Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(1), 794-800 | ISSN:2251-6727

Quantitative Assessment of Caffeine Content in Blended Roasted Ground Coffee using Near Infrared Hyperspectral Imaging

[1] Woranitta Sahachairungrueng, [2] Anthony Keith Thompson, [3] Sontisuk Teerachaichayut

[1] Department of Food Science, School of Food-Industry, King Mongkut's Institute of Technology Ladkrabang, Chalongkrung Road, Ladkrabang, Bangkok, Thailand.
[2] Department of Postharvest Technology, Cranfield University, College Road, Cranfield, Bedford, MK430AL, United Kingdom.
[3] Department of Food Process Engineering, School of Food-Industry, King Mongkut's Institute of Technology Ladkrabang, Chalongkrung Road, Ladkrabang, Bangkok 10520, Thailand.

(Received: 27 October 2023 Revised: 22 November Accepted: 26 December)

KEYWORDS

regression, model, prediction, spectra.

ABSTRACT:
Caffeine is an important component in coffee. As well as its stimulatory effects on the consumer, caffeine has many negative effects, which has led to the popularity of caffeine-free coffee throughout the world. Decaffeinating coffee is a laborious and expensive process and only about 40% to 60% successful, but it is often not possible to tell whether or not the label on a jar or packet of coffee has been successfully decaffeinated or not. Therefore, it is possible to fraudulently mislabel decaffeinated coffee. The objective of this study was therefore to assess the caffeine content in blended roasted ground coffee that was made from ground roasted soybeans adding to ground roasted Robusta coffee. Blended roasted ground coffee samples were prepared by adding ground roasted soybean to ground roasted Robusta coffee over the range from 1% to 99% (w/w). All the blended samples were then scanned using a reflectance near infrared hyperspectral imaging (NIR-HSI) spectrophotometer in the wavelength range of 935–1720 nm. Samples (N = 202) were divided into a calibration set (N = 142) and a prediction set (N = 60). A regression model was established for determining caffeine content using partial least squares regression (PLSR). The PLSR model achieved coefficients of determination (R^2_p) in prediction samples of 0.88 and root mean square error in prediction (RMSEP) of 1.36 mg/g, which indicates a good predictive accuracy for detecting caffeine content in blended roasted ground coffee using NIR-HSI technique. Predictive images were created by interpreting every pixel of spectral image from caffeine content to colors using the model. The visualization of this research revealed that caffeine content could be rapidly predicted related to the color scale. This non-destructive method would be useful to manufacturers for producing new products of blended roasted ground coffee since it would provide a monitoring system thus allowing the level of caffeine to be controlled.

1. Introduction

Coffee is among the most popular beverages throughout the world, due to its unique flavor and taste [1]. Every year, 500 billion cups of coffee, or 9.4 million tons of coffee, are consumed [2]. Drinking coffee has both benefits and harms, in small amounts, it helps create alertness [3], but too much, increases the risk of cardiovascular disease, nervousness, restlessness, disturbance of sleep, palpitation of the heart, nervous tension, headaches, exacerbate migraines and many chronic diseases [4]–[7]. These symptoms are mainly due to the caffeine content of the coffee [6]. Caffeine is an alkaloid of the methylxanthine family, it is an intensely bitter white powder, a chemical formula of $C_8H_{10}N_4O_2$, and a systematic name of 1,3,5-trimethylxanthine [5], [8]. A cup of coffee has an average of 100–150 mg of caffeine [9]. An article appeared recently in a British consumer magazine [10] that showed the wide variation in caffeine levels in cups of coffee purchased in high street cafés. Their findings showed that Caffè Nero had 1.5, Costa 0.3, Greggs 2.7, Pret a Manger 6.0 and Starbucks 1.3 mg caffeine per mL of single-shot expresso coffee, respectively. BaSalamah et al. [11] recommended that in a day adults should not excess up to 400 mg, otherwise it will be a health risk.

