Chemometric quantification of peaberry coffee in blends using UV–visible spectroscopy and partial least squares regression

Diding Suhandy, Meinilwita Yulia, and Kusumiyati

Citation: AIP Conference Proceedings **2021**, 060010 (2018); doi: 10.1063/1.5062774
View online: https://doi.org/10.1063/1.5062774
View Table of Contents: http://aip.scitation.org/toc/apc/2021/1
Published by the American Institute of Physics

Chemometric Quantification of Peaberry Coffee in Blends Using UV–Visible Spectroscopy and Partial Least Squares Regression

Diding Suhandy[1,a)], Meinilwita Yulia[2,b)] and Kusumiyati[3,c)]

[1]*Spectroscopy Research Group (SRG), Laboratory of Bioprocess and Postharvest Engineering, Department of Agricultural Engineering, Faculty of Agriculture, The University of Lampung, Jalan Prof. Dr. Soemantri Brojonegoro No.1, Bandar Lampung, 35145, Indonesia.*
[2]*Department of Agricultural Technology, Lampung State Polytechnic, Jalan Soekarno Hatta No. 10, Rajabasa Bandar Lampung, 35141, Indonesia.*
[3]*Laboratorium of Plant Production Technology, Department of Agronomy, Faculty of Agriculture, Padjadjaran University, Jalan Raya Jatinangor KM 21, Bandung, Indonesia.*

[a)]Corresponding author: diding.sughandy@fp.unila.ac.id
[b)]meinilwitayulia@polinela.ac.id
[c)]kusumiyati@unpad.ac.id

Abstract. The aim of the present study is to quantify peaberry coffee in blends using UV–visible spectroscopy and partial least squares (PLS) regression. A total of 210 ground roasted peaberry coffee in blends (pure and adulterated with degree of adulteration 0%–90%) were used as samples. After the extraction process, spectral data of 3 mL aqueous samples was acquired using a UV–visible spectrometer in the range of 190–1100 nm (Genesys 10s, Thermo Scientific, USA). The PLS regression was used to quantify the peaberry content in blends (peaberry-to-normal blends). The best PLS model was achieved using Savitzky-–Golay first derivative spectra in the interval of 190–450 nm with low root mean square error of calibration (RMSEC = 1.165430%) and high determination coefficient (R^2 = 0.99). The calibration model also had high a RPD, 11.88. This analytical method is simple, easy to use, of low cost, and has excellent sensitivity.

Keywords: chemometri, peaberry coffee, regression, UV-Visible Spectroscopy.

INTRODUCTION

Coffee is an important crop that guarantees a sustainable economy to farmers in tropical regions, and Indonesia is no exception. It is produced in over 50 developing countries around the world. Several specialty coffees have a particularly high economic value, including civet coffee (‘*kopi luwak*’ in the Indonesian language) and peaberry coffee (‘*kopi lanang*’ in Indonesian). A peaberry is a natural mutation of the coffee bean inside the cherry. Normally two coffee beans grow in a fruit (dicotyledonous)—flat against each other like halves of a peanut; however, on rare occasions a single bean is produced (monocotyledon). The production of peaberry coffee is very limited, with only about 7% of any given coffee crop containing peaberry beans.[1] The higher price for peaberry beans arises from its supposedly more concentrated flavor compared with normal beans. Due to the rarity and unique taste of this peaberry coffee, substitution by cheaper normal beans (not peaberry) may give rise to fraudulent substitution.[1,2]

The quality and authenticity of ground roasted coffee is an important issue since it has been the target of fraudulent admixtures with a variety of cheaper materials, including spent coffee grounds, coffee husks, and other roasted grains.[2] In order to satisfy the quality requirements of the consumer, the authentication of specialty coffees (such as civet and peaberry) is also one of the major challenges that has become increasingly important in the coffee

The 8th Annual Basic Science International Conference
AIP Conf. Proc. 2021, 060010-1–060010-7; https://doi.org/10.1063/1.5062774
Published by AIP Publishing. 978-0-7354-1739-7/$30.00

trade, as a result of the significant increase in the price gap between the specialty coffees and regular (non-specialty) coffees in the past few years.

