD NIRS *iptek* i16 was designed and developed in 2016 and already tested for spectral data acquisitions in wavelength range from 1000 to 2500 nm for most agricultural product samples like mangoes, bananas, apples, grapes, tomatoes, oranges, garlics, onions, cocoa beans and powder, coffee beans and powder, soil samples, rice, honeys and patchouli oils. The instrument has been calibrated using standard manufactured NIRS instrument (Thermo Nicolet Antaris TM II) in Georg-August University, Goettingen, Germany. The device characteristics including wavelength range 1000–2500 nm, Photodiode sensors with optical gain from 2 to maximum 16x. Spectra data can be acquired in from of absorbance, reflectance, and transmittance. Spectra file can be saved in two different extension file formats as *.spa and *.csv. In addition, the comparisons of spectra correction approaches were also analyzed. Moreover, Indonesia's original cultivar was utilized, and the models developed for unfermented and fermented cocoa beans.

Subsequently, several analyses have already been carried out on the cocoa samples. Most of them are based on standard manufactured instruments which were used to carry out a wide variety of studies related to the prediction of fermentation levels, inner quality, and review of near-infrared spectroscopy application for the assessment of cocoa products [20, 21, 22, 23, 23, 24, 25, 26, 27]. This research utilized a self-developed instrument designed in 2016. One advantage of this instrument is that irrespective of the fact that the size is smaller compared to the commonly used NIRS equipment with a dimension of 21 cm × 17 cm x 2 cm, it has a similar wavelength range in the near-infrared region (approximately 780–2500 nm) and a maximum 8x optical gain. Moreover, the instrument is cheaper and portable compared to the standard NIRS.

Therefore, this research aims to examine and apply the rapid, robust, and simultaneous method in determining the fat and moisture contents of raw intact cocoa bean samples using a self-developed NIRS instrument. Also, the prediction accuracy and robustness of spectra data correction was also analyzed.

2. Materials and methods

2.1. Cocoa bean samples

A total of 110 bulk cocoa bean samples cv. *Lindak*, harvested from June to August in the same cocoa plantations in East Java, Indonesia, were used to carry out this research. Each bulk contains approximately 54g of intact cocoa beans with cocoa added to both the unfermented and fermented types in different levels (1, 3, 5, and 7 days).

2.2. Spectral data acquisition

The near-infrared spectral data of all samples were taken in the form of diffuse reflectance spectrum using a portable sensing device near-infrared spectroscopy (PSD-NIRS i16 *iptek*). Spectral data were

obtained within the wavelength range of 1000–2500 nm with a resolution of 0.2 nm and co-added 32 scans per acquisition [28].

2.3. Reference fat and moisture content measurement

After the spectral data collection was completed, all cocoa bean samples were immediately taken to measure their inner quality parameters in the form of fat content (FC) and moisture content (MC). Initially, FC was measured using Soxhlet method. With 10 g of the sample mixed in a tube containing a maximum of 150 ml n-hexane and extracted in a soxhlet apparatus at a temperature of 95 °C for 6 h. It was further determined by evaporating the solvent using a rotary evaporator till only the liquid fat was left in the tube. The FC was then expressed in percentage (%) fat content [20, 29]. Conversely, MC was measured using a gravimetric method based on duplicated ISO 6673 and averaged. A forced-air electric oven (Thermicon type UT6120, Heraeus Instruments GmbH, Hanau, Germany) was used to dry approximately 15 g whole intact beans in open glass petri dishes (diameter: 14 cm, height: 2.3 cm) at 120 C for 18 h. After the drying process was completed, the dishes were immediately closed with glass lids to avoid exposure and stored in desiccators for one hour to equilibrate samples into ambient temperature. Also, the moisture content was expressed in percentage (%) dry bulb.

2.4. Calibration models

The next step after the spectra data acquisition and actual FC and MC measurement was the development of calibration models used to predict the quality parameters. At first, it was attempted to develop calibration models using original untreated or uncorrected spectral data. Two regression approaches, namely principal component regression (PCR) and partial least square regression (PLSR), were used to develop the models [30]. The results from the predictions were compared, and the ideal one between them was selected.

2.5. Spectral data correction and enhancement

Spectral data contains noise due to light scattering, which tends to interfere with the accuracy of the prediction. Therefore, it is recommended that they are corrected in order to improve prediction accuracy and robustness. In this study, four different spectra correction methods were employed, namely multiplicative scatter correction (MSC), detrending (DT), standard normal variate (SNV), and orthogonal signal correction (OSC).