Journal of Chemical Health Risks
www.jchr.org
JCHR (2024) 14(1), 794-800 | ISSN:2251-6727

Considering the current consumer trends in coffee consumption, some people avoid the negative effects of excessive coffee consumption by using coffee substitutes as a healthy alternative [6]. In recent years, agricultural products have been used to prepare coffee substitutes, such as using quinoa (Chenopodium quinoa) to produce caffeine-free and acceptable coffee substitutes [6], [7] or preparing good quality coffee from 15% soybean powder (Glycine max), which can avoid the disadvantages of coffee due to its highly nutritious beverage that is also good for health [4], [7]. Roasted ground soybeans can be used as coffee substitutes, they look like instant coffee, and have a coffee-like aroma, but retain the unique flavor of roasted soybeans [7], [12]. Soybeans help to control high blood pressure, aid digestion, lower blood cholesterol levels, promote gastrointestinal health, fight heart disease, and prevent cancer, particularly colon and rectal cancer, and they also provide eight essential amino acids that the body cannot produce [4], [7]. Blended roasted ground coffee is not only suitable for consumers who are interested in having a healthy diet and concerned about caffeine, but also for regular coffee drinkers. Thus, soybeans were used as a substitute for coffee in this study due to the benefits discussed above. Therefore, fast, and accurate techniques are needed to determine the caffeine content in blended roasted ground coffee.

Near infrared hyperspectral imaging (NIR-HSI) spectroscopy is a combination of near-infrared wavelength information coupled with image information. It can obtain both spatial and spectral information from food products [13]. The chemical composition of food products can be evaluated in a rapid and non-destructive, thus providing information on the spatial distribution of major chemical constituents across a sample [14]. Several workers have shown that NIR-HSI spectroscopy can be used for successfully detecting the quantity of chemical constituents in certain foods and agricultural products [15]–[17], including and for determining the caffeine content in beverages, including: matcha (green tea) [18] and black tea [19] as well as coffee [20], [21]. Therefore, a NIR-HSI spectroscope was tested to determine the caffeine content in samples of ground roasted Robusta coffee blended beans with ground roasted soybeans in order to ensure that the final product meets consumers' requirements.

2. Materials and Methods

A. Sample Preparation

Roasted Robusta coffee beans were purchased from Chumphon, in southern of Thailand, and soybeans were purchased from a local market in Bangkok, Thailand. The soybeans were then roasted in a brass pan at 170 °C for 25 minutes until they became brown. Then, both the roasted coffee beans and roasted soybeans were crushed separately in a blender (Blender 480, Kenwood, Thailand), and then the ground samples were combined with the soyabeans added to the coffee to give different combinations from 1 to 99% w/w coffee to soybeans, with increments every 1% w/w.

B. NIR-HSI

Each sample was scanned using the NIR-HSI spectrophotometer (Specim FX17e, Spectral Imaging Ltd., Oulu, Finland) in the reflectance mode. The speed of tray movement was set to 20 mm/s, and the integration time was set to 5.5 ms. This system captured 224 spectral bands from 935–1720 nm with a spectral resolution of approximately 3.5 nm, which provided an image size of 640 × 1187 pixels. Both a dark reference and a white reference were also scanned before every measurement. The dark image was acquired when the shutter was closed, and the white image was acquired by scanning a Spectralon bar (Fig. 1).

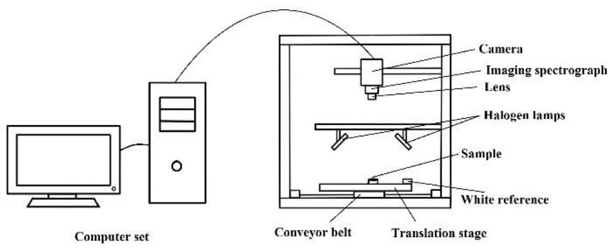

Fig.1 Sample presentation and NIR-HSI spectrophotometer.