For peaberry coffee, adulteration is done by adding cheaper normal (not peaberry) ground roasted coffee into authentic ground roasted peaberry coffee. This kind of adulteration is both frequent and imperceptible to the naked eye. The potential of several advanced analytical methods has been explored in the detection of adulteration and estimating the authenticity of ground roasted coffee. These include high performance liquid chromatography (HPLC),[3] ultra-high performance liquid chromatography coupled with high resolution mass spectrometry (UPLC-HRMS),[4] electrophoresis-tandem mass spectrometry (CE-MS),[5] nuclear magnetic resonance (NMR) spectroscopy,[6,7] near-infrared spectroscopy (NIRS),[8,9] UV–Vis spectroscopy,[1,10-12] mid-infrared spectroscopy,[13–15] Raman spectroscopy[16,17] and fluorescence spectroscopy.[18] A fusion of these methods for ground roasted coffee authentication has also been reported.[19]

Among these available methods, the detection of ground roasted coffee adulteration using UV–visible spectroscopy is preferable since the UV–visible instrumentation is easily obtainable by most developing countries' laboratories to carry out the routine analysis of detecting adulterants. Previously, Suhandy and Yulia[1] showed the possibility of using UV–visible spectroscopy combined with soft independent modelling of class analogy (SIMCA) and partial least squares discriminant analysis (PLS-DA) to discriminate between pure peaberry and pure normal ground roasted coffee. However, as far as the author's knowledge extends, there is no published quantitative study of estimating the concentration of peaberry in blends (peaberry-to-normal blends). For this reason, in this research, we demonstrate that the spectral information in the UV–visible region of ground roasted coffee can be used to quantitatively measure the peaberry-to-normal ratio in a blend. The main purpose of the work is to establish a new and simple analytical method with minimal and free-chemical sample preparation and relatively fast analysis that allows accurately determining the peaberry content in a blend.

MATERIALS AND METHODS

Coffee Samples

Peaberry and normal bean roasted coffee samples were purchased from the local market (Hasti Coffee, Lampung, Indonesia). All coffee samples were ground using a home coffee-grinder (Sayota). In this research, 210 blends of ground roasted coffee samples belonging to both peaberry and normal coffee were used (Table 1). The samples had eleven different levels of peaberry content (10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, and 100% (w/w), Table 1). The sample preparation, including the sieving and extraction procedures.[1,10–12]

Spectral Acquisition using UV-Visible Spectrometer

The spectrum of each coffee sample was measured immediately after the extraction procedures. The measurements were carried out with a UV–Visible spectrometer (GENESYS 10S UV–Vis, Thermo Scientific, USA) equipped with a quartz cell with optical path of 10 mm and spectral resolution of 1 nm at 27°C–29°C in the range of 190–1100 nm (full wavelength). The blank spectrum was recorded with distilled water.

Data Analysis

The samples were randomly divided into two subsets: calibration (147) and prediction (63) sets. The calibration set was used for the development of the model. The prediction set was used to estimate the performance of the model. The concentration of ground roasted peaberry coffee in blends (%, w/w) was calculated by using a PLS regression analysis. The PLS model was developed using modified spectra (using Savitzky–Golay first derivative with ordo: 2 and window: 9) in the range of 190–450 nm. The PLS calibration model was calculated using commercial multivariate analysis software (The Unscrambler® X (30 days trial version–CAMO Software, Oslo, Norway)).

The following parameters were used to assess the quality of the PLS model: root mean square error of calibration (RMSEC) and the coefficient of determination in calibration ($R^2_{calibration}$). The optimum number of latent variables (LV) for the PLS model was estimated by a full cross-validation method based on the smallest root mean square error of cross validation (RMSECV).[20]

The performance of PLS model was evaluated by using root mean square error of prediction (RMSEP), the coefficient of determination in prediction ($R^2_{prediction}$), the bias and the residual prediction deviation (RPD) of the PLS model. A large discrepancy between the RMSECV and RMSEP values indicates an over-fitted model. For RPD, higher value is desirable and a value greater than 3 corresponds to excellent prediction accuracy.[21–23]

TABLE 1. Peaberry content (% w/w) in coffee samples of calibration (a) and prediction set (b).