2.6. Performance evaluation

Prediction model performance was evaluated based on calibration and cross-validation results in accordance with the correlation coefficient (r), the root means square error (RMSE), and the residual predictive deviation (RPD) index. The RPD index was obtained by dividing the standard deviation of reference data with the RMSE value. In addition, the range to error ratio (RER) was also used to evaluate the model's performance. Subsequently, the number of latent variables (LVs) required to develop the prediction models was also considered. The ideal and robust models need to possess a higher r coefficient and RPD index, as well as lower RMSE and fewer LVs. The prediction performances for fat and moisture contents were then systematically compared based on those statistical indicators [31].

The best spectra correction method was selected and used to predict other external cocoa bean samples. Moreover, cross-validation was also carried out using K-fold where K = 10, which means 7 randomized samples were excluded during calibration and used to test the models. These samples were assumed to be similar to the 'mini' external validation. This step was repeated 10 times until all the folds were completed with different samples each time the cross-validation was performed. Furthermore, to obtain a certain precision, external validation was

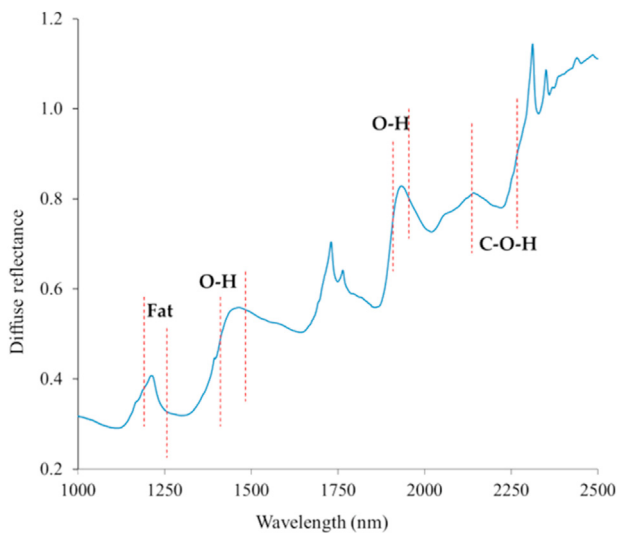


Figure 1. Near-infrared reflectance spectra feature of intact cocoa bean sample within the wavelength range of 1000–2500 nm.

Table 1. Descriptive statistics of actual fat and moisture contents using standard laboratory measurement for 72 samples in the calibrated dataset.

Statistical indicator	Actual fat content	Actual moisture content
N	72	72
Mean	40.71	9.12
Max	44.32	12.08
Min	35.26	6.74
Range	9.06	5.34
Std Deviation	2.19	1.30
Variance	4.79	1.69
RMS	40.77	9.21
Skewness	-0.08	0.70
Kurtosis	-0.94	-0.28
Median	40.72	8.96

carried out using 38 unknown cocoa bean samples obtained from local farmers in the Aceh province.

3. Result and discussion

3.1. Spectra features of cocoa beans

The recorded reflectance spectra for intact cocoa bean samples in the near-infrared region (1000–2500 nm) are shown in Figure 1. This infrared spectrum correlates with the related attributes were derived from the bands, thereby causing interaction between electromagnetic radiation and organic material. These bands correspond to specific molecular bonds of O–H, C–H, C–O, and N–H [20, 32].

Inner quality parameters of intact cocoa beans, such as fat and moisture contents, are formed by molecular bonds of C–H–O and O–H, respectively. Figure 1 shows that the moisture contents (O–H bonds) are probably predicted in the wavelength region of 1460–1490 nm and 1920–1980 nm. On the contrary, the fat content of intact cocoa bean samples is strongly determined to be within the 2100 and 2290 nm wavelength region. This is similar to a previous study that also stated that moisture or water absorption bands were observed at 1440 nm and 1935 nm due to O–H bands' combination and its overtone [33]. According to similar findings reported in some studies, strong water absorbance in organic materials such as fruit and other agricultural products was discovered to be in the wavelength region between 1420 and 1480 nm as well as 1920 and 1960 nm. Moreover, absorption bands within the wavelength range of 1240 and 1270 nm, 1750 and 1785 nm, 2210 and 2340 nm, including 2350 and 2380 nm, were related to the fat content. Similar findings were also reported in the research carried out on cocoa powder, cocoa beans, and chocolate bars, which stated that the optimum and effective predictions of wavelengths for fat content are within the range of 1230, 1740, 1755, 1768, 2200 and 2300, 2377, and 2385 nm [6, 16].

3.2. Fat and moisture content prediction

This study aims to simultaneously predict the fat and moisture contents of intact cocoa beans. The descriptive statistics for the actual reference data of fat and moisture contents obtained from the standard laboratory methods are shown in Table 1.