C. Caffeine determination

After scanning each sample using the NIR-HSI spectroscopy, their caffeine content was determined using the method described by [22] with modifications. The 5 g blended roasted ground coffee sample was placed in a beaker, and 5 g of sodium carbonate was added to together with 35 mL of distilled water. The mixture was then boiled on a hot plate for 5 minutes and

Journal of Chemical Health Risks
www.jchr.org
JCHR (2024) 14(1), 794-800 | ISSN:2251-6727

filtered, using a funnel and coffee filter paper. The filtrate obtained was transferred to the separatory funnel where 10 mL of dichloromethane (CH_2Cl_2) was added, sealed and allowed to separate into layers. The lower layer is drawn off and placed in an Erlenmeyer flask. The extraction was repeated with 10 mL of dichloromethane and then 7 ml of saturated sodium chloride was added to the Erlenmeyer flask containing dichloromethane and gently shaken. The mixture was then poured into a separator funnel, and the layers were allowed to separate completely. The lower layer containing dichloromethane was drained off and collected in an Erlenmeyer flask. Anhydrous powdered sodium sulfate was added to the Erlenmeyer flask containing dichloromethane, then the mixture was poured into a separator funnel to separate the dichloromethane, which was then placed in a beaker, which was placed on a hot plate and evaporated until it was dry, allowed to cool to room temperature, and then weighed and the caffeine content was determined as in (1).

Caffeine content (mg/g) = (weight of caffeine from extraction (g) × 1000) / weight of sample (g) (1)

D. Data processing and statistical analysis

The procedure of analysis in this study is shown in a process block diagram (Fig. 2). The spectral images of each sample, acquired using NIR-HSI spectroscopy, were averaged and used for statistical analysis. The average spectra of 202 samples were divided into a calibration set (142 samples) and a prediction set (60 samples) that is a ratio of 70:30. Partial least squares regression (PLSR) can analyze spectral data as independent variables and create a calibration model for predicting concentrations of component as dependent variables [23]. In this study, the PLSR was applied to correlate the spectral data and caffeine content in blended roasted ground coffee. The spectra of samples in the calibration set were pre-processed using the following spectral pretreatment methods: smoothing, the first and second derivative using Savitzky-Golay, multiplicative scatter correction (MSC), standard normal variate (SNV) and combinations of these methods. From these the optimal conditions were selected for establishing the calibration model. The coefficient of determination (R^2) and root mean **square** error (RMSE) are considered to be the most important factors for evaluating the performance of the model in order to achieve high predictive efficiency and overcome potential mistakes in evaluating the accuracy of the model. The optimum calibration model for determining caffeine content was selected based on the highest R^2 and the lowest RMSE. The accuracy of the calibration model was evaluated by using samples in the prediction set.

E. Predictive image processing

The ability to create predictive images for predicting chemical components in every pixel of spectral images is one of the advantages of NIR-HSI spectroscopy. The calibration model was used to predict the chemical component in every pixel and interpreted to colors based on the color scale that related to the concentration of the chemical component. In this study, the predictive images would display the concentration and the distribution of the caffeine content in blended roasted ground coffee by using the PLSR model.

The Unscrambler X 10.4 software (COMO, Trondheim, Norway) and the Prediktera Evince software (Prediktera Evince version 2.7.9, Prediktera AB, Umea, Sweden) were used for analysis in this study.

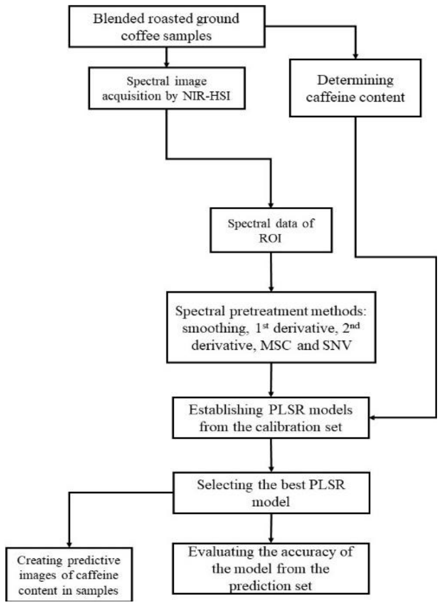

Fig. 2 Diagram of spectral analysis for creating the model and predictive images.

