

Article

Classification of *Lampung robusta* Specialty Coffee According to Differences in Cherry Processing Methods Using UV Spectroscopy and Chemometrics

Diding Suhandy ^{1,*}  and Meinilwita Yulia ² 
¹ Department of Agricultural Engineering, Faculty of Agriculture, The University of Lampung, Jl. Soemantri Brojonegoro No.1, Bandar Lampung 35145, Indonesia

² Department of Agricultural Technology, Lampung State Polytechnic, Jl. Soekarno Hatta No. 10, Rajabasa Bandar Lampung 35141, Indonesia; meinilwitayulia@polinela.ac.id

* Correspondence: diding.sugandy@fp.unila.ac.id; Tel.: +62-0813-7334-7128

Abstract: The postharvest processing factors including cherry processing methods highly influence the final quality of coffee beverages, especially in the composition of several coffee metabolites such as glucose, fructose, the amino acid (glutamic acid), and chlorogenic acids (CGA) as well as trigonelline contents. In this research, UV spectroscopy combined with chemometrics was used to classify a ground roasted *Lampung robusta* specialty coffee according to differences in the cherry processing methods. A total of 360 samples of *Lampung robusta* specialty coffee with 1 g of weight for each sample from three different cherry processing methods were prepared as samples: 100 samples of pure dry coffee (DRY), 100 samples of pure semi-dry coffee (SMD), 100 samples of pure wet coffee (WET) and 60 samples of adulterated coffee (ADT) (SMD coffee was adulterated with DRY and WET coffee). All samples were extracted using a standard protocol as explained by previous works. A low-cost benchtop UV-visible spectrometer (Genesys™ 10S UV-Vis, Thermo Scientific, Waltham, MA, USA) was utilized to obtain UV spectral data in the interval of 190–400 nm using the fast scanning mode. Using the first three principal components (PCs) with a total of 93% of explained variance, there was a clear separation between samples. The samples were clustered into four possible groups according to differences in cherry processing methods: dry, semi-dry, wet, and adulterated. Four supervised classification methods, partial least squares–discriminant analysis (PLS-DA), principal component analysis–linear discriminant analysis (PCA-LDA), linear discriminant analysis (LDA) and support vector machine classification (SVMC) were selected to classify the *Lampung robusta* specialty coffee according to differences in the cherry processing methods. PCA-LDA is the best classification method with 91.7% classification accuracy in prediction. PLS-DA, LDA and SVMC give an accuracy of 56.7%, 80.0% and 85.0%, respectively. The present research suggested that UV spectroscopy combining with chemometrics will be highly useful in *Lampung robusta* specialty coffee authentication.

Keywords: UV spectroscopy; PLS-DA; PCA-LDA; LDA; SVMC; dry processing; wet processing; semi-dry processing; Lampung specialty coffee



Citation: Suhandy, D.; Yulia, M. Classification of *Lampung robusta* Specialty Coffee According to Differences in Cherry Processing Methods Using UV Spectroscopy and Chemometrics. *Agriculture* **2021**, *11*, 109. <https://doi.org/10.3390/agriculture11020109>

Academic Editor: Massimo Cecchini

Received: 28 December 2020

Accepted: 26 January 2021

Published: 1 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Coffee (*Coffea* sp.) is a widely consumed beverage with two popular varieties being planted worldwide: arabica (*Coffea arabica*) with 57% of global coffee production and robusta coffee (*Coffea canephora*) with 43% of global coffee production [1]. According to the USDA (United States Department of Agriculture) [2], in 2019/2020 Indonesia shared 6.10% of world coffee production with a total production of 0.642 million tons. In Indonesia, coffee beans, mostly robusta coffee, were mainly produced in six important provinces: South Sumatera, Lampung, Bengkulu, North Sumatera, East Java, and South Sulawesi.

In Lampung, three different robusta coffee cherry processing methods are used: dry or natural (DRY), semi-dry (SMD), and wet or washed processing (WET). In dry process-

ing, bean processing including cherry sortation (removing defective and immature fruit), cherry drying (usually using open sun drying for 3–9 days), hulling or peeling, green bean sortation (removing a defective green bean), and packing. In wet processing, bean processing includes cherry sortation, pulping, fermentation (12–48 h equipped with a controlled water flow), washing, drying, peeling and polishing, green bean sortation, and packing [3–5]. Semi-dry processing is a relatively new processing that combines the best aspects of the dry and wet processes. In semi-dry processing, bean processing includes cherry sortation, pulping (without removing the mucilage), drying, peeling and polishing, green bean sortation, and packing [6].

The postharvest processing factors including cherry processing methods highly influence the final quality of coffee beverages, especially in the composition of several coffee metabolites such as glucose, fructose, the amino acid (glutamic acid), and chlorogenic acids (CGA) as well as trigonelline contents. For example, in general, coffees with wet processing methods have higher acidity and full aroma as a result of fermentation [7]. The glucose and fructose contents with dry processing are higher than the wet one [5]. It is noted that the final content of glucose and fructose was reduced drastically (up to 80–90%) in wet processing coffee. In the term of caffeine content, several works have shown that caffeine contents are not significantly different between dry and wet processing [5]. CGA contents in wet processing coffee were significantly higher than those processed by the semi-dry method [8]. In contrast, trigonelline was found to be decreased by wet processing. According to Bytof et al. [9], bean processing influences the amino acid contents. In an individual compound, during dry processing, glutamic acid decreases rapidly. During the roasting process, coffee cherry processing was also reported to influence the formation of furfuryl alcohol, a carcinogen compound (Group 2B), with dry processing showing a lower production of furfuryl alcohol as opposed to the wet processing [10]. In terms of cup quality, it was reported that the wet processed coffee was regarded as having a better quality compared to the dry [8]. However, semi-dry processed coffee is now becoming more popular with the most expensive coffees, especially for espresso coffee [11].