Initially, two different regression approaches, namely principal component regression (PCR) and partial least square regression (PLSR), were adopted to predict these quality parameters using 72 samples in the calibrated datasets. The prediction models were established by regressing raw untreated spectra data as an independent variable (X) and fat and moisture contents (Y). Prediction results for both quality parameters are shown in Table 2. Generally, the fat and moisture contents of intact cocoa bean samples are predicted to be quite satisfactory, and the maximum coefficient of determination was 0.67 and 0.72, respectively.

The maximum residual predictive deviation (RPD) indexes for fat and moisture contents prediction were 1.81 and 1.97, respectively. Based on certain literature, the RPD index, which was obtained to be between 1.5 and 2.0, was categorized as coarse, therefore, it needs to be improved. Furthermore, it was discovered that the PLSR regression approach led to a better prediction than the PCR, as shown in Figure 2. Therefore, this approach was adopted for the next data analysis, which is centered on exploring the impact of spectra enhancement and correction method on the prediction performances. PLSR seems to be better than PCR in terms of accuracy and robustness in the prediction of inner quality parameters. This is primarily based on the fact that the PLSR seeks to discover the best correlation between the reference and infrared spectra data during its transformation to latent variables (LVs) in the regression process. On the contrary, PCR only transforms spectra data to latent variables without involving the reference (fat and moisture contents).

Table 2. Prediction performance for fat and moisture contents using PCS and PLSR regression approaches.

Quality parameters	Method	Statistical indicators			
		R ²	RMSE	RPD	RER
Fat content	PCR	0.67	1.23	1.76	7.36
	PLSR	0.67	1.19	1.81	7.62
Moisture content	PCR	0.71	0.68	1.85	7.83
	PLSR	0.72	0.64	1.97	8.35

PCR: principal component regression, PLSR: partial least square regression, R²: coefficient of determination, RER: range to error ratio, RMSE: root mean square error, RPD: residual predictive deviation.

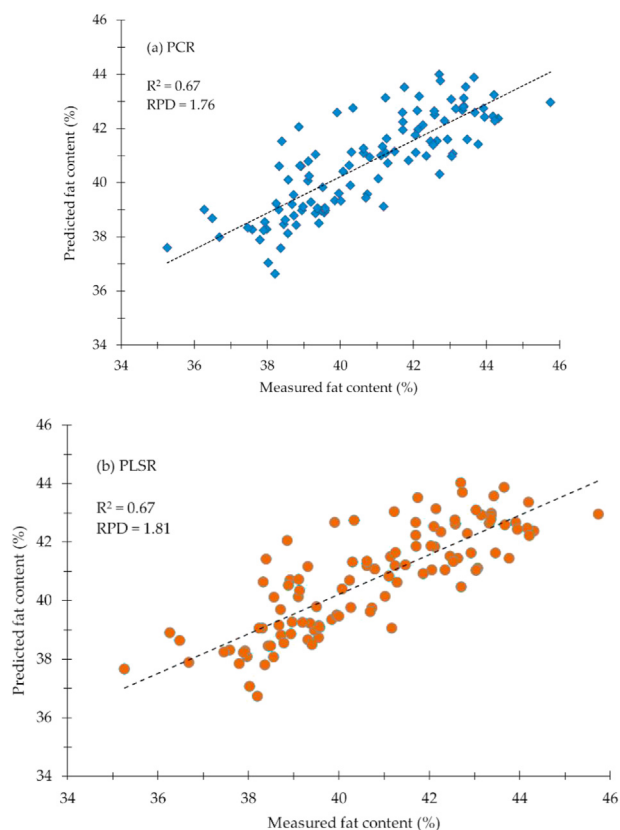


Figure 2. PCR (a) and PLSR (b) calibration used to predict fat content of intact cocoa beans.

In accordance with prediction of the fat content prediction, the PLSR was used to realize 0.67 as the coefficient of determination and 1.81 as the RPD index. Conversely, both the PLSR and PCR utilized 5 latent variables during the calibration to achieve prediction accuracy and robustness. Meanwhile, for moisture content, PLSR was better than PCR with a 0.72 coefficient of determination and 1.97 RPD index, as shown in Figure 3. Cocoa beans are biological objects that interfere with the inner quality parameters such as fat and moisture contents during ripening, storage, and distribution phases. External factors such as temperature and relative humidity also affect cocoa beans' inner quality and other agricultural products. Therefore, it tends to interfere with the accuracy and robustness of the predicted model. These effects need to be treated in order to achieve more robust and accurate prediction results. Therefore, it is strongly recommended to pre-process or enhance spectra data before the calibration of the prediction models.

The PLSR method was selected because it was discovered that this approach provides better prediction results than the PCR. This is also consistent with other findings, which reported that partial least square aids in achieving a more accurate and robust prediction than the principal component regression [34, 35].