Sample number	Peaberry (% w/w)	Sample number	Peaberry (% w/w)	Sample number	Peaberry (% w/w)	Sample number	Peaberry (% w/w)	Sample number	Peaberry (% w/w)
(a)									
L1B901a	10	L3B701b	30	L5B501a	50	L7B305b	70	L9B107a	90
L1B901b	10	L3B702b	30	L5B504a	50	L7B306a	70	L9B107b	90
L1B902b	10	L3B703a	30	L5B504b	50	L7B307a	70	L9B108b	90
L1B903a	10	L3B704a	30	L5B505b	50	L7B307b	70	L9B109a	90
L1B904a	10	L3B704b	30	L5B506a	50	L7B308b	70	L9B1010a	90
L1B904b	10	L3B705b	30	L5B507a	50	L7B309a	70	L9B1010b	90
L1B905b	10	L3B706a	30	L5B507b	50	L7B3010a	70	L95B0501	95
L1B906a	10	L3B707a	30	L5B508b	50	L7B3010b	70	L95B0502	95
L1B907a	10	L3B707b	30	L5B509a	50	L8B201a	80	L95B0504	95
L1B907b	10	L3B708b	30	L5B5010a	50	L8B201b	80	L95B0505	95
L1B908b	10	L3B709a	30	L5B5010b	50	L8B202b	80	L95B0507	95
L1B909a	10	L3B7010a	30	L6B401a	60	L8B203a	80	L95B0508	95
L1B9010a	10	L4B601a	40	L6B401b	60	L8B204a	80	L95B0510	95
L1B9010b	10	L4B601b	40	L6B402b	60	L8B204b	80	LA01a	100
L2B801a	20	L4B602b	40	L6B403a	60	L8B205b	80	LA01b	100
L2B801b	20	L4B603a	40	L6B404a	60	L8B206a	80	LA02b	100
L2B802b	20	L4B604a	40	L6B404b	60	L8B207a	80	LA03a	100
L2B803a	20	L4B604b	40	L6B401a	60	L8B207b	80	LA04a	100
L2B804a	20	L4B601a	40	L6B405b	60	L8B208b	80	LA04b	100
L2B804b	20	L4B605b	40	L6B406a	60	L8B209a	80	LA05b	100
L2B805b	20	L4B606a	40	L6B407a	60	L8B2010a	80	LA06a	100
L2B806a	20	L4B607a	40	L6B407b	60	L8B2010b	80	LA07a	100
L2B807a	20	L4B608b	40	L6B408b	60	L9B101a	90	LA07b	100
L2B807b	20	L4B609a	40	L6B409a	60	L9B101b	90	LA08b	100
L2B807b	20	L4B6010a	40	L7B301a	70	L9B102b	90	LA09a	100
L2B808b	20	L4B6010b	40	L7B301b	70	L9B103a	90	LA010a	100
L2B809a	20	L5B501a	50	L7B302b	70	L9B104a	90	LA010b	100
L2B8010a	20	L5B501b	50	L7B303a	70	L9B104b	90		
L2B8010b	20	L5B502b	50	L7B304a	70	L9B105b	90		
L3B701a	30	L5B503a	50	L7B304b	70	L9B106a	90		
(b)									
L1B902a	10	L3B703b	30	L5B505a	50	L7B306b	70	L9B108a	90
L1B903b	10	L3B705a	30	L5B506b	50	L7B308a	70	L9B109b	90
L1B905a	10	L3B706b	30	L5B508a	50	L7B309b	70	L95B0503	95
L1B906b	10	L3B708a	30	L5B509b	50	L8B202a	80	L95B0506	95
L1B908a	10	L3B709b	30	L6B402a	60	L8B203b	80	L95B0509	95
L1B909b	10	L4B602a	40	L6B403b	60	L8B205a	80	LA02a	100
L2B802a	20	L4B603b	40	L6B405a	60	L8B206b	80	LA03b	100
L2B803b	20	L4B605a	40	L6B406b	60	L8B208a	80	LA05a	100
L2B805a	20	L4B606b	40	L6B408a	60	L8B209b	80	LA06b	100
L2B806b	20	L4B608a	40	L6B409b	60	L9B102a	90	LA08a	100
L2B808a	20	L4B609b	40	L7B302a	70	L9B103b	90	LA09b	100
L2B809b	20	L5B502a	50	L7B303b	70	L9B105a	90		
L3B702a	30	L5B503b	50	L7B305a	70	L9B106b	90		

