Journal of Physics: Conference Series

PAPER • **OPEN ACCESS**

The Use of SIMCA Method and NIR Spectroscopy with Hand-Held Spectrometers Equipped with Integrating Sphere for Classification of Two Different Indonesian Specialty Coffees

To cite this article: D Suhandy *et al* 2021 *J. Phys.: Conf. Ser.* **1751** 012080

The Use of SIMCA Method and NIR Spectroscopy with Hand-Held Spectrometers Equipped with Integrating Sphere for Classification of Two Different Indonesian Specialty Coffees

D Suhandy[a,c,1], M Yulia[b,c] , S Kuroki[d] , K Nakano[e]
[a] Department of Agricultural Engineering, Faculty of Agriculture, The University of Lampung, Jl. Prof. Dr. Soemantri Brojonegoro No.1, Bandar Lampung, 35145, Lampung, Indonesia
[b] Department of Agricultural Technology, Lampung State Polytechnic, Jl. Soekarno Hatta No. 10, Rajabasa Bandar Lampung, 35141, Lampung, Indonesia
[c] Spectroscopy Research Group (SRG), Laboratory of Bioprocess and Postharvest Engineering, Faculty of Agriculture, The University of Lampung, Bandar Lampung, 35145, Lampung, Indonesia
[d] Graduate School of Agricultural Science, Kobe University, 1-1 Rokkodai, Nada, Kobe, 657-8501, Japan.
[e] The United Graduate School of Agricultural Science, Gifu University, 1-1 Yanagido, Gifu City, Gifu, 501-1193, Japan

E-mail: diding.sughandy@fp.unila.ac.id

Abstract. The objective of this study is to evaluate the SIMCA method and NIR spectroscopy for the non-invasive and non-destructive classification of Indonesian specialty coffees that come from two geographical origins: Gayo coffee from Aceh 10 samples and Wamena from Papua 10 samples. All samples were roasted at the same condition (medium roasting at a temperature of 200°C for 10 minutes) and were ground using a home coffee grinder and then sieved using 50 mesh to obtain a homogenous particle size of 297 micrometers. Spectral data in the short and long near infrared range of 650–1650 nm was measured in a diffuse reflectance mode using two handheld spectrometers equipped with an integrating sphere (ISP-REF, Ocean Optics, USA). The result demonstrated that the classification was satisfied with 100% accuracy, sensitivity, and specificity.

1. Introduction

Indonesia is one of the most important coffee producers in the world (especially for *Coffea canephora* or Robusta coffee) with a total production of about 9.4 thousand bags of 60 kg in 2018/2019 [1]. Not only Robusta coffee, several Arabica coffees (*Coffea arabica*) are planted in specific geographical environment in several places in Indonesia and protected as Indonesian specialty coffee such as Arabica Gayo (Aceh), Arabica Wamena (Papua), Arabica Kalosi (Sulawesi), Arabica Java Preanger (West Java). Those specialty coffees have great flavor as well as a high score of sensory properties and

[1] To whom any correspondence should be addressed.

meet the requirement of SCAA (Specialty Coffee Association of America) specialty coffee with a score greater than 80 [2]. For this reason, Indonesian specialty coffee with labeled as specific geographical origins are traded at higher prices compared to normal non-specialty coffee.

Recently, confirmation of the geographical origin of specialty coffee is one of the most important food quality analysis [3]. There are two methods to do this task, in general: targeted and non-targeted analysis [4]. In a targeted analysis, first, we determine a specific property of the product as a marker to decide the authenticity of the tested product. For non-targeted analysis, broad information coming from the spectrum, chromatogram, or fingerprints of both adulterated and non-adulterated samples was measured. This set of confirmed members was used to characterize the adulterated and non-adulterated class. The new unknown samples were measured, and then a classification or discrimination method was used to classify the samples into adulterated and non-adulterated class. Several examples of non-targeted analysis have been published such as the authentication of peaberry coffee using UV-visible spectroscopy [5], authentication of wild salmon using ^{1}H Nuclear magnetic resonance (NMR) spectroscopy [6], geographic origin of wines using UV-visible-near and mid infrared spectroscopy [7], authentication and discrimination of green tea samples using UV–vis, FTIR and HPLC techniques [8], geographical origin of honey using near and mid infrared spectroscopy [9-10], discrimination of Chinese rice using ^{1}H NMR-based metabolomics [11], etc.

