International Food Research Journal 27(2): 287 - 294 (April 2020)

Journal homepage: *http://www.ifrj.upm.edu.my*

Near-infrared spectroscopy with linear discriminant analysis for green 'Robusta' coffee bean sorting

[1,2]*Khuwijitjaru, P., [1]Boonyapisomparn, K. and [2]Huck, C. W.

[1]*Department of Food Technology, Faculty of Engineering and Industrial Technology, Silpakorn University, Nakhon Pathom 73000, Thailand*
[2]*Institute of Analytical Chemistry and Radiochemistry, Centre for Chemistry and Biomedicine, University of Innsbruck, Innrain 80/82, 6020 Innsbruck, Austria*

Article history

Received: 13 November 2019
Received in revised form:
2 April 2020
Accepted:
4 April 2020

Keywords

NIR spectroscopy (NIRS), 'Robusta' coffee, defective beans, contaminated beans, principal component analysis (PCA), linear discriminant analysis (LDA)

Abstract

The present work investigated the feasibility of near-infrared (NIR) spectroscopy for separation of good quality green 'Robusta' coffee beans from defective (broken beans, beans with parchment, and beans with husk) and contaminated beans (faecal matter and soil) by single bean measurement. Linear discriminant analysis using principal components from principal component analysis (PCA-LDA) as variables was used as a supervised method for the classification. It was found that smoothing pre-treatment applied to the spectra was suitable for the classification, with the highest classification accuracy of 97.5%. The present work indicated that NIR spectroscopy coupled with appropriate chemometric methods could be an efficient tool for coffee bean sorting.

Introduction

Coffee is a popular drink around the world. Therefore, coffee cultivation, coffee bean processing, and coffee product manufacturing such as roasted beans, instant coffee, and packaged coffee drinks are economically important in many countries. The total world production of green coffee beans continuously increased from 7.5 million tons in 2000 to 10.3 million tons in 2018 (FAO, 2020). The quality of green coffee beans is a crucial factor determining the quality of brewed coffee drink. Therefore, they are normally sorted to remove the defective ones, including blackened, mouldy, broken, insect-damaged, and dried cherries and other foreign matter (National Bureau of Agricultural Commodity and Food Standards, 2009).

Automatic optical sorting machines are commercially available for large-scale production of coffee beans. For small-scale producers, however, sorting by hand is still being practiced due to the high cost of machines. Near-infrared (NIR) spectroscopy, which is a non-invasive analytical technique for coffee bean sorting for small-scale coffee growers, is a very promising technology since its versatility has been shown in numerous applications and the cost of equipment is recently becoming lower (Khuwijitjaru, 2018). NIR spectroscopy has been studied in different processing stages of coffee, from differentiating 'Arabica' and 'Robusta' coffee beans, identifying origin of beans, monitoring roasting process, to prediction of coffee drink quality (Downey and Boussion, 1996; Alessandrini *et al.,* 2008; Barbin *et al.,* 2014; Santos *et al.,* 2016; Okubo and Kurata, 2019; Giraudo *et al.,* 2019). Even though computer vision based on image analysis is possible to classify some coffee bean defects, NIR spectroscopy has an advantage since it can also be used for analysis of various major and minor internal chemical components of coffee beans without destroying any sample, such as moisture content, proteins, caffeine, lipids, chlorogenic acid and other phenolic compounds, total sugars, sucrose (Fox *et al.,* 2013; Scholz *et al.,* 2014), theobromine (Huck *et al.,* 2005), and ochratoxin A contamination (Taradolsirithitikul *et al.,* 2017). The detection of adulteration of barley in roasted coffee by NIR spectroscopy has been reported by Ebrahimi-Najafabadi *et al.* (2012). Although the evaluation of green 'Arabica' and 'Robusta' coffee beans with different defects including insect-damaged, broken, sour, partially black, black, mouldy, silver skin, faded, and immature beans has been investigated in the past (Santos *et al.,* 2012), the work was not based on single bean measurement.

