UV spectroscopy for discrimination of two arabica coffee cultivars in West Java Indonesia: a feasibility study

M Yulia¹, K R Ningtyas¹, S Suharyatun², W Rahmawati³ and D Suhandy³

¹ Department of Agricultural Technology, Lampung State Polytechnic, Jl. Soekarno Hatta No. 10, Rajabasa Bandar Lampung, 35141, Indonesia
² Department of Agricultural Engineering, The University of Lampung, Jl. Prof. Dr. Soemantri Brojonegoro No.1, Bandar Lampung, 35145, Indonesia
³ Spectroscopy Research Group (SRG), Laboratory of Bioprocess and Postharvest Engineering, Department of Agricultural Engineering, Faculty of Agriculture, The University of Lampung, Bandar Lampung, Lampung, Indonesia

Corresponding Author: meinilwitayulia@polinela.ac.id

Abstract. Arabica coffee variety is one of the key parameters that highly influence the cup quality of arabica coffee. The objective of this current research is to evaluate the possible application of UV spectroscopy for the discrimination of arabica coffee with different cultivars. Green beans from two arabica coffee cultivars (Coffea arabica) of Typica and Sigarar Utang were collected from the same origin in Papandayan mountain West Java, Indonesia. The samples were subjected to the same postharvest treatments (wet cherry processing method). All samples were roasted in medium roasting with 200°C for 16 minutes using a portable roasting machine. A total of 40 samples of Typica and Sigarar Utang were provided by weighing 1 gram of coffee powder (mesh 40) for each sample. The extraction of coffee samples was performed based on previously reported work. The UV-visible spectral data of aqueous coffee samples were acquired by using a benchtop of UV-vis spectrometer in the range of 190-1100 nm. The first two PCs (principal components) with 94% of CEV (cumulative explained variance) could be used to separate between the Typica and Sigarar Utang samples. The Typica samples were situated on the left of PC1 (PC1<0) and Sigarar Utang samples were on the right of PC1 (PC1>0). In the future, it is promising to apply UV spectroscopy for simple and reliable discrimination of arabica coffee cultivar.

1. Introduction
Coffee is one of the important agricultural products in Indonesia with total area production of 1.3 million hectares (ha) in 2020. The development of better farming practices and proper coffee genotype selection has promoted an improvement in the final coffee cup quality and thus increase the popularity of specialty coffee products, especially for Arabica coffee [1-2]. In Indonesia, several arabica cultivars are available and planted widely such as arabica Typica, Yellow Bourbon, Catimor, and Sigarar Utang. In general, there are two types of analytical methods for coffee cultivar identification: traditional laboratory-based chemical methods and spectroscopy-based analytical methods. The first method including instrumental neutron activation analysis (INAA) [3], liquid chromatography coupled with UV spectrophotometry [4], and sensory data [5]. The spectroscopy-based analytical methods have been reported for coffee variety and cultivar identification from near-infrared spectroscopy to fluorescence spectroscopy [1, 6-14].
However, there is no reported work on the Indonesian cultivar coffee identification using UV spectroscopy. Coffee authentication both for arabica and robusta coffee through UV spectroscopy has been well reported. Yulia et al. [15] reported the promising application of UV spectroscopy and the SIMCA (soft independent modeling of class analogy) method for the classification of ground roasted decaffeinated coffee. UV spectroscopy and PLS-DA (partial least-squares discriminant analysis) method were demonstrated to be suitable for authentication of organic Lampung robusta ground roasted coffee [16]. Suhandy and Yulia [17] utilized support vector machine regression (SVMR) coupled with UV spectroscopy for Luwak coffee content measurement in coffee blends. A similar approach had been used to identify and discriminate between fresh and expired ground roasted robusta coffee [18].

The objective of the research was to develop a simple and relatively fast analytical method using UV spectroscopy and PCA (principal component analysis) for the discrimination of different cultivars of ground-roasted arabica coffee from Indonesia.

2. Materials and methods

2.1. Samples
Two cultivars of arabica coffee of Sigarar Utang and Typica were prepared in this research. Both cultivars were Java Preanger arabica coffee which was harvested in 2020 from the same origin in Papandayan mountain West Java, Indonesia. Roasting of coffee samples was conducted using a portable roaster at a medium level with 200°C for 16 minutes. A total of 40 samples of Typica and Sigarar Utang were provided by weighing 1 gram of coffee powder (mesh 40) for each sample. The extraction of coffee samples was performed based on previously reported work [19]. Figure 1 shows arabica Sigarar Utang and arabica Typica coffee samples before extraction. It is hard to directly discriminate between the two samples by using human naked eyes.

