Infrared spectroscopy technique for quantification of compounds in plant-based medicine and supplement

Effan Cahyati Junaedi[1,3],
Keri Lestari[2], Muchtaridi Muchtaridi[1]

Departments of [1]Pharmaceutical Analysis and Medicinal Chemistry, and [2]Pharmacology and Clinical Pharmacy Faculty of Pharmacy, Universitas Padjadjaran, [3]Department of Pharmaceutical Chemical Analysis, Faculty of Mathematics and Natural Science, Universitas Garut, Langensari, Indonesia

J. Adv. Pharm. Technol. Res.

Abstract

Quality control of plant-based medicine and supplements must be carried out to ensure uniformity in quality and safety in their use, resulting in the need for effective and accurate analytical methods. Infrared spectroscopy is a method of qualitative and quantitative analysis that is fast, time-saving, cost-effective, accurate, and nondestructive. This method has been applied for quantitative analysis of compounds in complex matrices such as plant-based medicine and supplements supported by chemometrics techniques. The success of infrared spectroscopy applications for quantitative analysis of phytochemicals and adulterants content in plant-based medicine and supplement can happen by several factors. This article highlights the effect of spectral preprocessing and variable selection on quantitative analysis of phytochemical and adulterant in plant-based medicine and supplements using infrared spectroscopy. Literature search was conducted with PubMed, Google Scholar, and Science Direct by selecting quantitative analysis research on plant-based medicines and supplements that utilize spectral preprocessing techniques and variable selection in processing data analysis. The preprocessing spectra and variables selection can affect the accuracy and precision of infrared spectroscopy methods. The variable selection can be done using the wavenumber point technique, the wavenumber interval, or a combination thereof. Variable selection is more commonly used for near-infrared data than for IR data. The optimization of the preprocessing spectra and variables selection technique will be useful in increasing the ability of infrared spectroscopy in predicting compound levels.

Key words: Adulterant, herbal, phytochemical, preprocessing, variable selection

INTRODUCTION

The use of plant-based medicine and supplements had been increased in the past few decades. This is in line with the evolution of self-medication so that there is a tendency to return to traditional and natural products. Consumers choose to use herbal and natural products because natural herbs are safe and more likely to minimize the side effects of using chemical drugs, improve health, and reduce treatment costs.[1] However, the use of herbal products is not yet acceptable in some countries due to counterfeit products, uneven quality, and safety of their use. It may cause negative effects to the consumer, product quality assurance can be assessed from both qualitative and quantitative aspects.

Address for correspondence:
Prof. apt. Muchtaridi Muchtaridi,
Department of Pharmaceutical Analysis and Medicinal Chemistry, Faculty of Pharmacy, Universitas Padjadjaran, Jl. Bandung-Sumedang KM 21, Jatinangor, 45363, Sumedang, Indonesia.
E-mail: muchtaridi@unpad.ac.id

Access this article online

Quick Response Code:

Website:
www.japtr.org

DOI:
10.4103/japtr.JAPTR_96_20

How to cite this article: Junaedi EC, Lestari K, Muchtaridi M. Infrared spectroscopy technique for quantification of compounds in plant-based medicine and supplement. J Adv Pharm Technol Res 2021;12:1-7.

Quantitative assessment of herbal products quality is focused on phytochemical components that are naturally contained in sample or adulterants that should not be contained therein.[2,3]

Various analytical methods can be used to examine phytochemical content and adulterants in herbal medicines such as high performance liquid chromatography (HPLC), ultra-high-HPLC, liquid chromatography (LC)-mass spectrometry (MS), gas chromatography (GC)-MS, nuclear magnetic resonance (NMR), and thin-layer chromatography. However, the methods have several disadvantages. HPLC method is still costly and unpracticed in sample handling, needs long-time analysis, and has a high consumption of solvents. NMR method is insensitive and requires a relatively large amount of sample to make a measurement.[4] Furthermore, in herbal medicines analysis, one of the obstacles is the efficiency of the analysis, because the complex components of herbal medicines can complicate the process. In addition, the analysis process can also damage the material.[5]