*Corresponding author.
Email: khuwijitjaru_p@su.ac.th

The use of discriminant analysis techniques on NIR spectral data for classification has been studied for several applications. Linear discriminant analysis (LDA) is one of the supervised classification techniques that uses information from training set of data to allocate unknown new samples into correct groups (Miller and Miller, 2000). Biancolillo and Marini (2018) discussed the limitation of LDA with data from spectroscopic methods which normally contain much larger number of variables than number of samples and resulted in non-invertible covariance matrix. Several variable reduction/selection methods were therefore tested to improve LDA prediction. Brito *et al.* (2013) reported that classification of three types of cereal bars using LDA with wavelength selection by genetic algorithm gave the best result of 90% classification. Diniz *et al.* (2014) showed that only 12 wavenumbers selected by successive projections algorithm for LDA could correctly classify all 84 samples of tea into five classes. The use of principal components from principal component analysis for LDA (PCA-LDA) has also been reported by several studies which showed very promising results. Jakubíková *et al.* (2016) showed that the application of PCA-LDA on NIR spectra could classify four kinds of fruit spirits at 100% correctness. Li *et al.* (2014) also reported that PCA-LDA could predict brands, flavour styles, ages, and alcohol contents of Chinese liquor at higher than 95% correctness from Vis/NIR spectra. The classification of green coffee beans including common defects using NIR spectroscopy has scarcely been reported. In addition, NIR spectroscopy study on detection of soil and faecal contaminations in coffee beans has never been reported in the literature.

In the present work, therefore, the potential of NIR spectroscopy with the PCA-LDA method for sorting good quality green 'Robusta' coffee beans from common defective beans as well as from soil or faecal contaminated beans was investigated.

Materials and methods

Coffee beans

'Robusta' coffee cherries were harvested and processed into green beans by a local producer (Thamsing Coffee, Chumphon, Thailand). The dried beans were obtained as a gift at the producer's premises. Following receipt, the samples were sorted manually according to the criteria used by the producer into four groups, *i.e.* good beans (G), broken beans (B), beans with parchment (P), and beans with husk (H). In addition, two groups of beans with contaminants were artificially prepared in laboratory by smearing good beans with soil (S) and chicken faecal matter (F) on the dorsal side of the beans (Figure 1). Each group consisted of 200 beans.

NIR spectral acquisition

NIR spectra of coffee beans in the wavenumber range 12489 - 3895 cm^{-1} (800 - 2567 nm) were obtained using an FT-NIR spectrometer (MPA, Bruker Optik, Germany) at the resolution of 16 cm^{-1} and 32 scans. All beans were scanned on their dorsal side using reflectance mode on a custom-made sample placement for individual bean measurement. OPUS software version 7.0 (Bruker Optik) was used for controlling the instrument and recording spectra.

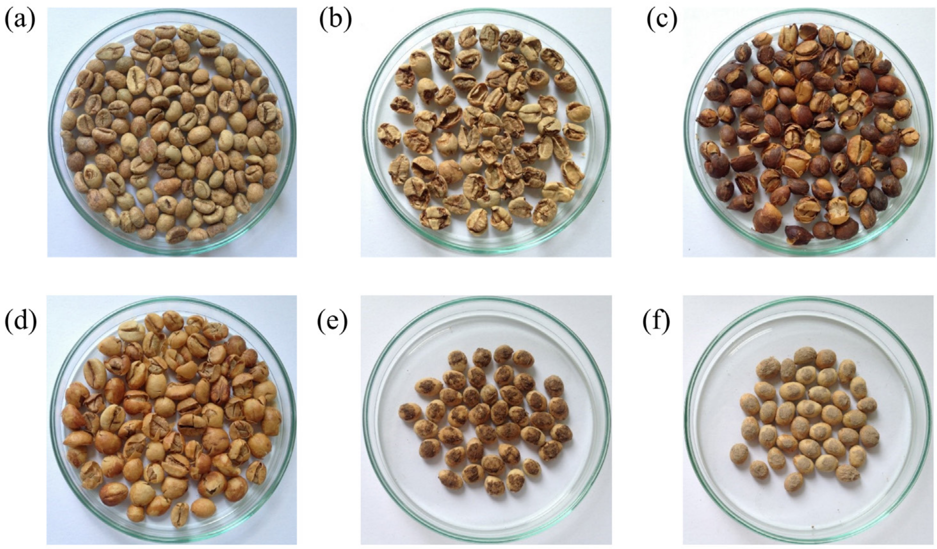

Figure 1. Pictures of green 'Robusta' coffee bean samples. Good quality beans (a), broken beans (b), beans with husk (c), beans with parchment (d), chicken faecal contaminated beans (e), and soil contaminated beans (f).

