Determination of Catechin Content in Gambir Powder from Dried Gambir Leaves Quickly using FT NIR PLS Model

Andasuryani[#], Y.A. Purwanto[*], I W Budiastra[*], K. Syamsu[**]

[#] Department of Agricultural Engineering Faculty of Agricultural Technology Andalas University Kampus Limau Manis Padang 25163,
West Sumatera Indonesia
E-mail: andasuryani@gmail.com

*Department of Mechanical and Biosystem Engineering Faculty of Agricultural Engineering and Technology Bogor Agricultural University
Kampus IPB Darmaga Bogor 16680, West Java Indonesia
E-mail: arispurwanto@gmail.com, wbudiastra@yahoo.com

**Department of Agro-Industrial Technology Faculty of Agricultural Engineering and Technology Bogor Agricultural University
Kampus IPB Darmaga Bogor 16680, West Java Indonesia
E-mail: khaswar@yahoo.com

Abstract— Conventional production process of gambir often produces gambir with low content of catechin. Engineering of production processes of gambir leaves to produce gambir powder has been developed by previous researchers. The objective of this study was to develop a calibration model to predict the content of catechin in gambir powder from dried gambir leaves quickly using FT-NIR PLS model. Reflectance spectra of gambir powder from dried gambir leaves obtained at a wavelength of 1000 to 2500 nm. Spectra pre-processing treatment method used was a combination between normalization between 0 and 1 (n01) method and first derivative Savitzky-Golay 9 points (dg1). The result showed that correlation coefficient and standard error of prediction (SEP) were 0.99 and 2.10 % respectively, obtained when used 6 factors of partial least square (PLS). Calibration model developed has shown higher accuracy and precision to predict catechin content of gambir powder from dried gambir leaves.

Keywords— gambir, catechin, FT-NIR spectroscopy, PLS.

I. INTRODUCTION

Leaves and young twigs of gambir (Uncaria gambir Roxb.) are part of the plant were processed into gambir. Gambir is used as a complement to eat betel, medicines, raw materials of pharmaceutical and food industry [1]. In addition, gambir widely used by other industries such as batik fabric, paint, leather tanning, biopesticides, growth hormone, and as a ingredient of supplementary foods [2].

Processing of gambir done simply through several stages which are boiling, pressing, sedimentation, molding and drying. However this production process generated raw gambir with low quality and unstandard. Some of the factors that cause the low quality of gambir produced from traditional processing are 1) raw material: harvesting the leaves and twigs with different harvesting time, leaf damage due to pests and diseases and delay the processing of the leaves and twigs after harvest, 2) human resources: lack of knowledge, awareness and commitment to maintain the quality of gambir and prioritize to quantity of production 3)

environment: the presence of multiple sources of impurities in producing such as from ground floor, area of compression, sap reservoirs, ash and soot from the furnace and drying area, 4) facilities: facilities used in producing gambir cannot guarantee low levels of impurities in gambir and 5) working methods: method does not produce high quality of gambir [3].

In order to meet the quality requirements, gambir production process should obtain catechin extract as much as possible. Reference [4] has studied the production of gambir powder from dried leaves which meet export quality standard using a spray dryer. The study found that the highest level of catechin obtained was 93.4% and the lowest was 24.36%. Gambir powder produced from the processing of dried gambir leaves can replace the raw gambir production.

To fulfill the standard of export quality, it is necessary to test the quality of chemical gambir that can be done by measuring its chemical components. Near Infrared (NIR) spectroscopy is one of measurement methods that can

analyze chemical components quickly, precisely and accurately. This method has been successfully used to determine the chemical components of the material, including determination of alkaloids and phenol compounds in green tea leaves [5], the determination of carbohydrates, total acid, tartaric acid, malic acid, pH and poliphenol content in wine [6], determination of condensed tannins and dry matter in red grape homogenates [7], determination of the major catechin content in green tea [8], and determination of catechin content in raw gambir in the form of powder [9] and solid [10].