Several methods have been reported to discriminate between natural arabica, washed arabica, and robusta varieties using NIR spectroscopy and electronic devices (e-nose and e-tongue) with a satisfactory result (the percentage of correctly classified samples was almost 100% for green and roasted coffee) [12]. Flambeau et al. [13] utilized e-nose/e-tongue combined with principal component analysis (PCA) and discriminant factorial analysis (DFA) to discriminate ground roasted coffee with different cultivar, origin, and processing with an acceptable result. Those reported works involved a relatively expensive device to perform the analysis.

A relatively affordable analytical method based on UV spectroscopy is preferred to discriminate *Lampung robusta* coffee according to its cherry processing methods. UV spectroscopy has been used for specialty coffee authentication, especially for ground roasted coffee. Suhandy and Yulia [14–17] studied the application of UV spectroscopy and chemometrics for authentication of Luwak coffee, peaberry coffee, several Indonesian specialty coffees with a geographic indication (GIs) and discriminating Gayo wine and normal coffee.

However, there is no report on the use of UV spectroscopy for discrimination between dry, semi-dry, and wet processing methods and their adulteration. Therefore, this study aimed to evaluate the practical application of UV spectroscopy and several classification methods for classification of *Lampung robusta* coffee according to differences in cherry processing methods. Four classifications of PLS-DA (partial least squares–discriminant analysis), PCA-LDA (principal component analysis–linear discriminant analysis), LDA (linear discriminant analysis), and SVMC (support vector machines classification) were tested and its comparative performance was evaluated.

2. Materials and Methods

2.1. Samples

A total of 360 samples of *Lampung robusta* coffee with 1 g of weight for each sample from three different cherry processing methods were prepared as samples: 100 samples of pure dry coffee (DRY), 100 samples of pure semi-dry coffee (SMD), 100 samples of pure wet coffee (WET) and 60 samples of adulterated coffee (ADT) (SMD coffee was adulterated with DRY and WET coffee). The samples were collected from the same harvest season in Sumber Jaya coffee plantation, West Lampung, Lampung (5°00'28.5" S 104°28'37.4" E). The samples were belonging to premium grade (first grade) by maintaining the number of defective beans as low as 11 scores according to Indonesian National Standard for coffee bean (ISN No. 01-2907:2008). The composition of pure and adulterated samples was shown in Table 1 along with its standard deviation.

Table 1. Composition of pure and adulterated samples with three different cherry processing methods.

Samples	Samples Code	DRY (g)	WET (g)	SMD (g)	Number of Samples
DRY	DRY1–DRY100	1 ± 0.0001	0	0	100
WET	WET1–WET100	0	1 ± 0.0001	0	100
SMD	SMD1–SMD100	0	0	1 ± 0.0001	100
ADT	ADT1–ADT10	0.05 ± 0.0001	0	0.95 ± 0.0003	10
	ADT11–ADT20	0.05 ± 0.0003	0.05 ± 0.0003	0.90 ± 0.0002	10
	ADT21–ADT30	0.075 ± 0.0003	0.075 ± 0.0002	0.85 ± 0.0003	10
	ADT31–ADT50	0.1 ± 0.0001	0.1 ± 0.0002	0.80 ± 0.0002	20
	ADT51–ADT60	0.125 ± 0.0002	0.125 ± 0.0002	0.75 ± 0.0002	10

It has been well reported that the quality of coffee flavor was highly affected by roasting conditions [18,19]. For this reason, in this research, all samples were roasted in the same condition: 200 °C for 20 min using a portable roasting machine. All samples have a homogenous particle size of 0.297 mm by grinding and sieving using 50 mesh [20].

2.2. Coffee Extraction using Distilled Water

All samples were extracted using a protocol as explained by previous works [14–17]. For each sample, 50 mL of hot distilled water (98 °C) was added and then well stirred for 10 min using CiBlanc magnetic stirrer. The extracted samples were filtered and diluted with distilled water with the proportion 1:20 mL. About 3 mL of diluted samples were prepared for spectral measurement.

2.3. Spectral Measurement using UV-Visible Spectrometer

A low-cost benchtop UV-visible spectrometer (Genesys™ 10S UV-Vis, Thermo Scientific, Waltham, MA, USA) was utilized to obtain UV spectral data in the interval of 190–400 nm using fast scanning mode. Reference was measured using a distilled water. The absorbance data was used for further analysis [21].