Data analysis

NIR spectra from the OPUS software were exported into comma separated value (CSV) files for data analysis using R software version 3.5.1 (R Core Team, 2019). To remove noise from meaningful data, the spectra were processed with various mathematical pre-treatment methods including Savitzky-Golay smoothing (SG, second order, 17 points), standard normal variate (SNV), first derivative (SG-1D, 17 points), and second derivative (SG-2D, 17 points) using the respective functions in the package 'prospectr' (Stevens and Ramirez-Lopez, 2013). Multiplicative scatter correction (MSC) was performed by *msc* function in the package 'pls' (Mevik *et al.,* 2016). The PCA was performed using *prcomp* function with scaled and centred data, and then LDA was analysed by *lda* function from the package 'MASS' (Venables and Ripley, 2002). Spectra from each group were separated into training set (160 spectra) and test-set (40 spectra) according to Kennard-Stone algorithm using *kenStone* function with Euclidean distance criteria in the package 'prospectr'.

Results and discussion

NIR spectra of coffee beans

Green 'Robusta' coffee bean typically contains 48.6 - 55.0% of carbohydrates, 10% of proteins, 10% of lipids, and 4.4% of minerals (Viani and Petracco, 2016). The moisture content of green coffee bean is usually lower than 12.5% due to the safety standard. Normalised and averaged spectra for each group of the coffee beans used in the present work are shown in Figure 2. From visual inspection, it could be seen that the spectrum from soil contaminated beans (S) was obviously different from the others, particularly at wavenumbers around 4750 cm^{-1} (2105 nm). Spectra in the short-wavelength region (9000 - 12500 cm^{-1} or 800 - 1111 nm), however, showed no significant band. The first distinctive band in long-wavelength region appeared around 8200 - 8300 cm^{-1} (1204 - 1219 nm) and corresponded to second overtone stretching of C-H bonds. Two large bands at 6900 cm^{-1} (1449 nm) and 5200 cm^{-1} (1923 nm) represent moisture absorption while the band at 4750 cm^{-1} (2105 nm) corresponds to band of protein (Manley, 2014). Peaks around 4200 - 4400 cm^{-1} (2272 - 2380 nm) represent absorption bands of carbohydrates (Workman and Weyer, 2012). Major bands found were similar to those for 'Arabica' and 'Robusta' beans reported by Scholz *et al.* (2014) and Giraudo *et al.* (2019). Detection of faecal contamination in food product using NIR spectroscopy has been reported by only a few studies. Chao *et al.* (2008) studied the detection of faecal matter on stainless steel and found that the featured bands of faecal contaminated sample were 980, 1195, and 1450 nm (10204, 8368, and 6896 cm^{-1}). Liu *et al.* (2007) employed visible/NIR hyperspectral imaging (450 - 950 nm) to detect faecal matter on apples, and found that the ratio between 725 and 811 nm could effectively detect faecal contamination.

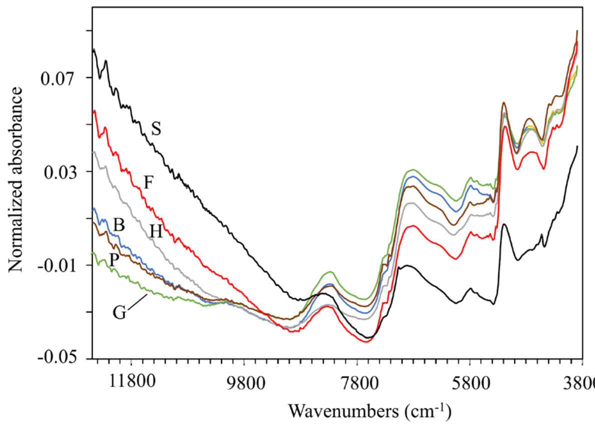

Figure 2. Vector normalised and averaged spectra of green 'Robusta' coffee bean samples for good quality beans (G), broken beans (B), beans with husk (H), beans with parchment (P), chicken faecal contaminated beans (F), and soil contaminated beans (S).