3.3. The impact of spectra enhancement on prediction performance

In order to examine the impact of spectra correction and enhancement method on the prediction performance, four different spectra correction techniques, namely multiplicative scatter correction (MSC), de-trending (DT), standard normal variate (SNV), and orthogonal signal correction (OSC) were systematically compared and combined with PLSR regression approach to enhance prediction accuracy and robustness. The predicted models for both quality parameters (fat and moisture contents) were established using 72 spectral data corrected by those four spectra enhancement methods. The best spectra correction method was selected

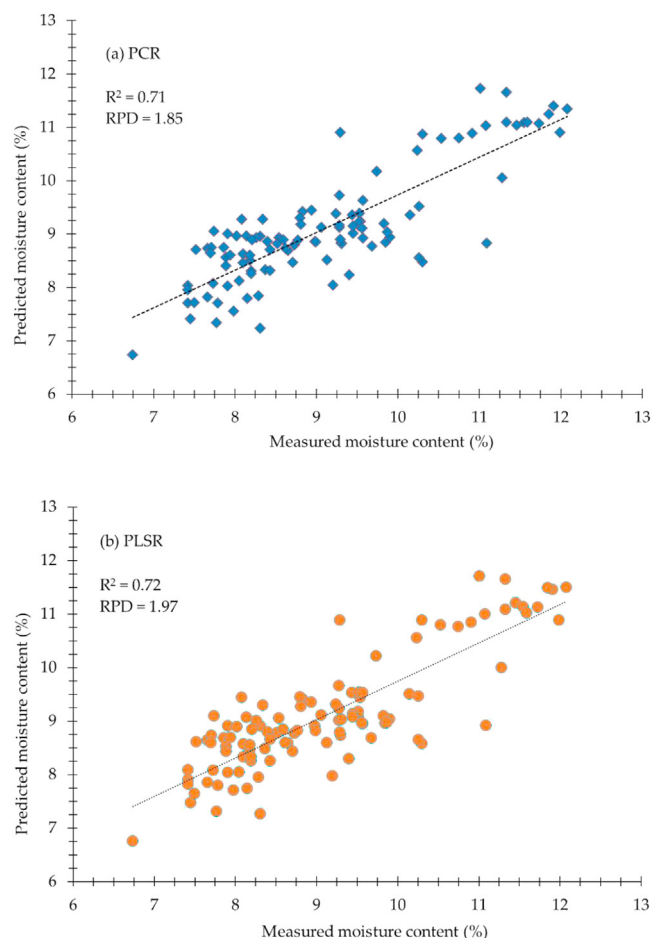


Figure 3. PCR (a) and PLSR (b) calibration to predict the moisture content of intact cocoa beans.

based on their prediction performances. Firstly, the raw untreated spectra data of intact cocoa beans was enhanced using the multiplicative scatter correction (MSC) method. This technique seeks to enhance the spectra by removing its multiplicative effects due to physical error during acquisition. The predicted results for fat and moisture contents using MSC spectra correction is shown in Table 2. It is evident that the coefficient of determination and RPD index were significantly improved after the MSC correction. The coefficient of determination for the fat content was predicted to be increased to 0.81 and RPD index to 2.81, while the root mean square error (RMSE) was decreased to 0.78. Similar findings were also reported during the moisture content prediction, where the MSC correction method enhanced its correlation to 0.85 and RPD index to 2.98 while the RMSE index decreased to 0.43.

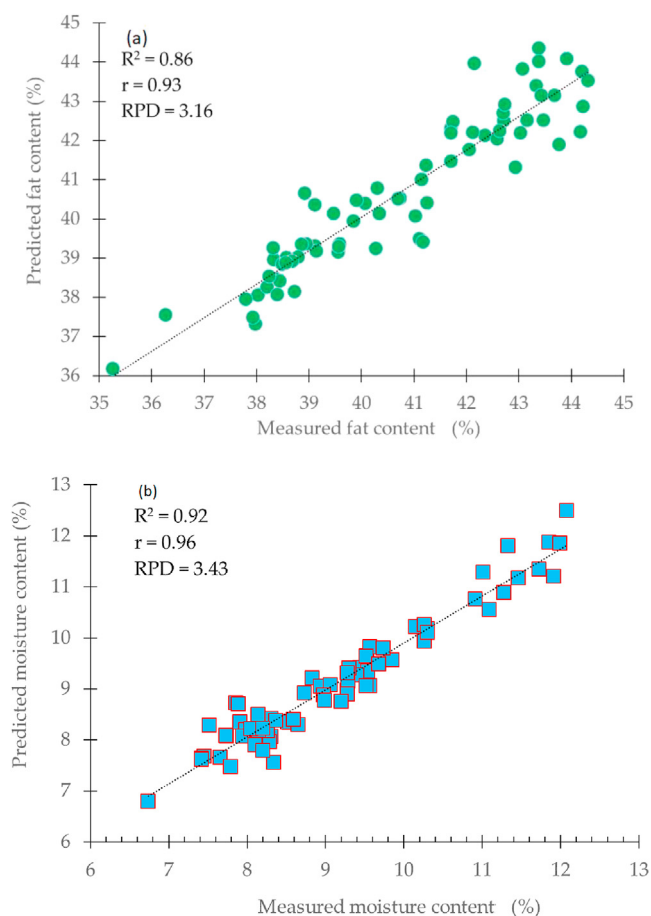
Moreover, the SNV correction method generated similar results as the MSC. The coefficient of determination and RPD achieved from SNV were similar for moisture content prediction, while for fat content, MSC was proven to be slightly better than SNV, as shown in Table 2. This is mostly because the MSC and SNV generally function based on an ideal spectrum. MSC obtains ideal spectrum from its mean spectra data for all samples, with SNC realized from the scaling algorithm. This is consistent with the research results, which stated that MSC and SNV provide similar or slightly better prediction accuracy and robustness.