RESULTS AND DISCUSSION

Spectral Analysis

Fig. 1 shows the average spectra obtained for ground roasted coffee blends with different levels of peaberry content (low, middle and high). The spectra are similar: most of the significant wavelengths are concentrated in the range of 190–450 nm. In general, the absorbance values were higher for this range, while in the range of 450–1100 nm, the absorbance values are close to zero.

Several sharp wavelengths at 236 nm, 270 nm, 288 nm, 320 nm, and 345 nm can be clearly seen in the modified spectrum corresponding to ground roasted coffee. These wavelengths have been previously reported to be present in the spectra of ground roasted arabica and robusta coffee samples. The wavelength at 275 nm is related to the C=O chromophore absorption of caffeine,[1,10–12, 24] The wavelengths at 288 nm and 320 nm are closely related to the absorbance of chlorogenic acids and trigonelline, respectively.[10–12, 24]

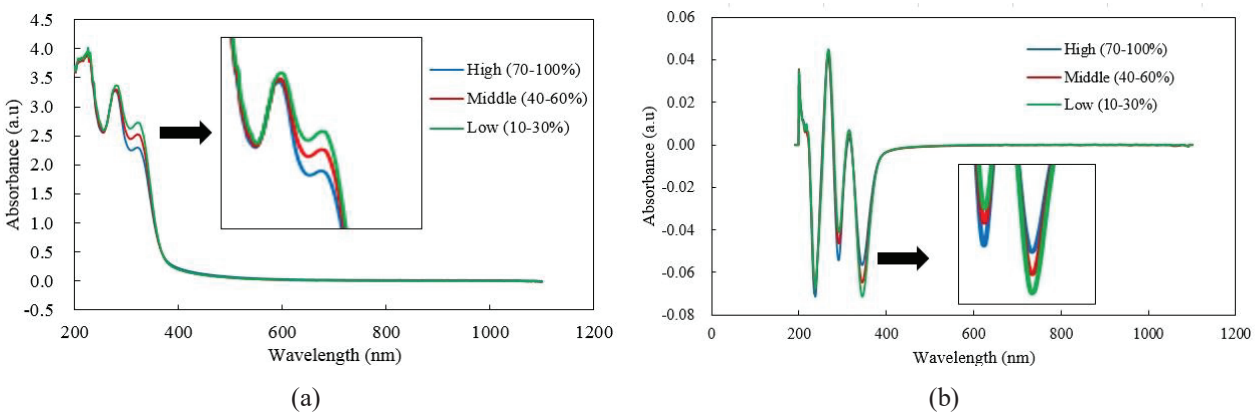
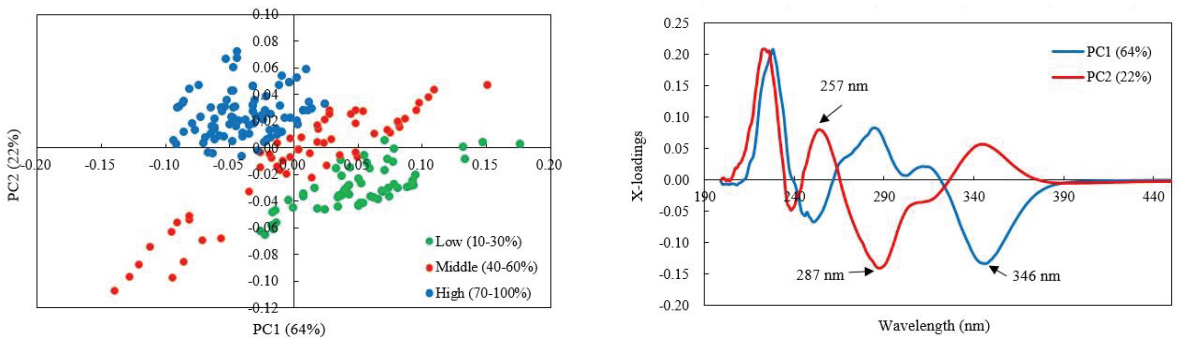