Principal component analysis

Principal component analyses (PCA) were separately performed using short- (12489 - 9025 cm^{-1} or 800 - 1108 nm), long- (9025 - 3895 cm^{-1} or 1108 - 2567 nm), and full-wavelength (12489 - 3895 cm^{-1} or 800 - 2567 nm) raw spectra to visualise all samples. Figure 3 shows the plots of the first two principal component (PC) scores obtained from each spectral range. Results showed that the explained variances by the first two PCs were up to 99, 97, and 96% for short-, long-, and full-wavelength spectra, respectively. Short-wavelength spectra gave an incomplete separation of good beans from other defective beans. Interestingly, beans with husk were obviously separated from others. Faecal and soil contaminated beans were grouped together but incompletely separated from uncontaminated beans. On the other hand, long- and full-wavelength spectra gave rather similar plots in which faecal and soil contaminated beans were well separated from other groups while good beans were closely grouped with other defective beans. This suggested that long-wavelength gave more information on chemical components from faecal matter and soil which were largely different from other coffee beans. The group of soil contaminated beans was the most isolated one from the other groups which agreed with the observation of the spectra in Figure 2. This agreed with the fact that long-wavelength NIR (1300 - 2500 nm) provides stronger absorptivity from the first overtone and combination bands but shallower penetration depth than short-wavelength region (Abbas and Baeten, 2017).

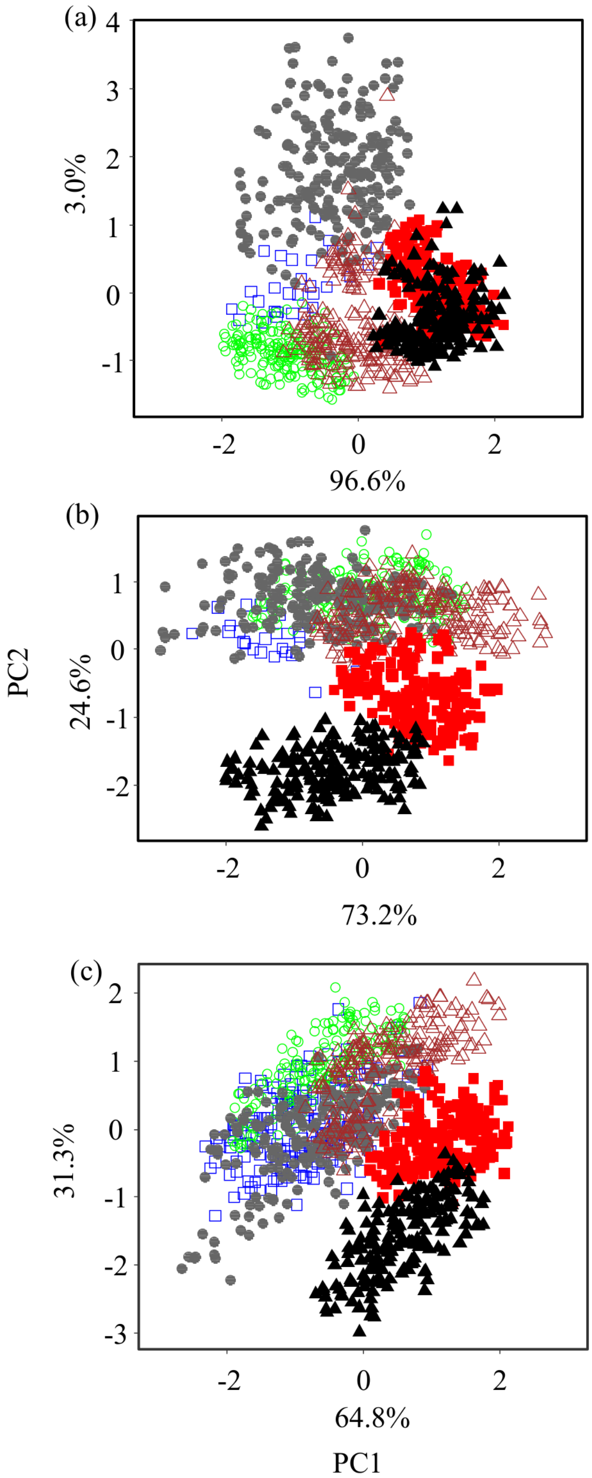

Figure 3. Score plot of the first two principal components from short- (a), long- (b), and full-wavelength (c) raw spectra for good quality beans (○), broken bean (□), beans with husk (●), bean with parchment (△), chicken faecal contaminated beans (■), and soil contaminated beans (▲).