NIR spectroscopy method requires a calibration process that NIR wave data obtained can be used to determine the chemical content of material being measured. Therefore, calibration methods are needed that can provide information related to the chemical content of the material being measured. One of the methods is partial least squares (PLS). The objective of this research was to develop a calibration model to predict the catechin content in gambir powder from dried gambir leaves quickly using Fourier Transform (FT) - NIR PLS model.

II. MATERIAL AND METHOD

A. Sample preparation

Samples of gambir powder were made from dried gambir leaves obtained from Siguntur Village, District Koto XI Tarusan, Pesisir Selatan District, West Sumatera. Other materials that were needed for chemical analysis of materials were catechin standard ($\geq$ 98%) (Sigma Aldrich), aquades and ethyl acetate.

Production process of powder gambir from dried gambir leaves samples refered to procedure performed by [4] which were: 1) drying fresh leaves, 2) size reduction, 3) extraction with solvents (solid: solvent, 1: 5), 4) shaker waterbath (T =500C, t = 90 minutes), 5) maceration (3 x 24 hours), 6) filtration, and 7) drying. Gambir powder produced from the dried gambir leaves subsequently used for analysis. References method used to determine the catechin content of gambir powder from dried gambir leaves was spectroscopic ultraviolet method that refers to the Indonesia National Standard (SNI) 01-3391-2000 [11] by using an instrument spectrophotometer U-2010, Hitachi.

The data were used as reference data while build NIR calibration model to predict the catechin content in gambir powder from dried gambir leaves. The number of gambir powder from dried gambir leaves samples was 73. Based on the reference value of the laboratory, the sample was divided into two parts randomly. The first group consists of two thirds of the total sample used for the development of the model called the calibration set (C-set) and the second group as many as one third of the total samples were used for testing the model called the validation set (V-set).

B. Spectra acquisition

Measurement method of gambir powder from dried gambir leaves sample spectra was similar to the method used by [9] for gambir powder from raw gambir. This method started by placing the sample in a petri dish using a spectrophotometer Buchi NIRFlex N-500 solids. Software NIRWare 1.2 (Büchi Labortechnik AG, Flawil, Switzerland)

used to operate the instrument and colect spectra data at room temperature of 25 ^{0}C. Original spectra data were collected by measuring the reflectance of the sample with a data interval of 4 cm^{-1} on a scale of 10000 to 4000 cm^{-1} (1000-2500 nm) (See Fig.1). Measurements were taken three times at different positions for each sample. Spectrum produced by each sample were averaged and then used for analysis and stored as reflectance (R).
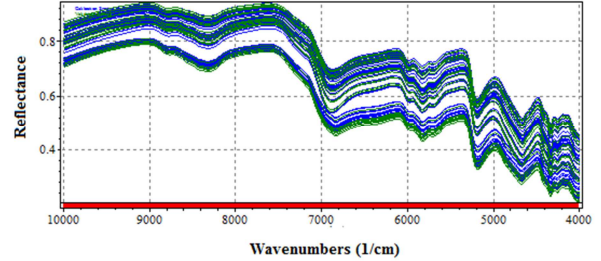

Fig.1 The original spectra of gambir powder from dried gambir leaves

C. Chemometric analysis

Chemometric analysis done using software NIRCal 5.2 (Büchi Labortechnik AG, Flawil, Switzerland). Calibration model developed using PLS algorithm with 1-15 PLS factors. Meanwhile, selection of the optimum number of PLS factors based on the value of Predicted Residual Error Sum Square (PRESS) in validation set (V-set-PRESS) and the consistency value that have V-PRESS set as small as possible and the value of consistency between 80-110% [12]. Pre-processing of spectra data used in this study was combination of normalization between 0 and 1 (n01) and first derivative Savitzky-Golay 9 point (dg1).