A benchtop NIR spectrometer is relatively costly and inapplicable for on-site measurement. For coffee authentication, NIR benchtop-based spectrometer has been reported [12-15]. To perform coffee authentication with low-cost analytical methods, it is desirable to use a handheld NIR spectrometer. In this research, a possible application of soft independent modeling of class analogy (SIMCA) method and NIR spectroscopy with two different handheld spectrometers equipped with an integrating sphere for non-targeted classification of two different geographical origins of Indonesian specialty coffees was presented.

2. Research methodology

2.1. Coffee samples

In this research, a total of 20 samples of Indonesian specialty coffee from two geographical origins were prepared (Gayo and Wamena). The samples of green bean coffee were purchased from a local trusted coffee shop in Bandar Lampung, Indonesia. All samples were roasted at the same condition (medium roasting at a temperature of 200°C for 10 minutes). Samples were ground using a home coffee grinder and sieved using 50 mesh to obtain a homogenous particle size of 297 micrometers in order to minimize the influence of coffee particle size to the quality of spectral data [16]. For spectral acquisition, ground roasted coffee samples were directly placed on the sample holder about 1.33 grams in weight on a flat surface, as depicted in Figure 1.

Sample holder

Figure 1. Sample holder for direct NIR spectral acquisition of ground roasted coffee using two handheld spectrometers equipped with an integrating sphere in the range of 650-1650 nm.

2.2. System for coffee spectral acquisition

Figure 2 showed the system for coffee spectral data acquisition. This system consists of two handheld spectrometers from Ocean Optics (USA): NIR Flame-S operated in the range of 650–900 nm, and NIRQuest512 operated in the range of 900–1650 nm (total the range was 650–1650 nm). The sample was placed at the top of the integrating sphere, and a built-in tungsten-halogen light source was located at the bottom of the integrating sphere. Spectral data was measured in interactance mode simultaneously by using the two spectrometers via fiber optics, and then the spectral data were transferred into the computer through a USB cable. The original spectral data (no data pre-processing) in the full range of 650–1650 nm (short and long near infrared wavelength) was used for chemometrics calculation, including principal component analysis (PCA) and soft independent modelling of class analogy (SIMCA).

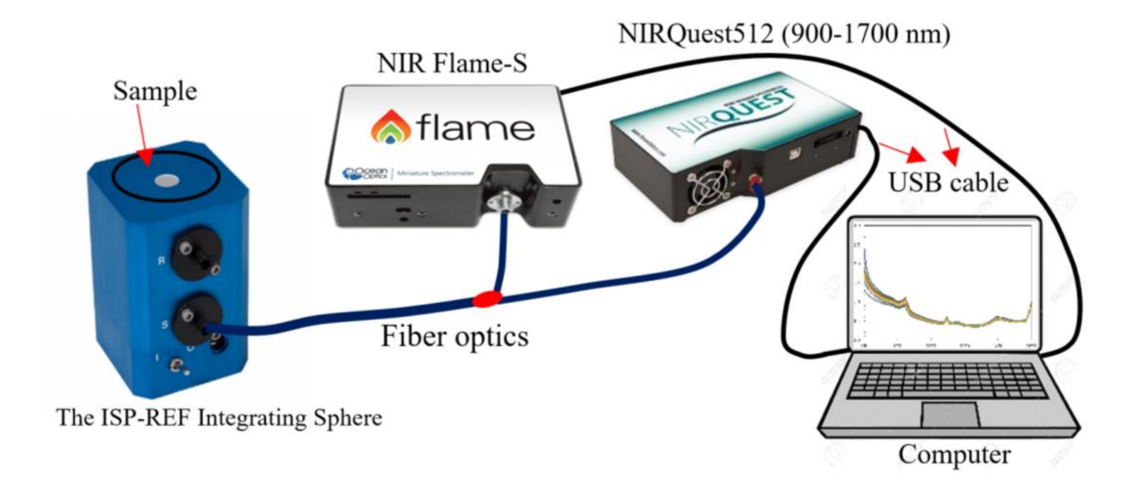

Figure 2. Spectral acquisition of coffee samples from two different geographical origins (Gayo and Wamena) using two handheld spectrometers equipped with an integrating sphere in the range of 650-1650 nm.

2.3. PCA and SIMCA method

The obtained spectral data in the range of 650–1650 nm contained 1206 predictor variables (X variables). To easier understand this complex data, the original spectral data was decomposed into a few new variables, the so-called principal components (PCs) data using PCA [17-18]. PCA is one of the unsupervised classification methods and frequently used for authentication research. One important result of PCA is the score plot of principal components (PCs). PC1 and PC2 score plots were usually presented to map the samples and to see the grouping potential of the samples. In the present research, PCA was applied using original spectral data set in the range of 650–1650 nm (1206 variables).