2.4. Chemometrics

PCA (principal component analysis) was used to perform unsupervised pattern recognition. The calculation of PCA using the following parameters: 10 number of principal components (PCs) and leave-one-out cross validation. Four supervised classifications were performed using PLS-DA, PCA-LDA, LDA, and SVMC. The reliability of each classification model was validated using validation procedure. PLS-DA works based on a PLS regression algorithm which searches for latent variables (LVs) with a maximum covariance with the Y-variables. It was chosen because it has been satisfactorily applied in the field of food analysis, as mentioned in previous works [22,23]. LDA and PCA-LDA are popular classical statistical methods for feature extraction and dimension reduction and mostly employed among many supervised pattern recognition methods [24]. In LDA and PCA-LDA, the variance between the categories to be maximized and the variance within the categories to be minimized [25]. The main drawback for LDA and PCA-LDA is only well working when the number of variables is fewer than the number of samples. It was mentioned by

Harvey et al. [26] that for LDA and PCA-LDA, in order to avoid model over-fitting, it is required that the number of samples have to be at least twice as many as the number of variables. SVM is one of the machine-learning methods that can be operated with relatively small datasets. It has recently become popular and widely used and investigated because of its ability in prediction for both, classification and regression [27]. Two SVMC types are available in the Unscrambler: type 1 (C-SVMC) and type 2 (nu-SVMC). In this study, the SVM classification type 2 was used as this type minimizes the error function. The nu value (lower bound on correct classified support vectors and an upper bound on misclassified samples) was set to 0.5 (default value), and the linear function kernel was applied as the optimal method. To select the appropriate gamma value (γ), a grid search was used. A detailed explanation of those methods can be found in several reported works [28–31]. The accuracy of each classification methods was calculated using the following equation [32]:

$$\text{Accuracy (\%)} = \frac{\text{Number of correct classification}}{\text{Number of total samples}} \times 100\% \quad (1)$$

2.5. Software

All chemometrics were performed using the Unscrambler ver. 9.8 and ver. 10.4 (CAMO, Oslo, Norway).

3. Results and Discussion

3.1. Spectral Data of Coffee Samples with Different Cherry Processing

Figure 1 shows the original spectral data of all samples in the interval of 190–400 nm obtained directly from spectral acquisition system. As reported by Shawky and Selim [33], the typical feature of original spectra is rich in unrelated information such as background information and systematic noise coming from the influences of light scattering, different in path length, sample particle size, and other factors. In general, the obtained original spectral data were overlapped and it was hard to discriminate the spectral data according to differences in cherry processing methods. Spectral data with high noise levels were identified at the beginning of wavelengths in the interval of 190–230 nm. The source of such noise might be coming from the low lamp intensities in that interval used in the spectral acquisition. The spectral data with very low absorbance were identified after a wavelength of 350 nm. For this reason, spectral data in the interval of 230–350 nm with relatively containing low noise levels were used for further analysis.

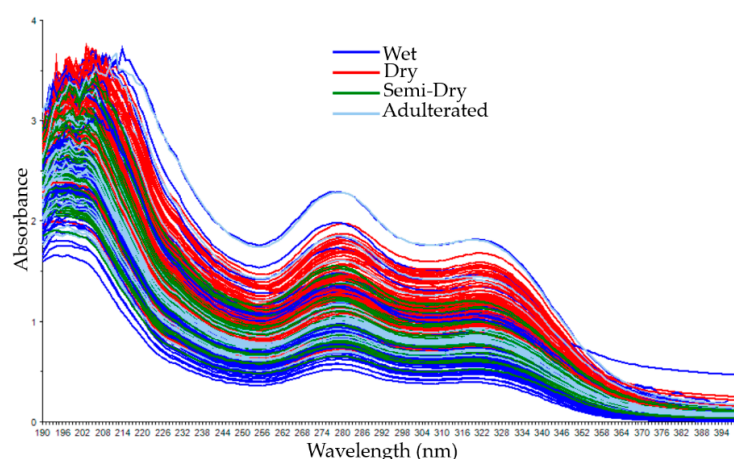


Figure 1. Original spectral data of all samples with different cherry processing methods in the interval of 190–400 nm.

To improve the quality of original spectral data, three spectral pre-treatments were applied simultaneously: Savitzky–Golay smoothing with smoothing points: 5 segments (SGS), standard normal variate (SNV), and Savitzky–Golay first derivative with a second-

order polynomial and a window size of 5 points (SG 1d). According to Santos et al. [34], SGS is effectively improving the signal-noise ratio (SNR) while SNV is similar to multiplicative signal correction (MSC). It can minimize the effects of light scattering. SG 1d is used to correct baseline offsets and to enhance small spectral differences [33,34]. Due to similarity in cherry processing methods—especially for the wet and semi-dry method—it was expected that the spectral difference in coffee samples due to differences in cherry processing methods was small. This is the main reason to use SG 1d: to enhance those small spectral differences. However, at the same time, as a consequence of derivation, the noises were also enhanced. To avoid this, the spectra were first smoothed using SGS pre-treatment as recommended by previous work [33,35]. Instead of selecting the best pre-treatments, in order to optimize the effect of spectral pre-treatment, the combination of several spectral pre-treatment was often used. Therefore, in this present study we utilized three sequentially spectral pre-treatments: SGS, SNV and SG 1d (SGS + SNV + SG 1d). Our approach was previously used by Shawky and Selim [33] and Zhang et al. [35]. Figure 2 shows spectral data of all samples after pre-treatments using combination of SGS, SNV and SG 1d in the interval of 230–350 nm. Several wavelength peaks with high absorbance intensities were identified as an artifact of spectral pre-treatments at 270 and 315 nm (positive absorbance) and 290 and 340 nm (negative absorbance). Our spectral features were in line with previously reported work by Souto et al. [36].

