

Non Destructive and Rapid Analysis of Catechin Content in Gambir (*Uncaria gambir* Roxb.) Using NIR Spectroscopy

Andasuryani, Yohanes Aris Purwanto, I Wayan Budiastara, Khaswar Syamsu

Abstract— Quality of gambir is mainly determined by its catechin content. While trading in traditional market, catechin content of gambir is predicted by a treatment of expert experience which leads to subjective judgement. Meanwhile, quantitative analysis is carried out by using chemical method. However, this method is destructive, high cost and time consuming, thus not suitable for real time measurement. The objective of this study was to develop calibration model to predict catechin content in gambir non-destructively and rapidly using near infrared (NIR) spectroscopy with difference spectra pre-processing treatment and Partial Least Square (PLS) methods. Spectra pre-processing treatment method used were normalization between 0 and 1 (n01), first derivative Savitzky-Golay 9 points (dg1), second derivative Savitzky-Golay 9 points (dg2), combination between n01 and dg1, and combination between n01 and dg2. Determination of the optimum number of PLS factors was conducted based on value of consistency and Predicted Residual Sum Square Error (PRESS) of the validation set (V-set-PRESS). The best spectra pre-processing treatment method was n01 in combination with dg1. Evaluation of model demonstrated that the model could predict catechin content in gambir. The model had high value of correlation coefficient ($r = 0.95$), low values of SEC and SEP (3.56 and 3.27 respectively), and high value of ratio of prediction to deviation (RPD = 3.60). This study demonstrated that NIR spectroscopy had excellent potential as non destructive and rapid analysis to determine catechin content in gambir.

Index Terms— catechin, first derivative Savitzky-Golay 9 points, gambir, NIR spectroscopy, non destructive, normalization between 0 and 1, PLS.

1 INTRODUCTION

GAMBIR extracted product from the leaves and young twigs of gambir (*Uncaria gambier* Roxb.) plant. It is a member of family Rubiaceace and genus *Uncaria* [16]. Gambir is one of the Indonesian export commodities. It contributes to around 80 % of the gambir trading in the world [15]. It contains several chemical components, which one of them is catechin. The catechin is a bioactive compound which can be found a lot in gambir [1], [2], [28] and is potential as a raw material in various industries, particularly pharmaceutical and cosmetic industries. Catechin content in gambir is used as one of the quality parameters of gambir in accordance with the Indonesia gambir trading standard, SNI 01 - 3391-2000. Traditionally, determining of the gambir catechin content is conducted qualitatively based on experience of the assessor which led to subjective measurement. Meanwhile, there is chemical method which can be used to determine catechin content of gambir quantitatively. However, this method is not efficient since it requires expensive chemical reagents and time consuming. In addition, this method is destructive. Therefore,

to determine gambir catechin content.

Near Infrared (NIR) spectroscopy is a non destructive technique, accurate, and fast which can replace chemical method [7], [12]. Several studies have shown the ability of NIR spectroscopy technology to determine bioactive component of agricultural products such as: a) determining of alkaloid and phenol compound in green tea leaves using pre-processing method of weighted of multiple scatter correction (WMCS) to counter the effect of interferences of scatter and transforming individually through first to third derivatives [26], b) determining procyanidins in cocoa using pre-processing of standard normal variate (SNV) as well as detrending scatter correction and transforming individually by first derivative [30], c) determining of total antioxidant capacity in green tea by applying offset correction pre-processing method, detrend, multiplicative scatter correction (MSC), SNV, first and second derivatives [34], d) determining of carbohydrate, total acid, tartaric acid, malic acid, pH and polyphenol content in wine using pre-processing combination of normalization through derivative [14], e) determining of condensed tannin and dry matter in red grape homogenates using SNV pre-processing [9], f) determining of the main catechin content in green tea by applying three pre-processing methods, namely mean centering (MC), MSC and SNV [7], g) determining of flavanols in grape seeds using pre-processing method of MSC, SNV and detrend [13], h) monitoring of red wine fermentation using a pre-processing method of SNV, MSC, MSC combination with derivatives [10], and i) determining of phenol compounds of grape skins during ripening using pre-processing method of MSC, SNV and detrend [12].

