talenta PUBLISHER

Journal of Sustainable Agriculture and Biosystems Engineering

Journal homepage: https://talenta.usu.ac.id/jsabe

SMLR Method to Obtain the Representative NIR Wavelength in Determining the CGA and Trigonelline Content in Coffee Beans

Putri Chandra Ayu[1*], I Wayan Budiastra[2,3], Sutrisno[2,3], Sukrisno Widyotomo[4]

[1]*Study Program of Agricultural and Biosystem Engineering, Faculty of Agriculture, Universitas Sumatera Utara, Medan, Indonesia.*
[2]*Department of Mechanical and Biosystem Engineering, Faculty of Agricultural Engineering and Technology, IPB University, Bogor, Indonesia.*
[3]*Center for Research on Engineering Application in Tropical Agriculture (CREATA), IPB University, Bogor, Indonesia.*
[4]*Indonesian Coffee and Cocoa Research Institute, Jember, Indonesia*

*Corresponding Author: pputricandra@usu.ac.id

ARTICLE INFO

Article history:
Received 28 May 2024
Revised 31 August 2025
Accepted 31 August 2025
Available online

E-ISSN : 3026-4065

How to cite :
P C Ayu, I W Budiastra, Sutrisno, S Widyotomo, "SMLR Method to Obtain the Representative NIR Wavelength in Determining the CGA and Trigonelline Content in Coffee Beans," Journal of Sustainable Agriculture and Biosystems Engineering, vol. 3, no.2, August 2025.

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International.

DOI: 10.32734/jsabe.v3i02.16549

ABSTRACT

Coffee from diverse place will serve a diverse flavor, especially the accidity and bitternes, could be easily detected by human's sense of taste. This is due to the differences of the chemical content, in this case the chlorogenic acid (CGA) and trigonelline, inside the green beans. In coffee industry, the assesment of coffee quality usually done by the sertified quality grader (Q-grader). However, the Q-grader could not provide the exact concentration percentage inside and the information only could be determined using the chemical method such as the LCMS and HPLC method. These method are expensive, time consuming and destructive the opposite with the need of industries. Near Infrared Spectroscopy (NIRS) method could be used to answer the requirement. In this study, the CGA and trigonelline concentration in coffee beans could be determined by several most representative NIR wavelenght. A hundred samples (96 grams coffee beans) were subjected to NIR measurement in the wavelength of 1000 – 2500 nm, followed by the chemical analysis by LCMS method, then calibrated by the Stepwise multiple linear regression (SMLR) method to obtain the representative NIR wavelength. Result showed that the NIR spectroscopy could be used to determine the CGA concentration by using 24 NIR wavelenghts indicated by high r, RPD and consistency value of 0.949, 2.59 and 88.78%, respectively. However, the trigonelline could not be assesed by the SMLR method due to its small amount and its chemical bounding inside the beans.

Keywords: CGA, Coffee Bean, SMLR, Trigonelline

1. Introduction

Coffee quality become the most important aspect in coffee industries that could be assesed by the bean color and size, percentage of damaged seeds, and flavor. In coffee industry, the evaluation is done by the sertified quality grader (Q-grader). Besides that, the chemical content of coffee bean is actually needed to be known, since it contains more than 700 compounds [1], which are responsible and used to form the coffee flavor and quality. The concentration of chemical content in coffee beans is diverse according to its origin and also the postharvest processing. Chemical content of coffee bean reflects the coffee flavor, also there are several compounds that can exert beneficial biological activities for human health, such as the caffeine and chlorogenic acids

(CGA) [2], with concentration of 0.78-1.55% (w/w) and 3.30-7.65% (w/w) [3], respectively. Besides that, there is trigonelline as the second abundant alkaloid in coffee bean, range from 0.62 to 1.27% (w/w) [3], which also contributes to the overall bitterness in roasted coffee [4]. CGA could exhibit a radical scavenging effect that is similar to ascorbic acid, can chelate transition metals such as Fe^{2+} to scavenge free radicals and disrupt chain reaction. Therefore, the CGA content can act as antioxidant and beneficial for health [5].

