Near Infrared Technology as a Fast and Non-destructive Method for the Prediction of Quality Parameters in Intact Coffee Beans

Yusmanizar[1], Imas Siti Setiasih[2], Sarifah Nurjanah[2], Mimin Muhaemin[2], Bambang Nurhadi[2], Santi Rosniawaty[2], Agus Arip Munawar[1]*

[1]Department of Agricultural Engineering, Syiah Kuala University, Banda Aceh – Indonesia
[2]Padjadjaran University, Jatinangor, Bandung – Indonesia
*Corresponding author: aamunawar@unsyiah.ac.id

Abstract— The main purpose of this present study is to apply the near infrared technology in predicting and determining two main quality attributes namely moisture content (MC) and caffeine of intact coffee bean samples. This study also investigated the impact of spectra enhancement method to the prediction accuracy and robustness. Near infrared spectral data in form of diffuse reflectance spectrum were acquired in wavelength range from 1000 to 2500 nm with co-added of 64 scans per acquisition. Spectral data were corrected and enhanced using two different methods: standard normal variate (SNV) and de-trending (DT). Prediction models, used to predict MC and caffeine content of intsct coffee beans were developed using partial least square regression (PLSR) approach. Prediction accuracy and robustness were evaluated using statistical indicators namely correlation coefficient (r) and residual predictive deviation (RPD) index. The results showed that both quality parameters can be predicted satisfactory with maximum r coefficient and RPD index: 0.98 and 4.21 for MC prediction whilst 0.93 and 3.57 for caffeine prediction respectively. Obtained results also noted that spectra enhancement using SNV and DT can improve overall prediction performance. Thus, it may conclude that infrared technology can be applied as a fast and non-destructive method to determine quality parameters of intact coffee beans.

Index Terms— Infrared, NIRS, prediction, coffee, quality.

---

1 Introduction

TIn general, coffee is one of the most popular agricultural products for people around the world and traded immensely in international markets [1]. It become popular beverages among people due to its flavor, taste, and several benefits values. Coffee can be consumed and served as daily beverages, material additives and processed to other form of consumed products [2], [3]. Moreover, consumers willing to pay an affordable price as long as they obtained coffee with good qualities. They need to be ensured that coffee which are purchased were consistently high quality coffee products. Therefore, quality control in coffee, especially in raw materials coffee beans plays a major important role in modern coffee industry [4].

Moisture contents (MC) and caffeine content (CC) are two important inner quality materials of raw coffee materials that are contributed to the overall quality of consumed coffee products [5], [6]. Caffeine has been tested and proved affected an important role in human health, as well as contribute to the flavor, acidity and bitterness [7]. On the other hand, moisture content is considered as one of the main quality standards for raw green coffee beans requested by the importing countries [8]. Based on literatures, the safety range for moisture comtent in coffee beans is 8.0 to 12.5% [5], [8]. Moisture content below 8% may cause shrunken and unwanted appearance in raw coffee beans [3], [7], whereas moisture content above 12.5% generates fungal growth and facilitates mycotoxin production [3], [8]. As we knew that raw coffee beans contain a wide range of different chemical propereties, concentrations and chemical compounds, which react and interact at all stages of coffee roasting, resulting in greatly diverse final coffee products [9], [10]. Therefore, it is necessary and crucial to determine chemical contents and othe related coffee beans parameters, so that processed coffee will provide a high quality and best taste of consumed coffee.

In order to determine moisture content, caffeine and other-coffee quality parameters, several methods have been already used and employed. However, most of those methods are based on solvent extraction followed with other standard laboratory procedures [9], [11]. In general, they are time consuming, destructive, laborious, complicated sample preparations and may cause pollutions [12], [13]. Therefore, it is not suitable for the coffee products industry whose needed fast, simple processing, and real-time quality measurements of raw coffee beans and processed coffee products.

More than two decades, huge efforts were conducted by researchers to find an alternative fast and robust method that can be used to predict and determine several quality parameters of foods and agricultural products with high speed of analysis, simple sample preparation, non-destructive in nature and without involving chemical materials [3], [14], [15]. Near infrared reflectance spectroscopy (NIRS) is among of those methods which gain more attentions and widely employed for the analysis of foods and agricultural products.