- Andasuryani is currently pursuing PhD degree program in agricultural engineering in Bogor Agricultural University, Indonesia, PH-6281374317783. E-mail: andasuryani@gmail.com
- Yohanes Aris Purwanto, is senior lecturer, E-mail: arispurwanto@gmail.com
- I Wayan Budiastara, is senior lecturer, E-mail: wbudiastara@yahoo.com
- Khaswar Syamsu, is professor, E-mail: khaswar@yahoo.com

an efficient and non-destructive method should be developed

Rinnan et al. (2009) have reviewed some pre-processing methods which are most commonly used in the NIR and categorized them into two parts, namely scatter correction and derivative method. MSC, Inverse MSC, Extended MSC (EMSC), Extended Inverse MSC, de-trending, SNV and normalization included in the first group while the derivative Norris-Williams (NW) and derivative Savitzky-Golay polynomial filter (SG) included in the second group. However, since researchers do not have sufficient information regarding characteristics of pre-processing methods and have difficulties to decide the most appropriate methods, NIR spectroscopy pre-processing method still involve trial and error process [32].

The objective of this study was to develop calibration model to predict of catechin content of gambir non-destructively and rapidly using NIR spectroscopy with different pre-processing spectra treatment and partial least square (PLS) method.

2 MATERIAL AND METHOD

2.1 Sample Preparation

The samples were raw gambir obtained from Siguntur village, Koto 11 Tarusan sub district, Pesisir Selatan District, West Sumatera, Indonesia. The number of samples were 162 divided randomly into two groups, namely two thirds (108 samples) as calibration set and one third (54 samples) as validation set.

2.2 Chemical Analysis

Chemical analysis carried out to obtain data of gambir catechin content. The data were used as reference data while developing NIR calibration model. Determination of catechin content in gambir was conducted by referring to Indonesia National standard SNI 01-3391-2000. Instrument of Spectrophotometer U-2010, Hitachi was used to measure gambir catechin content. Catechin standard were purchased from Sigma Co., USA.

2.3 Spectra Acquisition

This study used instrument Buchi NIRFlex N-500 Fiber optics solids which had resolution of 4 cm^{-1} and 8 scans, to collect sample spectra. The instrument was operated using software NIRWare 1.2 (Büchi Labortechnik AG, Flawil, Switzerland). Spectra data were collected from 10000 to 4000 cm^{-1} (1000 to 2500 nm) with interval of 4 cm^{-1} . Measurements of each sample were conducted three times at different positions in a room temperature of $25\text{ }^{\circ}\text{C}$. Spectrum produced by each sample, was averaged prior to establishment a calibration model.

2.4 Chemometric

Chemometric is a science dealing with measurement of the chemical system or process through application of mathematical or statistical methods. Several stages in chemometric are pre-processing of spectra data, building calibration model and transfer the model [5]. NIR spectroscopy analysis needs chemometric in order to extract the hidden chemical infor-

mation of NIR spectra data. NIR spectra shows wide spectra features more than the sharp features. It occurs due to overtone and combination band.

The wide spectra feature requires chemometric method to obtain accurate and precision calibration model of NIR [25]. In addition, Nicolai, et al. (2007) acknowledged that chemometric is needed because spectra can be complicated by wavelength dependent scattering effects, tissue heterogeneities, instrumental noise, ambient effects and other sources of variability [21]. As a consequence, it is difficult to determine the specific absorption band of specific functional group especially chemical components. Furthermore, chemometric analysis requires a regression method in order to explain the relationship between NIR spectra data with a single component of analyzed agricultural product. Chemometric analysis in this study used NIRCal 5.2 (Büchi Labortechnik AG, Flawil, Switzerland) by PLS algorithm.

2.4.1 Spectra Pre-processing

NIR spectra obtained from NIR instrument do not only contain information of the sample but also information of the background and noise. Therefore, spectra pre-processing treatment method is needed before modelling to obtain reliable, accurate and stable calibration models [5]. The physical properties of solid samples also affect the NIR spectrum which can cause some problems while evaluating sample aspects such as identification and composition determination of raw materials. To cope with this circumstance, spectra pre-processing treatment method should be used to minimize the contribution of information which is not relevant with spectrum in order to develop reliable model [3].

This study did not only use the original spectra (original), but also several spectra pre-processing methods such as normalization between 0 and 1 (n01), first derivative Savitzky-Golay 9 points (dg1), second derivative Savitzky-Golay 9 points (dg2), combination of n01 with dg1 (n01,dg1) and combination of n01 with dg2 (n01,dg2). Normalization is used to eliminate multiplicative spectra effects [27]. The algorithms for derivative are direct differentiation and Savitzky-Golay [5]. The first derivatives can be used to eliminate additive baseline effects while second derivatives are used to eliminate sloped additive baselines [27]. The disadvantage of direct differentiation is bring noises thus can decrease signal to noise ratio. Savitzky-Golay algorithm is often used to cope with this problem [5].

2.4.2 Calibration Method

NIR spectra usually contain unselective and widely overlapped bands. Therefore, quantification analysis requires chemometrics analysis such as PLS, principal component regression (PCR), or multiple linear regressions (MLR) [18]. In recent years, the PLS regression method is a very popular method for chemometric analysis [4].