Infrared spectroscopy can be examined by qualitative and quantitative analysis. This method provides information about the compound content in complex samples with small levels. The complexity of the information provided can be resolved by chemometric techniques. Chemometric techniques can assist in extracting information from spectra through multivariate analysis. Spectroscopy is considered fast, time-saving, cost-effective, accurate, and nondestructive analytical tool.[6] Infrared spectroscopy has been used successfully for quantitative analysis in various fields including pharmaceutical, food industry, agriculture, and biological evaluation. The success of this method is assessed from the value of root mean square error of calibration (RMSEC), root mean square error of prediction (RMSEP), root mean square error of cross-validation (RMSECV), and determination coefficient (R^2).[7] Good linearity and accepted standard errors of a model of good linearity and standard errors of a prediction model are influenced by several factors. The preprocessing spectra and variables selection can affect the accuracy and precision of infrared spectroscopy methods. We performed the PubMed, Google Scholar, and Science Direct databases to find articles providing information on the effect of preprocessing spectra and variables selection. Some of the research done on quantitative analysis using infrared spectroscopy has been presented in Table 1. This article will discuss several preprocessing techniques and variable selection that contribute to improving the quantitative analysis results of chemical components (in the form of adulterant) and phytochemicals (second metabolites) in plant-based medicine and supplements to standardize products with green analytical chemistry methods.

QUANTITATIVE ANALYSIS USING INFRARED SPECTROSCOPY

Quantitative analysis with infrared spectroscopy is an indirect method because it still requires reference methods such as chromatographic (HPLC, GC and capillary electrophoresis), and spectroscopic (NMR, ultraviolet). To support the quantification of compound information in the sample, the relationship between Fourier transform infrared spectroscopy (FTIR) spectrum as a predictor variable, and the value of measurement level results with the reference method (response variable) will be translated by chemometrics techniques. In general, content analysis with infrared spectroscopy combined with chemometrics is carried out through several stages. First, the samples of spectra data are scanned using near-infrared (NIR) and mid-infrared (MIR) spectrometers. The sample is also analyzed by the reference method to get the actual value. This reference method is chosen based on the ability to quantify components in a sample accurately. Furthermore, infrared spectra can be treated further with preprocessing and variable selection. This aims to reduce the dimensions of the data to simplify calibration modeling without losing important information.[8]

Calibration is made by plotting the prediction variable (variable of infrared spectra) as *x* with the actual variable (compound content obtained from the reference method) as *y*. Calibration models for quantitative analysis are divided into two categories: linear and nonlinear. Generally, linear calibration that can be used is partial least square (PLS), principal component regression, and step-wise multiple linear regression, while nonlinear calibration is artificial neural network.[8] After that, the calibration model formed is evaluated to testify the ability in analyzing the levels of compounds in the sample.

QUANTIFICATION OF PHYTOCHEMICAL COMPOUND AND ADULTERANT

An analysis with infrared spectroscopy is ideal if the evaluation results show R^2 values approaching a value of 1, low standard error (RMSEC, RMSECV, RMSEP, or predicted residual error sum square [PRESS]), and the relative percent difference (RPD) values should be at least 2.4 or greater.[28,29] R^2 calibration values describe the linearity of the calibration curve formed from infrared spectrum data (*x* value) with the concentration data measured using the reference method (*y* value). R^2 values close to 1 indicate that the infrared spectrum data are able to explain the concentration of compounds as the dependent variable. R^2 validation value illustrates the accuracy of the measured concentration value by the infrared method. A validation value of R^2 close to 1 indicates that the measurement of compound concentrations using the FTIR method yields a value that is proportional to the measurement results

Table 1: Quantitative analysis using infrared spectrometry coupled with chemometrics