The NIRS works based on the principles of interaction between infrared radiation with biological object. Near infrared radiation covers the range of the electromagnetic spectrum from 780 to 2500 nm [16]. In NIRS, the biological samples are irradiated with near infrared radiation, and the diffuse reflected or transmitted radiation is acquired. When the infrared radiation penetrates onto the object, their electromagnetic spectral properties and characteristics changes along near infrared wavelength. Those changes depend on the chemical composition of the studied ssamples [17], [18].

Numerous studies have been performed and reported regarding with the potential ability and the application of NIRS method in determining main quality attributes of foods and agricultural products, such as: cocoa beans and chocolate products [19], [20], meat and dairy products [21], mangos and oranges [22]–[24]. The increasing numbers of research and publications and also fact that now many industries implementing NIRS method shows that this non-destructive method is obviously important.

Based on that point of view, this study attempted to apply NIRS method to predict and determine two main quality parameters (moisture content and caffeine) of intact coffee beans samples. This paper also attemopted to investigate the impact of spectra data enhancement to the prediction accuracy and robustness.

2 Materials and Methods

2.1 Coffee bean samples

Coffee samples (Arabica and Robusta) were collected from four different geographical origins in Indonesia: North Sumatera, Aceh, West Kalimantan, Bali and West Java provinces. A total of 98 bulk of coffee bean samples, containing 50 g of intact coffee beans were used in this study. Before measurement and analysis, coffee bean samples were stored for two days in room temperature (26^{o}C) to equilibrate [25].

2.2 Infrared spectral data acquisition

Infrared spectral data, in form of diffuse reflectance spectrum were acquired and recorded using FT-IR instrument Nicolet Antarais TM II (Thermo Scientific, Waltham, USA) equipped with an integrating sphere. Spectra data of each coffee bean sample was collected and acquired in near infrared wavelength range from 1000 to 2500 nm. Those data were consisted of an average of 32 scans and recorded in the integrated local computer [4].

2.3 Reference moisture and caffeine content measurements

After spectral data acquisitions were completed, coffee bean samples were taken to the laboratory for actual reference quality parameters namely moisture content (MC) and caffeine content (CC). Moisture content was determined using hermogravimetry method according to ISO 6673 [3]. Around 15 g of intact green coffee bean samples were placed in an open petri dishes glass with diameter 14 cm and 2.3 cm height. Heraeus Oven (Heraeus GmbH, Germany) was used to measure the MC of coffee samples. A forced air at 105 ± 1 ^{o}C was transferred for about 16 h. On the other hand, actual caffeine content (CC) was measured and determined using the High Performance Liquid Chromatography (HPLC). Both reference MC and CC were measured in triplicate and averaged.

2.4 Spectra data enhancement

Diffuse reflectance spectra data of coffee samples were corrected and enhanced using two spectra enhancement approaches namely standard normal variate (SNV) and detrending (DT). Both approaches have particular characteristics to minimize and eliminate noises due to light scattering and amplification [4]. Enhanced spectra data were then used to predict moisture and caffeine of intact coffee bean samples by establishing prediction models.

2.5 Moisture and caffeine prediction models

The core part in NIRS method is to develop prediction models that can be used to predict moisture content (MC) and caffeine content (CC) of coffee bean samples. Prediction models were developed using partial least square regression (PLSR) by regressing spectral data as X variable and reference MC and CC as Y variables. To achieve accurate and robust prediction results, validation was performed by means of cross validation approach [26], [27]. Prediction performance were quantified and evaluated for their accuracies and robustness using these following statistical parameters: the coefficient of determination (R^2), coefficient of correlation (r), the root mean square error (RMSE), and the residual predictive deviation (RPD) index [16], [28]. The maximum latent variables (LVs) required by the PLSR algorithms were set to 10 LVs and the prediction results were compared.