FIGURE 1. UV–Visible spectra of the average 210 coffee samples with low, middle and high peaberry content (a) original spectra, (b) modified spectra (using Savitzky-Golay first derivative with ordo: 2 and window: 9).

PCA Overview of UV–Visible Spectral Data

In order to see mapping of the coffee sample data set, PCA has been applied to the extracted the meaningful PCs. The results of PCA on the whole samples (210 samples) using wavelength in the range of 190-450 nm is shown in Fig. 2. The samples were divided into three different groups according to concentration of peaberry coffee in blends (%, w/w), low groups (10%–30%), middle group (40%–60%) and high group (70%–100%). There are several different strategies to select number of adequate PCs. In this research, percentages of variance in the data matrix is explained by each PC and the cumulative percentages of variance (CPV) are reported for PCA analysis.

As seen in Fig. 2 (a), the first two principal components can explain 86% of the variance in the dataset. In general, the results of the PCA showed that there was satisfactory discrimination between low, middle and high levels of peaberry content. The results of PCA showed that peaberry coffee samples (pure and adulterated) clustered into three different groups according to their concentration of peaberry (authenticity), changing along the direction of PC1. In Fig. 2 (b), we can see a plot of *x*-loadings versus wavelength for PC1 and PC2. This plot (Fig. 2 (b)) shows the wavelengths that had a significant contribution to the variation described by PC1 and PC2. This could be identified by their higher loading values (absolute values). Wavelengths with a higher loading indicate carrying more information about the difference of peaberry content in the coffee samples. As seen in Fig. 2 (b) there are several wavelengths which had higher loading values: 257 nm, 287 nm and 346 nm. The wavelengths 257 nm and 287 nm correspond with the absorbance of caffeine[1,24] while wavelength 346 nm corresponds with the absorbance of chlorogenic acid (CGA).[1]



(a) (b)

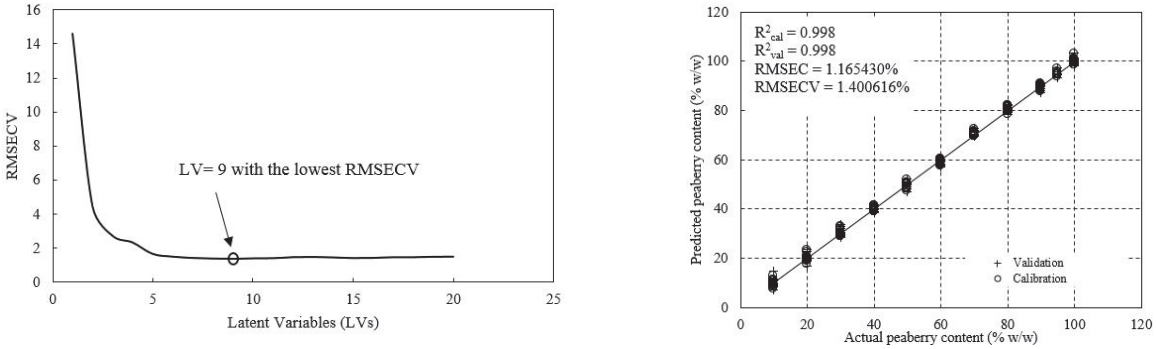
FIGURE 2. PCA results based on absorbance spectra (190-450 nm) of ground roasted coffee aqueous samples (a) Plot PC1 vs. PC2 (b) X-loadings vs. wavelength.