![Arabica Sigarar Utang and Arabica Typica](image)

**Figure 1.** The visual appearance of arabica Sigarar Utang and arabica Typica samples before extraction.

2.2. Spectral data acquisition and preprocessing
The UV-visible spectra of aqueous coffee samples were measured by using a benchtop of a UV-visible spectrometer in the window of 190-1100 nm. Distilled water was used as a reference and it was measured before sample measurement. Three spectral preprocessing were used: moving averaging smoothing (MAS) to improve the signal-noise ratio (SNR), standard normal variate (SNV) to minimize the effects of light scattering, and Savitzky-Golay 1st derivation to enhance a small difference between spectral data of arabica Sigarar Utang and arabica Typica.
2.3. Data analysis

Discrimination of the two cultivars of arabica Sigarar Utang and arabica Typica was conducted using PCA (principal component analysis). First, the PCA was calculated using original spectral data in the window of 190-1100 nm and an influence plot was plotted. This plot is leverage in the x-axis and residual x-variance in the y-axis. Residual x-variance is used to measure the variation that is not taken into account by the PCA model. Leverage is used to measure how extreme a data point is compared to the majority. We assigned a potential outlier for any samples with both high residual x-variance and high leverage value. The second plot is Hotelling’s $T^2$ ellipse. This plot is also useful to identify and remove potential outlier samples [20-21]. Using those two plots, detection of outlier samples was possible. Further PCA calculation was performed using all samples of arabica Sigarar Utang and arabica Typica without outlier samples. Scores plot of the first two PCs (principal components) was created. To select the appropriate number of PC included in the calculation of PCA, the percentages of cumulative explained variance (CEV) of the PC were used (CEV should be 90% or more [18]). Software of The Unscrambler 9.8 (CAMO AS, Oslo, Norway) was used to calculate spectral preprocessing, outlier detection, and PCA analysis.

3. Results and discussion

3.1. Spectra of arabica Sigarar Utang and Typica

Spectral data of all samples of arabica Sigarar Utang and arabica Typica in the window of 190-1100 nm was depicted in Figure 2. The absorbance intensity of arabica Sigarar Utang is higher than that of arabica Typica especially in the window of 250-400 nm which is in line with the result of the previously reported study [19]. Figure 3 shows the preprocessed spectra of the samples in the window of 250-400 nm. In this window, several wavelengths with high peaks were observed. Those peaks were in line with the findings of the previously reported works. For instance, the positive peak at a wavelength of 270 nm was in correspondence with the absorbance of caffeine [22]. The positive peak at 315 nm and the negative peak at 345 nm is related to the absorbance of caffeic acid or trigonelline [23].

![Figure 2. Spectra of arabica Sigarar Utang and Typica coffee samples obtained directly from UV-visible spectral measurement (without preprocessing) in the window of 190-1100 nm.](image-url)
3.2. Detection of outlier sample

Figure 4 shows the plot of influence showing leverage in the $x$-axis and residual $x$-variance in the $y$-axis coming from PCA calculation using original spectra in the window of 190-1100 nm. This plot can be used to identify an occurrence of potential outlier samples. Any samples with high values both in leverage and residual $x$-variance can be associated with a potential of outlier samples. One sample of arabica Sigarar Utang sample (SU1a) has a high value both in leverage and residual $x$-variance. Two samples of arabica Typica (AT1a and AT1b) have a high value of leverage but an acceptable value for residual $x$-variance. However, the result of Hotelling’s $T^2$ ellipse as depicted in Figure 5 showed that the two samples of AT1a and AT1b were lied outside the confidence ellipse and hence confirmed as potential outlier samples. All potential outlier samples were excluded from further PCA calculation.
3.3. **PCA calculation**

The potential outlier samples (one sample of arabica Sigarar Utang and two samples of arabica Typica) were removed from data analysis. The result of PCA calculation without outlier samples was plotted in Figure 6. The calculation was performed using preprocessed spectral data of 37 samples of arabica Sigarar Utang and arabica Typica in the window of 250-400 nm. The first two PCs (principal components) with 94% of CEV (cumulative explained variance) could be used to separate between the Sigarar Utang and Typica samples. The Sigarar Utang samples were on the right of PC1 (PC1>0) and Typica samples were situated on the left of PC1 (PC1<0) resulted in 100% of accuracy (calculated at 5% of significance level). PCA and UV spectroscopy was effective to do the unsupervised classification of Indonesian specialty arabica coffee concerning the origin of the cultivar. Our finding was in line with several previously reported works. Edelmann et al. [24] utilized UV-visible spectroscopy and chemometrics to discriminate Austrian red wines with different cultivars with an acceptable result. Zhang et al. [9] used mid-infrared spectroscopy and PCA to distinguish four coffee varieties including Typica and Catimor from China with a promising result.