De-trending spectra correction was taken into account because it was stated in several studies that this spectra correction method was fit and need to be applied when dealing with bulk samples. In this research, the second-order DT correction method was used to improve the PLSR prediction accuracy for fat and moisture content. As shown in Table 3, the

Table 3. Prediction performance for fat and moisture contents using different spectra correction methods.

Quality parameters	Spectra correction	Statistical indicators			
		R ²	RMSE	RPD	RER
Fat content	MSC	0.81	0.78	2.81	11.57
	DT	0.76	0.81	2.70	11.12
	SNV	0.79	0.79	2.79	11.49
	OSC	0.86	0.70	3.16	13.01
Moisture content	MSC	0.85	0.43	2.98	12.42
	DT	0.81	0.45	2.86	11.87
	SNV	0.85	0.43	2.97	12.42
	OSC	0.92	0.37	3.43	14.43

MSC: multiplicative scatter correction, DT: De-trending, SNV: standard normal variate, OSC: orthogonal signal correction, R2: coefficient of determination, RER: range to error ratio, RMSE: root mean square error, RPD: residual predictive deviation.

**Figure 4.** Scatter plot derived from OSC correction method for fat (a) and moisture contents (b) prediction of intact cocoa beans.

DT correction seems to be less accurate than the other two (MSC and SNV). Nevertheless, when compared to the raw un-corrected spectra, it obviously improved prediction accuracy and robustness for both quality parameters of intact cocoa beans.

The ideal prediction performance for fat and moisture contents was achieved when the spectra data were corrected and enhanced using the orthogonal signal correction (OSC) method. The coefficient of determination and RPD index for fat content prediction increased to 0.86 and 3.16. Moreover, OSC also improved the prediction accuracy for moisture content, whereas the coefficient of determination and RPD index was increased to 0.92 and 3.43, respectively. The scatter plot of the actual measured quality parameters and predicted ones derived using the OSC

Table 4. Descriptive statistics of actual fat and moisture contents using standard laboratory measurement for 38 samples in the calibrated dataset.

Statistical indicator	Actual fat content	Actual moisture content
N	38	38
Mean	40.56	9.03
Max	45.75	11.59
Min	36.49	7.42
Range	9.26	4.17
Std Deviation	2.16	1.19
Variance	4.68	1.41
RMS	40.61	9.11
Skewness	0.15	0.70
Kurtosis	-0.43	-0.42
Median	40.44	8.79

correction approach to predict fat and moisture contents are shown in Figure 4.

Based on these results, it was argued that the spectra approach significantly improved the prediction accuracy and robustness for both inner quality parameters of intact cocoa bean samples. Its correction and enhancement are used to remove irrelevant data such as noises and background information improperly handled by the regression techniques (PCR and PLSR).

External validation was carried out in order to ascertain the accuracy of the findings in this research, with cross-validation conducted at the calibration phase to overcome over fittings. Since OSC was discovered to be the best spectra correction method, the unknown fat and moisture content of 38 intact cocoa beans in a bulk of 54 g per sample was predicted. The descriptive statistics of actual fat and moisture content measurement and its validation performance are shown in Tables 4 and 5, respectively. A combination of the NIRS method and OSC spectra correction was also used to predict the coefficient of determination for fat, and moisture contents, which are 0.79 and 0.81, respectively.

Moreover, the residual predictive deviation (RPD) indexes for fat and moisture contents were 2.72 and 2.85, with RER of 11.65 and 10.07. This indicates that these models are robust because the REP values are above 9 as reported in certain literature. In addition, the scatter plot derived from OSC-PLSR methods is shown in Figure 5.