The determination of chemical content could be measured by the chemical method, such as HPLC and LCMS method [6]. This method offers accuracy; however not with the cost, time and its character that need a smaller size of samples, that is not suitable for industries. Therefore, a new method that could solve those problems is needed, one of which is by using the near infrared spectroscopy (NIRS) method that could assess the chemical concentration inside the coffee beans [6-16]. NIRS is a technology that is for the most part based on the actual number of molecules of individual compounds. Chemically, all of the organic material, including coffee beans, consists of atoms, mainly carbon, oxygen, hydrogen, nitrogen, phosporus and sulphur, with minor amounts of other elements. These atoms combine to form molecules, and because of the nature of bonds, and the electrostatic charges on the atoms and molecules themselves, the molecules are constantly in motion and vibrate in frequencies corresponding to wavelengths in the infared region or the electromagnetic spectrum. This is the reason NIRS could be used to assess the chemical information inside material since it affords a method to translate those vibrations of molecules into simple, rapid and environmentally friendly analytical procedures [7, 17].

Coffee could have a different flavor, especially the accidity and bitternes, which can be detected by human's sense of taste. This is due to the differences of the chemical content inside the green beans [3,12], in this case the chlorogenic acid (CGA) and trigonelline, for accidity and bitterness, respectively. NIR study combine with chemical data and chemometrics method has been widely used in determining the chemical concentration in green bean coffee [6-15]. Beside that, the use of machine learning to enhance the NIR model performance has also expanded recently [16, 18-19]. The calibration of both chemical and NIR data could be done by several method such as partial least square (PLS), principle component analysis (PCA), artificial neural network (ANN), stepwise multiple linear regression (SMLR) method, etc. The PLS model of CGA and trigonelline concentration have been developed and showed a good model prediction indicated by the high value of coefficient correlation in both components of 0.936 and 0.975, respectively [6]. However, the actual representative wavelength for CGA and trigonelline in coffee beans have not been obtained, which can be used to build a faster detection and sortation system for coffee industry. SMLR method could solve the problem, since this method could reduce the wavelength numbers which also could reduce the running time compare to PLS [20]. In this study, the representative wavelength for both CGA and trigonelline is assesed.

2. Methods

A hundred samples of Arabica Java Preanger coffee bean were prepared to be subjected to NIR measurement (96 grams per sample) using NIRFlex N-500 Merk Buchi and chemical measurement (0.2 grams per each sample) by LCMS. After that, followed by the calibration method using SMLR method that conducted using unscrambler sofware v.10.3 (CAMO, Norway).

3. Results and Discussion

Figure 1 show the NIR absorbance of Java Preanger green beans coffee, the absorbance pattern in this case is quite similar to the others such as Lintong, Mandailing and Sidikalang bean. The difference could be seen at the amount of the absorbance [7-10]. Development of NIR model by PLS method for trigonelline and CGA have been conducted actually and the result showed a good model with high r value of 0.975 and 0.936, respectively [6]. However, to reduce the amount of wavelength, another data processing is needed to obtain the most representative wavelength. In fact, range of wavelength from 1000 to 2500 nm not only contain variables that include informations about the component but also contain noise or variable that is not contain information that could affect the strength of prediction result and could build an overfitting regresion model [20-21]. Therefore, selection of wavelength could improve the performance of the model with no preprocessing data needed.

Table 1 explain the concentration of trigonelline and CGA content in arabica Java Preanger coffee beans. The concentration is in accordance to the percentage range of both components that has been reported [3]. It is also shown that both of the components have abundant percentage in coffee beans; therefore, they can be well predicted using all NIR wavelength combined with the PLS method [6]. This percentage will absolutely affect the selection of the representative NIR wavelength as shown in Table 1.

.