There are hundreds or thousands of wavelengths (variables) in a NIR spectrum that allow two successive wavelengths provide similar information which can distract the process of building a calibration model [19]. Therefore, there is a need to reduce the variable of original data into several cor-

related variables and only contains relevant information to the sample. Principal component analysis is the most famous and widely used method for data reduction [3]. In this study, selection the number of optimum PLS factors was done based on value of consistency and value of Predicted Residual Sum Square Error (PRESS) of the validation set (V-set-PRESS). Selected PLS factor has consistency value between 80-110% and as smallest as possible V-Set PRESS [22]. The number of PLS factors was 1 to 15. To develop the best calibration model, the value of consistency are used to ensure that the number of chosen PLS factors is optimal [11].

Statistical parameters used to evaluate the resulted model were bias, standard error of calibration set (SEC), standard error of validation set (SEP), coefficient correlation (r), coefficient of variation (CV) and ratio of prediction to deviation (RPD). The Bias, SEC, SEP, r, CV, RPD, PRESS and consistency were calculated by the following equations:

$$\text{Bias}(\%) = \frac{1}{N} \sum (x_n - y_n) \quad (1)$$

$$\text{SEC}(\%) = \sqrt{\frac{1}{N-1} \sum (x_n - y_n)^2} \quad (2)$$

$$\text{SEP}(\%) = \sqrt{\frac{1}{N-1} \sum (x_n - y_n - \text{Bias})^2} \quad (3)$$

$$r = \frac{\sum (x_n - \bar{x}_n)(y_n - \bar{y}_n)}{\sqrt{\sum (x_n - \bar{x}_n)^2 \sum (y_n - \bar{y}_n)^2}} \quad (4)$$

$$\text{CV}(\%) = \frac{\text{SEP}}{\bar{x}} \times 100\% \quad (5)$$

$$\text{RPD} = \frac{\text{SD}}{\text{SEP}} \quad (6)$$

$$\text{PRESS} = \sum (x_n - y_n)^2 \quad (7)$$

$$\text{Consistency}(\%) = \frac{\text{SEC}}{\text{SEP}} \times 100 \quad (8)$$

Where N is the number of samples; x_n is value of reference catechin; y_n is the value of NIR prediction catechin.

3 RESULT AND DISCUSSION

3.1 Data of Gambir Catechin

Table 1 shows the statistics summary for all samples catechin content in the calibration and validation set. Values of catechin content in the validation set ranges in calibration set [11], [33]. Standard deviation between the calibration set and validation set showed insignificant differences. As a result, the variance of data in the calibration and validation set are equal.

TABLE 1
MEAN REFERENCE VALUE, RANGE AND STANDARD DEVIATION
OF CATECHIN CONTENT IN CALIBRATION AND VALIDATION SET (% W/W)

Sample Set	N	Min.	Max.	Mean	S.D.
Calibration Set	108	40.71	84.79	65.65	11.88
Validation Set	54	41.54	83.91	67.21	11.76

N = number of sample; Min. = Minimum; Max. =Maximum; S.D. = standard deviation

3.2 Spectra Investigation

The original reflectance spectra of some gambir samples (Fig. 1) reveal some valleys in the region of 1000-2500 nm. However, NIR spectra in the region of 2222.22 - 2500 nm showed high noise. These valleys exist because molecule structures of catechin contain has many hydric groups (such as: C-H and O-H). As a consequence, NIR spectra in the region of 2222.22 - 2500 nm was eliminated and only spectra in the region 1000-2222.22 were used for analysis to develop the model.