3. Results and Discussion

A. NIR spectral data

Original spectra of blended roasted ground coffee from NIR-HSI spectroscopy showed the absorption bands in

Journal of Chemical Health Risks
www.jchr.org
JCHR (2024) 14(1), 794-800 | ISSN:2251-6727

the NIR region that were complex and had broad and overlapping bands. These overlapping bands were difficult to interpret. Therefore, the 2nd derivative spectral pretreatment method was used for identification of the chemical composition in the blended roasted ground coffee sample, which showed the peaks of caffeine content that were associated C-H stretching and C-H deformation of caffeine at the wavelength range of 1360 and 1446 nm (Fig. 3). This effect supports the findings of [24], [25]. The average 2nd derivative spectra between 100% ground roasted coffee and 100% ground roasted soybean were considered, as shown in Fig. 4. The wavelength range of 1360 and 1446 nm has a clearly different absorbance.

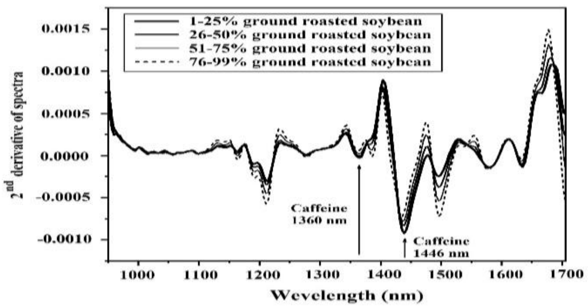

Fig. 3 The second derivative spectra of blended roasted ground coffee containing different levels of roasted soybeans.

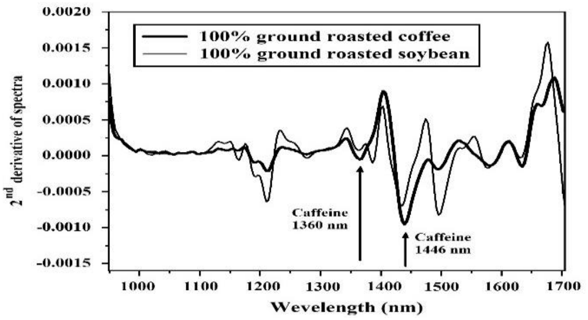

Fig. 4 The second derivative spectra between 100% roasted ground coffee and 100% roasted ground soybean.

B. The calibration model

All 202 samples were divided into two groups: calibration and prediction at a ratio of 70:30. The range of caffeine content in each group is shown in Table 1.

The PLSR was used for creating models for determination of the caffeine content and cross-validated by samples in the calibration set. Spectral data were pretreated with smoothing, the 1st derivative, the 2nd derivative (Savitzky-Golay), multiplication scatter correction (MSC), standard normal variate transformation (SNV), and combined methods. The model from MSC spectral pretreatment showed the most accurate results for caffeine content with highest coefficient of determinations (R^2_c = 0.89 and R^2_{cv} = 0.89) and lowest root mean square errors (RMSEC = 1.30 mg/g and RMSECV = 1.32 mg/g) (Table 2). Therefore, the optimum model from MSC spectral pretreatment was selected for use in this study. The model was tested by samples in the prediction set for evaluating the accuracy of the model. The results show a good prediction with R^2_p of 0.88 and RMSEP of 1.36 mg/g (Table 3), which indicates that NIR-HSI technique has good potential to use for predicting caffeine content in blended roasted ground coffee. Therefore, this acquired model was used to determine the predicted values of the caffeine content compared with the actual values of caffeine content. The results showed the plots were close to the 45° line (Fig. 5), indicating that the model was a good fit and was reliable in predicting the caffeine content of blended roasted ground coffee.