To perform a supervised classification method, a SIMCA method which focuses on interclass similarities was calculated [19]. First, the samples were divided randomly into calibration (12 samples) and prediction set (8 samples). PCA was developed using calibration samples for each class (Gayo class model and Wamena class model). The prediction samples were projected onto each SIMCA model, and from this process, scores and residuals of each prediction samples were obtained. The membership of the prediction samples set was then assigned using two parameters: object-to-model distance and leverage. The object-to-model distance was calculated using combined residuals, and leverage was measured by using the scores. To evaluate the performance of SIMCA, three figures of merit were used: accuracy, sensitivity, and specificity. The calculation of those parameters was performed, according to Luna *et al.* [20]. All chemometric data pre-processing was performed by using The Unscrambler X version 10.4 (64-bit) (Camo Software AS, Oslo, Norway).

3. Results and discussions

3.1. Analysis of spectra of coffee samples with two different geographical origins

Original spectral data in the full range of 650–1650 nm was presented in Figure 3. There was a high variation in spectral data obtained by the NIR Flame-S spectrometer in the range of 650–900 nm. On the other hand, a low variation was observed in the range of 900–1650 nm obtained by NIRQuest512. Clear peaks were observed between 900–1650 nm at 1150 nm, 1200 nm, and 1450 nm. The peaks at around 1150 nm and 1200 nm were related to second overtone of C-H bonding in coffee [21]. The peak at 1450 nm was contributed to the first overtone of O-H stretching in coffee due to the presence of water in the coffee samples [21].

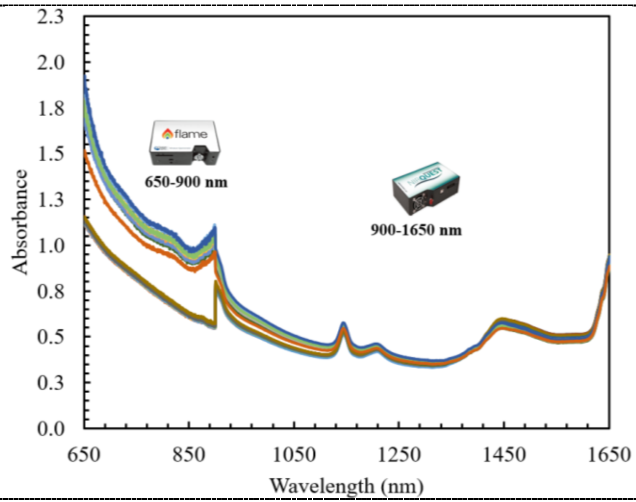

Figure 3. Spectra of coffee samples from two geographical origins (Gayo and Wamena) for original spectra in the range of 650–1650 nm.

3.2. *Result of calculation of PCA*

Figure 4 showed the score plot of the first and the second PC using short and long original near infrared spectral data in the range of 650-1650 nm. The first two PCs (PC1 and PC2) could explain 100% of the total variance of the original spectral data. A clear separation was observed between Gayo and Wamena specialty coffees. All Gayo samples were clustered in the positive PC1 (PC1 > 0), whereas all Wamena samples were in negative PC1 (PC1 < 0). There were no any overlapped samples indicated that Gayo and Wamena specialty coffee could be distinguished well using NIR spectral data in the range of 650–1650 nm.

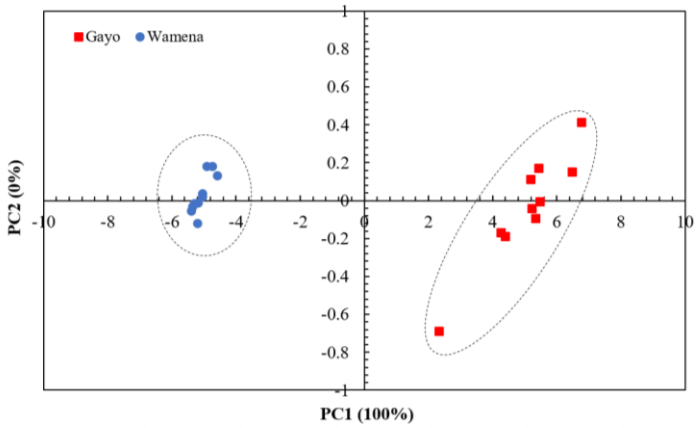

Figure 4. PCA score plot of coffee samples with different geographical origins (Gayo and Wamena) calculated based on original spectra in the range of 650-1650 nm.