Quantitative analysis by infrared spectroscopy	Description
Morfin and Thebain in Poppy head (flower or fruit of family Papaveraceae) by FTIR[9]	Mean-centred scalling technique as preprocessing spectra followed by PCA as variable selection gives a better predictive ability, but the model still needs an optimation due to high % RSD for morfin prediction
Quersetin, Quersitrin, Rutin, and total flavonoid in genus Fagopyrum by FTIR[10]	Several spectra pretreatments are employed. Wavelet transform gives the best result to preprocessed high resolution spectra, while wavelet transform together with derivative has good linearity for low resolution spectra. Spectra at wavenumber 500-1700 cm^{-1} and between 2950-3500 cm^{-1} are chosen to build a PLS calibration model
Total flavonoid in 20 leaves extract and commercial herbal supplement by FTIR and NIR[11]	PLS regression built by full spectra of NIR with no spectra pretreatment shows a better prediction than FTIR spectra
3,4-dihydroxy-phenyl lactic acid (danshensu) in Pill containing *Salviae miltiorrhizae* Radix et Rhizoma *Notoginseng* Radix et Rhizoma and by NIR[12]	Spectral of wavenumber 7994-7258, 6135-5399 cm^{-1} is chosen to build PLS modeling after preprocessing of spectra. The 1st derivative spectra produce the best calibration model with a good evaluation result
Puerarin in Puerariae radix by NIR and UV spectrofotometry[13]	The best PLS model obtained from NIR-UV fusion data with MSC with the1st derivative and smoothing as preprocessing technique
Alkaloids BB, CO, PA, EP and JA in Coptidis rhizoma by NIR[14]	The best model for BB, CO and PA model built by ANN, while EP and JA by PLS. Different preprocessing spectra fits to different alkaloid
Isoflavon and saponin in Soybeans by NIR[15]	The MLR regression model of derivative spectra has a good evaluation value for total isoflavones, but the results are not satisfactory for total saponin
Stevioside, rebaudioside, and sum of them in *Stevia rebaudiana* Bertoni by NIR[16]	NIR spectra is preprocessed by smoothing, and then followed by the 1st and the 2nd derivative. The PLS regression is built in wavenumber region 4760-5016 cm^{-1}. The results show that model evaluation values are not satisfactory due to low linearity and different value of RMSEC and RMSEP
Five total antraquinone (emodin, chrysophanol, rhein, aloeemodin, and physcion) in Rhei radix et rhizome by NIR[17]	Different preprocessing is applied to NIR spectra with 16 cm^{-1} resolution, and then the wavenumber regions at 4242-5581 cm^{-1}, 5885-6233 cm^{-1}, and 6394-7011 cm^{-1} are chosen to build a multivariate regression, namely SMLR, PCR, and PLS. The best model is obtained using PLS with the 1st derivative
Icariin in Epimedium species from Berberidaceae family by NIR[18]	NIR spectra with 8 cm^{-1} resolution is processed by the 1st derivative and smoothing, followed by CARS as variable selection technique. After that, the variable builds a PLS regression and results in good calibration and validation value
CO and SFO as adulterant in Extra virgin olive oil by FTIR[19]	The FTIR with resolution at 4 cm^{-1} spectra is processed by derivative. The best SFO model is obtained using PLS with the 1st derivative, while the best CO model is obtained using a PLS with no preprocessing. The models have good linearity, but significant differences in RMSEC and RMSEP
Millet (*Eleusine cora-cana*) and Buckwheat (*Fagopyrum esculentum*) as adulterant in Blackpepper (*Piper nigrum*) by NIR[20]	The best model is obtained by PLS regression of NIR with MSC, followed by the1st derivative preprocessing spectra. It has good R^2 and RPD, but significant differences in RMSEC and RMSEP
Sibutramine and phenolptalein as adulterant in Weightloss herbal medicine by FTIR[21]	SNV and a Savitzky-Golay 2nd derivative are used as a preprocessing technique. The best model is obtained by FTIR. $R^2>0.93$. Sibutamine's model produces 0.8% RMSEC and RMSECV, and phenolphthalein 2.2%
Total flavone content in snow lotus (*Saussurea involucrate*) by NIR[22]	SNV is choosing as a preprocessing spectrum. Before constructing the calibration model, the spectral split into several intervals by iPLS than the variable of the split interval is choosing by GA. Comparing to other variable selection, iPLS-GA produces the lowest RMSEC and RMSEP
EGCG in green tea by FT-NIR[23]	SNV as preprocessing spectra and siPLS-GA as variable selection give the best result of a calibration model with few variables, lowest RMSEC and RMSEP, and highest R^2
Alpha-mangostin in mangosteen exctract by FTIR[24]	Wave number 3825-937 cm^{-1} was selected using TQ analyst™ to build a prediction model using PLS. The validation model produces accurate and precise results for the determination of alpha-mangostin levels
Rosmarinic acid in rosmarini leaves by ATR-IR and NIR[25]	The whole spectrum is using to constructed PLS model pretreated by MSC followed with 2nd derivative method give the best evaluation value
BB in rhizoma coptidis by IR and NIR fusion[26]	SNV followed by 2nd derivative is the best spectra preprocessing method for analysis of BB with fusion strategy

Contd...

Table 1: Contd...