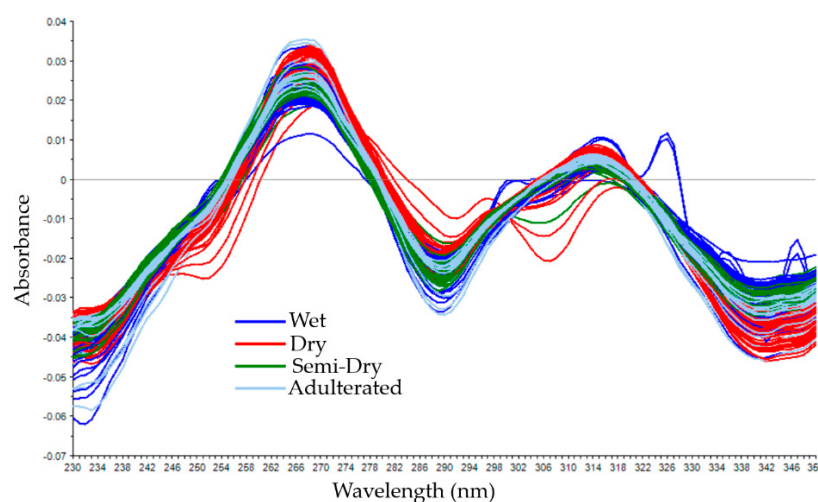


Figure 2. Combined pre-treated spectral data of all samples with different in cherry processing methods in the interval of 230–350 nm.

3.2. Principal Component Analysis (PCA)

Figure 3 shows the scores plot of the first three PCs ($PC1 \times PC2 \times PC3$) from PCA analysis calculated for all samples using the combined pre-treated spectral data in the interval of 230–350 nm. The cumulative percent variance (CPV) for 10 PCs in calibration and validation are presented in Table 2. The first three PCs could explain 93% of the total variances of spectral data which meets the general requirements of $CPV > 70\text{--}85\%$ for PCA analysis as mentioned by Hu et al. [37]. Using these three PCs, there was a clear separation between samples. The samples were clustered into four possible groups according to differences in cherry processing methods: dry, wet, semi-dry, and adulterated. The most of dry coffee samples were clustered in the negative of $PC1$ ($PC1 < 0$). The adulterated coffee samples were mostly located in the middle part of $PC1$ and $PC2$ (close to 0 both for $PC1$ and $PC2$). The most of wet and semi-dry coffee samples were located in the positive of $PC1$ ($PC1 > 0$). However, some of the wet and semi-dry coffee samples were overlapped considering the similarity of the cherry processing methods between the two, as reported by Duarte et al. [8].

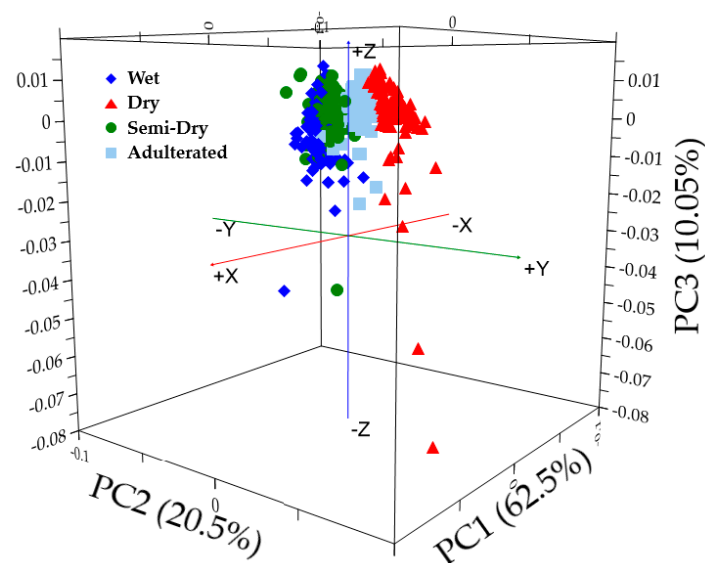


Figure 3. Scores plot of PC1, PC2 and PC3 from PCA analysis in the interval of 230–350 nm of combined pre-treated spectral data.

Table 2. Number of principal components and its cumulative percent variance (CPV) chosen for principal component analysis (PCA) using combined pre-treated spectral data in the interval of 230–350 nm.

Principal Components (PCs)	Cumulative Percent Variance (%)	
	Calibration	Validation
PC1	62.474	61.569
PC2	82.972	81.596
PC3	93.018	92.312
PC4	96.275	95.200
PC5	97.734	96.867
PC6	98.634	98.309
PC7	99.065	98.697
PC8	99.464	99.340
PC9	99.661	99.587
PC10	99.756	99.704

Figure 4 shows the contribution of each wavelength in the interval of 230–350 nm for separating the coffee samples according to different in cherry processing methods. There are six contributive wavelengths with high x-loadings identified at 255, 270, 290, 310, 315 and 320 nm. Those wavelengths are associated with the absorbance of several important chemical compounds in ground roasted coffee [36]. In previous work, Yulia and Suhandy [38] reported four influential wavelengths at 263, 297, 330 and 350 nm for discrimination between fresh and expired *Lampung robusta* coffee.