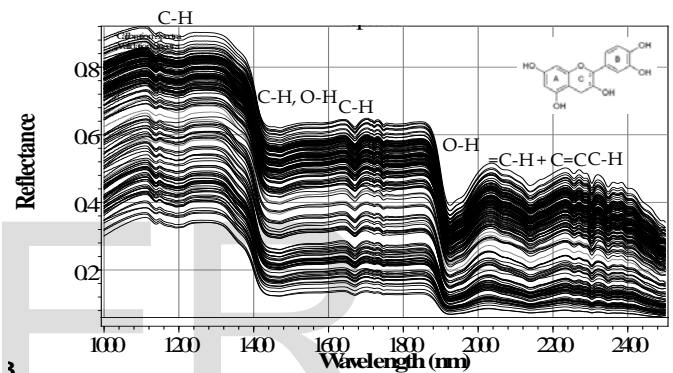


Fig. 1. The original reflectance spectra of some gambir samples

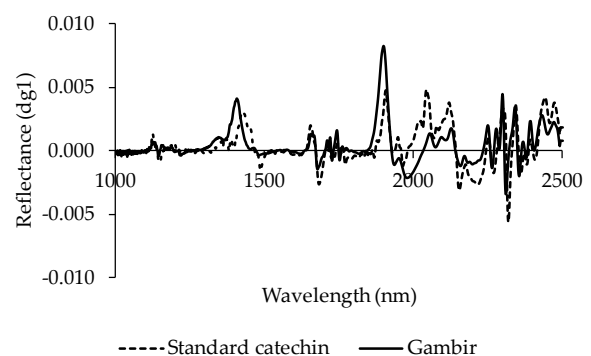


Fig.2. First derivative spectrum of gambir and catechin standard

One of the basic problems of NIR spectra was existence of overlapping spectra. Derivative approach would separate the overlapping and throw away unwanted component. To ensure the analysis using catechin spectrum, gambir spectra were compared to standard catechin spectra where both of them were treated by derivative pre-processing (Fig. 2). The comparison showed that both spectra have similar form. It proved that the analyzed spectrum was spectrum catechin in gambir.

3.3 Effect of Different Pre-processing Methods to Gambir NIR Spectra

While developing calibration model with PLS algorithm, the determination of the spectra pre-processing method and the number of PLS factors were very important stage. Fig.3 shows the form of average gambir spectra on different spectra pre-processing methods. Effect of n01 to the original spectrum was widened the range of reflectance value. If spectra pre-processing was not applied, the range of reflectance value was from 0.24 to 0.70. However, if n01 was applied, the range of reflectance value was from 0.00 to 0.99. If dg1 was applied, valleys of original spectrum would be clearer and could separate the overlapping spectrum. This effect would be more obvious when dg2 was applied. Spectra pre-processing method of n01 in combination with dg1 produced the same spectrum shape with dg1. Nevertheless, the range of reflectance value was widened. If dg1 was applied, the range of reflectance value was from -0.0022 to 0.0082 and if n01 in combination with dg1 was applied, the range of reflectance value was from -0.046 to 0.0173. It also occurred for the effect of pre-processing of n01 in combination with dg2 which produced the same spectrum shape with dg2. If dg2 was applied, the range of reflectance value was from -0.0004 to 0.0006 and if n01 in combination with dg2 was applied, the range of reflectance value was from -0.0009 to 0.0012.

3.4 The influence of different pre-processing methods and number of factors to accuracy of PLS Model

PLS is an effective method to reduce dimension in near-infrared spectroscopy analysis. The spectra information of sample component is showed by the latent variables. Furthermore, the selection of number of PLS factors is essential to reduce noise and use the whole of spectra information [6]. The model will depend on the data set and provide poor prediction if too much number of PLS factors is used. It occurred because the excess of variable X, which led over fitted solution. On contrary, the model will not be large enough to explain the data variability if too few number of PLS factors were used [8]. Fig.4 shows consistency value and V-set PRESS value, plotted with the number of PLS factors.

Decision to select optimal PLS factors was based on consistency value from 80% to 110% and the smaller V-set PRESS. If spectra pre-processing treatment was not applied, a minimum of validation error was observed in the 15 PLS factors. It provided consistency value 98.69 % and V-set PRESS value 948.51. Meanwhile, if dg1, dg2, n01 in combination with dg1, and n01 in combination with dg2 were used, the numbers of PLS-factors were 6, 5, 6 and 5, respectively. It provided consistency value 108.27 %, 100.55 %, 108.96 % and 100.07 %, respectively and V-set PRESS value 702.76, 836.43, 566.23, and 787.27, respectively. It showed that if spectra pre-processing method was applied, the number of PLS-factors could be reduced. Furthermore, derivative spectra pre-processing method could overcome the overlapping spectra and throw other components except catechin. This process remained information about catechin only.