Quantitative analysis by infrared spectroscopy	Description
Total flavonoid content in *Ginkgo biloba* leaf by NIR[27]	SNV is the superior pretreatment spectra technique among several other techniques applied in this research with the lowest RMSECV value

PCA: Principal component analysis, RSD: Relative standard deviation, PLS: Partial least square, NIR: Near-infrared, UV: Ultraviolet, MSC: Multiplicative scatter correction, BB: Berberine, CO: Coptisine, PA: Palmatine, EP: Epiberberine, JA: Jatrorrhizine, ANN: Artificial neural network, MLR: Multiple linear regression, SMLR: Stepwise MLR, RMSEC: Root mean square error of calibration, RMSEP: Root mean square error of prediction, RMSECV: Root mean square error of cross validation, CARS: Competitive adaptive reweighted sampling, CO: Corn oil, SFO: Sunflower oil, SNV: Standard normal variate, iPLS: Interval PLS, GA: Genetic algorithm, EGCG: Epigallocatechin-3-gallate, siPLS: Synergy PLS, FTIR: Fourier Transform Infrared, RPD: Relative percent difference, IR: Infrared

of the reference method. The RPD values were used to evaluate the fitting and prediction capacities of the models by calculating the standard deviation divided by standard error prediction. The RMSEC, RMSECV, RMSEP, or PRESS values are used to assess the precision. The analysis process using infrared spectroscopy can be seen in Figure 1.

Spectra preprocessing

Spectra preprocessing in quantitative analysis aims to minimize the noise and the physical phenomena so that the resulting signal correlates with concentration that increases the predictive ability.[30] There are several ways of preprocessing the spectra. In this review, Savitzy-Golay-based derivatization, standard normal variate (SNV), multiplicative scatter correction (MSC), and scaling techniques are commonly used.[9,21,30]

Most of the analysis with plant materials using infrared spectroscopy is done on solid samples, making it susceptible to scattering.[30] The scattering causes irrelevant variations in the spectra data. If there are no preprocessing spectra, a mix of information and noise will occur, and this can cause a decrease in the predictive ability of a model. The scattering usually results in fewer MIR measurements than NIR. It is due to inhomogeneity of particle size in NIR measurement.[30] In research conducted by Otsuka, it was shown that the particle size was inversely proportional to the scattering coefficient.[31]

MSC and SNV are the examples of useful scattering correction techniques in reducing physical variability due to scattering. MSC and SNV often produce similar evaluation values.[13,17,20] MSC works to correct the spectrum by estimating the correction coefficient of the raw spectrum and calculating it with the average reference spectrum used in the calibration model. Since it uses the average spectrum, correction errors might occur if there is a dominant spectrum used as a reference spectrum.

Therefore, repeated application of MSC generally reduces spectrum correction errors. Nevertheless, it does not mean that the more MSC processes are corrected for the better, the repetition of MSC can reduce the difference between the spectra data sets used.[32] In using SNV, spectra correction is done by calculating the mean and deviation of the sample spectrum. Unlike MSC, spectra preprocessing

using SNV does not require a reference spectrum to make corrections, so the process is simpler.[33] SNV works by normalizing the spectrum due to physical differences but not chemically.[34] SNV success occurs when the uniform scatters between samples in the full spectrum. If this condition does not occur, then the spectral correction will not be optimal.[35]

Savitzky-Golay is a derivative technique which has a mode for spectral smoothing.[36] The derivatization can provide a more detailed picture of the structure in the spectra to increase sensitivity. Spectra derivatization that is widely used in quantitative analysis of compounds in plants includes the 1st derivatives and the 2nd derivatives. The 1st derivative will remove baseline variation between the data, and the 2nd derivative will improve the resolution of spectra to remove any slope effect on the data.[18,21] Although derivatization generally provides a detailed description, not all analyzes with IR spectroscopy require derivatization for preprocessing spectra. Such is the case with research conducted by Kokalj Ladan *et al.*, who conducted the derivatization of spectra with different resolutions. From these studies, it is known that derivatization is useful for improving prediction models in spectra with a resolution of 16 cm^{-1} but not at a resolution of 4 cm^{-1}.[10] In this regard, lower resolutions can cause loss of useful information for quantitative analysis; therefore, derivatization will help in describing the spectra in more detail.[37] However, the spectra contain more noise in higher resolution so that when derivatization is carried out, it will cause too much noise in the data and decrease the signal to noise ratio. Thus, the standard error in making the model also does not improve or even increase.[10]

