Determination of trigonelline and chlorogenic acid (CGA) concentration in intact coffee beans by NIR spectroscopy

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

(1. Department of Mechanical and Biosystem Engineering, Bogor Agricultural University, Bogor, Indonesia;
2. Indonesian Coffee and Cocoa Research Institute, Jember, Indonesia;
3. Study Program of Agricultural Engineering, Universitas Sumatera Utara, Medan, Indonesia)

Abstract: Trigonelline and chlorogenic acid (CGA) are important quality indicators of coffee. Commonly, trigonelline and CGA concentration are determined using chemical method. This method is time consuming and destructive so it is not suitable for coffee industries which need a fast measurement. The objective of this study was to assess near infrared spectroscopy for predicting trigonelline and CGA concentration in intact coffee beans. Coffee beans samples of 96 g (n=100) were placed in a petri dish. The reflectance of samples was measured by fourier transform near infrared spectrometer in the wavelengths of 1000-2500 nm. Subsequently, the trigonelline and CGA content of samples were determined using Liquid Chromatography Mass Spectrometry (LCMS). Spectra data processing such as first and second derivative, multiple scatter correction (MSC), standard normal variate (SNV) and the combination of them were carried out to reduce scattering, to eliminate overlapped absorption bands, and to optimize the best data input in calibration process. After that, these spectra (n=67) were calibrated to chemical data using partial least square (PLS) to find the best calibration models. Then the calibration models were applied to predict trigonelline and CGA in another set of samples (n=33). The results showed that NIR spectra data processing of second derivative combined with four factors of PLS was the best model for predicting CGA concentration of coffee (r=0.94, coefficient of variation=2.75%, residual predictive deviation=2.27). For trigonelline, however, the best model was the combination of second derivative and MSC of spectra data processing and four factors of PLS (r=0.98, CV=1.63%, RPD=2.98). These results indicated that NIR spectroscopy could be used as a fast and nondestructive method for determining trigonelline and CGA in intact coffee beans accurately.

Keywords: chlorogenic acid, intact coffee bean, NIRS, PLS, trigonelline

Citation: Budiastra, I. W., M. Sutrisno, S. Widyotomo, and P. C. Ayu. 2020. Determination of trigonelline and chlorogenic acid (CGA) concentration in intact coffee beans by NIR spectroscopy. Agricultural Engineering International: CIGR Journal, 22(1): 162-168.

1 Introduction

In coffee market, ensuring good quality and customer satisfaction is the most important thing for the coffee business continuation. There are some functional quality characteristics of coffee which effect on its flavor; they are trigonelline, caffeine, chlorogenic acid (CGA), water content, carbohydrate and protein. Trigonelline is one of alkaloid compound present in coffee. In fact, during thermal degradation in roasting process, trigonelline generates the non-volatile alkylpiridiniums which contributes to antioxidative activity and aroma of coffee beans. Besides that, trigonelline have more medicinal properties such as antimigraine, anticarcinogenic and antidiabetic activities (Farah, 2012). On the other hand, CGA in coffee bean not only contributes to the acidity, quality and acceptance of final cup of coffee, but also plays an important role in preventing various diseases such as cancer, aging and cardiovascular (Belay and Gholap, 2009). By knowing the concentration of

---

Received date: 2017-08-22 **Accepted date:** 2019-11-13

***Corresponding author: I Wayan Budiastra,** Senior Lecturer, Departement of Mechanical and Biosystem Engineering, Bogor Agricultural University, IPB Campus Darmaga, PO Box 220 Bogor 16002, Indonesia. Email : wbudiastra@yahoo.com, Tel/Fax: 622518623026.

trigonelline and CGA in coffee beans, we can predict about 40% of coffee flavor after brewed.

At present, Liquid Chromatography Mass Spectrometry (LCMS) is used for the quantification of coffee concentration such as trigonelline and CGA. However, this method has number of limitations such as complicated sample preparation, destructive, high cost, and time consuming, so it is not appropriate for coffee industries which need a rapid measurement. Therefore, an alternative method should be developed to substitute LCMS method, one of the potential methods is near infrared (NIR) spectroscopy.