2.6 Prediction performance evaluation

Prediction accuracy and robustness were evaluated based on mentioned statistical parameters. It is obvious that excellent prediction performance desired to have high value of R^2 and r (closes to 1), lower value of RMSE and high value of RPD index. Based on literatures, RPD ≤ 1.5 is categorized as a coarse prediction model, while 1.5 < RPD < 2 is categorized as sufficient NIRS prediction model, yet still need some improvements. Furthermore, 2 ≤ RPD ≤ 3 and RPD > 3 are categorized as good and excellent prediction models respectively [16], [29]–[31].

3 Results and Discussion

3.1 Infraed spectra features

Spectra features of intact coffee bean sample in near infrared wavelength region (1000 - 2500 nm) is shown in Fig.1. This spectrum indicates the presence of several quality parameters of coffee bean samples. Spectra peak and valley as resulted from the interaction between infrared radiation and coffee bean samples. They indicate the absorbance bands resulted

• *Yusmanizar is currently pursuing doctoral degree program in Padjadjaran University, Bandung - Indonesia, PH: +6281220743572. E-mail: yusmanizar@unsyiah.ac.id*

from the molecular vibration of certain organic bonds such as O-H, C-H, C-O, C-C, C-O-H and N-O with the near infrared radiation.

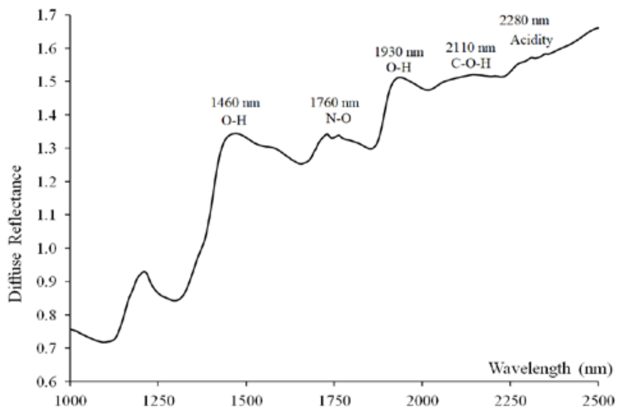


Fig. 1. Spectra features of intact coffee bean sample in near infrared wavelength region.

When the light goes through coffee bean samples, there are several reactions given by the samples. Some of light are reflected, absorbed and transmitted. The different reactions depend on main chemical composition, cell structure and other chemical properties of the samples. As shown in Fig. 1, O-H bands clearly absorbed in wavelength 1460 and 1930 nm. This mean that moisture contents of coffee bean samples can be predicted that respective wavelengths area. Similar findings also noted that O-H bands, with related to moisture contents can be determined obviously in those wavelength regions [25], [27}.

Moreover, acidities and caffeine of intact green bean samples are observed in wavelength area between 2285 nm and 2345 nm. Moreover, C-H-O and N-O bands absorbed more NIR radiation in wavelength range of 2120 nm and 1760 nm respectively. We may argue that caffeine concentrations can be predicted in those wavelength region. Beside two mentioned quality parameters, surely there are some other chemical constituents of coffee beans that can be determined. Nonetheless, this present study aimed on the prediction of moisture and caffeine contents from which related to the consumer and market preferences.

3.2 Moisture content prediction

Firstly, coffee quality parameter in form of moisture content was predicted by establishing prediction models by means of partial least square regsreesion (PLSR) approach. Spectral data used to establish these models are original unenhanced spectra known as raw spectrum, and enhanced spectra data using standard normal variate (SNV) and detrending (DT) methods. Moisture contents prediction performance using those three different spectra data is shown in Table 1.

As described in Table 1, moisture content of coffee bean samples can be predicted very well even using raw un-treated spectra data. The coefficient of correlation between predicted MC and actual reference MC is 0.982 and robustness index (RPD) of prediction model is 3.482 which categorized as excellent prediction performance. Scatter plot derived from prediction model using raw spectra data for moisture content is shown in Fig.2a.