Variable selection

The selection of spectra preprocessing techniques is very important to determine a successful calibration model formation, because it can emerge or eliminate important information connected with the content measurement. However, the successful of spectra effect preprocessing on the calibration model formation can only be known after validating the model.[38] It is important to compare several spectral preprocessing techniques and their combinations to determine a more suitable technique for the analyzed data. Preprocessing techniques are appropriate if it decreases standard errors by reducing or minimally maintaining the complexity of the model.
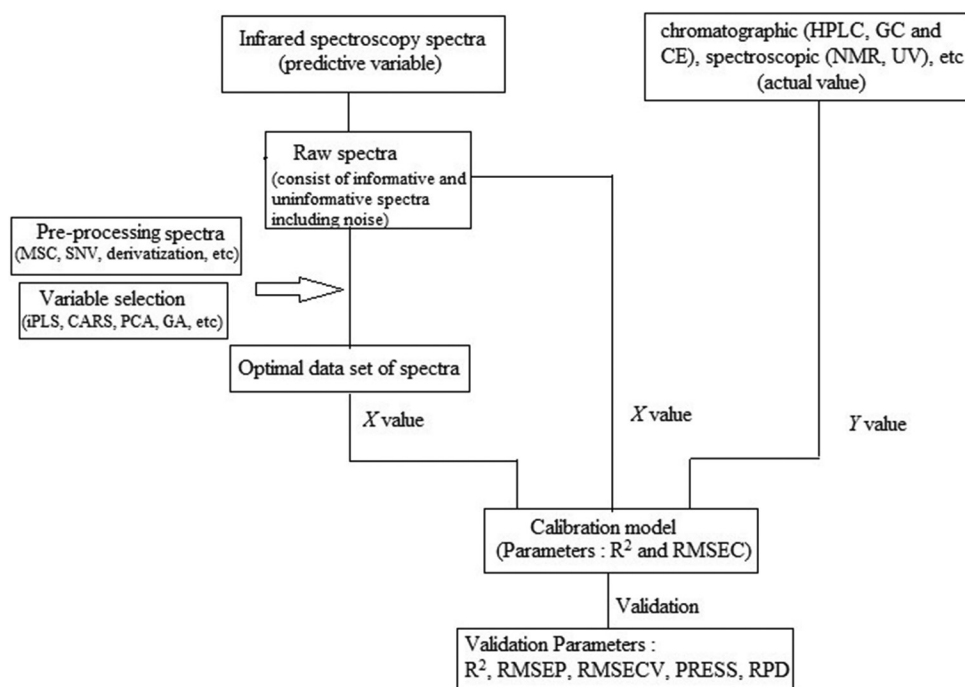
Figure 1: Quantitative analysis process using infrared spectroscopy

Besides using spectral preprocessing, reducing the complexity of the model can also be done by selecting variables. From the summary of the research, it can be seen that the selection of variables using principal component analysis (PCA),[9] interval PLS (iPLS),[13] Competitive adaptive reweighted sampling (CARS),[18] or genetic algorithm (GA)[39] reduce the number of variables by eliminating irrelevant variables. PCA will change the original variables correlated into new variables that contain a combination of the original variables. This combination can reduce the number of variables, but still can explain most of the original variable information.[40,41]

CARS variable selection has three stages. The first step is a random sampling of data sets, and then a number of these variables will be reduced again using the exponential decrease function technique where the number of variables rapidly decreased. The remaining variables that produce low RMSECV will be selected as informative variables included in the calibration model.[42] This method can make the calibration model simpler and effective.[43] While GA is done by randomly selecting a subset of variables and then calibrating the PLS model to select the variable with the greatest influence. Furthermore, crossover and mutations are carried out to form new variables, and re-calibration is carried out to decide which variable is the most suitable.[44,45]

In addition to the selection based on wave number points, variable selection can also be done by separating the wave numbers at certain intervals. In this article, iPLS and synergy interval (siPLS) have been used. iPLS works by dividing the

spectrum into several parts with the same interval. For each subinterval, submodels will be made and the standard error will be calculated. The submodel with the lowest standard error contains important information correlated with the objectives of the analysis. At this point, it will be chosen as a subinterval of wavenumbers, which will be used to form the calibration model.[46] While siPLS is a modification of iPLS to optimize variable selection with interval combinations containing informative variables.[47] siPLS can contain more informative variables because it does not only select one spectrum region but combines several.[48]