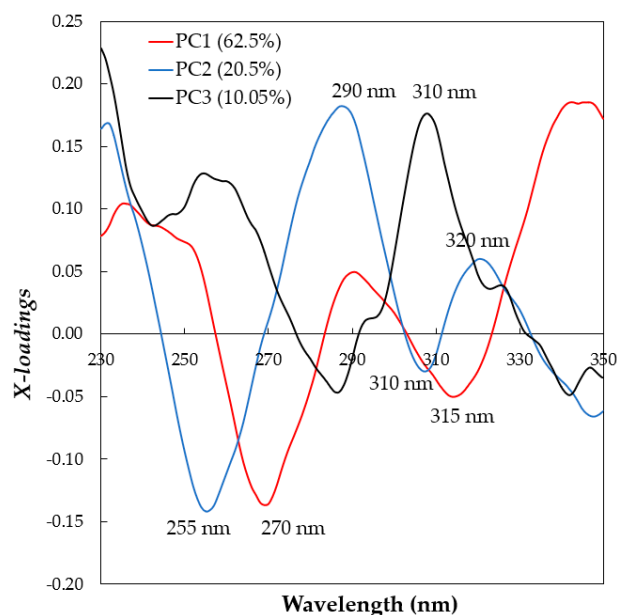


Figure 4. Plot of x-loadings versus wavelength from PCA analysis in the interval of 230–350 nm of combined pre-treated spectral data.

3.3. Supervised Classification Results

For supervised classification purposes, samples were divided randomly into two groups: 83.3% of samples for calibration and validation set (300 samples) or training set and the remaining 16.7% for prediction or test set (60 samples). For PLS-DA, a classification model was developed using pre-treated spectral data using all wavelengths in the interval of 230–350 nm. The PLS-DA model has 9 latent variables (LVs) with a classification accuracy of 67.60% for calibration. PLS-DA model was validated using the leave-one-out cross validation. The typical analytical information from PLS-DA was overlapped and more interference problem which resulted in lower accuracy in classification [39]. Using full-spectrum, the accuracy was improved by using SVM classification. The SVM model was developed using type 2 (nu-SVM classification) as this type minimizes the error function and was validated using 10-fold cross-validation. The linear kernel type was selected as the best model with the following parameters: $\nu = 0.5$ and $\gamma = 1$ which were adjusted through a grid search function. It resulted in a training accuracy of 88.67% and validation accuracy of 83.33%. To improve classification accuracy, two classification methods with fewer variables were also investigated.

In general, LDA and PCA-LDA is belong to supervised classification technique where the number of variables is smaller than the number of samples. In this study, the variable selection for LDA and PCA-LDA was performed in different way. For the LDA classification model, 6 wavelengths with high x-loadings from PCA results were selected as input variables: 255, 270, 290, 310, 315 and 320 nm. The developed LDA model has 81.0% of accuracy. The classification accuracy was improved comparing to PLS-DA but not for SVM classification. LDA has fewer variables comparing to PLS-DA. However, LDA with 6 wavelengths may still suffer a collinearity problem. For PCA-LDA, the classification model was developed using the PCA sample scores on 10 principal components (PC1 to PC10) in the range of 230–350 nm of the modified spectral data as input variables. The PCA-LDA model was developed using a training sample set (total 300 samples). During the PCA-LDA training, the calibration set was composed of 180 samples (including 51 dry, 50 wet, 49 semi-dry and 30 adulterated samples). The model was verified with the validation set of 120 samples (including 33 dry, 33 wet, 34 semi-dry and 20 adulterated samples) after the establishment of the PCA-LDA model. Figure 5 shows the PCA-LDA model with 93.33% of accuracy. As expected, it is noted that variables selection using PCA scores was appropriate to improve classification accuracy. The typical feature of PCs data is

uncorrelated. This is the main reason for the significant improvement of classification using the PCA-LDA. There was a clear separation of the most samples according to differences in cherry processing methods. However, as seen in Figure 5, some of wet, semi-dry and adulterated samples are still overlapped and fail to be discriminated by using the developed PCA-LDA model. In this model, 7 wet samples were misclassified as semi-dry, 6 semi-dry samples were misclassified as wet, 2 semi-dry samples were misclassified as adulterated and 5 adulterated samples were misclassified as semi-dry samples resulted in 93.33% of accuracy.

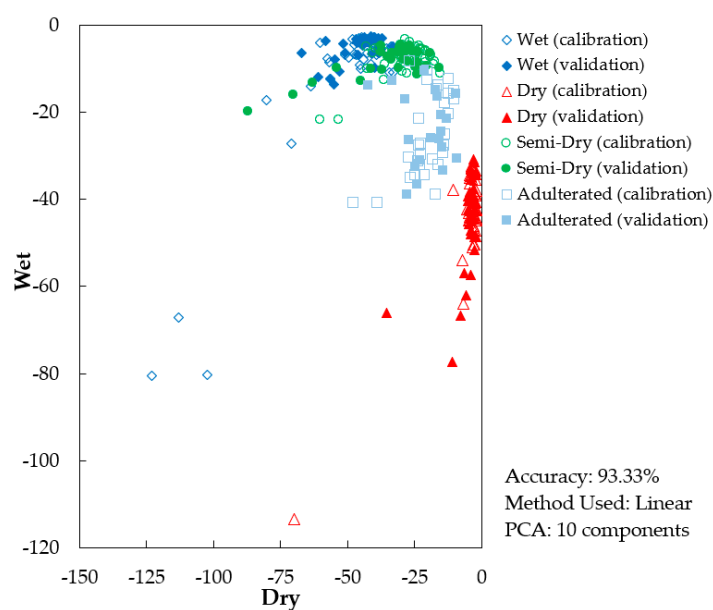


Figure 5. Principal component analysis–linear discriminant analysis (PCA-LDA) classification model developed using 10 PCs.