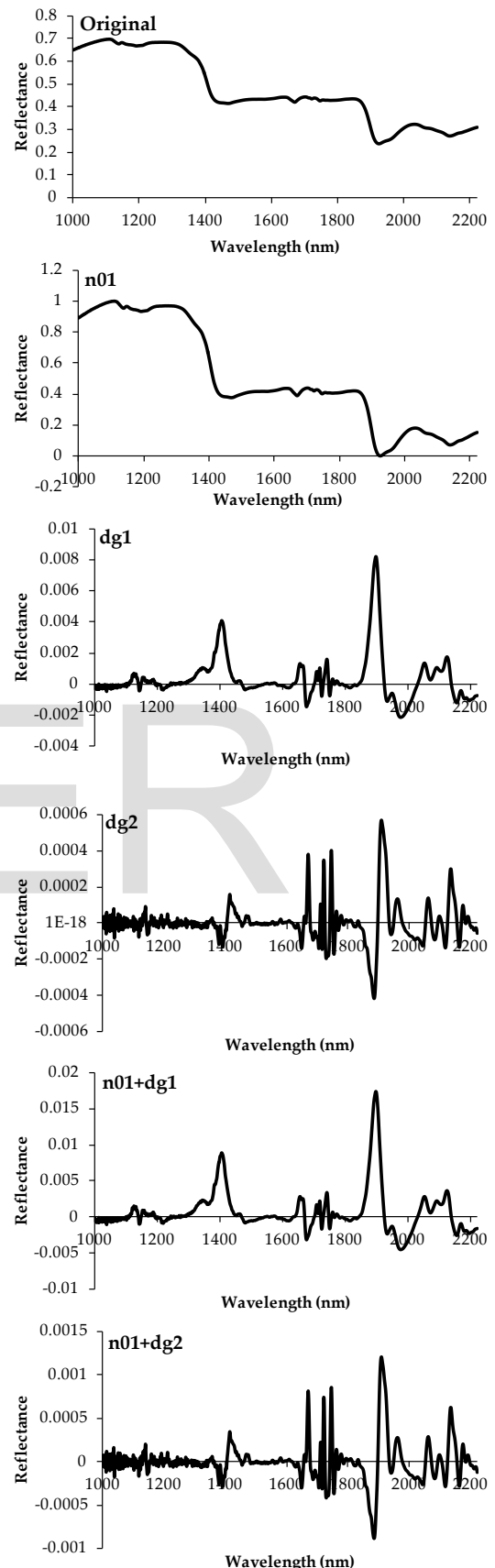


Fig. 3. Effect of pre-processing methods to form of gambir NIR spectra

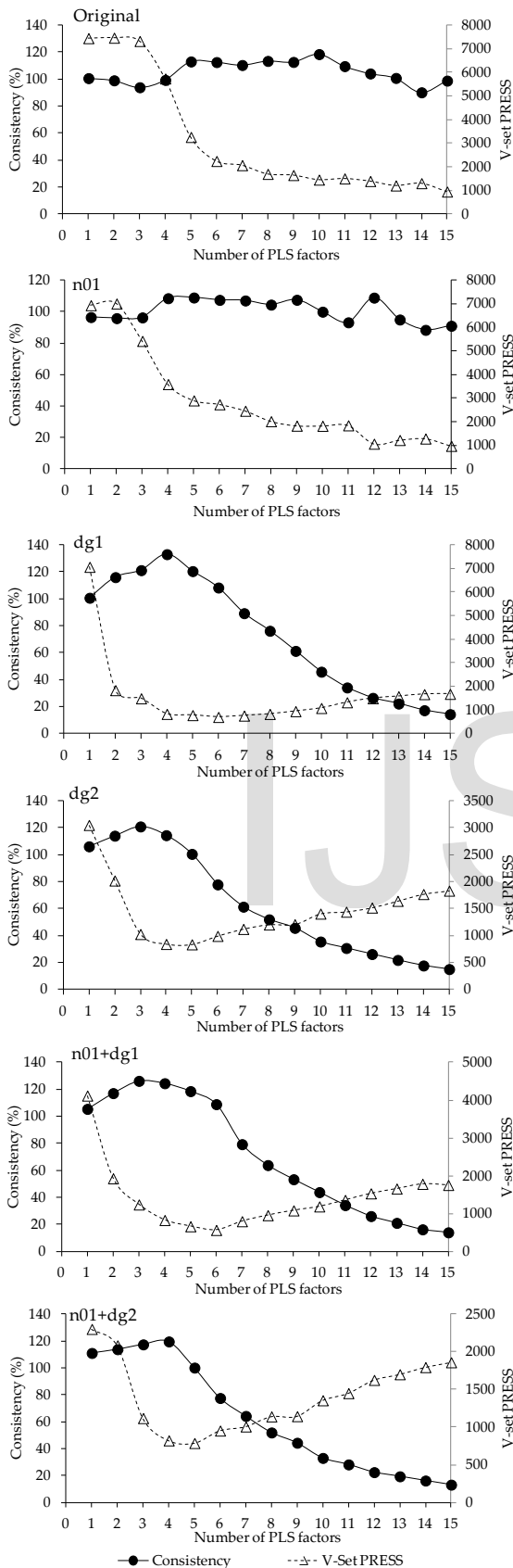


Fig.4. The influence of the number of PLS factors to consistency and V-Set-PRESS of different spectra pre-processing

The optimal number of PLS factors using n01 was 15. It provided consistency value 90.94 % and V-set PRESS value 974.74. This large number indicated that few numbers of factor could not explain the complex correlation between NIR spectra and catechin content. Chen et al., (2009) also obtain the large optimum number of PLS factors for the four types of catechin in green tea, which are from 9 to 14 factors when using pre-processing method of the MC, MSC and SNV. Cozzolino et al. (2008) in their study also obtain 13 PLS factors to measure condensed tannins in red grape homogenates using SNV. It could be concluded that the group of scatter correction pre-processing methods such as SNV, MSC and normalization required a large number of PLS factors.