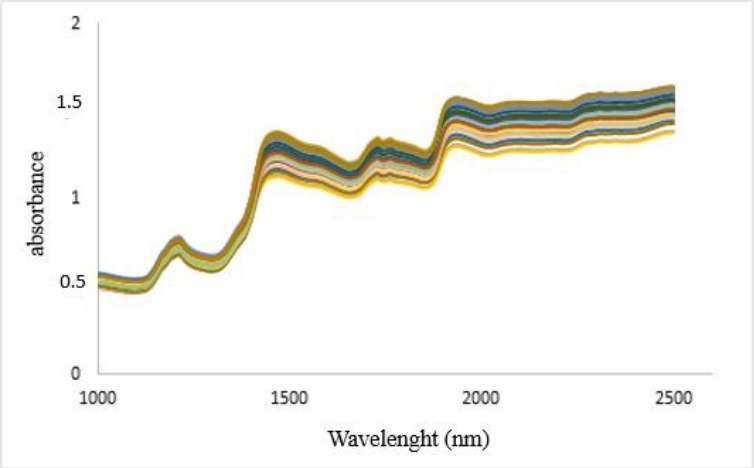

Figure 1. NIR absorbance for coffee beans

Table 1 shows the result of calibration by SMLR method for CGA and trigonelline content. This method will select the wavelengths that exhibit the best single term calibration as the first independent variable, then finds the best wavelength to add as a second variable in a two terms regression, and so on until some stopping criterias are met [17]. Moreover, in SMLR method the data pretreatment process is not required. Results showed that by using 24 selected NIR wavelength, the CGA content can be well predicted, this indicated by a high coefficient correlation (r) value and standard error calibration (SEC) of 0.949 and 0.129, respectively (Figure 2). Besides that, this model also showed a high consistency of 88.78% and high residual prediction deviation (RPD) of 2.59. Result of r value shows that this model indicated a strong positive linear relationship between both of NIR data predictions and references [22]. In addition, the obtained RPD value also indicated an excellent model since it was higher than 2 [23,24].

Table 1. Result of calibration by SMLR method.

Parameter	Concentration (%)	Selected wavelength (nm)	r	SEC (%)	SEP (%)	CV (%)	RPD	Consistency (%)
Trigonelline	0.87 – 1.14	-	-	-	-	-	-	-
CGA	5.42 – 6.89	7496, 7556, 7820, 8505, 8540, 8796, 8812, 8900, 8908, 8956, 9088, 9148, 9196, 9224, 9272, 9380, 9408, 9460, 9492, 9272, 9380, 9408, 9460, 9492, 9724, 9380, 9408, 9460, 9492, 9724, 9744, 9780, 9804, 9808	0.949	0.129	0.145	2.41	2.59	88.78

On the other hand, the wavelength selection for trigonelline can not be done (Table 1), while in PLS method which use a large range of wavelength from 1000 – 2500 nm, the model determination for the trigonelline in Java Preanger coffee has been obtained with r value of 0.975 [6]. Furthermore, it is shown that the number of the selected wavelengths also determined by the observed concentration percentage and the bond relation with the other chemical components in the molecules [25], which caused not all of the chemical content can be determined by SMLR model. In fact, wavelength from 1000 to 2000 nm could contain variables that bring information or not, and could influence and lead the prediction result to an overfitting model [26]. Therefore, wavelength selection is a method used to increase the model performance with no data pretreatment needed.

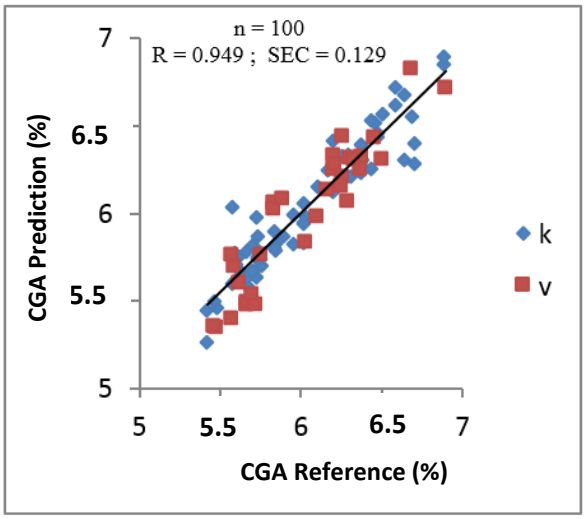

Figure 2. CGA prediction using 24 selected NIR wavelength
(k is calibration data; v is validation data)

The selected 24 wavelength is a combination of the CGA and several wavelengths that accommodate the scatter and intercorrelation effects between chemical components in the material [25]. As in [26], the MLR method has been used to examine the roasting degree of coffee beans, which resulted 5 selected wavelength of water content of 1940, 1725, 1225, 1111, and 1450 nm. The wavelength of 1940 and and 1225 are indicated as water content wavelength [26], and the other 3 wavelengths are used to accommodate the scattering influence and the relation between chemical compounds inside the material.