Variable selection can also be done by combining the selection of wave number points and wave number intervals such as siPLS-GA. Variable selection using combination techniques can produce a simpler model with fewer variables than without a combination. However, the smaller number of variables cannot be ascertained to have the best predictive ability, because of the loss of informative data.[49,50] Variable selection is useful for reducing noise and interfering variables in data so as to make prediction model formation more efficient with fewer variables but reliable prediction ability.[51]

A successful method in completing multivariate analysis is assessed by calibration and validation parameters. R^2 values that are close to 1 indicate that the predictive variables used in making the model correlate linearly with the levels of compounds. In addition, a calibration model is good if there is no large difference between the RMSEC and root mean square error P values.[17] If this happens, it will be overfitting or underfitting. Overfitting is a situation where

the calibration model produces a low standard error but unable to predict new samples and vice versa.

CONCLUSION

Infrared spectroscopy is useful in the quantification of phytochemical components and adulterants in plant-based medicines and supplements. This analysis can help further study to determine the quality of nutritious plants to ensure the safety and effectiveness of plant-based products. It is found that chemometrics can overcome the complexity of the chemical content in herbs. Optimization of preprocessing techniques and selection of variables and a combination of both are generally useful in improving predictive ability of models. This is indicated by an increase in linearity value (in calibration and validation) and a decrease in the standard error in the model by the spectrum that is experiencing preprocessing, variable selection, and a combination of both compared to raw spectra. The high R^2 validation value (close to 1) indicates that the measurement of compound concentrations using the FTIR method yields a proportional accuracy to the measurement results of the reference method.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