Based on the results obtained, it was therefore reported that NIRS serves as an alternative rapid and robust method in predicting inner quality parameters of intact cocoa bean samples. However, supposing it is arranged in order of relevance, the OSC correction method is more superior to the MSC, SNV, and DT. Spectra correction needs to be carried out before regression in order to achieve accurate and robust prediction results. The proper regression and spectra correction approaches lead to an effective outcome even when unknown intact cocoa bean samples are utilized.

Table 5. Validation performance for the prediction of fat and moisture contents using OSC spectra correction method.

Quality parameters	Statistical indicators			
	R ²	RMSEP	RPD	RER
Fat content	0.79	0.79	2.72	11.65
Moisture content	0.81	0.41	2.85	10.07

R²: coefficient of determination, RER: range to error ratio, RMSEP: root mean square error for prediction, RPD: residual predictive deviation.

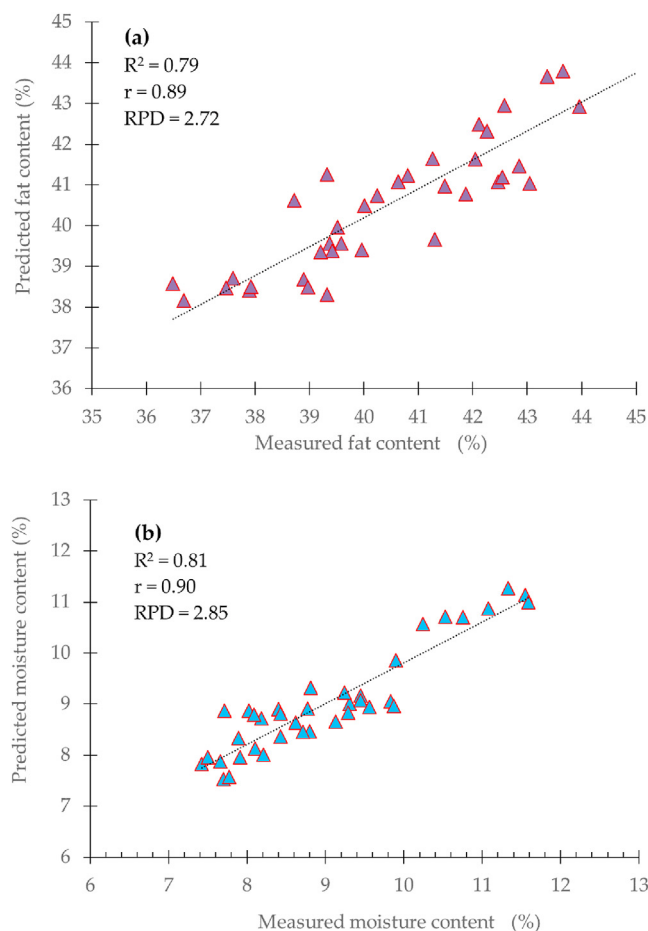


Figure 5. External validation performance for fat (a) and moisture contents (b) prediction using OSC spectra data on 38 independent cocoa bean samples.

Furthermore, it was previously reported that the proposed novelty in this study is the use of a portable NIRS instrument. Its advantage is that the size is smaller and cheaper than the common type. It is also comparable in terms of prediction performances. However, besides these advantages, there are also some limitations, such as when acquiring spectra data, the samples need to be rotated manually according to the operator's interference. In addition, during measurement, typical noises occur regularly on a specific wavelength range, particularly from 1405 to 1460 nm and between 1820 and 1865 nm. This is due to instability and overheating of the electronic components inside the NIRS instrument. Subsequently, based on the prediction performance, it is comparable and consistent with other related studies. The achieved RPD index is for fat, and moisture contents are 3.16 and 3.43, respectively, and it is categorized as good model performance based on certain literature [6, 36, 37].

4. Conclusion

This study demonstrated that near-infrared spectroscopy (NIRS) is a rapid and robust method used to predict intact cocoa bean samples' inner

quality parameters. The results showed that a combination of NIRS with partial least square regression (PLSR) approach can predict the fat and moisture contents of intact cocoa beans. The maximum coefficient of determination (R²) and residual predictive deviation (RPD) index for fat content were 0.86 and 3.16, while for moisture, 0.92 and 3.43 were obtained, respectively. Moreover, spectra correction significantly improves prediction accuracy and robustness. Therefore, spectra correction and enhancement prior to the development of the calibrated models were recommended. Orthogonal signal correction (OSC) was the best correction method and can be coupled with the PLSR regression technique.

Declarations

Author contribution statement

Rita Hayati: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Zulfahrizal: Analyzed and interpreted the data.