4. Conclusion

Result showed that by using 24 NIR wavelengths, the CGA content in coffee beans could be well predicted by NIR spectroscopy indicated by the high value of r, RPD and consistency of 0.949, 2.59 and 88.78%, respectively. However, the trigonelline content can not be determined using SMLR method.

References

[1] M. Jeszka-Skowron, A. Sentkowska, K. Pyrzynska, M. P. De Pena, “Chlorogenic acids, caffeine content and antioxindant properties of green coffee extracts: influence of green coffee bean preparation,” *Eur Food Res Technol,* vol. 242, pp. 1403-1409, 2016.

[2] R. J. Buldak, T. Hejmo, M. Osowski, L. Buldak, M. Kukla, R. Polaniak, E. Birkner, “The impact of coffee and its selected bioactive compounds on the development and progression of colorectal cancer in vivo and in vitro,” *Molecules,* vol.23, no. 3309, 2018.

[3] K. Wale, K. Tolessa, M. Atlabachew, B. Mehari, M. Alemayehu, D. A. Mengistu, B. Kerisew, “Level of caffeine, trigonelline and chlorogenic acids in green coffee (*Coffea arabica* L.) beans from Amhara region, Ethiopia,” *Journal of Agriculture and Food Research,* vol. 16, pp. 1-6, 2024.

[4] H. Ashihara, W. Deng, C. Nagai, "Trigonelline biosynthesis and the pyridine nucleotide cycle in coffea arabica fruit: metabolic fate of [*carboxyl*-^{14}C]nicotinic acid riboside," *Phytochemistry Letters,* vol. 4, no. 3, pp 235-239, 2011.

[5] A. Rojas-Gonzales, C. Y. Figueroa-Hernandez, O. Gonzales-Rios, M. L. Suarez-Quiros, R. M. Gonzales-Amaro, Z. J. Hernandez-Estrada, P. Rayas-Duarte, "Coffee chlorogenic acids incorporation for bioactivity enhancement of foods: A review," *molecules,* vol 27, no. 11, 2022.

[6] I.W. Budiastra, M. Sutrisno, S. Widyotomo, P.C. Ayu, "Determination of trigonelline and chlorogenic acid (CGA) concentraton in intact coffe beans by NIR spectroscopy," *CIGR Journal*, vol. 22, no.1, 2020.

[7] P. C. Ayu, N. U. W. Sebayang, K. N. Sinamo, "Environmentally friendly method in determining the water content of coffee by NIR spectroscopy,*" IOP Conf. Ser.: Earth Environ. Sci*, 2020. vol. 1241, pp 1-4, 2024.

[8] P.C. Ayu, I. W. Budiastra, A. Rindang, "NIR spectroscopy application for determination caffeine content of arabica green bean coffee," *IOP Conf. Ser.: Earth Environ. Sci*, vol. 454, pp 1-4, 2020.

[9] A. Rindang, P.C. Ayu, Sahfitri, "NIR spectroscopy as an environmentally friendly method to determine the lipid and carbohydrate concentration in Lintong Arabica coffe bean," *IOP Conf. Ser.: Earth Envirno. Sci*, 2022.

[10] S. D. Siregar, A. Rindang, P.C Ayu, "Principle component analysis (PCA) classification of Arabica green bean coffee of North Sumatera using FT-NIRS," *IOP Conf. Ser.: Earth Environ. Sci*, 2020.