To evaluate the performance of the developed classification models, a prediction was performed using 60 unknown samples in the prediction sample set which had not been used in the model training: 16 dry processing samples, 17 wet processing samples, 17 semi-dry processing samples, and 10 adulterated samples. The superiority of the PCA-LDA model was observed in the prediction result. As demonstrated in Table 3, PCA-LDA resulted in the highest accuracy rate of 91.7%. In PCA-LDA, all prediction samples of dry and wet processing were properly classified into their respective classes. For semi-dry samples, four samples were misclassified as wet class. For adulterated samples, only one sample was misclassified as a semi-dry class.

Table 3. Confusion matrix for classification results using different classifiers in prediction set.

Classifiers	Samples		Actual				Accuracy Rate
			DRY	WET	SMD	ADT	
PLS-DA	Predicted	DRY	10	4	3	1	56.7%
		WET	6	8	0	0	
		SMD	0	5	9	2	
		ADT	0	0	5	7	
SVMC	Predicted	DRY	16	0	0	0	85.0%
		WET	0	15	5	0	
		SMD	0	2	12	2	
		ADT	0	0	0	8	
LDA	Predicted	DRY	16	0	0	0	80.0%
		WET	0	12	5	1	
		SMD	0	5	12	1	
		ADT	0	0	0	8	
PCA-LDA	Predicted	DRY	16	0	0	0	91.7%
		WET	0	17	4	0	
		SMD	0	0	13	1	
		ADT	0	0	0	9	

In previous works, several reports also reported the effective improvement of classification results using PCA-LDA. Dankowska et al. [40] used synchronous fluorescence and UV-Vis spectra combining with PCA-LDA to discriminate between arabica and robusta coffee with various mixtures. Khuwijitjaru et al. [41] reported the highest classification accuracy of 97.5% using PCA-LDA with smoothing pre-treatment of NIR spectral data for the discrimination of green robusta coffee. Diniz et al. [39] used several classification methods for tea classification with different geographical and varieties. It was reported that PCA-LDA significantly gave acceptable results with 92% and 100% accuracy rate.

4. Conclusions

A classification of *Lampung robusta* specialty coffee with different cherry processing methods using UV spectroscopy and chemometrics was proposed. It has been demonstrated that the spectral data of dry, wet, semi-dry and adulterated coffee were overlapped. A full spectrum-based classification using PLS-DA with highly-correlated variables resulted in low-classification accuracy. Using fewer uncorrelated variables based on PCA-LDA resulted in the best classification accuracy of 93.33% in the calibration and 91.7% in the prediction. In terms of the number of variables, it can be concluded that the LDA and PCA-LDA models with fewer variables tend to produce a more robust classification model. In terms of the delta accuracy between training and prediction (delta accuracy = accuracy in training – accuracy in prediction), the LDA and PCA-LDA models also resulted in a smaller delta accuracy of 1% and 1.63% compared to the SVMC and PLS-DA models. The proposed analytical method based on UV spectroscopy provides a simpler method with water extraction (chemical free) and a more affordable device for authentication of *Lampung robusta* specialty coffee according to differences in the cherry processing methods.

Author Contributions: Conceptualization, D.S. and M.Y.; methodology, D.S. and M.Y.; software, D.S.; validation, M.Y.; formal analysis, D.S. and M.Y.; investigation, M.Y.; resources, M.Y.; data curation, D.S.; writing—original draft preparation, D.S. and M.Y.; writing—review and editing, D.S. and M.Y.; visualization, M.Y.; supervision, D.S.; project administration, M.Y.; funding acquisition, M.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Research and Technology/National Agency for Research and Innovation, Republic of Indonesia (KEMENRISTEK/BRIN), PKPT grant 2020–2022 and The APC was funded by The University of Lampung.