Table I Caffeine content in samples in the calibration set and prediction set of the blended samples.

Datasets	Number of samples	Range (mg/g)	Mean (mg/g)	Standard deviation (mg/g)
Calibration set	142	0–17.50	9.14	3.99
Prediction set	60	0–15.23	9.11	3.94

Table II Results of the partial least squares regression (PLSR) models for determination of the caffeine content using various spectral pretreatment methods in the calibration set.

Pre-treatment methods	F	PLSR models			
		R^2_c	RMSEC (mg/g)	R^2_{cv}	RMSECV
Original	2	0.89	1.33	0.88	1.36
Smoothing	3	0.89	1.33	0.89	1.36
1st derivative	1	0.88	1.36	0.88	1.40
2nd derivative	2	0.89	1.33	0.87	1.44
MSC	4	0.89	1.30	0.89	1.32
SNV	1	0.89	1.32	0.89	1.34

Journal of Chemical Health Risks
www.jchr.org
JCHR (2024) 14(1), 794-800 | ISSN:2251-6727

1st derivative + MSC	2	0.88	1.39	0.87	1.42
1st derivative + SNV	1	0.88	1.38	0.88	1.41

F = factors

Table III The accuracy of the PLSR model for predicting caffeine content.

Pre-treatment method	F	Calibration set			Prediction set		
		N	R^2_c	RMSEC (mg/g)	N	R^2_P	RMSEP (mg/g)
MSC	4	142	0.89	1.30	60	0.88	1.36

F = factors
N = number of samples
MSC = Multiplicative scatter correction
R^2_c = correlation coefficients of calibration
RMSEC = root mean square error of calibration
R^2_p = correlation coefficients of prediction
RMSEP = root mean square error of prediction

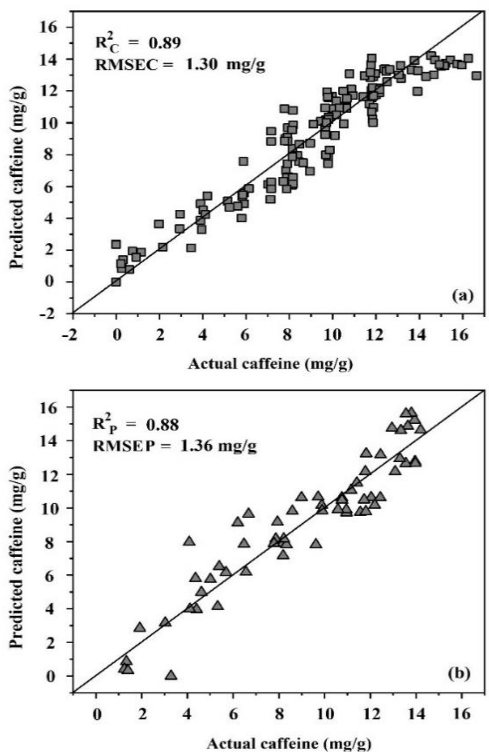

Fig. 5 Scatter plots of actual values versus predicted values of caffeine content in (a) the calibration set and (b) the prediction set.

C. Predictive images

The acquired model that MSC was used for spectral pretreatment was the best PLSR model for predicting caffeine contents in blended roasted ground coffee. This model was used to create predictive images for predicting the caffeine content in each pixel of spectral images of blended roasted ground coffee as shown in Fig. 6. The pixel color of the image with the high caffeine content was displayed in red, while the image with the low caffeine content was shown in blue. The visualization of caffeine content showed the concentration and distribution of caffeine contents in blended roasted ground coffee.