To date, near infrared spectroscopy (NIRS) has been successfully used for evaluating the defect in intact green coffee beans (Santos et al., 2012), predicting the concentration of caffeine, theobromine and theophylline in ground green coffee bean (Huck et al., 2005), predicting caffeine concentration in ground roasted coffee (Zhang et al., 2013) and predicting caffeine in intact coffee bean (Rosita et al., 2016). Moreover, NIRS has already applied in quantifying caffeine, lipid, sucrose, trigonelline and CGA concentration in ground coffee beans (Davrieux et al., 2008), quantifying CGA in ground roasted coffee beans (Shan et al., 2014), in line monitoring of the coffee roasting process based on sucrose content and coffee color (Santos et al., 2016), and prediction of coffee roasting degrees based on water content (Shan et al., 2015). However, there is no research on the application of NIR spectroscopy for the prediction of trigonelline and CGA concentration in intact coffee beans. Thus, the objective of this research was to assess NIR spectroscopy for determining trigonelline and CGA concentration in intact coffee beans.

2 Materials and methods

2.1 Sample preparation

A batch of samples of Arabica Java Preanger green coffee bean (water content 11% – 12%) from Indonesian Coffee and Cocoa Research Institute (ICCRI) were used for this experiment. Samples of intact coffee bean of 96 g were poured and arranged into four layers in a petri dish to comply diffuse reflectance measurement (Rosita, 2016). One hundred set of samples were prepared for testing NIR spectroscopy.

2.2 NIR measurement

The reflectance spectra of samples were measured by fourier transform near infrared Spectrometer type NIRFlex N-500 (BUCHI Labortechnic AG Switzerland). Each spectra was acquired in diffuse reflectance mode with a 4 cm^{-1} interval using wavenumber from 10000 cm^{-1} to 4000 cm^{-1} (1000 – 2500 nm) and a scan speed of 3 scans s^{-1}. NIR spectra measurement was conducted at temperature around $22^{o}C – 25^{o}C$.

2.3 Chemical analysis

After NIR measurement, subsequently, 10 g out of 96 g was grounded using grinder, and then the ground coffee bean were used for trigonelline and CGA measurement using LCMS procedure (Shimadzu LC-MS 2020).

Firstly, 2 ppm of trigonelline and CGA standard respectively was injected into the LCMS. The condition in the LCMS were: C-18 waters as column (T $40^{o}C$), mobile phase were divided into two, that was 40% water/formic acid 0.1% and 60% acetonitrile, as mobile phase A and B respectively, and the flow rate was 0.2 mL min^{-1}. Subsequently, the peak of each concentration was obtained at any given time. After that, 0.2 g of each coffee sample was injected into LCMS in the same condition. After the similar time with the standard was obtained, the peak of each concentration will be obtained. The concentration of each component can be calculated from the area and the height of peak, using Equation 1 and 2.

$$\text{Concentration (\%)} = C_s \times \frac{100}{1000} \times \frac{100}{W_s} \times C_f \qquad (1)$$

$$C_f = \frac{100}{100 - \text{water content}} \qquad (2)$$

Where,

C_s = Concentration of component in sample (ppm)

W_s = Weight of sample injected (mg)

C_f = Correction factor

Water content in Equation 2 was determined using gravimetric method (AOAC, 2000).

2.4 Data processing and analysis

Samples were divided into 67 samples for calibration and 33 samples for validation based on equally chemical data distribution. The spectra data, trigonelline and CGA concentration variables were input into the Unscramble software v10.3 (CAMO, Norway) to carry out spectra data processing, calibration and validation process. Several spectra data processing method such as first and second derivative (dg1 and dg2), multiple scatter correction (MSC), standard normal variate (SNV) and combination of them were conducted to develop the best data input for calibration model of trigonelline and CGA. Subsequently, partial least square (PLS) method was used to build calibration models. Cross validation was used to test the calibration model's performance. The calibration model performance was evaluated based on their correlation coefficient (r), standard error (SE), residual predictive deviation (RPD), and coefficient of variation (CV) (Andasuryani et al., 2013).