Acknowledgments: We acknowledge Spectroscopy Research Group (SRG) The University of Lampung for technical support during experiment and data analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rocchetti, G.; Braceschi, G.P.; Odello, L.; Bertuzzi, T.; Trevisan, M.; Lucini, L. Identification of markers of sensory quality in ground coffee: An untargeted metabolomics approach. *Metabolomics* **2020**, *16*, 127. [CrossRef]
2. Coffee: World Markets and Trade. Available online: <https://apps.fas.usda.gov/psdonline/circulars/coffee.pdf> (accessed on 27 December 2020).
3. Belitz, H.D.; Grosch, W.; Schieberle, P. *Food Chemistry*, 4th ed.; Springer: Berlin/Heidelberg, Germany, 2009; pp. 939–940.
4. Ghosh, P.; Venkatachalapathy, N. Processing and drying of coffee—A review. *Int. J. Eng. Res. Technol.* **2014**, *3*, 784–794.
5. Selmar, D.; Kleinwachter, M.; Bytof, G. Metabolic responses of coffee beans during processing and their impact on coffee flavor. In *Cocoa and Coffee Fermentations*; Schwan, R.F., Fleet, G.H., Eds.; CRC Press: Boca Raton, FL, USA, 2015; pp. 431–476.
6. Toledo, P.R.A.B.; Pezza, L.; Pezza, H.R.; Toci, A.T. Relationship between the different aspects related to coffee quality and their volatile compounds. *Compr. Rev. Food Sci. Food Saf.* **2016**, *15*, 705–719. [CrossRef]
7. Santos, J.R.; Lopo, M.; Rangel, A.O.S.S.; Lopes, J.A. Exploiting near infrared spectroscopy as an analytical tool for on-line monitoring of acidity during coffee roasting. *Food Control* **2015**, *60*, 408–415. [CrossRef]
8. Duarte, G.S.; Pereira, A.A.; Farah, A. Chlorogenic acids and other relevant compounds in Brazilian coffees processed by semi-dry and wet post-harvesting methods. *Food Chem.* **2010**, *118*, 851–855. [CrossRef]
9. Bytof, G.; Knopp, S.E.; Schieberle, P.; Teutsch, I.; Selmar, D. Influence of processing on the generation of g-aminobutyric acid in green coffee beans. *Eur. Food Res. Technol.* **2005**, *220*, 245–250. [CrossRef]

10. Lachenmeier, D.W.; Teipel, J.; Scharinger, A.; Kuballa, T.; Walch, S.G.; Grosch, F.; Bunzel, M.; Okaru, A.O.; Schwarz, S. Fully automated identification of coffee species and simultaneous quantification of furfuryl alcohol using NMR spectroscopy. *J. AOAC Int.* **2020**, *103*, 306–314. [\[CrossRef\]](#)
11. Teixeira, A.A.; Brando, C.H.J.; Thomaziello, R.A.; Teixeira, R. The raw bean. In *Espresso Coffee the Science of Quality*; Illy, A., Viani, R., Eds.; Elsevier Academic Press: Berkeley, CA, USA, 2005; pp. 87–95.
12. Buratti, S.; Sinelli, N.; Bertone, E.; Venturello, A.; Casiraghi, E.; Geobaldo, F. Discrimination between washed arabica, natural arabica and robusta coffees by using near infrared spectroscopy, electronic nose and electronic tongue analysis. *J. Sci. Food Agric.* **2015**, *95*, 2192–2200. [\[CrossRef\]](#)
13. Flambeau, K.J.; Lee, W.J.; Yoon, J. Discrimination and geographical origin prediction of washed specialty bourbon coffee from different coffee growing areas in Rwanda by using electronic nose and electronic tongue. *Food Sci. Biotechnol.* **2017**, *26*, 1245–1254. [\[CrossRef\]](#)
14. Suhandy, D.; Yulia, M. Peaberry coffee discrimination using uv-visible spectroscopy combined with SIMCA and PLS-DA. *Int. J. Food Prop.* **2017**, *20*, S331–S339. [\[CrossRef\]](#)
15. Suhandy, D.; Yulia, M. The use of partial least square regression and spectral data in uv-visible region for quantification of adulteration in Indonesian palm civet coffee. *Int. J. Food Sci.* **2017**, *2017*, 6274178. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Suhandy, D.; Yulia, M. Unsupervised classification of three specialty coffees from java based on principal component analysis and uv-visible spectroscopy. *IOP Conf. Ser. Earth Environ. Sci.* **2020**, *537*, 012034. [\[CrossRef\]](#)
17. Suhandy, D.; Yulia, M. The classification of arabica gayo wine coffee using uv-visible spectroscopy and PCA-DA method. *MATEC Web Conf.* **2018**, *197*, 09002. [\[CrossRef\]](#)
18. Petisca, C.; Pérez-Palacios, T.; Farah, A.; Pinho, O.; Ferreira, I.M.P.L.V.O. Furans and other volatile compounds in ground roasted and espresso coffee using headspace solid-phase microextraction: Effect of roasting speed. *Food Bioprod. Process.* **2013**, *91*, 233–241. [\[CrossRef\]](#)
19. Yergenson, N.; Aston, D.E. Monitoring coffee roasting cracks and predicting with in situ near-infrared spectroscopy. *J. Food Process Eng.* **2020**, *43*, e13305. [\[CrossRef\]](#)
20. Suhandy, D.; Yulia, M. The development of classification model of peaberry and civet ground roasted coffee using uv-visible spectroscopy and two different compensation methods. *J. Phys. Conf. Ser.* **2019**, *1341*, 022007. [\[CrossRef\]](#)
21. Ball, D.W. *The Basics of Spectroscopy*; SPIE Publications: Bellingham, WA, USA, 2001; pp. 59–60.
22. Jiménez-Carvelo, A.M.; González-Casado, A.; Bagur-González, M.G.; Cuadros-Rodríguez, L. Alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity—A review. *Food Res Int.* **2019**, *122*, 25–39. [\[CrossRef\]](#)
23. Medina, S.; Perestrelo, R.; Silva, P.; Pereira, J.A.M.; Câmara, J.S. Current trends and recent advances on food authenticity technologies and chemometric approaches. *Trends Food Sci. Technol.* **2019**, *85*, 163–176. [\[CrossRef\]](#)
24. Jia, S.; Yang, L.; An, D.; Liu, Z.; Yan, Y.; Li, S.; Zhang, X.; Zhu, D.; Gu, J. Feasibility of analysing frost-damaged and non-viable maize kernels based on near infrared spectroscopy and chemometrics. *J. Cereal Sci.* **2016**, *69*, 145–150. [\[CrossRef\]](#)
25. Kennard, R.W.; Stone, L.A. Computer aided design of experiments. *Technometrics* **1969**, *11*, 137–148. [\[CrossRef\]](#)
26. Harvey, T.J.; Gazi, E.; Henderson, A.; Snook, R.D.; Clarke, N.W.; Brown, M.; Gardner, P. Factors influencing the discrimination and classification of prostate cancer cell lines by FTIR microspectroscopy. *Analyst* **2009**, *134*, 1083–1091. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Olivier, D.; Cyril, R.; Alexandra, D.; Ludovic, D.; Jean-Pierre, H. Support vector machines (SVM) in near infrared (NIR) spectroscopy: Focus on parameters optimization and model interpretation. *Chemometr. Intell. Lab. Syst.* **2009**, *96*, 27–33. [\[CrossRef\]](#)
28. Lee, L.C.; Liong, C.Y.; Jemain, A.A. Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: A review of contemporary practice strategies and knowledge gaps. *Analyst* **2018**, *143*, 3526–3539. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Tharwat, A.; Gaber, T.; Ibrahim, A.; Hassanien, A.E. Linear discriminant analysis: A detailed tutorial. *AI Commun.* **2017**, *30*, 169–190. [\[CrossRef\]](#)
30. Gislum, R.; Nikneshan, P.; Shrestha, S.; Tadayyon, A.; Deleuran, L.; Boelt, B. Characterisation of Castor (*Ricinus communis* L.) seed quality using Fourier transform near-infrared spectroscopy in combination with multivariate data analysis. *Agriculture* **2018**, *8*, 59. [\[CrossRef\]](#)
31. Marfo, T.D.; Datta, R.; Vranov, V.; Ekielski, A. Ecotone dynamics and stability from soil perspective: Forest-agriculture land transition. *Agriculture* **2019**, *9*, 228. [\[CrossRef\]](#)
32. Basri, K.N.; Hussain, M.N.; Bakar, J.; Sharif, Z.; Khir, M.F.A.; Zoofakar, A.S. Classification and quantification of palm oil adulteration via portable NIR spectroscopy. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2017**, *173*, 335–342. [\[CrossRef\]](#)
33. Shawky, E.; Selim, D.A. NIR spectroscopy-multivariate analysis for discrimination and bioactive compounds prediction of different Citrus species peels. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2019**, *219*, 1–7. [\[CrossRef\]](#)
34. Santos, K.M.; Moura, M.F.V.; Azevedo, F.G.; Lima, K.M.G.; Raimundo, I.V.; Pasquini, C. Classification of Brazilian coffee using near-infrared spectroscopy and multivariate calibration. *Anal. Lett.* **2012**, *45*, 774–781. [\[CrossRef\]](#)
35. Zhang, Z.; Wang, Y.; Yan, H.; Chang, X.; Zhou, G.; Zhu, L.; Liu, P.; Guo, S.; Dong, T.T.X.; Duan, J. Rapid geographical origin identification and quality assessment of angelicae sinensis radix by FT-NIR spectroscopy. *J. Anal. Methods Chem.* **2021**, *2021*, 8875876. [\[CrossRef\]](#)