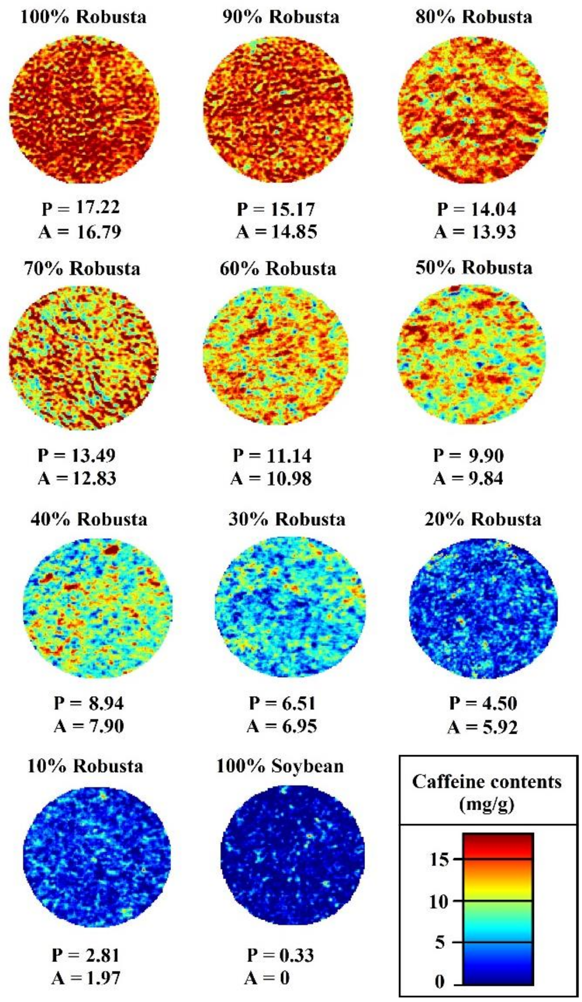

Fig. 6 Predictive images of caffeine content of blended roasted ground coffee (P is the predicted value; A is the actual value).

Journal of Chemical Health Risks

www.jchr.org

JCHR (2024) 14(1), 794-800 | ISSN:2251-6727

4. Conclusion

The objective was to test whether it was possible to determine, non-destructively, the level of caffeine in blended roasted ground coffee by using a model that was established using the acquired spectra from NIR-HSI spectroscopy measurements. Various preprocessing methods of the spectra were tested, and MSC spectral pretreatment was found to be the most suitable for developing the model, giving good accuracy for prediction of R^2_p = 0.88 and RMSEP = 1.36 mg/g. It was therefore concluded that NIR-HSI spectroscopy could be used to determine caffeine content in blended roasted ground coffee. This technique has potential to be useful to commercial companies for developing new products of blended roasted ground coffee enabling them to determine and predict caffeine content and to control caffeine content based on consumers' preferences.

5. Acknowledgement

This work is supported by King Mongkut's Institute of Technology Ladkrabang. The authors are grateful Prof. Panmanas Sirisomboon for technical help.

References:

