The Classification of Ground Roasted Decaffeinated Coffee Using UV-VIS Spectroscopy and SIMCA Method

The Classification of Ground Roasted Decaffeinated Coffee Using UV-VIS Spectroscopy and SIMCA Method

M Yulia[1], A R Asnaning[1], and D Suhandy[2]

[1] Department of Agricultural Technology, Lampung State Polytechnic, Indonesia.
[2] Department of Agricultural Engineering, The University of Lampung, Indonesia

E-mail: meinilwitayulia@polinela.ac.id

Abstract. In this work, an investigation on the classification between decaffeinated and non-decaffeinated coffee samples using UV-VIS spectroscopy and SIMCA method was investigated. Total 200 samples of ground roasted coffee were used (100 samples for decaffeinated coffee and 100 samples for non-decaffeinated coffee). After extraction and dilution, the spectra of coffee samples solution were acquired using a UV-VIS spectrometer (Genesys™ 10S UV-VIS, Thermo Scientific, USA) in the range of 190-1100 nm. The multivariate analyses of the spectra were performed using principal component analysis (PCA) and soft independent modeling of class analogy (SIMCA). The SIMCA model showed that the classification between decaffeinated and non-decaffeinated coffee samples was detected with 100% sensitivity and specificity.

1. Introduction

Coffee is the most frequently consumed functional food worldwide [1]. Caffeine is one of the functional ingredients in coffee and one of the main alkaloid in coffee beans, accounting for 1 to 4% (dry basis), with large variation within cultivars and among them [2-4]. Caffeine contents are strongly related to the quality of coffee beverages because it contributes to its bitterness [5].

It has been reported that caffeine gives benefits to our health. Caffeine was reported well for increasing alertness, through stimulation of the central nervous system, rising blood circulation and respiration, is probably the main reason for coffee popularity [2, 6]. Other possible benefits of caffeine include mood enhancement, better exercise performance and reaction time, and reduction of symptoms associated with Parkinson's disease and tremors [7]. For more detailed report on positive effects of moderate consumption of caffeine has been published [8].

However, caffeine has also some negative effects such as sleeplessness and mild addiction, which has prompted development of a decaffeinated coffee industry (estimated for around 10–15% of the total amount of coffee consumed in the world). High doses of caffeine also cause anxiety, restlessness, tension, nervousness, and psychomotor agitation [9], while long-term use of this alkaloid may increase the risk of cardiovascular diseases, with individual differences in caffeine response, probably related to genetic factors [10].

The adverse side effects of caffeine [11] have increased the market for decaffeinated coffee to about 10% of coffee consumption worldwide (www.ncausa.org). Recently, the need for decaffeinated coffee authentication is increasing due to the popularity of decaffeinated coffee and its higher price than that of non-decaffeinated coffee. The purpose of this authentication is to guarantee the purity of

decaffeinated coffee and to determine whether it has been mixed with non-decaffeinated coffee in order to satisfy food quality and safety requirements [12].

Several methods have been conducted to determine caffeine content in coffee beverage [13] and in roasted coffee [14-15]. Most of the previously reported studies for caffeine content determination, however, involved the use of high-cost devices such as FT-NIR (Fourier transform near infrared). On the other hand, a relatively low cost analytical based method using UV-VIS spectroscopy which utilizes wavelength region 200-700 nm has been used for coffee authentication in civet coffee [16] and in peaberry coffee [17]. The results were satisfied with coefficient of determinations were more than 0.90 and RPDs were more than 3.0 [16-17].

However, to our knowledge, no studies are available on the application of low-cost analysis based on UV-VIS spectroscopy to discriminate between decaffeinated and non-decaffeinated coffee samples. Therefore, the objective of this study was to investigate the feasibility of using UV-VIS spectroscopy and SIMCA method to classify decaffeinated and non-decaffeinated coffee samples.

2. Materials and Methods

2.1. Samples

In this research, a number of 200 samples of ground roasted decaffeinated and non-decaffeinated coffee was provided (1 gram weight for each sample). Since particle size in coffee powder has a significant influence on the spectra obtained, we sieved all coffee samples through a nest of U. S. standard sieves (mesh number of 40) on a Meinzer II sieve shaker (CSC Scientific Company, Inc. USA) for 10 minutes to obtain a particle size of 420 μm. These experiments were performed at room temperature (around 27-29°C). An aqueous extraction procedure of the coffee samples was conducted based on reference [16-17].