[11] Andasuryani, A. A. Munawar, Kusumiyati, Adrizal, Yusmanizar, "Non-invasive NIR spectroscopy for precise water content determiantion in Sumatran coffee beans," *International Journal of Design & Natrue and Ecodynamics,* vol. 19, No. 5, pp 1785-1791, 2024.

[12] L. Munyendo, K Schuster, W. Armbruster, M. Babor, D. Njoroge, Y. Zhang, A. von Wrochem, A. Schaum, B. Hitzmann, "Monitoring a coffee roasting process based on Near-Infrared and Raman Spectroscopy coupled with chemometrics," *Journal of Chemometrics,*pp 1-14, 2025.

[13] S. Green, E. Fanning, J Sim, G. T. Eyres, R. Frew, B. Kebede, "The potential of NIR spectroscopy and chemometrics to discriminate roast degrees and predict volatiles in coffee," *Molecules,* vol. 29, no. 318, pp 1-13, 2024.

[14] A. Guerro-Pena, L. Vazquez-Hernandez, A. Bucio-Galindo, V. Morales-Ramos, "Chemical analysis and NIR spectroscopy in the determination of the origin, variety and roast time of Mexican coffee," *Heliyon,* vol. 9, pp 1-9, 2023.

[15] S. Grassi, A. Giraudo, C. Novara, N. Cavalini, F. Geobaldo, E. Casiraghi, F. Savorani, "Monitoring chemical changes of coffee beans during roasting using real-time NIR spectroscopy and chemometrics," *Food Analytical Methods,* vol 16, pp. 947-960, 2023.

[16] J. Sim, C. McGoverin, I. Oey, R. Frew, B. Kebede, "Non-linier machine learning coupled near infrared spectroscopy enhanced model performance and insights for coffee origin traceability," *Journal of Near Infrared Spectroscopy,* vol. 32(4-5), pp. 127-139, 2024.

[17] P. William and K. Norris, *Near Infrared Technology in The Agricultural and Food Industries.* St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc., 1987.

[18] S. Veerasakulwat, A. Sitorus, V. Udompetaikul, "Rapid classification of sugarcane nodes and internodes using near-infrared spectroscopy and machine learning techniques," *Sensors,* vol. 24, no. 7102, pp 1-18, 2024.

[19] A. Sitorus, R. Lapcharoensuk, "Exploring deep learning to predict coconut milk adulteration using FT-NIR and Micro-NIR spectroscopy," *Sensors,* vol. 24, no. 2362, pp 1-23, 2024.

[20] I. W. Budiastra, Sutrisno, S. Widyotomo, P. C. Ayu, "Prediction of caffeine content in Java Preanger coffee beans by NIR Spectroscopy using PLS and MLR method," *IOP Conf. Ser.: Earth Environ. Sci*, 2018.

[21] J. Shan, T. Suzuki, Y. Ogawa and N. Kondo, "Coffee roasting degrees prediction in terms of weight loss with selected wavebands based on near infrared spectroscopy" *Engineering in Agriculture, Environment and Food*, vol. 8, pp 195-199, 2015.

[22] B Ratner, "The correlation coefficient: Its values range between +1/-1, or do they ?," *Journal of Targetting, Measurement and Analysis for Marketing,* vol 17, pp 139-142, 2009.
[23] K. Ncama, U.L. Opara, S.Z. Tesfay, O.A. Fawole, 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 Engineering.* Vol. 193, pp. 86-94, 2017.
[24] N. Caporaso, M.B. Whitworth, S. Grebby, I.D. Fisk, "Rapid prediction of single green coffee bean moisture and lipid content by hyperspectral imaging," *Journal of Food Engineering,* vol 227, pp 18-29, 2018.
[25] I.W. Budiastra, Y. Ikeda, T. Nishizu, "Prediction of individual sugars and malica acid concentration of apples and mangoes by developed NIR reflectance system," *Journal of JSAM,* vol. 60, no. 3, pp 117-127, 1998.
[26] J. Shan, T. Suzuki, Y. Ogawa, N. Kondo, "Coffee roasting degrees prediction in terms of weight loss with selected wavebands based on near infrared spectroscopy," *Engineering in Agriculture,* vol. 8, pp. 195-199, 2015.