Performance of parameters of the model using different pre-processing methods to predict catechin content in gambir is shown in Table 2. Spectra pre-processing of n01 in combination with dg1 produced the best calibration model compared to other spectra pre-processing. This combination increased correlation coefficient between catechin references and predicted from 0.94 to 0.95 whilst SEP value decreased from 4.14 % to 3.27 % and number of PLS factors decreased from 15 to 6. In addition, the fairly high catechin content in gambir (Table 1) also caused the application of n01 in combination with dg1 showed a better result than n01 in combination with dg2 method.

TABLE 2
 THE RESULT OF CALIBRATION AND VALIDATION OF GAMBIR CATECHIN CONTENT BY VARIOUS DIFFERENCE PRE-PROCESSING METHODS

Pre-processing	PLS Factor	Consistency (%)	Calibration Set		Validation Set				
			SEC (%)	r	SEP (%)	r	CV (%)	Bias (%)	RPD
Original	15	98.70	4.09	0.94	4.14	0.94	6.16	-0.86	2.84
n01	15	90.94	3.88	0.95	4.26	0.93	6.34	-0.48	2.76
dg1	6	108.27	3.94	0.94	3.64	0.95	5.42	-0.02	3.23
dg2	5	100.55	3.99	0.94	3.97	0.94	5.91	-0.09	2.96
n01+dg1	6	108.96	3.56	0.95	3.27	0.96	4.86	0.10	3.60
n01+dg2	5	100.07	3.86	0.95	3.85	0.95	5.73	0.03	3.05

The model with application of pre-processing n01+dg1 method showed higher accuracy and precision than the original model. It indicated that pre-processing was important prior to develop a model as it would improve the accuracy and precision of calibration model [23], [26], [29].

3.5 Prediction of Catechin Content in Gambir

The best calibration model for predicting catechin content was $y = 0.91x + 5.90$. It is presented by scatter plots in Fig.5. The model provided low value of SEC (3.56 %) and SEP (3.27 %) compared to other models. Since SEP value was not greater than two times of SEC, over fitting can be prevented [17]. The slight different values of SEC and SEP indicated that the model had high precision.

Bias values closed to zero indicated that model has high accuracy [31]. It led the model would be maintained for the prediction of catechin content in gambir. Furthermore, the model obtained high correlation between the reference catechin content and NIR prediction catechin ($r = 0.95$). The good model has high correlation coefficient ($r > 0.90$) [31]. The model obtained CV value $< 5\%$, which indicated that the model

was appropriate to predict catechin content in gambir in the new data set.

The model had RPD value 3.60. RPD is ability of model to predict component efficiently [31]. RPD values below 1.5 indicated that the calibration could not be used, between 1.5 and 2.0 is possible to distinguish high and low values, between 2.0 and 2.5 is possible for quantitative prediction, while values between 2.5 and 3.0 and above 3.0 is considered good and very good respectively [20]. Based on obtained RPD values, the resulted model was very good to predict the gambir catechin content.

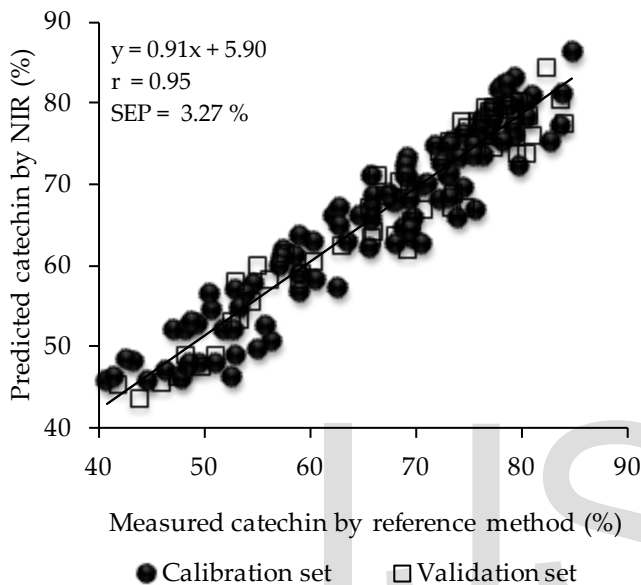


Fig. 5. Scatter plot of measured versus predicted catechin content using n01 in combination with dg1 and 6 PLS factors