PCA-LDA classification performance

In the present work, principal components (PCs) from PCA were used as variables for LDA. Short-, long-, and full-wavelength NIR spectra were also separately used for PCA-LDA analysis. LDA analysis was performed using up to six PCs. Because the main objective was to separate all defective beans from good beans (G), positive classification (assigning G into correct group) and false-positive classification (assigning other defective beans as G) were reported. Correctness of the classification model was evaluated with leave-one-out cross validation and test-set validation (40 beans).

Table 1 shows the classification results using short-wavelength NIR spectra. The number of PC in each model was selected from the one that gave the highest positive and lowest false-positive classification from the test-set validation. The use of raw spectra resulted in a positive classification of the test-set samples as high as 100% with six PCs, but the false-positive classification was also very high, which was not appropriate for the purpose of bean sorting. Smoothing, 1st derivative, and 2nd derivative pre-treatments adversely affected the prediction performance with lower positive classifications but higher false-positive classifications. MSC and SNV pre-treatments gave similar results. They gave positive classifications around 50% without any false-positive classification. Nevertheless, the performances were too low for real applications.

Classification results from long-wavelength region are also shown in Table 1. It was found that using six PCs from raw spectra could give a positive classification of 75% without false-positive classification. Furthermore, after smoothing, similar prediction results were obtained. However, other pre-treatments gave lower classification performances.

Full-spectra gave the best result for classification after smoothing pre-treatment with a positive classification of 97.5% without false-positive classification using six PCs. Table 2 shows that classifications of other beans were also highly effective. Figure 4 shows linear discriminant scores plot to illustrate the separation of each coffee bean group using LD1-3. It can be seen that good beans are grouped together and separated well from other beans. The first three PCs explained 99% of the spectral variation in coffee bean samples while the first six PCs explained 99.9% of the variation. PC1 explained 65% of the total variation of coffee beans spectra while PC2 explained 31%, and PC3 explained 3% of the total variation. In the present work, PCA eigenvectors were used to explain more about the separation basis of the beans (Cozzolino *et al.,* 2005).

Table 1. Good quality coffee beans classification results obtained with PCA-LDA.

Spectral region	Spectra pre-treatment	Number of PCs	Cross validation		Test-set validation	
			Positive (%)	False-positive (%)	Positive (%)	False-positive (%)
Short	None	6	87.5	21.5	100.0	30.0
	Smoothing	3	81.9	32.5	97.5	35.0
	1st derivative	2	80.6	29.4	80.5	37.5
	2nd derivative	2	22.5	107.5	28.1	94.4
	SNV	6	93.1	6.9	57.5	0
	MSC	6	93.1	6.3	52.5	0
Long	None	6	88.1	7.5	75.0	0
	Smoothing	6	89.4	8.1	75.0	0
	1st derivative	6	94.4	11.9	97.5	2.5
	2nd derivative	5	83.7	20.0	95.0	12.5
	SNV	3	81.2	30.0	82.5	17.5
	MSC	6	83.1	27.5	65.0	10.0
Full	None	3	86.9	20.6	95.0	5.0
	Smoothing	6	93.7	3.7	97.5	0
	1st derivative	6	93.7	13.1	95.0	2.5
	2nd derivative	5	83.7	21.2	92.5	15.0
	SNV	5	98.7	10.0	100.0	25.0
	MSC	2	61.9	30.6	85.0	60.0

Table 2. PCA-LDA classification of green ‘Robusta’ coffee bean samples using smoothing pre-treated full spectra with 6 PCs (assigned group in row).