3.3. Classification result using SIMCA method

Table 1 showed the result of the classification of prediction samples (8 samples) using the SIMCA model. All Wamena samples (4 samples) were classified correctly. Three samples of Gayo could be properly classified, but only one Gayo sample did not belong to any defined class (Gayo and Wamena). As presented in Table 1, accuracy, sensitivity, and specificity were 100% both for Gayo and Wamena class. This result indicated a potential application of low-cost handheld NIR spectrometers to establish authentication of Indonesian specialty coffee. However, in the future, more samples have to be measured to draw a more reliable conclusion.

Table 1. The result of classification based on SIMCA using original spectral data in the range of 650-1650 nm.

Models	Predicted Gayo class	Predicted Wamena class
Actual Gayo class	3	0
Actual Wamena class	0	4

4. Conclusion

As presented in this research, a potential application of low-cost handheld NIR spectrometers to establish authentication of Indonesian specialty coffee had been demonstrated. Using the unsupervised method of PCA on original spectral data in the range of 650-1650 nm, a clear separation was observed between Gayo and Wamena coffee. Using supervised classification based on the SIMCA method, accuracy, sensitivity, and specificity of 100% could be obtained. To draw a more reliable conclusion, it is highly desired to measure more samples.

Acknowledgment

We would like to thank the Ministry of Research and Technology/National Agency for Research and Innovation, Republic of Indonesia for financial support (Basic Research Grant 2020-2022 No.139/SP2H/LT/DRPM/2020).

References

[1] ICO 2020 http://www.ico.org/historical/1990%20onwards/PDF/1a-total-production.pdf
[2] Guimarães E R, Leme P H M V, De Rezende D C, Pereira S P and Dos Santos A C 2019 *J. Food Prod. Market.* **25** 49–71.
[3] Rodionova O Y and Pomerantsev A L 2020 *Food Chem*. **317** 126448.
[4] McGrath T F, Haughey S A, Patterson J, Fauhl-Hassek C, Donarski J, Alewijn M, van Ruth S and Elliott C T 2018 *Trends Food Sci. Technol*. **76** 38–55.
[5] Suhandy D and Yulia M 2017 *Int. J. Food Prop.* **20** S331–S339.
[6] Capuano E, Lommen A, Heenan S, de la Dura A, Rozijn M and van Ruth S 2012 *Lipid Technol.* **24** 251–253.
[7] Cozzolino D, Cynkar W U, Shah N and Smith P A 2011 *Food Chem.* **126** 673–678.
[8] Aboulwafa M M, Youssef F S, Gad H A, Sarker S D, Nahar L, Al-Azizi M M and Ashour M L 2019 *J. Pharm. Biomed. Anal.* **164** 653–658.
[9] Bisutti V, Merlanti R, Serva L , Lucatello L, Mirisola M, Balzan S, Tenti S, Fontana F, Trevisan G, Montanucci L, Contiero B, Segato S and Capolongo F 2019 *J. Near Infrared Spec*. **27** 65–74.
[10] Lekova S and Tsankova D 2017 *J. Chem. Technol. Metall.* **52** 52–57.
[11] Huo Y, Kamal G M, Wang J, Liu H, Zhang G, Hu Z, Anwar F and Du H 2017 *J. Cereal Sci*. **76** 243–252.
[12] Suhandy D, Yulia M, Ogawa Y and Kondo N 2018 *IOP Conf. Ser.: Earth Environ. Sci.* **147** 012011.

[13] Esteban-Díez I, González-Sáiz J M, Sáenz-González C and Pizarro C 2007 *Talanta* **71** 221–229.
[14] Monteiro P I, Santos J S, Brizola V R A, Deolindo C T P, Koot A, Boerrigter-Eenling R, van Ruth S, Georgouli K, Koidis A and Granato D 2018 *Food Control* **91** 276–283.
[15] Belchior V, Botelho B G, Casal S Oliveira L S and Franca A S 2020 *Food Anal. Methods* **13** 275–283.
[16] Yulia M and Suhandy D 2018 *IOP Conf. Ser.: Mater. Sci. Eng*. **334** 012062.
[17] Suhandy D and Yulia M 2018 *MATEC Web Conf.* **197** 09002.
[18] Yulia M and Suhandy D 2018 *MATEC Web Conf*. **197** 09003.
[19] Firmani P, Bucci R, Marini F and Biancolillo A 2019 *J. Food Compost. Anal*. **82** 103235.
[20] Luna A S, Pinho J S A and Machado L C 2016 *Anal. Methods* **8** 7204–7208.
[21] Zhang C, Jiang H, Liu F and He Y 2017 *Food Bioproc. Tech.* **10** 213–221.