Statistical parameters used for the analysis the model were:

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

(b) Standard error of calibration set (SEC)
$$\text{SEC} = \sqrt{\frac{1}{N-1}\sum(x_n - y_n - \text{Bias})^2}$$

(c) Standard error of validation set (SEP)
$$\text{SEP} = \sqrt{\frac{1}{N-1}\sum(x_n - y_n - \text{Bias})^2}$$

(d) Coeffisien of correlation (r)
$$r = \frac{\sum(x_n - \bar{x}_n)(y_n - \bar{y}_n)}{\sqrt{\sum(x_n - \bar{x}_n)^2 \sum(y_n - \bar{y}_n)^2}}$$

(e) Coefficient of variance (CV)
CV in C-Set: $\text{CV} = \frac{\text{SEC}}{\bar{x}} \times 100$

CV in V-Set: $\text{CV} = \frac{\text{SEP}}{\bar{x}} \times 100$

(f) Ratio of prediction to deviation (RPD).

$$RPD = \frac{SD}{SEP}$$

N = number sample; x_n = reference value; y_n = predicted value by NIR.

III. RESULT AND DISCUSSION

Table 1 shows the statistics of catechin content in gambir powder from dried gambir leaves in the calibration set and validation set were measured with a reference method. The result of measurements of catechin content in validation set was within the range of value of catechin content in calibration set. These results were similar to study by [13] that acknowledged that the set of calibration had the highest and lowest value of the properties.

TABLE I
STATISTICS OF CATECHIN CONTENT IN GAMBIR POWDER
FROM DRIED GAMBIR LEAVES

Sample set	Mean (%)	Standard deviation (%)	Maximum (%)	Minimum (%)	SN
Calibration set	82.16	12.89	99.69	61.79	49
Validation set	82.88	13.47	99.36	62.56	24

A. Effect of Pre-processing

A reliable, accurate and stable calibration model could explain relationship between NIR spectrum data and the desired chemical content. Pre-processing is required in order to obtain a good calibration model, because data obtained from NIR spectrophotometer sample contains not only information but also noise. Pre-processing of normalization was designed to reduce baseline variations caused by differences of particle size [14]. Pre-processing of derivatives may reduce the effect of the baseline and improve the absorption peaks [12].
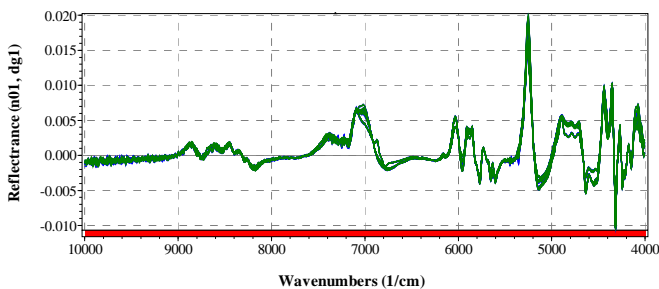

Fig.2 NIR spectra of gambir powder from dried gambir leaves obtained from combination pre-processing of n01 and dg1

The shape of gambir powder from dried gambir leaves reflectance spectrum with pre-processing method combination between n01 and dg1 (n01, dg1) is shown in Fig. 2 with the reflectance values range from -0010 - 0.020. It can be seen that the original spectra (shown in Fig. 1) has become close to one another and baseline variations caused by differences in particle size has been discarded.

B. Calibration Development

To obtain a large amount of spectra information provided by NIR spectrophotometers were needed in building the model calibration as well as to obtain a high correlation in the spectra, a method was needed to reduce variables are not correlated. It only left variables contained relevant information to the sample. Principal Component Analysis (PCA) is a well-known method and has been widely used [15]. Reference [16] acknowledged that PCA is a technique to reduce the amount of spectrum data that is intended to find the main component (PC) which linear combination with the original variables. Furthermore, the result of the analysis showed that 6 components or factors were the optimum number of PLS factors to predict catechin content in gambir powder from dried gambir leaves. The number of PLS factors provided value consistency (comparison between SEC and SEP) of 100.27% and the value of V-sets PRESS (sum of squared error on the validation set) of 101.93%.