3 Results and discussions

3.1 Trigonelline and CGA concentration in Java Preanger coffee

Trigonelline and CGA concentration in Java Preanger green coffee bean were around 0, 9% and 6%, respectively (Table 1).

Table 1 Trigonelline and CGA concentration of coffee beans sample

Parameter	Concentration (%)	Mean (%)	Standard deviation (%)
Trigonelline	0.87 - 1.14	0.96	0.059
CGA	5.42 - 6.89	6.04	0.403

This result is similar with Clarke and Macrae (1985) that the concentration of trigonelline and CGA in Arabica green coffee bean were around 1% – 1.2% and 5.5%-8% respectively. Trigonelline concentration was much lower than CGA, indicating that determination of trigonelline in coffee bean may be more difficult than CGA.

3.2 NIR absorbance spectra

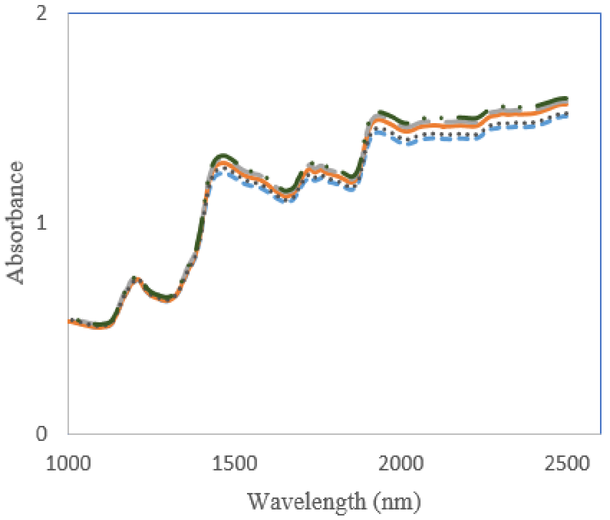
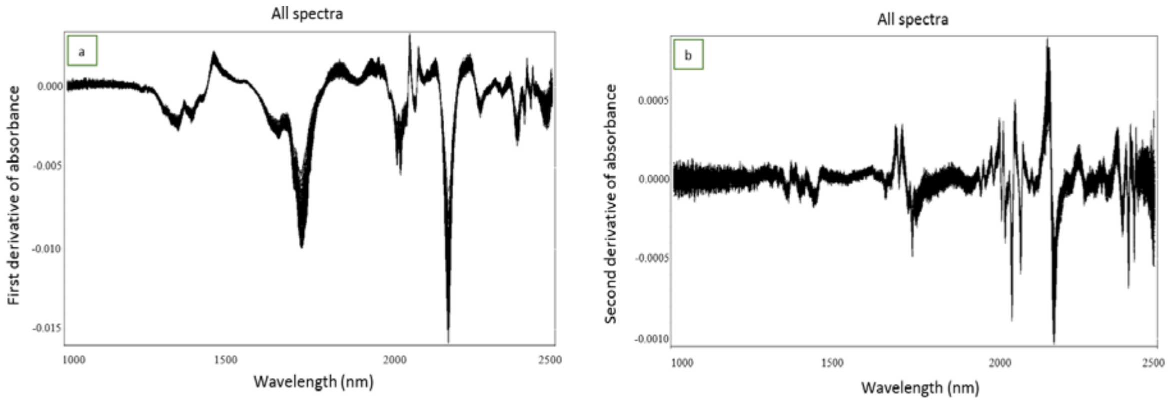


Figure 1 NIR absorbance spectra of intact coffee bean

Figure 1 shows the original absorbance spectra of intact coffee bean of Arabica Java Preanger. The shape of spectra was similar. The difference in absorbance values was high in the wavelength of 1500-1900 nm, and the highest in the wavelength of 2000-2500 nm. It means that in this region, there is a high possibility of NIR method for determining chemical content of coffee. The spectra shape is influenced by particle size, chemical component concentration and wavelengths. The water absorption is indicated by the peak of spectra at wavelengths of 1450 and 1940 nm. The CGA absorption is indicated by the peak at the wavelength of 1477 nm. However, the peak of trigonelline 1128 nm did not appear. This may be caused by a low concentration of trigonelline in coffee bean or influence of other major components (water, carbohydrate etc.). Thus the spectra data processing such as derivative may be required.