2.2. UV-VIS spectral acquisition

The UV-VIS spectral data of aqueous coffee samples were obtained in the range of 200-700 nm by using a UV-VIS spectrometer (Genesys™ 10S UV-Vis, Thermo Scientific, USA). This spectrometer was equipped with a quartz cell with optical path of 10 mm. The spectral acquisition was done at a spectral resolution of 1 nm at a room temperature. The original spectra (without any preprocessing) were used for further analysis. All spectral acquisitions were conducted at room temperature (around 28°C).

2.3. Chemometric analysis

The chemometric analysis was performed by using software The Unscrambler 9.8 (CAMO, AS, Norway). In order to verify the repeatability of the UV-VIS spectral data, we took 2 measurements for each sample and obtained mean values for further analysis. All data acquired were then classified into decaffeinated or non-decaffeinated class based on result from algorithm of 'soft independent modeling class analogy' (SIMCA). For SIMCA, the samples were randomly divided into two groups: calibration set (70 samples for decaffeinated and non-decaffeinated, respectively) and prediction set (30 samples for decaffeinated and non-decaffeinated, respectively).

3. Results and Discussion

3.1. Spectral data of decaffeinated and non-decaffeinated coffee in the range of 200-700 nm

Figure 1 shows spectral data of 200 samples of decaffeinated and non-decaffeinated coffee in the range of 200-700 nm. In the left part of figure 1, it is difficult to see any difference between the two types of coffee since that all spectra have similarity in shape. For this reason, the average of 100 samples of decaffeinated and 100 samples of non-decaffeinated was calculated and the result is shown in right part of figure 1. It is noted that the overall intensity of decaffeinated and non-decaffeinated coffee samples was very close each other except in the region 280-350 nm. In this part, there is a

significant difference between decaffeinated and non-decaffeinated coffee samples. The average non-decaffeinated coffee sample has higher absorbance than that of the average decaffeinated coffee sample. For this reason, the region 280-350 nm may give us a good separation between decaffeinated and non-decaffeinated coffee samples. The wavelength at around 280 nm may refer to the absorbance of caffeine, while 310 and 350 nm are closely related to the absorbance of caffeic acid and chlorogenic acids (CGA), respectively [18-19].

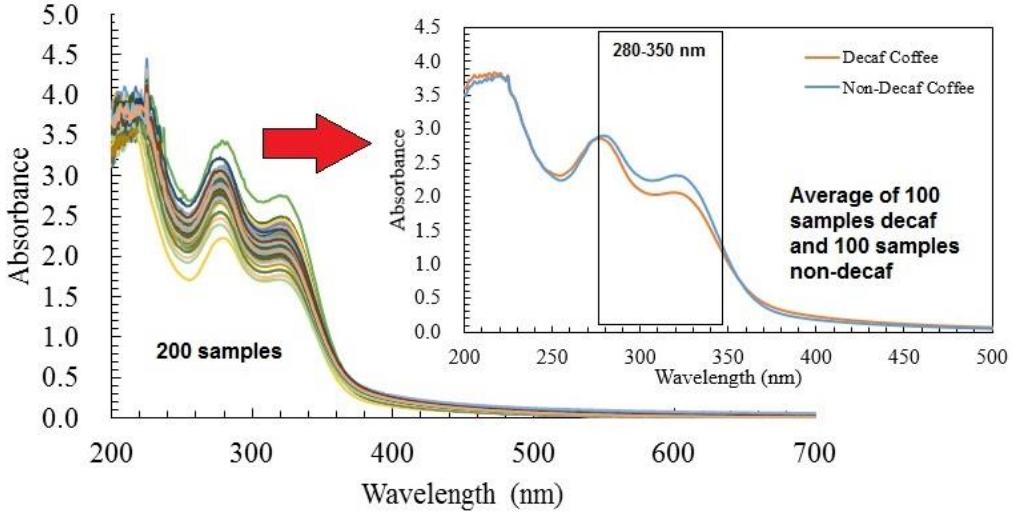

Figure 1. Spectral data of decaffeinated and non-decaffeinated coffee samples in the range of 200-700 nm.