TABLE 1
PREDICTION PERFORMANCES FOR MOISTURE CONTENT PARAMETER USING DIFFERENT SPECTRAL DATA

Spectral data	R^2	r	RMSE	RPD
Raw	0.905	0.951	0.013	3.244
SNV	0.968	0.984	0.007	4.211
DT	0.965	0.982	0.008	4.192

DT: de-trending, R^2: coefficient of determination, r: coefficient of correlation, RMSE: root mean square error, RPD: residual predictive deviation index, SNV: standard normal variate.

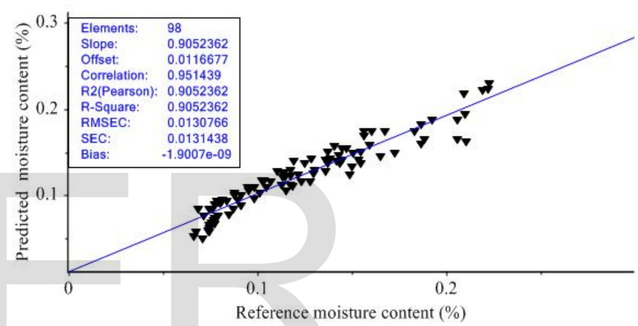


Fig. 2a. Scatter plot between predicted and reference moisture content using raw spectral data and partial least square regression.

Prediction performance for moisture content was slightly increased and improved when the model was constructed using enhanced spectra data by means of SNV method (Table 1). The correlation coefficient for MC prediction was improved to 0.984 whilst prediction error become lower to 0.007 respectively. Scatter plot derived from predicted moisture content and the accrual reference using SNV approach is shown in Fig.2b.

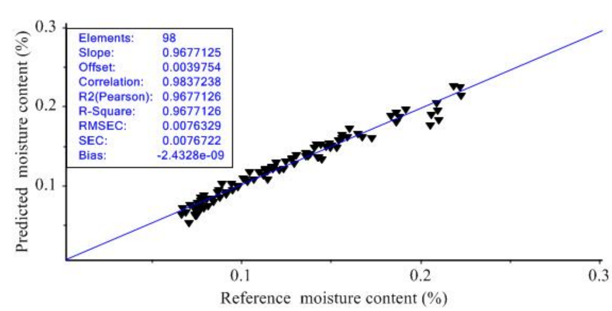


Fig. 2b. Scatter plot between predicted and reference moisture content using SNV spectral data and partial least square regres-

As a result, the RPD index was also increased to 3.594. This resulst indicates that spectra correction can significanlt improve prediction performance of studied quality parameters. Moreover, prediction model for moisture content was also developed using another spectra enhancement method namely de-trending (DT) as presented in Fig. 2c.

In general, the DT spectral data also can improve prediction performance comared to raw spectra data. However, superior prediction result can be obtained when the model is constructed using SNV approach.

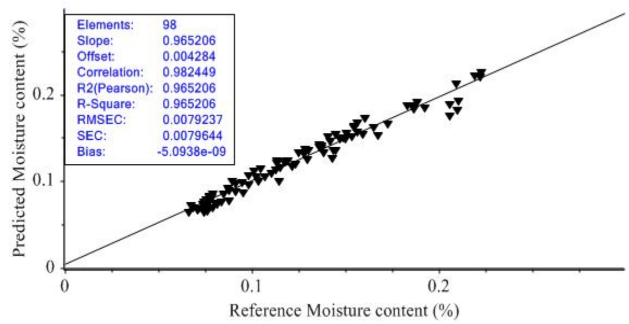


Fig. 2c. Scatter plot between predicted and reference moisture content using DT spectral data and partial least square regression.

Similar findings also noted by other researchers that spectral data enhancement can improve and increase prediction accuracy and robustness. Proper spectra enhacement and regression approaches can significantly improved prediction performacnes of coffee bean samples and other agricultural products.

3.3 Caffeine content prediction

Beside moisture content, caffeine content is also one the most important coffee quality parameter considered in coffe industry. Prediction models for caffeine content were also constructed using three different spectral data; raw, SNV and DT spectrum, and established by regressing those spectral data with actual caffeine content obtained by standatd laboratory procedures. Prediction model performances of these three spectral data is presented in Table 2.