3.3 Results of spectra data processing



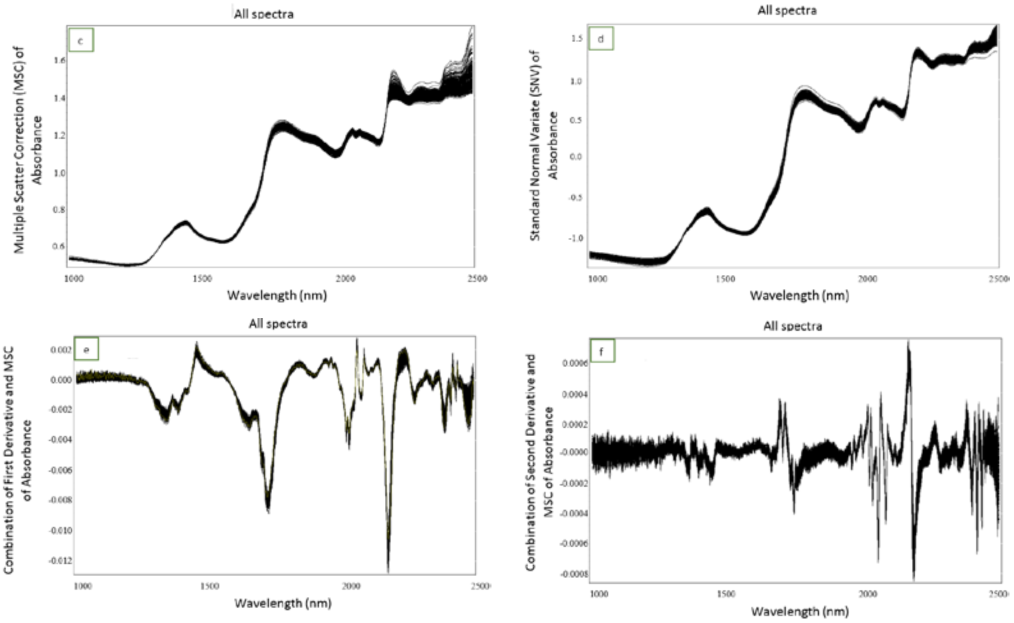


Figure 2 Result of several spectra data processing (a. first derivative of absorbance; b. second derivative of absorbance; c. MSC of absorbance; d. SNV of absorbance; e. combination of first derivative and MSC of absorbance; f. combination of second derivative and MSC of absorbance)

In fact, the first derivative data processing produces different shape of spectra from the original spectra (Figure 2a). The first derivative data processing produces more peaks and valleys than the original spectra due to its ability to separate the overlapped absorption bands. The second derivative data processing (Figure 2b) produces more peaks than the first derivative, indicating that more chemical components can be identified, for instance, the peak of trigonelline appears (1128 nm) and the peak of CGA be sharpen (1477; 1726; 1934; 2128 nm) (Ribeiro et al., 2011). Beside overcoming the overlapped absorption bands, the first and second derivative data processing could remove the linear baseline variation (Hruschka, 1990). MSC data processing reduces the scattering effect in the wavelength of 1000 – 1800 nm (Figure 2c). It means that MSC may be useful for predicting the chemical composition of coffee beans using this region. In contrast, SNV can reduce the scattering effect in all regions of wavelengths (Figure 2d). MSC and SNV data processing could remove the multiplicative interference effect in the data which caused by scattering, particle size, path length, and the air cavities/gap between coffee beans. The difference of MSC and SNV data processing is in the determination of standard while centering the data in calculation. In SNV, value of each spectrum is centered to zero, while in MSC, all data is centered to the mean value (Arnita and Sutarman, 2011). The combination of first or second derivative and MSC (Figure 2e, 2f) produced similar spectra with that in the first or second derivative, but the values slightly decreased due to the effect of MSC that corrected scattering factor of the samples (intact coffee bean).