	G	B	H	P	F	S	Classification (%)
Good beans (G)	39	1	0	0	0	0	97.5
Broken beans (B)	0	40	0	0	0	0	100
Beans with husk (H)	0	1	39	0	0	0	97.5
Beans with parchment (P)	0	0	0	40	0	0	100
Faecal contaminated beans (F)	0	0	0	0	40	0	100
Soil contaminated beans (S)	0	0	0	0	0	40	100

Figure 5(a) and (b) show an example of NIR spectrum after smoothing pre-treatment and the eigenvectors of the first three components, respectively. The eigenvectors of PC1 showed that the spectra from 12489 to 7500 cm^{-1} equally contributed to the coffee bean discrimination. In addition, there were downward peaks at 6900, 5200, and 4750 cm^{-1} which corresponded to peaks on NIR spectra as previously discussed. Downward eigenvectors of PC2 and 3 were found at the region around 8200 cm^{-1} while several upward eigenvectors were also found at the wavelengths associated with NIR peaks. These results indicated that these eigenvectors correlated positively or negatively with physical characteristics and chemical components of coffee beans.

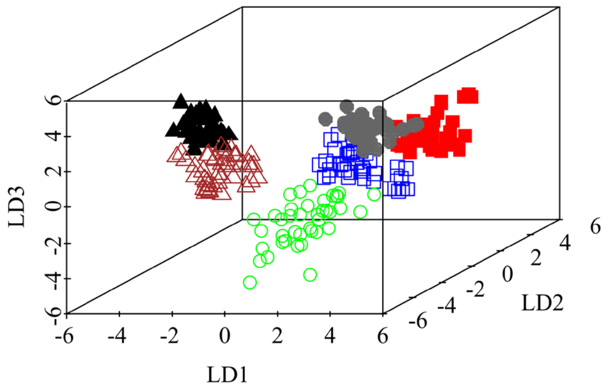

Figure 4. Linear discriminant scores plot for good quality beans (○), broken beans (□), beans with husk (●), beans with parchment (△), chicken faecal contaminated beans (■), and soil contaminated beans (▲).

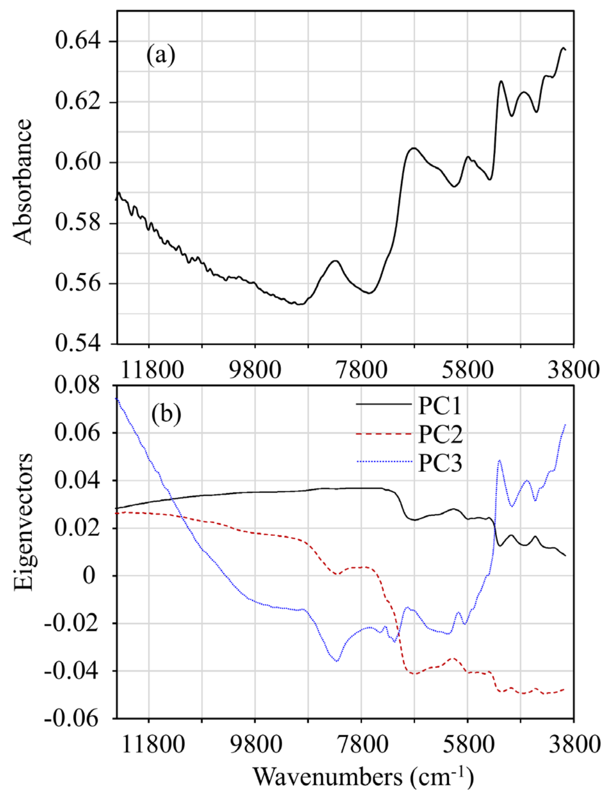

Figure 5. NIR spectrum of green 'Robusta' coffee bean samples after smoothing pre-treatment (a) and eigenvectors of PC1-3 (b).

Although the PCA-LDA method successfully classified each group of coffee beans, considering that the ultimate goal of the present work is building a cheap bean sorting machine for small-scale coffee bean producers, a lower cost NIR spectroscopy system is required.

Therefore, various wavelength selection techniques to find only meaningful variables (Yun *et al.,* 2019) may be applied before performing the discriminant analysis. In addition, portable and affordable NIR sensors which have been developed recently (Oliveira *et al.,* 2020; Savoia *et al.,* 2020) might accelerate the implementation of NIR spectroscopy for coffee bean producers as well.