Model with pre-processing combination n01, dg1 showed a good calibration model to predict catechin content in gambir powder from dried gambir leaves. It also showed a better performance of calibration model compare to the original model (Table 2). The model showed a low value of SEC and SEP, high r value and small difference between SEC and SEP. The use of a combination pre-processing (n01, dg1) caused the value of the correlation coefficient between the reference catechin and predicted NIR increased from 0.96 into 0.99, the value of the SEC declined from 3.69% to 2.11%, the value of SEP decreased from 3.48% to 2.10%, the value of CV on the set calibration and validation sets successively decreased from 4.49% and 4.19% to 2.57% and 2.54%, RPD values increased from 3.88 to 6.40 and the number of PLS factors was reduced from 9 to 6. The value of r above 0.90 is high [17] and value of CV at calibration set <5% and at validation set <10% showed a reliable calibration model [18] and fit to predict catechin content in gambir powder from dried gambir leaves in the new data set of gambir. Based on RPD values obtained, the resulting model showed an excellent model in predicting catechin content in the gambir powder from dried gambir leaves. This is because the RPD values obtained was > 3. In accordance with that acknowledged by [19], if a model had a value of RPD > 3, the model was able to predict very well. RPD value obtained was highly dependent on the diversity of data. The greater the diversity the greater RPD values obtained.

TABLE II

PERFORMANCE PARAMETER MODEL TO PREDICT CATECHIN CONTENT IN GAMBIR POWDER FROM DRIED GAMBIR LEAVES

Pre-processing	F	Consistency (%)	Calibration Set			Validation Set				
			SEC (%)	r	CV (%)	SEP (%)	r	CV (%)	Bias (%)	RPD
Original	9	106.06	3.69	0.96	4.49	3.48	0.97	4.19	-0.03	3.88
n01, dg1	6	100.27	2.11	0.99	2.57	2.10	0.99	2.54	0.05	6.40

Fig. 3 shows a plot of actual data and data of NIR predicted of catechin content in gambir powder from dried gambir leaves. Based on the value of R2, the resulting model could predict catechin content in gambir powder from dried gambir leaves very well (R2 > 0.91) [19]. R2 indicated the percentage of of variance of variable Y specified by variable X. The results indicated that 97% of catechin content value of NIR predicted was determined by value of reference catechin content.
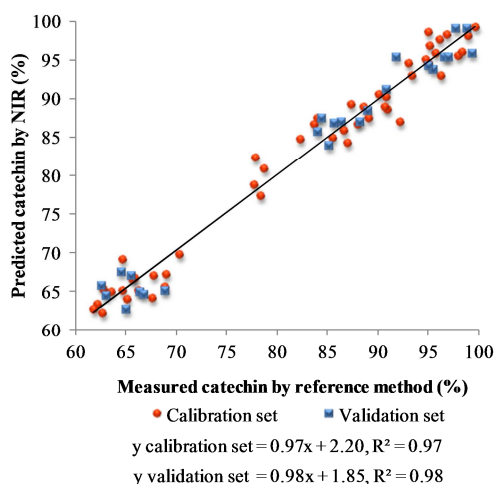

y calibration set $= 0.97x + 2.20$, $R^2 = 0.97$
y validation set $= 0.98x + 1.85$, $R^2 = 0.98$

Fig.3 Plot of actual data and data of NIR predicted of catechin content in gambir powder from dried gambir leaves

IV. CONCLUSIONS

This study demonstrated the capability of NIR spectroscopy technology to determine catechin content in gambir powder from dried gambir leaves using the PLS algorithm. The pre-processing methods combination n01 with dg1 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.