3.4 Result of calibration and validation

The best calibration and validation result of trigonelline concentration is shown in Table 2. For trigonelline, the best calibration model was found in combination of 2nd derivative and MSC data processing and four factors of PLS, indicated by the highest correlation of coefficient (0.98), highest RPD (2.98) and the lowest coefficient of variation (1.63%). This result is much better than the results (r = 0.87) from Davrieux et al. (2008) in predicting trigonelline concentration in coffee powder.

The highest accuracy of combination of second derivative and MSC in prediction of trigonelline may be caused as explained follows. The concentration of trigonelline in green coffee bean is small, so that by using 2nd derivative, it will enhance the spectra resolution (Stuart, 2004) and bring out the peak of trigonelline. Moreover, MSC method could improve the predictive ability by correcting the additive and multiplicative effects in the spectra (Abu-Khalaf et al., 2004), thus, the best calibration model was obtained. In Table 3, the data processing of second derivative (number 3) and the combination of second derivative and MSC (number 4) are having the same R value. But after applying MSC, the RPD of number 4 increased and the CV decreased. MSC could reduce error due to the scattering effect of particle size or air gap between coffee beans. The more data processing applied, then number of required PLS factors will decrease and the accuracy increase. If more PLS factors being applied, the model will be over fitting. Selecting number of PLS factors is important to remove noise and obtained optimum model is affected by number of PLS factors. (Chen et al., 2013).

Table 2 Results of calibration and validation for prediction of trigonelline concentration

No	Processing	Number of PLS factors	R	SEC (%)	SEP (%)	CV	RPD	Consistency (%)
1	Original	7	0.63	0.0496	0.0504	5.25	0.92	98.41
2	1st Derivative	6	0.96	0.0171	0.0175	1.82	2.66	97.71
3	2nd Derivative	4	0.98	0.0142	0.0185	1.93	2.51	76.76
4	2nd Derivative, MSC	4	0.98	0.0143	0.0156	1.63	2.98	91.67
5	MSC, 1st Derivative	6	0.97	0.0156	0.0171	1.78	2.72	91.23
6	MSC, 2nd Derivative	4	0.98	0.0141	0.0162	1.69	2.87	87.04

The best calibration and validation result for CGA concentration is shown in Table 3. For CGA, the best calibration model was found by only using 2nd derivative data processing and 4 factors of PLS, indicated by the high correlation coefficient (0.94), high RPD (2.27) and the low coefficient of variation (2.75%). This result is better than the Davrieux et al. (2008) and Shan et al. (2014) which showed that the r value in predicting CGA in ground coffee was 0.87. It may be caused by the difference of NIR instruments and chemical data distribution.

Table 3 Results of calibration and validation for prediction of CGA concentration

No	Processing	Number of PLS Factors	R	SEC (%)	SEP (%)	CV (%)	RPD	Consistency (%)
1	Original	7	0.70	0.2935	0.3182	5.28	1.18	92.23
2	1st Derivative	5	0.90	0.1823	0.1847	3.06	2.04	98.70
3	2nd Derivative	4	0.94	0.1454	0.1660	2.75	2.27	87.59
4	2nd Derivative, MSC	3	0.92	0.1649	0.1762	2.92	2.13	93.59
5	MSC, 1st Derivative	3	0.82	0.2373	0.2260	3.75	1.66	105
6	MSC, 2nd Derivative	3	0.92	0.1599	0.1748	2.90	2.15	91.48

The second derivative data processing can increase the accuracy of NIRS in predicting CGA concentration. The combination of 2nd derivative and MSC doesn’t give a better model than the single data processing (2nd derivative) indicated by the decreased of R and RPD. This might be caused by the concentration of CGA that it is large enough and the spectra was not affected by the scatter interference. In contrary, Shan et al. (2014) found that the combination of MSC and SNV was the best data processing to determine CGA in roasted coffee beans. This is caused by the sample which was in powder form (600 µm), that the water influence was relieved, so derivative data processing did not required. Moreover, the bigger particle size then the higher absorbance will be obtained. That is the reason in this research that MSC as processing data did not require.