As shown in Table 2, caffeine content can also be predicted using the NIRS method by means of raw spectrum with coeffient correlation is 863 and robustness RPD index is 2.816. This resuly confirmed that even using raw un-enhanced spectra data, caffeine content of intact coffee bean samples can be determined satisfactorz using NIRS technology. Scatter plot derived from PLSR regression approach for caffeine prediction using raw spectrum is shown in Fig.3a.

Like moisture content, prediction models for caffeine content is developed using enhanced spectral data (SNV and DT approaches). As also shown in Fig.3a, prediction performance of enhanced spectrum, both SNV and DT, can improve prediction accuracy and robustness.

TABLE 2
PREDICTION PERFORMANCES FOR CAFFEINE CONTENT PARAMETER USING DIFFERENT SPECTRAL DATA

Spectral data	R^2	r	RMSE	RPD
Raw	0.745	0.863	0.290	2.816
SNV	0.863	0.929	0.213	3.571
DT	0.840	0.917	0.229	3.128

DT: de-trending, R^2: coefficient of determination, r: coefficient of correlation, RMSE: root mean square error, RPD: residual predictive deviation index, SNV: standard normal variate.

Correlation coefficient between predicted caffeine and measured reference caffeine was increased to 0.917 when prediction model is constructed using DT spectral data. Moreover, excellent result is achieved when caffeine prediction model is established using SNV spectral data. Caffeine content can be determined with correlation coefficient between predicted and actual reference is 0.93 and robustness RPD index is 3.571.

This performance clearly shown that SNV correction method provide most accurate and robust prediction results for both coffee quality parameters. Scatter plot derived from PLSR approach for caffeine prediction using enhanced spectra data are shown in Fig.3b and Fig.3c for DT and SNV approaches respectively.

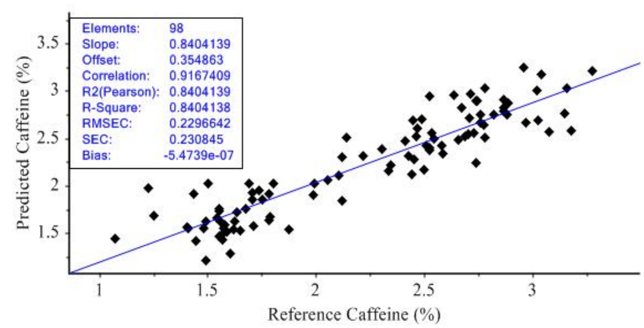
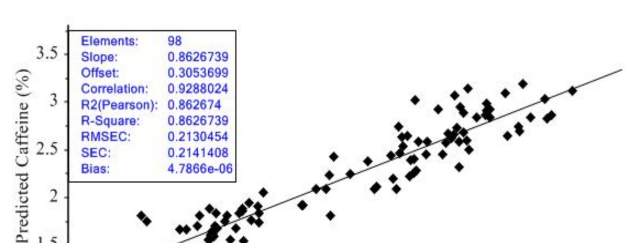


Fig. 3b. Scatter plot between predicted and reference caffeine content using DT spectral data and partial least square regression.



Judging from the prediction performances, it is obvious that prediction accuracy and robustness of moisture content and caffeine content are highly affected by the spectral enhancement and correction. Obtained results from this present study clearly demonstrated that enhanced spectral data provided and generated better prediction results than raw unenhanced spectral data. Thus, we may argue that it is important to pre-process and enhance spectral data prior to predin models development.

4 Conclussion

Based on obtained results, we may conclude that near infrared technology was able to predict and determine inner quality paramters in form of moisture content caffeine and of intact coffee bean samples simultaneously and rapidly. Maximum correlation obtained for moisture content prediction is 0.98 while for caffeine prediction is 0.93 respectively. Moreover, spectra enhancement was significantly improved prediction accuracy and robustness for both quality parameters. Standard normal variate (SNV) found to be the best spectra enhancement method to predict moisture and caffeine content of intact coffee beans.

Acknowledgment

The authors wish to thank Kmeristek DIKTI for providing research funding through doctoral scholarship.

References

[1] J. McNutt and Q. (Sophia) He, "Spent coffee grounds: A review on current utilization," *J. Ind. Eng. Chem.*, 2018.