Agus Arip Munawar: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Funding statement

This work was supported by Kementerian Pendidikan dan Kebudayaan BRIN-DIKTI and LPPM Universitas Syiah Kuala, Republic of Indonesia.

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

- [1] A. Krämer, et al., Fast and neat - determination of biochemical quality parameters in cocoa using near infrared spectroscopy, *Food Chem.* 181 (2015) 152–159.
- [2] E. Teye, et al., Estimating cocoa bean parameters by FT-NIRS and chemometrics analysis, *Food Chem.* 176 (2015) 403–410.
- [3] A.A. Munawar, Y. Yunus, Devianti, P. Satriyo, Calibration models database of near infrared spectroscopy to predict agricultural soil fertility properties, *Data Br.* 30 (Jun. 2020).
- [4] R. Hayati, A.A. Munawar, F. Fachrudin, Enhanced near infrared spectral data to improve prediction accuracy in determining quality parameters of intact mango, *Data Br.* (Apr. 2020) 105571.
- [5] I. Ichwana, Z. Nasution, A. Arip Munawar, Near-infrared spectroscopy as a rapid and simultaneous assessment of agricultural groundwater quality parameters, *INMATEH Agric. Eng.* (Apr. 2020) 233–240.
- [6] C. Pasquini, Near infrared spectroscopy: a mature analytical technique with new perspectives – a review, *Anal. Chim. Acta* 1026 (2018) 8–36.
- [7] T. Nordey, J. Joas, F. Davrieux, M. Chillet, M. Léchaudel, Robust NIRS models for non-destructive prediction of mango internal quality, *Sci. Hortic. (Amsterdam)* 216 (2017) 51–57.
- [8] A.A. Munawar, D. von Hörsten, J.K. Wegener, E. Pawelzik, D. Mörlein, Rapid and non-destructive prediction of mango quality attributes using Fourier transform near