Figure 3 shows the scatter plot of reference vs. prediction values of CGA concentration obtained from the best PLS model using 2nd derivative. Value of r was 0.94, CV was 2.75% and RPD was 2.27. Figure 4 shows the scatter plot of reference vs. prediction values of trigonelline concentration from the best PLS model in combination with 2nd derivative and MSC processing data. Value of r was 0.98, CV was 1.63% and RPD was

2.98. These results indicate that the calibration model can predict trigonelline and CGA concentration in intact coffee beans accurately.

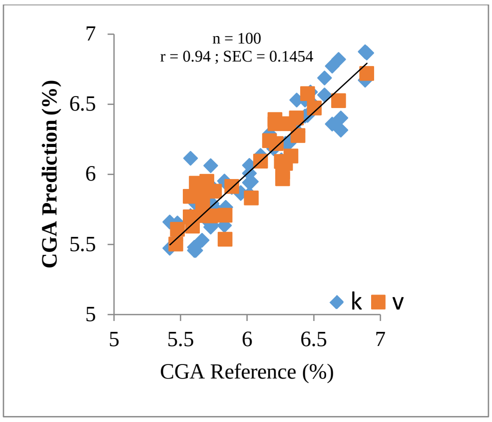


Figure 3 Plots of CGA reference vs. prediction (k = calibration data; v = validation data)

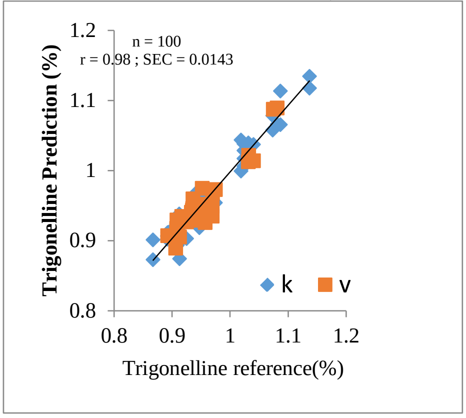


Figure 4 Plots of trigonelline reference vs. prediction (k = calibration data; v = validation data)

4 Conclusions

The second derivative of absorbance spectra' data processing combined with 4 factors of PLS was the best model for predicting CGA (r=0.94, CV=2.75%, RPD=2.27). For trigonelline, the best model was the combination of second derivative and MSC of absorbance spectra' data processing and 4 factors of PLS (r=0.98, CV=1.63%, RPD=2.98). The results indicate that NIR spectroscopy can be used as a fast and nondestructive method for determining trigonelline and CGA in intact coffee beans.

Acknowledgement

We would like to thanks the Directorate of Research, Ministry of Research, Technology and Higher Education for supporting this research (STRANAS Research Grant of FY 2016-2017), and Indonesian Coffee and Cocoa Research Institute (ICCRI) for its cooperation and supports.

References

AOAC. 2000. *Official Method of Analysis of The Association of Official Analytical of Chemist*. 17th edition. Gaithersburg, MD, USA: The Association of Official Analytical Chemist, Inc.

Abu-Khalaf, N., B. S. Bennedsen, and G. K. Bjørn. 2004. Distinguishing carrot's characteristic by near infrared (NIR) reflectance and multivariate data analysis. *Agricultural Engineering International: the CIGR Journal of Scientific Research and Development*, Manuscript FP 03 012, 6: 1-17.

Andasuryani, Y. A. Purwanto, I. W. Budiastra, and K. Syamsu. 2013. Determination of catechin as main bioactive component of gambir (*Uncaria gambir* Roxb) by FT-NIR Spectroscopy. *Journal of Medicinal Plant Research*, 7(41): 3076-3083.

Arnita, A., and S. Sutarman, 2011. Comparing the multiplicative scatter correction (MSC) and standard normal variate (SNV) methods in calibration model of multiple variables. *Bulletin of Mathematics*, 3(1): 25-38.