[2] J. Hu, X. Ma, L. Liu, Y. Wu, and J. Ouyang, "Rapid evaluation of the quality of chestnuts using near-infrared reflectance spectroscopy," *Food Chem.*, vol. 231, pp. 141–147, 2017.

[3] F. Mohammed, D. Guillaume, S. Dowman, and N. Abdulwali, "An easy way to discriminate Yemeni against Ethiopian coffee," *Microchem. J.*, vol. 145, no. June 2018, pp. 173–179, 2019.

[4] 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.*, vol. 506, p. 012033, 2019.

[5] C. Ciaramelli, A. Palmioli, and C. Airoldi, "Coffee variety, origin and extraction procedure: Implications for coffee beneficial effects on human health," *Food Chem.*, vol. 278, no. November 2018, pp. 47–55, 2019.

[6] A. H. Abdullah, Z. Ismail, A. S. Zainal Abidin, and K. Yusoh, "Green sonochemical synthesis of few-layer graphene in instant coffee," *Mater. Chem. Phys.*, vol. 222, no. July 2018, pp. 11–19, 2019.

[7] P. Veiga-Santos, L. T. Silva, C. O. de Souza, J. R. da Silva, E. C. C. Albuquerque, and J. I. Druzian, "Coffee-cocoa additives for bio-based antioxidant packaging," *Food Packag. Shelf Life*, vol. 18, no. August, pp. 37–41, 2018.

[8] J. R. Santos, O. Viegas, R. N. M. J. Páscoa, I. M. P. L. V. O. Ferreira, A. O. S. S. Rangel, and J. A. Lopes, "In-line monitoring of the coffee roasting process with near infrared spectroscopy: Measurement of sucrose and colour," *Food Chem.*, vol. 208, pp. 103–110, 2016.

[9] C. C. Couto *et al.*, "Coffea arabica and C. canephora discrimination in roasted and ground coffee from reference material candidates by real-time PCR," *Food Res. Int.*, vol. 115, no. August 2018, pp. 227–233, 2018.

[10] M. Defernez *et al.*, "Low-field1H NMR spectroscopy for distinguishing between arabica and robusta ground roast coffees," *Food Chem.*, vol. 216, pp. 106–113, 2017.

[11] M. Ventura, J. R. Silva, L. H. C. Andrade, R. P. Scorza Júnior, and S. M. Lima, "Near-near-infrared thermal lens spectroscopy to assess overtones and combination bands of sulfentrazone pesticide," *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.*, vol. 188, pp. 32–36, 2018.

[12] A. Giraudo, S. Grassi, F. Savorani, G. Gavoci, E. Casiraghi, and F. Geobaldo, "Determination of the geographical origin of green coffee beans using NIR spectroscopy and multivariate data analysis," *Food Control*, vol. 99, no. December 2018, pp. 137–145, 2018.

[13] B. G. Botelho, L. S. Oliveira, and A. S. Franca, "Fluorescence spectroscopy as tool for the geographical discrimination of coffees produced in different regions of Minas Gerais State in Brazil," *Food Control*, vol. 77, pp. 25–31, 2017.

[14] Y. B. Monakhova *et al.*, "Rapid approach to identify the presence of Arabica and Robusta species in coffee using 1H NMR spectroscopy," *Food Chem.*, vol. 182, pp. 178–184, 2015.

[15] T. V. Silva, D. M. B. P. Milori, J. A. G. Neto, E. J. Ferreira, and E. C. Ferreira, "Prediction of black, immature and sour defective beans in coffee blends by using Laser-Induced Breakdown Spectroscopy," *Food Chem.*, vol. 278, no. November 2018, pp. 223–227, 2019.

[16] B. M. Nicolaï *et al.*, "Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review," *Postharvest Biol. Technol.*, vol. 46, no. 2, pp. 99–118, 2007.

[17] F. Comino, V. Aranda, R. García-Ruiz, M. J. Ayora-Cañada, and A. Domínguez-Vidal, "Infrared spectroscopy as a tool for the assessment of soil biological quality in agricultural soils under contrasting management practices," *Ecol. Indic.*, vol. 87, no. January, pp. 117–126, 2018.