4 CONCLUSIONS

This study demonstrated the capability of NIR spectroscopy technology could replace traditional as well as chemical method to determine catechin content in gambir. The prediction model was developed using the PLS algorithm to explain relationship between NIR spectra data and catechin content. The pre-processing methods and number PLS factors could affect the accuracy of the NIR calibration models. The best calibration model was obtained when n01 in combination with dg1 pre-processing method and 6 PLS factors were applied. It had a low SEC and SEP, a high r and a slight difference between SEC and SEP values. This study showed that there is possibility to develop non-destructive method and rapid analysis to measure gambir catechin content.

ACKNOWLEDGMENTS

The authors express their appreciation for the financial support by Strategic Research Grant of DGHE Ministry of Education and Culture, The Republic of Indonesia No. 80/IT3.41.2/L1/SPK/2013 (02 Mei 2013).

REFERENCES

1. Anggraini, T., Tai, A., Yoshino, T. and Itani, T. (2011). Antioxidative activity and catechin content of four kinds of *Uncaria gambir* extracts from West Sumatera, Indonesia. *African Journal of Biochemistry Research* 5:33-38.
2. Apea-Bah, F.B., Hanafi, M., Dewi, R.T., Fajriah, S., Darmawan, A., Artanti, N., Lotulung, P., Ngady mang, P. and Minarti, B. (2009). Assessment of the DPPH and α -glucosidase inhibitory potential of gambir and qualitative identification of major bioactive compound. *J. Med. Plants* 3:736-757.
3. Blanco, M. and Villarroya, I. (2002). NIR Spectroscopy: a rapid-response analytical tool. *Trends in analytical chemistry* 21: 240-250.
4. Buddenbaum, H. and Steffens, M. (2012). The effects of spectral pre-treatment on chemometric analyses of soil profiles using laboratory imaging spectroscopy. *Applied and Environmental Soil Science*: 1-12.
5. Cen, H. and He, Y. (2007). Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends in Food Science & Technology* 18:72-83.
6. Chen, H., Song, Q., Tang, G., Feng, Q. and Lin, L. (2013). The combined optimization of Savitzky-Golay smoothing and multiplicative scatter correction for FT-NIR PLS models. *ISRN Spectroscopy*: 1-9.
7. Chen, Q.S., Zhao, J.W., Caitep, S. and Guo, Z.M. (2009). Simultaneous analysis of main catechin content in green tea (*Camellia sinensis* (L.)) by Fourier transform near infrared reflectance (FT-NIR) spectroscopy. *Food Chemistry* 113: 1272-1277.
8. Cozzolino, D., Cynkar, W.U., Shah, N., Damberg, R.G. and Smith, P.A. (2009). A brief introduction to multivariate methods in grape and wine analysis. *International Journal of Wine Research* 1: 123-130.
9. Cozzolino, D., Cynkar, W.U., Shah, N., Damberg, R.G., Mercurio, M.D. and Smith, P.A. (2008). Measurement of condensed tannins and dry matter in red grape homogenates using near infrared spectroscopy and partial least squares. *J. Agric. Food Chem.* 56: 7631-7636.
10. di Egidio, V., Sinelli, N., Giovallelli, G., Moles, A. and Casiraghi, E. (2010). NIR and MIR spectroscopy as rapid methods to monitor red wine fermentation. *Eur Food Res Technol* 230: 947-955.
11. Elfadl, E., Reinbrecht, C. and Claupein, W. (2010). Development of near infrared reflectance spectroscopy (NIRS) calibration model for estimation of oil content in a worldwide safflower germplasm collection. *International Journal of Plant Production* 4: 259-270.
12. Gallego, R.F., Hierro, J.M.H., Gonzalo, J.C.R., Bailón, M.T.E. (2011). Determination of phenolic compounds of grape skins during ripening by NIR spectroscopy. *LWT-Food Science and Technology* 44:847-853.
13. Gallego, R.F., Hierro, J.M.H., Gonzalo, J.C.R., Bailón, M.T.E. (2010). Feasibility study on the use of near infrared spectroscopy to determine flavanol in grape seeds. *Talanta*. 82:1778-1783