Belay, A., and A. V. Gholap. 2009. Characterization and determination of chlorogenic acids (CGA) in coffee beans by UV-Vis spectroscopy. *African Journal of Pure and Applied Chemistry*, 3(11): 234-240.

Chen, H., Q. Song, G. Tang, Q. Feng, and L. Lin. 2013. The combined optimization of Savitzky-Golay smoothing and multiplicative scatter correction for FT-NIR PLS models. *ISRN Spectroscopy*, 2013: Article ID 642190.

Clarke, R. J., and R. Macrae. 1985. *Coffee Volume 1: Chemistry*. 655 Avenue of the Americas, NY 10010, USA: Elsevier Science Publishing Co, Inc.

Davrieux, F., B. Guyot, E. Tardan, and F. Descroix. 2008. Coffee arabica var. *laurina* authentication using near infrared spectroscopy. Available at: https://Agritrop.cirad.fr:430-437. Accessed 20 June 2017.

Farah, A. 2012. Coffee Constituents. In *Coffee: Emerging Health Effects and Disease Prevention*, ed. Yi-Fang Chu, ch. 2, 21-58. New York, USA: John Wiley & Sons, Inc.

Hruschka, W. R. 1990. Data analysis: wavelength selection method. In *Near Infrared Technology in the Agricultural and Food Industries*, ed. P. William and K. Norris, ch. 3, 35-53, Minnesota, USA: American Association of Cereal Chemist, Inc.

Huck, C. W., W. Guggenbichler, and G. K. Bonn. 2005. Analysis of caffeine, theobromine, theophylline in coffee by near infrared spectroscopy (NIRS) compared to high performance liquid chromatography (HPLC) coupled to mass spectrometry. *Analytica Chimica Acta*, 538(1-2): 195-203.

Ribeiro, J. S., M. M. C. Ferreira, and T. J. G. Salva. 2011. Chemometric models for the quantitative descriptive sensory analysis of Arabica coffee beverages using near infrared spectroscopy. *Talanta*, 83(5): 1352-1358.

Rosita, R. 2016. Determination of chemical composition in Arabica Gayo coffee beans nondestructively using near infrared spectroscopy. M.S. thesis, Mechanical and Biosystem Engineering Department. Bogor Agricultural University, Indonesia.

Rosita, R., I. W. Budiastra and Sutrisno. 2016. Prediction of caffeine content of Arabica coffee bean by near infrared spectroscopy. *Jurnal Keteknikan Pertanian*, 4(2): 179-186.

Santos, J. R., M. C. Sarraguça, A. O. Rangel, and J. A. Lopes. 2012. Evaluation of green coffee beans using near infrared spectroscopy: A quantitative approach. *Food Chemistry*, 135(3): 1828–1835.

Santos, J. R., O. Viegas, R. N. M. J. Páscoa, I. M. P. L. V. O. Ferreira, A. O. S. S. Rangel, and J. A. Lopes. 2016. In-line monitoring of the coffee roasting process with near infrared spectroscopy: Measurement of sucrose and color. *Food Chemistry*, 208: 103-110.

Shan, J., T. Suzuki, Y. Ogawa, and N. Kondo. 2015. Coffee roasting degrees prediction in terms of weight loss with selected wavebands based on near infrared spectroscopy. *Engineering in Agriculture, Environment and Food*, 8(4): 195-199.

Shan, J., T. Suzuki, D. Suhandy, Y. Ogawa, and N. Kondo. 2014. Chlorogenic acid (CGA) determination in roasted coffee beans by near infrared (NIR) spectroscopy. *Engineering in Agriculture, Environment and Food*, 7(4): 139-142.

Stuart, B. H. 2004. *Infrared Spectroscopy: Fundamentals and Applications*. The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England: John Wiley & Sons, Ltd.

Zhang, X., W. Li, B. Yin, W. Chen, D. P. Kelly, X. Wang, K. Zheng, and Y. Du. 2013. Improvement of near infrared spectroscopic (NIRS) analysis of caffeine in roasted Arabica coffee by variable selection method of stability competitive adaptive reweighted sampling (SCARS). *Molecular and Biomolecular Spectroscopy*, 114: 350-356.