[18] H. Cen and Y. He, "Theory and application of near infrared reflectance spectroscopy in determination of food quality," *Trends Food Sci. Technol.*, vol. 18, no. 2, pp. 72–83, 2007.

[19] K. Slettengren, P. Heunemann, O. Knuchel, and E. J. Windhab, "Mixing quality of powder-liquid mixtures studied by near infrared spectroscopy and colorimetry," *Powder Technol.*, vol. 278, pp. 130–137, 2015.

[20] E. Teye *et al.*, "Estimating cocoa bean parameters by FT-NIRS and chemometrics analysis," *Food Chem.*, vol. 176, pp. 403–410, 2015.

[21] E. Wüst and L. Rudzik, "The use of infrared spectroscopy in the dairy industry," *J. Mol. Struct.*, vol. 661–662, no. 1–3, pp. 291–298, 2003.

[22] A. A. Munawar, D. von Hörsten, J. K. Wegener, E. Pawelzik, and D. Mörlein, "Rapid and non-destructive prediction of mango quality attributes using Fourier transform near infrared spectroscopy and chemometrics," *Eng. Agric. Environ. Food*, vol. 9, no. 3, pp. 208–215, 2016.

[23] P. Rungpichayapichet, B. Mahayothee, P. Khuwijitjaru, M. Nagle, and J. Müller, "Non-destructive determination of β-carotene content in mango by near-infrared spectroscopy compared with colorimetric measurements," *J. Food Compos. Anal.*, vol. 38, pp. 32–41, 2015.

[24] K. Ncama, U. L. Opara, S. Z. Tesfay, O. A. Fawole, and L. S. Magwaza, "Application of Vis/NIR spectroscopy for predicting sweetness and flavour parameters of 'Valencia' orange (Citrus sinensis) and 'Star Ruby' grapefruit (Citrus x paradisi Macfad)," *J. Food Eng.*, vol. 193, pp. 86–94, 2017.

[25] J. R. Santos, M. C. Sarraguça, A. O. S. S. Rangel, and J. A. Lopes, "Evaluation of green coffee beans quality using near infrared spectroscopy: A quantitative approach," *Food Chem.*, vol. 135, no. 3, pp. 1828–1835, 2012.

[26] V. Belchior, B. G. Botelho, L. S. Oliveira, and A. S. Franca, "Attenuated Total Reflectance Fourier Transform Spectroscopy (ATR-FTIR) and chemometrics for discrimination of espresso coffees with different sensory characteristics," *Food Chem.*, vol. 273, no. August 2017, pp. 178–185, 2019.

[27] D. Krakowian *et al.*, "Application of EPR spectroscopy to the examination of pro-oxidant activity of coffee," *Food Chem.*, vol. 151, pp. 110–119, 2014.

[28] A. A. Munawar, D. V. Hörsten, D. Mörlein, E. Pawelzik, and J. K. Wegener, "Rapid and non-destructive prediction of mango sweetness and acidity using near infrared spectroscopy," in *Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft fur Informatik (GI)*, 2013, vol. P-211.

[29] A. A. Munawar, D. von Hörsten, J. K. Wegener, E. Pawelzik, and D. Mörlein, "Rapid and non-destructive prediction of mango quality attributes using Fourier transform near infrared spectroscopy and chemometrics," *Eng. Agric. Environ. Food*, vol. 9, no. 3, 2016.

[30] T. Nordey, J. Joas, F. Davrieux, M. Chillet, and M. Léchaudel, "Robust NIRS models for non-destructive prediction of mango internal quality," *Sci. Hortic. (Amsterdam).*, vol. 216, pp. 51–57, 2017.

[31] E. J. N. Marques, S. T. De Freitas, M. F. Pimentel, and C. Pasquini, "Rapid and non-destructive determination of quality parameters in the 'Tommy Atkins' mango using a novel handheld near infrared spectrometer," *Food Chem.*, vol. 197, pp. 1207–1214, 2016.