1. K. Zhang, J. Cheng, Q. Hong, W. Dong, X. Chen, G. Wu, and Z. Zhang, "Identification of changes in the volatile compounds of robust coffee beans during drying based on HS-SPME/GC-MS and E-nose analyses with the aid of chemometrics," LWT - Food Sci. Technol., vol. 161, 113317, Mar. 2022, https://doi.org/10.1016/j.lwt.2022.113317
2. S. F. Taghizadeh, M. Azizi, G. Hassanpourfard, R. Rezaee, and G. Karimi, "Assessment of Carcinogenic and Non-carcinogenic Risk of Exposure to Metals via Consumption of Coffee, Tea, and Herbal Tea in Iranians," Biol. Trace Elem. Res., vol. 201, no. 3, pp. 1520–1537, Mar. 2023, https://doi.org/10.1007/s12011-022-03239-x
3. Felberg, R. Deliza, A. Farah, E. Calado, and C. M. Donangelo, "Formulation of A Soy-Coffee Beverage by Response Surface Methodology and Internal reference Mapping," J. Sens. Stud., vol. 25, pp. 226–242, Nov. 2009, https://doi.org/10.1111/j.1745-459X.2010.00278.x
4. S. B. Sataw, D. M. Choudhari, R. Jaybalan, and V. M. Chavarkar, "Sensory quality and cost of production of instant soy coffee," Food Sci. Res. J., vol. 2, no. 2, pp. 169–172, Sept. 2011, http://researchjournal.co.in/upload/assignments/2_169-172.pdf
5. N. Violeta, I. Trandafir, and I. M. Elena, "Quantitative determination of caffeine in carbonated beverages by HPLC method," J. Agroaliment. Processes Technol., vol. 14, pp. 123–127, Jan. 2008.
6. F. O. Abou-Zaid, "Evaluation of Coffee Substitute Produced from Quinoa," Asian Res. J. Curr. Sci., vol. 4, no. 1, pp. 125–133, Feb. 2022.
7. A. Kalne, K. A. Khan, N. A. Ghurde, and V. B. Kalmegh, "Optimization and Standardization of Process Technology for Preparation of Soycoffee," in Soybean Research, O. P. Joshi, Ed. Society for Soybean Research and Development Directorate of Soybean Research, Khandwa Road, Indore 452 001, Madhya Pradesh, India, 2014, pp. 219–224.
8. Mumin, K. F. Akhter, Z. Abedin, and Z. Hossain, "Determination and Characterization of Caffeine in Tea, Coffee and Soft Drinks by Solid Phase Extraction and High-Performance Liquid Chromatography (SPE-HPLC)," Malays. J. Chem., vol. 8, no. 1, pp. 045–051, Nov. 2006.
9. S. R. Dager, M. E. Layton, W. Strauss, T. L. Richards, A. Heide, S. D. Friedman, A. A. Artru, C. E. Hayes, and S. Posse, "Human Brain Metabolic Response to Caffeine and Effects of Tolerance," Am. J. Psychiatry, vol. 156, no. 2, pp. 299–237, Feb. 1999, https://doi.org/10.1176/ajp.156.2.229.
10. Which, 2023 [Online]. Available: https://www.which.co.uk/news/article/caffeine-levels-in-high-street-coffees-vary-significantly-which-finds-ay7cA4G1zh1S.
11. M. BaSalamah, R. AlMghamsi, A. AlTowairqi, K. Fouda, A. Mahrous, M. Mujahid, H. Sindi, and A. Aldairi, "The Effect of Coffee Consumption on Blood Glucose Levels," J. Biochem. Technol., vol. 13, no. 2, pp. 64–69, Jun. 2022, https://doi.org/10.51847/VOLNuKyp3c
12. L. Sharma, "Nutritional composition of a food product developed with combination of pulse and

cereal family," Int. j. physiol. nutr. phys. educ., vol. 4, no. 1, pp. 139–143, Dec. 2018, https://www.journalofsports.com/pdf/2019/vol4issue1/PartD/4-1-37-997.pdf