NOMENCLATURE

C-set   calibration set
CVc   coefficient of variation on calibration set
CVv   coefficient of variation on validation set
dg1   first derivative Savitzky-Golay 9 point
F   factor PLS
FT-NIR   Fourier transform-NIR
n01   normalization between 0 and 1
n01, dg1 combination between n01 and dg1
NIR   near infrared

nm   nanometer
PCA   principal componet analysis
PLS   partial least square
PRESS   predicted residual error sum square
r   correlation coefficient
R^2 determination coefficient
SD   standard deviation
SEC   standard error of calibration set
SEP   standard error of validation set
SG   Savitzky-Golay
SN   sample number
SNI   Standar Nasional Indonesia
V-set   validation set

REFERENCES

[1] Dhalimi A. 2006. Problems gambir (*Uncaria gambir* L.) in West Sumatra and alternative solutions. *Perspektif* 1: 46 – 59.
[2] Ermiati. 2004. The cultivation, processing and feasibility of farming gambir (*Uncaria gambir* Roxb.) In the District 50 City. *Bul TRO* 15: 50-63.
[3] 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.
[4] Anungputri PS, K Syamsu, Gumbira Sa'id E. 2010. Study of the manufacture powder from dried gambir leaves (*Uncaria gambir* Roxb.) using a spray dryer. Faculty of Agricultural Technology IPB. Bogor.
[5] Schultz H, Engelhardt UH, Wegent A, Drews HH, Lapczynski S. 1999. Application of near-infrared reflectance spectroscopy to the simultaneous prediction of alkaloids and phenolic subtances in green tea leaves. *J Agri Food Chem* 475:5064-5067.
[6] Guggenbichler W, Huck CW, Kobler A, Popp M, Bonn GK. 2006. Near infrared spectroscopy, cluster and multivariate analysis – contributions to wine analysis. *J Food Agric Environ* 4: 98 - 106.
[7] Cozzolino D, Cynkar WU, Shah N, Dambergs RG, Mercurio MD, Smith PA. 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.
[8] Chen QS, Zhao JW, Caitep S, Guo ZM. 2009. Simultaneous analysis of main catechin content in green tea (*Camellia sinensis* (L.)) by Fourier transform near infrared reflectance (FT-NIR) spectroscopy. *Food Chem* 113: 1272-1277.
[9] Andasuryani, YA Purwanto, I W Budiastra, K Syamsu. 2013a. Determination of catechin as main bioactive component of gambir (*Uncaria gambir* Roxb.) by FT-NIR Spectroscopy. *J Med Plants* 7(41):3076-3083.
[10] Andasuryani, YA Purwanto, I W Budiastra, K Syamsu. 2013b. Non destructive and rapid analysis of catechin content in gambir (*Uncaria gambir* Roxb.) using nir spectroscopy. IJSER 4: 383-389.
[11] [BSN] Badan Standarisasi Nasional. 2011. Gambir. http://sisni.bsn.go.id.
[12] Nircal 5.2 Manual. 2007. Büchi Labortechnik AG, CH-Flawil. Switzerland.
[13] Yan H, Chang WX, Wen DD. 2009. Rapid determination of moisture and protein contents in silver carp surimi by Fourier transform near-infrared (FT-NIR) Spectrometry. *Asian Fish Sci* 22: 337-345.

[14] Udelhoven T, Emmerling C, Jarmer T. 2003. Quantitative analysis of soil chemical properties with diffuse reflectance spectrometry and partial least-square regression: A feasibility study. *Plant Soil* 251: 319–329.

[15] Blanco M, Villarroya I. 2002. NIR Spectroscopy: a rapid-response analytical tool. *Trends Anal Chem* 21: 240-250.

[16] Miller JN, Miller JC. 2005. Chemometrics for analytical chemistry. Pearson Prentice Hall.

[17] Murray I, Williams PC. 1990. Chemical Principles of Near Infrared Techology. In: Williams P, Norris K, editor. *Near-infrared Technology in the Agricultural and Food Industries*. St. Paul, MN: American Association of Cereal Chemists Inc. page: 17-34.

[18] Mlček J, Šustová K, Simeonovová J. 2006. Application of FT NIR spectroscopy in the determination of basic chemical composition of pork and beef. *Czech. J. Anim* Sci 51: 361–368.

[19] Mouazen AM, Saeys W, Xing J, De Baerdemaeker J, Ramon H. 2005. Near infrared spectroscopy for agricultural materials: an instrument comparison. *J Near Infrared Spectrosc* 13: 87-97.