- infrared spectroscopy and chemometrics, *Eng. Agric. Environ. Food* 9 (3) (2016) 208–215.
- [9] Rahmaddiansyah Agussabti, P. Satriyo, A.A. Munawar, Data analysis on near infrared spectroscopy as a part of technology adoption for cocoa farmer in Aceh Province, Indonesia, *Data Br.* 29 (Apr. 2020).
- [10] A.A. Munawar, H. Syah, Yusmanizar, Fast and robust quality assessment of honeys using near infrared spectroscopy, *IOP Conf. Series: Earth Environ. Sci.* 365 (1) (Nov. 2019).
- [11] Y. Saputri, Yusriana, A.A. Munawar, Infrared spectroscopic features of turmeric powder, *IOP Conf. Ser. Earth Environ. Sci.* 365 (1) (Nov. 2019).
- [12] C.D. Iskandar, Zainuddin, A.A. Munawar, Rapid assessment of frozen beef quality using near infrared technology, *Int. J. Sci. Technol. Res.* 9 (5) (2020) 156–160.
- [13] Samadi, S. Wajizah, and A. A. Munawar, "Near infrared spectroscopy (NIRS) data analysis for a rapid and simultaneous prediction of feed nutritive parameters," *Data Br.*, vol. 29, Apr. 2020, doi: .
- [14] J.F. Xia, X.Y. Li, P.W. Li, M.A. Qian, X.X. Ding, Application of wavelet transform in the prediction of Navel orange vitamin C content by near-infrared spectroscopy, *Agric. Sci. China* 6 (9) (2007) 1067–1073.
- [15] E.T.S. Caramés, P.D. Alamar, R.J. Poppi, J.A.L. Pallone, Quality control of cashew apple and guava nectar by near infrared spectroscopy, *J. Food Compos. Anal.* 56 (2017) 41–46.
- [16] E. Arendse, O.A. Fawole, L.S. Magwaza, U.L. Opara, Non-destructive prediction of internal and external quality attributes of fruit with thick rind: a review, *J. Food Eng.* 217 (2018) 11–23.
- [17] O.O. Olarewaju, et al., Model development for non-destructive determination of rind biochemical properties of 'Marsh' grapefruit using visible to near-infrared spectroscopy and chemometrics, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* 209 (2019) 62–69.
- [18] D. Devianti, S. Sufardi, Z. Zulfahrizal, A.A. Munawar, Rapid and simultaneous detection of hazardous heavy metals contamination in agricultural soil using infrared reflectance spectroscopy, *IOP Conf. Ser. Mater. Sci. Eng.* 506 (1) (Apr. 2019).
- [19] Zulfahrizal Darusman, Y. Yunus, A.A. Munawar, Soil quality assessment by near infrared spectroscopy: predicting pH and soil organic carbon, *Int. J. Sci. Technol. Res.* 8 (10) (2019) 2512–2516.
- [20] C. Hue, et al., Near infrared spectroscopy as a new tool to determine cocoa fermentation levels through ammonia nitrogen quantification, *Food Chem.* 148 (2014) 240–245.
- [21] N.N. Batista, D.P. de Andrade, C.L. Ramos, D.R. Dias, R.F. Schwan, Antioxidant capacity of cocoa beans and chocolate assessed by FTIR, *Food Res. Int.* 90 (2016) 313–319.
- [22] N. Caporaso, M.B. Whitworth, M.S. Fowler, I.D. Fisk, Hyperspectral imaging for non-destructive prediction of fermentation index, polyphenol content and antioxidant activity in single cocoa beans, *Food Chem.* 258 (2018) 343–351. November 2017.
- [23] M.A. Quelal-Vásconez, M.J. Lerma-García, É. Pérez-Estevé, A. Arnau-Bonachera, J.M. Barat, P. Talens, Fast detection of cocoa shell in cocoa powders by near infrared spectroscopy and multivariate analysis, *Food Contr.* 99 (2019) 68–72. November 2018.
- [24] E. Teye, X. Huang, H. Dai, Q. Chen, Rapid differentiation of Ghana cocoa beans by FT-NIR spectroscopy coupled with multivariate classification, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* 114 (2013) 183–189.
- [25] J. Moros, F.A. Inón, S. Garrigues, M. de la Guardia, Near-infrared diffuse reflectance spectroscopy and neural networks for measuring nutritional parameters in chocolate samples, *Anal. Chim. Acta* 584 (1) (2007) 215–222.
- [26] A. Veselá, A.S. Barros, A. Synytsya, I. Delgadillo, J. Čopíková, M.A. Coimbra, Infrared spectroscopy and outer product analysis for quantification of fat, nitrogen, and moisture of cocoa powder, *Anal. Chim. Acta* 601 (1) (2007) 77–86.
- [27] R.N. D'Souza, S. Grimbs, B. Behrends, H. Bernaert, M.S. Ullrich, N. Kuhnert, Origin-based polyphenolic fingerprinting of Theobroma cacao in unfermented and fermented beans, *Food Res. Int.* 99 (2017) 550–559. March.
- [28] A.A. Munawar, Kusumiyati, D. Wahyuni, Near infrared spectroscopic data for rapid and simultaneous prediction of quality attributes in intact mango fruits, *Data Br.* 27 (Dec. 2019).
- [29] Y. Hu, et al., Determination of antioxidant capacity and phenolic content of chocolate by attenuated total reflectance-Fourier transformed-infrared spectroscopy, *Food Chem.* 202 (2016) 254–261.
- [30] D.N. de Oliveira, A.C.B. Camargo, C.F.O.R. Melo, R.R. Catharino, A fast semi-quantitative screening for cocoa content in chocolates using MALDI-MSI, *Food Res. Int.* 103 (2018) 8–11. July 2017.
- [31] Yusmanizar, et al., Fast and non-destructive prediction of moisture content and chologenic acid of intact coffee beans using near infrared reflectance spectroscopy, *IOP Conf. Ser. Mater. Sci. Eng.* 506 (1) (Apr. 2019).
- [32] F.Y.H. Kutsanedzie, Q. Chen, M.M. Hassan, M. Yang, H. Sun, M.H. Rahman, Near infrared system coupled chemometric algorithms for enumeration of total fungi count in cocoa beans neat solution, *Food Chem.* 240 (2018) 231–238. May 2017.
- [33] A.A. Munawar, D.V. Hörsten, D. Mörlein, E. Pawelzik, J.K. Wegener, "Rapid and non-destructive prediction of mango sweetness and acidity using near infrared spectroscopy," in *Lecture Notes in Informatics (LNI)*, Proc. - Series Gesellschaft für Informatik (GI) P-211 (2013) 219–222.
- [34] A.A. Munawar, D. Devianti, P. Satriyo, Y. Yunus, Rapid and simultaneous prediction of soil quality attributes using near infrared technology, *Int. J. Sci. Technol. Res.* 8 (9) (2019) 9–12 [Online]. Available: <http://www.ijstr.org/paper-references.php?ref=LJSTR-0919-22461>.
- [35] F. Comino, V. Aranda, R. García-Ruiz, M.J. Ayora-Cañada, A. Domínguez-Vidal, Infrared spectroscopy as a tool for the assessment of soil biological quality in agricultural soils under contrasting management practices, *Ecol. Indicac.* 87 (2018) 117–126. January.
- [36] H. Cen, Y. He, Theory and application of near infrared reflectance spectroscopy in determination of food quality, *Trends Food Sci. Technol.* 18 (2) (2007) 72–83.
- [37] B.M. Nicolai, et al., Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review, *Postharvest Biol. Technol.* 46 (2) (2007) 99–118.