REFERENCES

1. Zahn R, Perry N, Perry E, Mukaetova-Ladinska EB. Use of herbal medicines: Pilot survey of UK users' views. Complement Ther Med 2019;44:83-90.
2. Kucharska-Ambrożej K, Karpinska J. The application of spectroscopic techniques in combination with chemometrics for detection adulteration of some herbs and spices. Microchem J 2020;153:1-9.
3. Chan Fung VS. Integrating traditional Chinese (herbal) medicines into risk based regulation - With focus on non-clinical requirements to demonstrate safety. J Tradit Chinese Med Sci 2020;7:88-94.
4. Silva Elipe MV. Advantages and disadvantages of nuclear magnetic resonance spectroscopy as a hyphenated technique. Anal Chim Acta 2003;497:1-25.
5. Upton R, David B, Gafner S, Glasl S. Botanical ingredient identification and quality assessment: Strengths and limitations of analytical techniques. Phytochem Rev 2019:1-21.
6. Guo C, Luo X, Zhou X, Shi B, Wang J, Zhao J, *et al.* Quantitative analysis of binary polymorphs mixtures of fusidic acid by diffuse reflectance FTIR spectroscopy, diffuse reflectance FT-NIR spectroscopy, Raman spectroscopy and multivariate calibration. J Pharm Biomed Anal 2017;140:130-6.
7. Westad F, Marini F. Validation of chemometric models - A tutorial. Anal Chim Acta 2015;893:14-24.
8. Yun YH, Li HD, Deng BC, Cao DS. An overview of variable selection methods in multivariate analysis of near-infrared spectra. TrAC - Trends Anal Chem 2019;113:102-15.
9. Turner NW, Cauchi M, Piletska EV, Preston C, Piletsky SA. Rapid qualitative and quantitative analysis of opiates in extract of poppy head via FTIR and chemometrics: Towards in-field sensors. Biosens Bioelectron 2009;24:3322-8.
10. Kokalj Ladan M, Straus J, Tavčar Benković E, Kreft S. FT-IR-based method for rutin, quercetin and quercitrin quantification in different buckwheat (Fagopyrum) species. Sci Rep; 7:7226.
11. Wulandari L, Retnaningtyas Y, Nuri , Lukman H. Analysis of flavonoid in medicinal plant extract using infrared spectroscopy and chemometrics. J Anal Methods Chem 2016;2016:4696803.
12. Zhang S, Xiong H, Zhou L, Ju W, Yang Z, Yan K, *et al.* Development and validation of in-line near-infrared spectroscopy based analytical method for commercial production of a botanical drug product. J Pharm Biomed Anal 2019;174:674-82.
13. Wang Y, Yang Y, Sun H, Dai J, Zhao M, Teng C, *et al.* Application of a data fusion strategy combined with multivariate statistical analysis for quantification of puerarin in Radix puerariae. Vib Spectrosc 2020;108:1-7.
14. Jintao X, Yufei L, Liming Y, Chunyan L, Quanwei Y, Weiying W, *et al.* Rapid and simultaneous analysis of five alkaloids in four parts of coptidis rhizoma by near-infrared spectroscopy. Spectrochim Acta A Mol Biomol Spectrosc 2018;188:611-8.
15. Berhow MA, Singh M, Bowman MJ, Price NPJ, Vaughn SF, Liu SX. Quantitative NIR determination of isoflavone and saponin content of ground soybeans. Food Chem 2020;317:126373.
16. Martono Y, Trihandaru S, Rondonuwu FS. Determination of stevioside and rebaudioside a in stevia rebaudiana bertoni leaves using near infrared spectroscopy and multivariate data analysis. Indones J Chem 2018;18:664-71.
17. Zhan H, Fang J, Wu H, Yang H, Li H, Wang Z, *et al.* Rapid determination of total content of five major anthraquinones in rhei radix et rhizoma by NIR spectroscopy. Chinese Herb Med 2017;9:250-7.
18. Yang Y, Wu Y, Li W, Liu X, Zheng J, Zhang W, *et al.* Determination of geographical origin and icariin content of Herba Epimedii using near infrared spectroscopy and chemometrics. Spectrochim Acta A Mol Biomol Spectrosc 2018;191:233-40.
19. Rohman A, Che Man YB. Quantification and classification of corn and sunflower oils as adulterants in olive oil using chemometrics and FTIR spectra. Sci World J 2012;2012:250795.
20. McGoverin CM, September DJ, Geladia P, Manley M. Near infrared and mid-infrared spectroscopy for the quantification of adulterants in ground black pepper. J Near Infrared Spectrosc 2012;20:521-8.
21. Rooney JS, McDowell A, Strachan CJ, Gordon KC. Evaluation of vibrational spectroscopic methods to identify and quantify multiple adulterants in herbal medicines. Talanta 2015;138:77-85.
22. Chen Q, Jiang P, Zhao J. Measurement of total flavone content in snow lotus (Saussurea involucrate) using near infrared spectroscopy combined with interval PLS and genetic algorithm. Spectrochim Acta A Mol Biomol Spectrosc 2010;76:50-5.
23. Guo Z, Chen Q, Chen L, Huang W, Zhang C, Zhao C. Optimization of informative spectral variables for the quantification of EGCG in green tea using Fourier transform near-infrared (FT-NIR) spectroscopy and multivariate calibration. Appl Spectrosc 2011;65:1062-7.
24. Rohman A, Arifah FH, Irnawati, Alam G, Muchtaridi M. The application of FTIR spectroscopy and chemometrics for classification of mangosteen extract and its correlation with alpha-mangostin. J Appl Pharm Sci 2020;10:149–54.
25. Kirchler CG, Pezzei CK, Beć KB, Henn R, Ishigaki M, Ozaki Y, *et al.* Critical evaluation of NIR and ATR-IR spectroscopic quantifications of rosmarinic acid in rosmarini folium supported by quantum chemical calculations. Planta Med 2017;83:1076-84.
26. Qi L, Ma Y, Zhong F, Shen C. Comprehensive quality assessment for Rhizoma Coptidis based on quantitative and qualitative metabolic profiles using high performance liquid chromatography, Fourier

Junaedi, *et al.*: Infrared spectroscopy for quantitative analysis of herbal

transform near-infrared and Fourier transform mid-infrared combined with multivariate statistical analysis. J Pharm Biomed Anal 2018;161:436-43.

27. Zhao LJ, Liu W, Xiong SH, Tang J, Lou ZH, Xie MX, *et al.* Determination of total flavonoids contents and antioxidant activity of ginkgo biloba leaf by near-infrared reflectance method. Int J Anal Chem 2018;2018:1-7.

28. Shenk JS, Landa I, Hoover MR, Westerhaus MO. Description and evaluation of near infrared reflectance spectro-computer for forage and grain analysis. Corp Sci 1981;21:355-8.

29. Rohman A, Sudjadi, Devi, Ramadhani D, Nugroho A. Analysis of curcumin in curcuma longa and Curcuma xanthorriza using FTIR spectroscopy and chemometrics. Res J Med Plant 2015;9:179-86.

30. Rinnan A. Pre-processing in vibrational spectroscopy-when, why and how. Anal Methods 2014;6:7124-9.

31. Otsuka M. Comparative particle size determination of phenacetin bulk powder by using Kubelka-Munk theory and principal component regression analysis based on near-infrared spectroscopy. Powder Technol 2004;141:244-50.