**Figure 4.** A plot of leverage versus residual x-variance calculated using all samples of original spectra in the window of 190-1100 nm.

**Figure 5.** PCA scores plot of all samples (40 samples) of arabica Sigarar Utang and Typica using original spectral data in the window of 190-1100 nm with 95% confidence interval of the Hotelling’s T² ellipse. Samples outside the confidence ellipse are considered outliers.
Figure 6. The score plots of PCA analysis without outlier samples (PC1xPC2) calculated using preprocessed spectral data in the window of 250-400 nm.

4. Conclusion
Feasible discrimination of two arabica cultivar coffees from West Java, Indonesia has been well demonstrated using UV spectroscopy. One important step before PCA calculation is outlier sample detection. Here, we showed that plot of residual x-variance versus leverage and the plot of Hotelling’s $T^2$ ellipse is quite appropriate to be used for detecting a potential outlier sample. In the future, it is promising to apply UV spectroscopy for simple and reliable discrimination of arabica coffee cultivar. However, this present research used a small number of samples with the same geographical origin. Therefore, to clarify this finding, the next research should be conducted with more samples with more cultivars from different geographical origins.

Acknowledgements
The authors acknowledge the Ministry of Education, Culture, Research and Technology/BRIN, the Republic of Indonesia for financial support (PKPT Grant 2020-2022).

References
[1] Marquetti I, Link J V, Lemes A L G, Scholz M B S, Valderrama P and Bona E 2016 Comput. Electron. Agric. 121: 313–9
[2] Moreira I and Scarminio I S 2013 Talanta 107: 416–22
[3] Fernandes E A D N, Tagliaferro F S, Azevedo-Filho A and Bode P 2002 Accred. Qual. Assur. 7: 378–87
[4] Alonso-Salces R M, Serra F, Reniero F and Heberger K 2009 J. Agric. Food Chem. 57: 4224–35
[5] Wada K, Ohgama S, Sasaki H, Shimoda M and Osajima Y 1987 Agric. Biol. Chem. 51: 1745–52
[6] Bona E, Marquetti I, Link J V, Makimori G Y F, da Costa Arca V, Guimarães Lemes A L, Ferreira J M G, dos Santos Scholz M B, Valderrama P and Poppi R J 2017 LWT-Food Sci. Technol.
76: 330–6

[7] dos Santos Scholz M B, Kitzberger C S G, Pereira L F P, Davrieux F, Pot D, Charmetant P and Leroy T 2014 J. Near Infrared Spectrosc. 22: 411–21
[8] Esteban-Díez I, González-Sáiz J M, Sáenz-González C and Pizarro C 2007 Talanta 71: 221–9
[9] Zhang C, Wang C, Liu F and He Y 2016 J. Spectrosc. 2016 7927286
[10] Monakhova Y B, Ruge W, Kuballa T, Ilse M, Winkelmann O, Diehl B, Thomas F and Lachenmeier D W 2015 Food Chem. 182: 178–84
[11] Luna A S, da Silva A P, da Silva C S, Lima I C A and de Gois J S 2019 J. Food Compos. Anal. 76: 44–50
[12] Figueiredo L F, Borém F M, Almeida M R, de Oliveira L F C, Alves A P C, dos Santos C M and Rios P A 2019 Food Chem. 288: 262–7
[13] Zhang C, Shen T, Liu F and He Y 2018 Sensors 18 95
[14] Suhandy D and Yulia M 2018 IOP Conf. Ser.: Mater. Sci. Eng. 334 012059
[15] Yulia M, Asnaning A R and Suhandy D 2018 IOP Conf. Ser.: Earth Environ. Sci. 147 012010
[16] Yulia M and Suhandy D 2019 J. Phys.: Conf. Ser. 1341 022006
[17] Suhandy D and Yulia M 2019 J. Phys.: Conf. Ser. 1341 022006
[18] Yulia M and Suhandy D 2018 MATEC Web Conf. 197 09003
[19] Suhandy D and Yulia M 2018 MATEC Web Conf. 197 09002
[20] Karunathilaka S R, Yakes B J, He K, Chung J K and Mossoba M 2018 Heliyon 4 e00806
[21] Basri K N, Hussain M N, Bakar J, Sharif Z, Khir M F A and Zoolfakar A S 2017 Spectrochim. Acta A Mol. Biomol. Spectrosc. 173 335
[22] López-Martínez L, López-de-Alba P L, García-Campos R and León-Rodríguez L M D 2003 Anal. Chim. Acta 493: 83–94
[23] Suhandy D, Yulia M and Kusumiyati 2018 AIP Conf. Proc. 2021 060010
[24] Edelmann A, Diewok J, Schuster K C and Lendl B 2001 J. Agric. Food Chem. 49:1139–45