3.2. Principal component analysis (PCA)

Figure 2 shows the result of PCA for all samples (200 samples) in the range of 280-350 nm with full-cross validation method. PCA was performed in order to investigate the presence of outliers and any trend of discrimination between decaffeinated and non-decaffeinated coffee samples. PC1 accounted for 59% of the variation explanation and described most variation in the calibration spectra. PC2 represented 25% of the variation in the spectra. It can be seen that using the two PCs, there is a clear separation between decaffeinated and non-decaffeinated coffee samples. The 95% confidence ellipse is also presented in figure 2 and it is based on Hotelling's T-square statistics. Hotelling's T-square statistic is a measure of the variation of a sample within a PLS model and a high T-square level for a certain sample indicates a high influence of the sample on the model. A number of 25 samples found outside the ellipse (high T-square level) and it was identified as outliers in this study. We decided to remove those potential outliers for further analysis. Thus, the remaining samples were 175 samples and it was used for SIMCA. Figure 3 shows the result of PCA after removing 25 samples (using 175 samples). It can be seen that using two PCs (PC1=98% and PC2=2%), the separation between decaffeinated (red color) and non-decaffeinated (blue color) coffee samples are very clear. The decaffeinated coffee was mostly located in the left part of PC1 (PC1<0) while non-decaffeinated coffee was completely positioned on the right part of PC1 (PC1>0).

Figure 4 shows the loadings plot of first two PCs. Examining the loading plot, the main wavelength region responsible for the separation of the samples were 300-350 nm with the peak located at around 325 nm.

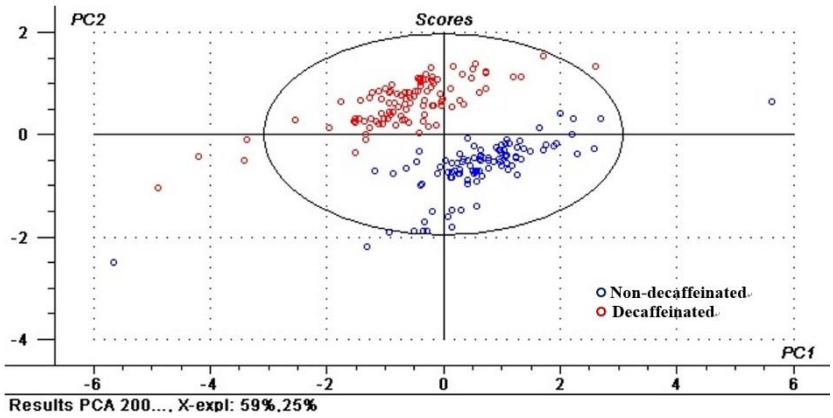

Figure 2. The result of PCA on all samples (200 samples decaffeinated and non-decaffeinated coffee) in the range wavelength of 280-350 nm.

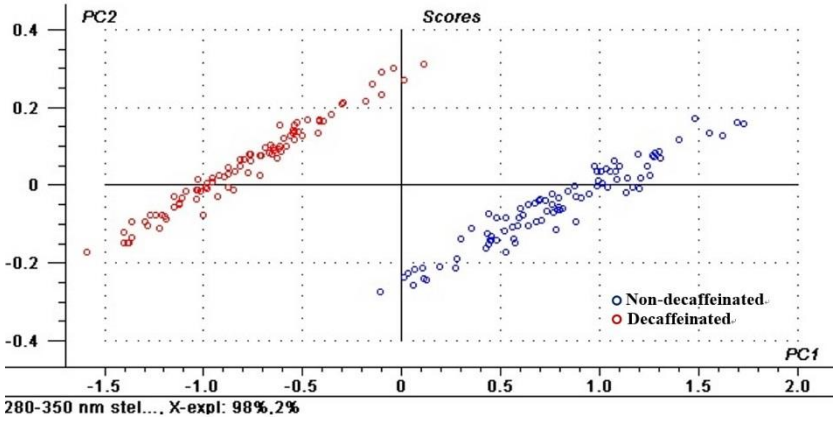

Figure 3. The result of PCA on remaining samples (175 samples decaffeinated and non-decaffeinated coffee) in the range wavelength of 280-350 nm.