Conclusion

The present work investigated the feasibility of applying NIR spectroscopy as a universal tool for sorting good quality green 'Robusta' coffee beans from defective beans and faecal and soil contaminated beans. The results showed that linear discriminant analysis using principal components from principal component analysis (PCA-LDA) was an effective method to classify each group of coffee beans. The highest classification performance for good coffee beans was obtained from full spectra with Savitzky-Golay smoothing pre-treatment at 97.5% correctness without false-positive classification.

Acknowledgement

The present work was financially supported by a grant from Silpakorn University Research and Development Institute (SURDI 58/02/66). The author gratefully acknowledges the ASEA-UNINET Staff Exchange, One Month Scholarship 2017 supported by Austrian Agency for International Cooperation in Education and Research (OeAD) and the Office of the Higher Education Commission, Thailand.

References

Abbas, O. and Baeten, V. 2017. Near-infrared spectroscopy. In Franca, A. S. and Nollet, L. M. L. (eds). Spectroscopic Methods in Food Analysis (1^{st} ed), p. 69-101. United States: CRC Press.

Alessandrini, L., Romani, S., Pinnavaia, G. and Della Rosa, M. 2008. Near infrared spectroscopy: an analytical tool to predict coffee roasting degree. Analytica Chimica Acta 625(1): 95-102.

Barbin, D. F., Felicio, A. L. S. M., Sun, D.-W., Nixdorf, S. L. and Hirooka, E. Y. 2014. Application of infrared spectral techniques on quality and compositional attributes of coffee: an overview. Food Research International 61: 23-32.

Biancolillo, A. and Marini, F. 2018. Chemometric methods for spectroscopy-based pharmaceutical analysis. Frontiers in Chemistry 6: article ID 576.

Brito, A. L. B., Brito, L. R., Honorato, F. A., Pontes, M. J. C. and Pontes, L. F. B. L. 2013. Classification of cereal bars using near infrared spectroscopy and linear discriminant analysis. Food Research International 51(2): 924-928.

Chao, K., Nou, X., Liu, Y., Kim, M. S., Chan, D. E., Yang, C. C., ... and Sharma, M. 2008. Detection of fecal/ingesta contaminants on poultry processing equipment surfaces by visible and near-infrared reflectance spectroscopy. Applied Engineering in Agriculture 24(1): 49-55.

Cozzolino, D., Chree, A., Scaife, J. R. and Murray, I. 2005. Usefulness of near-infrared reflectance (NIR) spectroscopy and chemometrics to discriminate fishmeal batches made with different fish species. Journal of Agricultural and Food Chemistry 53(11): 4459-4463.

Diniz, P. H. G. D., Gomes, A. A., Pistonesi, M. F., Band, B. S. F. and Araújo, M. C. U. 2014. Simultaneous classification of teas according to their varieties and geographical origins by using NIR spectroscopy and SPA-LDA. Food Analytical Methods 7: 1712-1718.

Downey, G. and Boussion, J. 1996. Authentication of coffee bean variety by near-infrared reflectance spectroscopy of dried extract. Journal of the Science of Food and Agriculture 71(1): 41-49.

Ebrahimi-Najafabadi, H., Leardi, R., Oliveri, P., Casolino, M. C., Jalali-Heravi, M. and Lanteri, S. 2012. Detection of addition of barley to coffee using near infrared spectroscopy and chemometric techniques. Talanta 99: 175-179.

Food and Agriculture Organization (FAO). 2020. Food and agriculture data - FAOSTAT. Retrieved from FAO website: http://www.fao.org/faostat/en/#home

Fox, G. P., Wu, A., Yiran, L. and Force, L. 2013. Variation in caffeine concentration in single coffee beans. Journal of Agricultural and Food Chemistry 61(45): 10772-10778.

Giraudo, A., Grassi, S., Savorani, F., Gavoci, G., Casiraghi, E. and Geobaldo, F. 2019. Determination of the geographical origin of green coffee beans using NIR spectroscopy and multivariate data analysis. Food Control 99: 137-145.