Results from PLS Regression

Fig. 3 shows the best PLS calibration model based on modified or pre-processed spectra (Savitzky–Golay first derivative with ordo: 2 and window: 9). This PLS model has 9 PLS factors with the lowest RMSECV as seen in Fig. 3 (a). As shown in Fig. 3 (b), R^2 was found to be 0.99 for both calibration and validation. Figure 4 shows the actual and predicted values of peaberry content in blends for the prediction set. The RMSEP was 2.485677% (w/w) and the determination coefficients were 0.99. The predicted bias was 0.397151% (w/w). The RPD of 11.88 was obtained using the best PLS model.

Our PLS model was quite superior to several previous reported studies. The determination coefficient of 0.97 and RMSECV of 4.58% (w/w) were obtained in the measurement of the arabica content in arabica-to-robusta blends.[8] Using this PLS model, a prediction of the robusta content in arabica-to-robusta blends was conducted and resulted in RMSEP = 4.34% (w/w), representing errors about 1.7 times higher than our result (RMSEP of 2.485677% w/w). Another work has developed a more sensitive analytical technique using ambient ionization mass spectrometry to predict percentage of robusta ground roasted coffee in blends and obtained RMSEP of 2.54%, which is very close to our result.[25] Using mid infrared spectroscopy, a quantification of robusta in blends has been reported using several wavelength selection methods. The best PLS model was obtained for the ordered predictors selection (OPS) method: it could predict the robusta content with RMSEP of 1.89% (w/w). [15]

It is noteworthy to mention here that the evaluation of the authenticity for ground roasted peaberry coffee samples using the proposed UV–Visible spectroscopy coupled with PLS regression analysis took around ten minutes for sample preparation (extraction process with hot distilled water) and an additional one minute for spectral data acquisition. Thus, it is established that our simple sample preparation and chemical free protocol, encapsulating good efficiency (rapid), accuracy and validity ($R^2_{prediction}$> 0.99), is a prospective alternative method to assess the authenticity of ground roasted peaberry coffee.



(a) (b)

FIGURE 3. PLS model development for peaberry content determination in blend using absorbance modified spectral data in the range of 190-450 nm (a) LVs vs. RMSECV (b) measured (actual) peaberry vs. predicted peaberry.

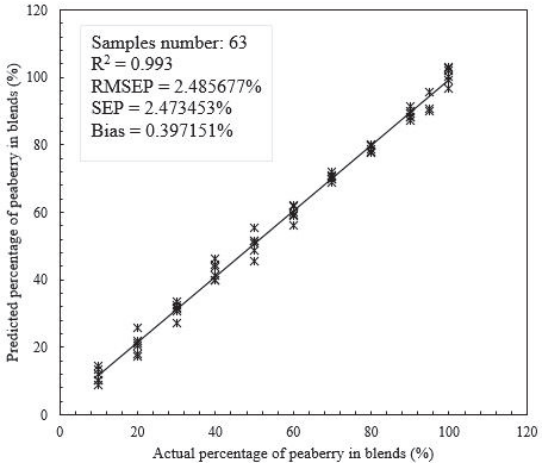


FIGURE 4. Actual vs. predicted peaberry content (% w/w) values for prediction sample set.

SUMMARY

In this research, we showed the potentialities of UV–Visible spectroscopy coupled with chemometrics analysis for detecting authenticity of ground roasted peaberry coffee and quantifying the content of ground roasted peaberry coffee in blends (peaberry-to-normal). PCA could clearly separate between low, middle and high levels of peaberry content. The best PLS model has nine PLS factors; it had the lowest RMSECV = 1.400616% (w/w), and it attained a determination coefficient (R^2) of 0.99 for both calibration and for validation. The RMSEP was 2.485677% (w/w) and the determination coefficients were 0.99. The results show that UV–visible spectroscopy combined with the PLS regression method is a prospective alternative method to assess the authenticity of ground roasted peaberry coffee.

ACKNOWLEDGEMENTS

The authors acknowledge Hasti Coffee Lampung for providing the peaberry and normal coffee samples. The authors also would like to express their sincere appreciation to Ministry of Research, Technology and Higher Education, Republic of Indonesia (*Kemenristekdikti*) for financial support under Research Grant *Penelitian Berbasis Kompetensi* (PBK) (Grant Number: 384/UN26.21/PN/2018).