13. W. Sahachairungrueng, C. Meechan, N. Veerachat, A. K. Thompson, and S. Teerachaichayut, "Assessing the Levels of Robusta and Arabica in Roasted Ground Coffee Using NIR Hyperspectral Imaging and FTIR Spectroscopy," Foods, vol. 11, p. 3122, Oct. 2022, https://doi.org/10.3390/foods11193122.
14. D. Liu, X. A. Zeng, and D. W. Sun, "Recent developments and applications of hyperspectral imaging for quality evaluation of agricultural products: A review," Crit. Rev. Food Sci. Nutr., vol. 55, no. 2, pp. 1744–1757, Aug. 2014, https://doi.org/10.1080/10408398.2013.777020
15. A. Tantinantrakun, S. Sukwanit, A. K. Thompson, and S. Teerachaichayut, "Nondestructive evaluation of SW-NIRS and NIR-HSI for predicting the maturity index of intact pineapples," Postharvest Biol. Technol., vol. 195, 112141, Jan. 2023, http://dx.doi.org/10.1016/j.postharvbio.2022.112141
16. J. Zhang, Z. Guo, Z. Ren, S. Wang, M. Yue, S. Zhang, X. Yin, K. Gong, and C. Ma, "Rapid determination of protein, starch and moisture content in wheat flour by near-infrared hyperspectral imaging," J. Food Compos. Anal., vol. 117, 105134, Jan. 2023, https://doi.org/10.1016/j.jfca.2023.105134
17. D. Saha, T. Senthilkumar S. Sharma, C. B. Singh and A. Manickavasagan, "Application of near-infrared hyperspectral imaging coupled with chemometrics for rapid and non-destructive prediction of protein content in single chickpea seed," J. Food Compos. Anal., vol. 115, 04938, Sept. 2022, https://doi.org/10.1016/j.jfca.2022.104938
18. Q. Ouyang, L. Wang, B. Park, R. Kang, and Q. Chen, "Simultaneous quantification of chemical constituents in match with visible-near infrared hyperspectral imaging technology," Food Chem., vol. 350, 129141, Feb. 2021, https://doi.org/10.1155%2F2021%2F5566612
19. Y. Mao, H. Li, Y. Wang, K. Fan, Y. Song, X. Han, J. Zhang, S. Ding, D. Song, H. Wang, and Z. Ding, "Prediction of Tea Polyphenols, Free Amino Acids and Caffeine Content in Tea Leaves during Wilting and Fermentation Using Hyperspectral Imaging," Foods, vol. 11, p. 2537, Aug. 2022, https://doi.org/10.3390/foods11162537
20. Zhang, H. Jiang, F. Liu, and Y. He, "Application of Near-Infrared Hyperspectral Imaging with Variable Selection Methods to Determine and Visualize Caffeine Content of Coffee beans," Food Bioproc. Tech., vol. 10, pp. 213–221, Oct. 2016, https://doi.org/10.1007/s11947-016-1809-8
21. N. Caporaso, M. B. Whitworth, S. Grebby, and I. D. Fisk, "Non-destructive analysis of sucrose, caffeine and trigonelline on single green coffee beans by hyperspectral imaging," Food Res. Int., vol. 106, pp. 193–203, Dec. 2017, https://doi.org/10.1016/j.foodres.2017.12.031
22. L. F. Fieser, and K. L. Williamson, "Extraction: Isolation of Caffeine from Tea and Cola Syrup," in Organic Experiments, 8th ed. Houghton Mifflin Company, New York, 1992, pp. 99–116.
23. G. W. Small, M. A. Arnold, and A. Marquardt, "Strategies for Coupling Digital Filtering with Partial Least-Squares Regression: Application to the Determination of Glucose in Plasma by Fourier Transform Near-Infrared Spectroscopy," Anal. Chem., vol. 66, no. 22, pp. 3279–3289, Nov. 1993, https://doi.org/10.1021/ac00070a019
24. Sanaeifar, X. Huang, M. Chen, Z. Zhao Y. Ji, X. Li, Y. He, Y. Zhu, X. Chen, and X. Yu, "Nondestructive monitoring of polyphrnols and caffeine during green tea processing using Vis-NIR spectroscopy," Food Sci. Nutr., vol. 8, no. 6, pp. 1–15, Aug. 2020, http://dx.doi.org/10.1002/fsn3.1861
25. Pizarro, I. Esteban-Diez, J. M. Gonzalez-Saiz, and M. Forina, "Use of Near-Infrared Spectroscopy and Feature Selection Techniques for Predicting the Caffeine Content and Roasting Color in Roasted Coffees," J. Agric. Food Chem., vol. 55, pp. 7477–7488, Aug. 2007, https://doi.org/10.1021/jf071139x