32. Windig W, Shaver J, Bro R. Loopy MSC: A simple way to improve multiplicative scatter correction. Appl Spectrosc 2008;62:1153-9.

33. Alrezj OA, Patchava KC, Benaissa M, Alshebeili S. Coupling scatter correction with bandpass filtering for preprocessing in the quantitative analysis of glucose from near infrared spectra. Annu Int Conf IEEE Eng Med Biol Soc 2017;2017:1800-3.

34. Sandak J, Sandak A, Meder R. Assessing trees, wood and derived products with near infrared spectroscopy: Hints and tips. J Near Infrared Spectrosc 2016;24:485-505.

35. Bi Y, Yuan K, Xiao W, Wu J, Shi C, Xia J, *et al.* A local pre-processing method for near-infrared spectra, combined with spectral segmentation and standard normal variate transformation. Anal Chim Acta 2016;909:30-40.

36. Savitzky A, Golay MJ. Smoothing and differentiation of data by simplified least square procedure. Anal Chem 1964;36:1627-39.

37. Chen Y, Thosar SS, Forbess RA, Kemper MS, Rubinovitz RL, Shukla AJ. Prediction of drug content and hardness of intact tablets using artificial neural network and near-infrared spectroscopy. Drug Dev Ind Pharm 2001;27:623-31.

38. Jonsson H, Gabrielsson J. Evaluation of preprocessing methods. Compr Chemom 2009;2:199-206.

39. Niazi A, Leardi R. Genetic algorithms in chemometrics. J Chemom 2012;26:345-51.

40. Abdi H, Williams LJ. Principal component analysis. Wiley Interdiscip Rev Comput Stat 2010;2:433-59.

41. Jolliffe IT, Cadima J. Principal component analysis: A review and recent developments. Philos Trans A Math Phys Eng Sci 016;374:1-16.

42. Li H, Liang Y, Xu Q, Cao D. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. Anal Chim Acta 2009;648:77-84.

43. Wang LL, Lin YW, Wang XF, Xiao N, Xu Y Da, Li HD, *et al.* A selective review and comparison for interval variable selection in spectroscopic modeling. Chemom Intell Lab Syst 2018;172:229-40.

44. Jiang H, Xu W, Chen Q. Comparison of algorithms for wavelength variables selection from near-infrared (NIR) spectra for quantitative monitoring of yeast (Saccharomyces cerevisiae) cultivations. Spectrochim Acta A Mol Biomol Spectrosc 2019;214:366-71.

45. Arakawa M, Yamashita Y, Funatsu K. Genetic algorithm-based wavelength selection method for spectral calibration. J Chemom 2011;25:10-9.

46. Norgaard L, Saudland A, Wagner J, Nielsen JP, Munck L, B. ES. Interval partial least-squares regression (iPLS): A comparative chemometric study with an example from near-infrared spectroscopy. Appl Spectrosc 2000;54:413-9.

47. Li Y, Guo M, Shi X, Wu Z, Li J, Ma Q, *et al.* Online near-infrared analysis coupled with MWPLS and SiPLS models for the multi-ingredient and multi-phase extraction of licorice (Gancao). Chin Med 2015;10:38.

48. Duarte LM, Paschoal D, Izumi CMS, Dolzan MD, Alves VR, Micke GA, *et al.* Simultaneous determination of aspartame, cyclamate, saccharin and acesulfame-K in powder tabletop sweeteners by FT-Raman spectroscopy associated with the multivariate calibration: PLS, iPLS and siPLS models were compared. Food Res Int 2017;99:106-14.

49. Zhu J, Ahmad W, Xu Y, Liu S, Chen Q, Hassan MM, *et al.* Development of a novel wavelength selection method for the trace determination of chlorpyrifos on Au@Ag NPs substrate coupled surface-enhanced Raman spectroscopy. Analyst 2019;144:1167-77.

50. Zhu J, Agyekum AA, Kutsanedzie FY, Li H, Chen Q, Ouyang Q, *et al.* Qualitative and quantitative analysis of chlorpyrifos residues in tea by surface-enhanced Raman spectroscopy (SERS) combined with chemometric models. Lwt Food Sci Technol 2018;97:760-9.

51. Guyon I, Elisseeff A. An introduction to variable and feature selection. J Machine Learn Res 2003;3:1157-82.