Huck, C. W., Guggenbichler, W. and Bonn, G. K. 2005. Analysis of caffeine, theobromine and 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.

Jakubíková, M., Sádecká, J., Kleinová, A. and Májek, P. 2016. Near-infrared spectroscopy for rapid classification of fruit spirits. Journal of Food Science and Technology 53: 2797-2803.

Khuwijitjaru, P. 2018. Near infrared spectroscopy research performance in food science and technology. NIR News 29(3): 12-14.

Li, Z., Wang, P.-P., Huang, C.-C., Shang, H., Pan, S.-Y. and Li, X.-J. 2014. Application of vis/nir spectroscopy for Chinese liquor discrimination. Food Analytical Methods 7(6): 1337-1344.

Liu, Y., Chen, Y.-R., Kim, M. S., Chan, D. E. and Lefcourt, A. M. 2007. Development of simple algorithms for the detection of fecal contaminants on apples from visible/near infrared hyperspectral reflectance imaging. Journal of Food Engineering 81(2): 412-418.

Manley, M. 2014. Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. Chemical Society Reviews 43(24): 8200-8214.

Mevik, B.-H., Wehrens, R. and Liland, K. H. 2016. pls: partial least squares and principal component regression. R package version 2.6-0. Retrieved from website: https://CRAN.R-project.org/package=pls

Miller, J. N. and Miller, J. C. 2001. Statistics and chemometrics for analytical chemistry. 4th ed. United States: Prentice Hall.

National Bureau of Agricultural Commodity and Food Standards. 2009. 5700-2552 - Robusta coffee bean. Thailand: Thai Agricultural Standard.

Okubo, N. and Kurata, Y. 2019. Nondestructive classification analysis of green coffee beans by using near-infrared spectroscopy. Foods 8(2): article ID 82.

Oliveira, M. M., Cruz-Tirado, J. P., Roque, J. V., Teófilo, R. F. and Barbin, D. F. 2020. Portable near-infrared spectroscopy for rapid authentication of adulterated paprika powder. Journal of Food Composition and Analysis 87: article ID 103403.

R Core Team. 2019. R: a language and environment for statistical computing. Austria: R Foundation for Statistical Computing.

Santos, J. R., Lopo, M., Rangel, A. O. S. S. and

Lopes, J. A. 2016. Exploiting near infrared spectroscopy as an analytical tool for on-line monitoring of acidity during coffee roasting. Food Control 60: 408-415.

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

Savoia, S., Albera, A., Brugiapaglia, A., Di Stasio, L., Ferragina, A., Cecchinato, A. and Bittante, G. 2020. Prediction of meat quality traits in the abattoir using portable and hand-held near-infrared spectrometers. Meat Science 161: article ID 108017.

Scholz, M. B. S., Kitzberger, C. S. G., Pereira, L. F. P., Davrieux, F., Pot, D., Charmetant, P. and Leroy, T. 2014. Application of near infrared spectroscopy for green coffee biochemical phenotyping. Journal of Near Infrared Spectroscopy 22(6): 411-421.

Stevens, A. and Ramirez-Lopez, L. 2013. An introduction to the prospectr package. Retrieved from website: https://cran.r-project.org/web/packages/prospectr/prospectr.pdf

Taradolsirithitikul, P., Sirisomboon, P. and Dachoupakan Sirisomboon, C. 2017. Qualitative and quantitative analysis of ochratoxin A contamination in green coffee beans using Fourier transform near infrared spectroscopy. Journal of the Science of Food and Agriculture 97(4): 1260-1266.

Venables, W. N. and Ripley, B. D. 2002. Modern applied statistics with S. 4th ed. New York: Springer-Verlag.

Viani, R. and Petracco, M. 2016. Coffee. In Elvers, B. (ed). Ullmann's Food and Feed (Volume 1), p. 281-312. Germany: Wiley-VCH.

Workman, J. and Weyer, L. 2012. Practical guide and spectral atlas for interpretive near-infrared spectroscopy. 2nd ed. United States: CRC Press.

Yun, Y.-H., Li, H.-D., Deng, B.-C. and Cao, D.-S. 2019. An overview of variable selection methods in multivariate analysis of near-infrared spectra. Trends in Analytical Chemistry 113: 102-115.