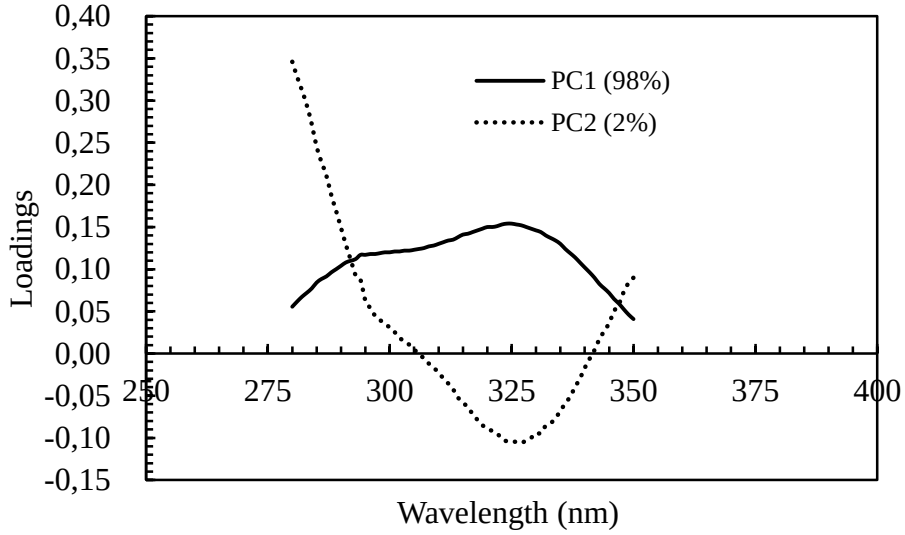

Figure 4. The loading plot of decaffeinated and non-decaffeinated coffee samples in the plane defined by the first two principal components.

3.3. Developing and evaluating SIMCA model for each class

For SIMCA, first, we developed SIMCA model for each class. For this purpose, we performed PCA for the calibration set for each class. SIMCA model decaffeinated coffee was developed using PCA with 60 samples. SIMCA model non-decaffeinated coffee was developed using PCA with 64 samples. The performance of SIMCA model was evaluated using three parameters: accuracy, sensitivity, and specificity using prediction sample set (23 samples for decaffeinated and 28 samples for non-decaffeinated). The accuracy is the proportion of both true positives (23 samples for decaffeinated) and true negatives (28 samples for non-decaffeinated) among the total number of cases examined (total 51 samples). The sensitivity is the SIMCA model ability to correctly classify the samples, relating the predicted samples to being in a class with the samples that actually are in this class. The specificity is the SIMCA model ability to correctly classify the samples, relating the predicted samples to not being in a class with the samples that actually are not in this class [20].Table 1 shows the confusion matrix of the result of classification using the developed SIMCA model. The confusion matrix presented in Table 1 enabled to detect if some samples of a given class are confused with another class. It can be seen that there were no any confusing samples. All prediction samples were correctly identified without confusion (no false positive or false negative) resulted in 100% of accuracy, sensitivity, and specificity.

Table 1. Confusion matrix of prediction samples calculated using developed SIMCA model.

Items	Decaffeinated Classified by using SIMCA model	Non-Decaffeinated Classified by using SIMCA model
Decaffeinated Class (actual)	23	0
Non-Decaffeinated Class (actual)	0	28

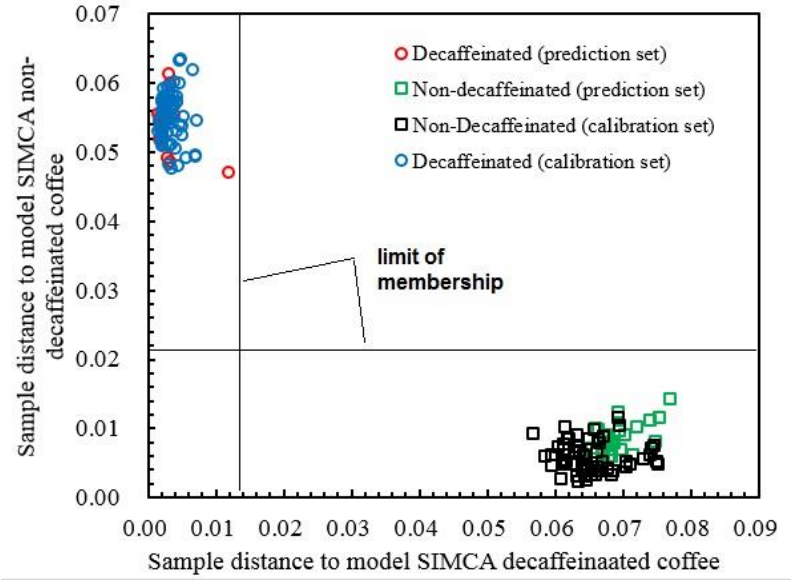

Figure 5. Coomans plot of decaffeinated and non-decaffeinated coffee samples calculated using developed SIMCA model.