14. Guggenbichler, W., Huck, C.W., Kobler, A., Popp, M., Bonn, G.K. (2006). Near infrared spectroscopy, cluster and multivariate analysis - contributions to wine analysis. *Journal of Food, Agriculture & Environment* 4: 98 - 106.
15. Gumbira-Sa'id, E. (2009). Review of agroindustrial strategic studies, researches and development in Indonesia: the case of oil palm, cacao and gambir. *j tek ind pert* 19: 45-55.
16. Heitzman, M.E., Neto, C.C., Winiarz, E., Vaisberg, A.J. and Hammond, G.B. (2005). Ethnobotany, phytochemistry and pharmacology of *Uncaria* (Rubiaceae). *Phytochem* 66: 5-29.
17. Hruschka, W.R. (1990). Data analysis: wavelength selection methods. In P. Williams and K. Norris (eds.) *Near-Infrared technology in the agricultural and food industries*. St. Paul, MN: American Association of Cereal Chemists Inc.: 35-55.
18. Jiang, B., Huang, Y.D., Li, W. and Liu, L. (2007). Non-destructive and rapid analysis of resin and volatile contents in carbon fibre/epoxy resin prepreg cloth by near-infrared spectroscopy. *Iranian Polymer Journal* 16: 319-326.
19. Marengo, E., Bobba, M., Robotti, E. and Lenti M. 2004. Hydroxyl and acid number prediction in polyester resins by near infrared spectroscopy and artificial neural networks. *Analytica Chimica Acta* 511: 313-322.
20. Mouazen, A.M., Saeys, W., Xing, J., De Baerdemaeker, J. and Ramon, H. (2005). Near infrared spectroscopy for agricultural materials: an instrument comparison. *J. Near Infrared Spectrosc* 13: 87-97.
21. Nicolari, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I. and Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. *Postharvest Biology and Technology* 46: 99-118.
22. Nircal 5.2 Manual. (2007). Büchi Labortechnik AG, CH-Flawil, Switzerland.
23. Ouyang, Q., Chen, Q., Zhao, J. and Lin, H. (2012). Determination of amino acid nitrogen in soy sauce using near infrared spectroscopy combined with characteristic variables selection and extreme learning machine. *Food Bioprocess Technol*. DOI 10.1007/s11947-012-0936-0.
24. Rinnan, Å., Van den Berg, F. and Engelsen, S.B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *Trends in Analytical Chemistry* 28:1201-1222.
25. Rodgers, J., Beck, K. (2009). NIR characterization and measurement of the cotton content of dyed blend fabrics. *Textile research journal* 79:675-686.
26. Schultz, H., Engelhardt, U.H., Wegent, A., Drews, H.H. and Lapczynski, S. (1999). Application of near-infrared reflectance spectroscopy to the simultaneous prediction of alkaloids and phenolic substances in green tea leaves. *J Agri Food Chem* 47:5064-5067.
27. Swierenga, H., de Weijer, A.P., van Wijk, R.J. and Buydens, L.M.C. (1999). Strategy for constructing robust multivariate calibration models. *Chemometrics and Intelligent Laboratory Systems* 49: 1-17
28. Taniguchi, S., Kuroda, K., Doi, K., Inada, K., Yoshikado, N., Yoneda, Y., Tanabe, M., Shibata, T., Yoshida, T. and Hatano, T. (2007). Evaluation of gambir quality based on quantitative analysis of polyphenolic components. *Yakugaku Zasshi*. 127(8):1291-1300.
29. Udelhoven, T., Emmerling, C. and Jarmer, T. (2003). Quantitative analysis of soil chemical properties with diffuse reflectance spectrometry and partial least-square regression: A feasibility study. *Plant and Soil* 251: 319-329.
30. Whitacre, E., Oliver, J., Van den Broek, R., Van Engelen, P., Kremers, B., Van der Horst, B., Stewart, M. and Jansen-Beuvinck A. (2003). Predictive analysis of cocoa procyanidins using near-infrared spectroscopy techniques. *Journal of Food Science* 68: 2618 - 2622.
31. Williams, P. and Norris, K. (1990). *Near-infrared technology in the agricultural and food industries*. American Association of cereal chemical, Inc. St. Paul, USA: 146.
32. Xu, L., Zhou, Y.P., Tang, L.J., Wu, H.L., Jiang, J.H., Shen, G.L. and Yu, R.Q. (2008). Ensemble preprocessing of near-infrared (NIR) spectra for multivariate calibration. *Analytica Chimica Acta*, 616: 138-143.
33. Yan, H., Chang, W.X. and Wen, D.D. (2009). Rapid determination of moisture and protein contents in silver carp surimi by Fourier transform near-infrared (FT-NIR) Spectrometry. *Asian Fisheries Science* 22: 337-345.
34. Zhang, M.H., Luypaert, J., Fernández Pierna, J.A., Xu, Q.S. and Massart, D.L. (2004). Determination of total antioxidant capacity in green tea by near-infrared spectroscopy and multivariate calibration. *Talanta* 62: 25-35.