36. Souto, U.T.C.P.; Barbosa, M.F.; Dantas, H.V.; de Pontes, A.S.; Lyra, W.S.; Diniz, P.H.G.D.; de Araújo, M.C.U.; da Silva, E.C. Identification of adulteration in ground roasted coffees using uv–vis spectroscopy and SPA-LDA. *LWT Food Sci. Technol.* **2015**, *63*, 1037–1041. [[CrossRef](#)]
37. Hu, L.; Yin, C.; Ma, S.; Liu, Z. Vis-NIR spectroscopy combined with wavelengths selection by PSO optimization algorithm for simultaneous determination of four quality parameters and classification of soy sauce. *Food Anal. Methods* **2019**, *12*, 633–643. [[CrossRef](#)]
38. Yulia, M.; Suhandy, D. Identification of fresh and expired ground roasted robusta coffee using uv-visible spectroscopy and chemometrics. *MATEC Web Conf.* **2018**, *197*, 09003. [[CrossRef](#)]
39. Diniz, P.H.G.D.; Barbosa, M.F.; de Melo Milanez, K.D.T.; Pistonesi, M.F.; de Araújo, M.C.U. Using uv–vis spectroscopy for simultaneous geographical and varietal classification of tea infusions simulating a home-made tea cup. *Food Chem.* **2016**, *192*, 374–379. [[CrossRef](#)] [[PubMed](#)]
40. Dankowska, A.; Domagała, A.; Kowalewski, W. Quantification of *coffea arabica* and *coffea canephora* var. *robusta* concentration in blends by means of synchronous fluorescence and uv-vis spectroscopies. *Talanta* **2017**, *172*, 215–220. [[CrossRef](#)]
41. Khuwijitjaru, P.; Boonyapisompan, K.; Huck, C.W. Near-infrared spectroscopy with linear discriminant analysis for green ‘robusta’ coffee bean sorting. *Int. Food Res. J.* **2020**, *27*, 287–294.