The result from the SIMCA analysis was also presented in a plot called the Coomans plot for decaffeinated and non-decaffeinated models, where distances between two classes were plotted against each other in a scores plot (figure 5). For this, first, each limit of membership (for decaffeinated and non-decaffeinated coffee) was calculated. Samples in the prediction set are projected into each SIMCA

model and their residual distance calculated. The residual standard deviation was used as a measure of the critical distance for classification (limit of membership). If the residual distance from the model was below their statistical limit, the sample was defined to that class and if the residual variance was higher, then the sample did not belong to that class [16]. The Coomans plots also indicated the large distances between the decaffeinated and non-decaffeinated classes with 99% confidence level.

4. Conclusion

In this research, we propose a rapid differentiation and classification method for decaffeinated and non-decaffeinated coffee using a relatively low-cost analysis method based on UV-VIS spectroscopy and SIMCA method. PCA and SIMCA are commonly used to analyze chemical data by displaying the predominant components to provide accurate evaluation. Within this research, our results confirmed high differentiation capacity of PCA and superior classification potential of SIMCA applied to discrimination between decaffeinated and non-decaffeinated coffee. The developed SIMCA model of decaffeinated and non-decaffeinated coffee could be used well to predict the class of new samples without any confusion. The obtained classification results were very good with 100% rate for accuracy, sensitivity, and specificity. This research has demonstrated that UV-VIS spectroscopy is found to be an efficient instrument for decaffeinated and non-decaffeinated classification and it can be recommended for practice.

Acknowledgement

The present research was supported by Kemenristekdikti Indonesia (Hibah Penelitian Dosen Pemula (PDP) 2017 (Nomor: 065.8/PL15.8/LT/2017)).

References

[1] Bisht S and Sisodia SS 2010 *J Nat Pharm* **1** 58–65.
[2] Belitz H D, Grosch W and Schieberle P 2009 *Food chemistry* (4th ed.). (Heidelberg: Springer).
[3] Dessalegn Y, Labuschagne M T, Osthoff G and Herselman L 2008 *J. Sci. Food Agric.* **88** 1726–30.
[4] Mazzafera P and Silvarolla M B 2010 *Seed Sci. Res.* **20** 163–7.
[5] Farah A, Monteiro M C, Calado V, Franca AS and Trugo L C 2006 *Food Chem.* **98** 373–80.
[6] Reich M S, Dietrich M S, Finlayson A J R, Fischer E F and Martin P R 2008 *Alcohol. Clin. Exp. Res.* **32** 1799–1806.
[7] Heckman M A, Weil J and González de Mejía E 2010 *J. Food Sci.* **75** R77–R87.
[8] Glade M J 2010 *Nutrition* **26** 932–8.
[9] Daly J W and Fredholm B B 1998 *Drug Alcohol Depend.* **51** 199–206.
[10] Yang A, Palmer A A and de Wit H 2010 *Psychopharmacology* **211** 245–527.
[11] James JE 1991 *Caffeine and Health* (Academic: San Diego).
[12] Pizarro C, Esteban-Diez I, Gonzalez-Saiz JM and Forina M 2007 *J Agric. Food Chem*. **55** 7477–88.
[13] Wang D, Min S G, Duan J, Xiong Y M and Li Q Q 2012 *Guang Pu Xue Yu Guang Pu Fen Xi/Spectroscopy and Spectral Analysis* **32** 982–4.
[14] Zhang X, Li W, Yin B, Chen W, Kelly DP and Wang X 2013 *Spectrochim Acta A Mol Biomol Spectrosc.* **114** 350–6.
[15] Garrigues J M, Bouhsain Z, Garrigues S and de la Guardia M 2000 *J Anal. Chem.* **366** 319–22.
[16] Suhandy D and Yulia M 2017 *Int. J. Food Prop.* **20** S331–339.
[17] Suhandy D and Yulia M 2017 *Int. J Food Sci.*
[18] Belay A, Ture K, Redi M and Asfaw A 2008 *Food Chem.* **108** 310–315.
[19] Souto U T C P, Barbosa M F, Dantas H V, Pontes A S, Lyra W S, Diniz P H G D, Araújo M C U and Silva E C 2015 LWT - *Food Sci. Technol.* **63** 1037–41.
[20] Bona E, Marquetti I, Link J V, Makimori G Y F, Arca V C, Lemes A L G, Ferreira J M G, Scholz M B S, Valderrama P and Poppi R J 2017 LWT - *Food Sci. Technol.* **76** 330–336.