REFERENCES

1. D. Suhandy and M. Yulia, Int. J. Food Prop. **20**, S331–S339 (2017).
2. N. Reis, A. S. Franca and L. S. Oliveira, LWT - Food Sci. Technol. **53**, 395–401 (2013).
3. E. D. Pauli, F. Barbieri, P. S. Garcia, T. B. Madeira, V. R. A. Junior, I. S. Scarminio, C. A. P. da Camara and S. L. Nixdorf, Food Res. Int. **61**, 112–119 (2014).
4. T. Cai, H. Ting and Z. Jin-lan, Food Chem.**190**, 1046–1049 (2016).
5. D. Daniel, F. S. Lopes, V. B. dos Santos and C. L. do Lago, Food Chem. **243**, 305–310 (2018).
6. M. V. M. Ribeiro, N. Boralle, H.R. Pezza, L. Pezza and A. T. Toci, J. Food Compos. Anal. **57**, 24–30 (2017).
7. L. R. Cagliani, G. Pellegrino, G. Giugno and R. Consonni, Talanta **106**, 169–173 (2013).
8. E. Bertone, A. Venturello, A. Giraudo, G. Pellegrino and F. Geobaldo, Food Control **59**, 683–689 (2016).
9. N. Reis, A. S. Franca and L. S. Oliveira, LWT-Food Sci. Technol. **50**, 715–722 (2013).
10. U. T. C. P. Souto, M. F. Barbosa, H. V. Dantas, A. S. de Pontes, W. S. Lyra, P. H. G. D. Diniz, M. C. U. Araújo and E. C. da Silva, LWT-Food Sci. Technol. **63**, 1037–1041 (2015).

11. M. Yulia and D. Suhandy, J. Phys.: Conf. Ser. **835**, 1–6 (2017).
12. D. Suhandy and M. Yulia, Int. J. Food Sci. 1-7 (2017).
13. J. Medina, D. C. Rodríguez, V. A. Arana, A. Bernal, P. Esseiva and J. Wist. Int. J. Anal. Chem. 1–8 (2017).
14. A. P. Craig, B. G. Botelho, Leandro S. Oliveira and A. S. Franca. Food Chem. **245**, 1052–1061 (2018).
15. C. Assis, L. S. Oliveira and M. M. Sena. Food Anal. Methods **11**, 578–588 (2018).
16. T. Wermelinger, L. D'Ambrosio, B. Klopprogge and C. Yeretzian. J. Agric. Food Chem. **59**, 9074–9079 (2011).
17. R. C. E. Dias and C. Yeretzian. Int J. Exp. Spectroscopic Tech. **1**, 1–5 (2016).
18. A. Dankowska, A. Domagała and W. Kowalewski, Talanta **172**, 215–220 (2017).
19. N. Reis, B. G. Botelho, A. S. Franca and L. S. Oliveira, Food Anal. Methods **10**, 2700–2709 (2017).
20. X. Sun, W. Lin, X. Li, Q. Shen and H. Luo. Anal. Methods **7**, 3939–3945 (2015).
21. D. Suhandy, T. Suzuki, Y. Ogawa, N. Kondo, H. Naito, T. Ishihara, Y. Takemoto and W. Liu, Eng. in Agric. Environ. Food **5**, 90–95 (2012).
22. D. Suhandy, M. Yulia, Y. Ogawa and N. Kondo, Eng. in Agric. Environ. Food **5**, 152–158 (2012).
23. D. Suhandy, M. Yulia, Y. Ogawa and N. Kondo, Eng. in Agric. Environ. Food **6**, 111–117 (2013).
24. G. Navarra, M. Moschetti, V. Guarrasi, M. R. Mangione, V. Militello and M. Leone, J. Chem. 1–8 (2017).
25. R. Garrett, B. G. Vaz, A. M. C. Hovell, M. N. Eberlin, C. M. Rezende, J. Agric. Food Chem. **60**, 4253–4258 (2012).