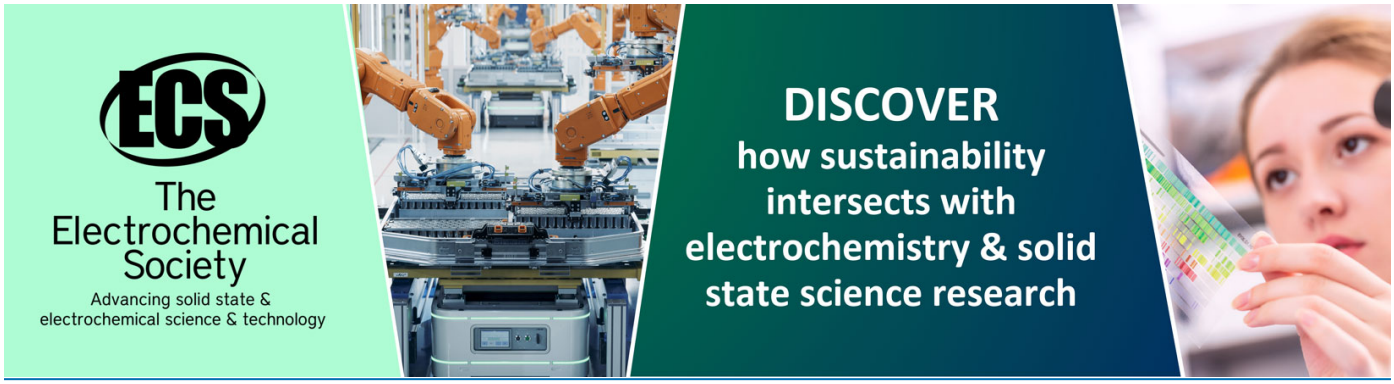
Determination of Chemical Content of Vanilla Pods (*Vanilla planifolia*) Non-destructively Using NIR Spectroscopy



Determination of Chemical Content of Vanilla Pods (*Vanilla planifolia*) Non-destructively Using NIR Spectroscopy

I W Budiastra[1,2], H Ramadhana[1], N Nurfadila[3]

1) Department of Mechanical and Biosystem Engineering, IPB University, Bogor, Indonesia
2) Center for Research on Engineering Application for Tropical Agriculture, IPB University, Bogor, Indonesia
3) Phytopathology Laboratory, SEAMEO BIOTROP, Bogor 16134, Indonesia.

Email: wbudiastra@apps.ipb.ac.id

Abstract. This study aims to assess Near infrared (NIR) spectroscopy to predict vanillin and moisture content non-destructively. Twenty-four samples of vanilla pods (@55 grams/sample) from Sumatra and Sulawesi Island were tested. The reflectance spectra of vanilla pod were measured using a NIRFlex N-500 spectrometer at the wavelength of 1000-2500 nm. After that, the vanillin and moisture content were measured using the reference method. The reflectance spectra were transformed to absorbance spectra and several pre-treatment methods of NIR spectra were applied, and the results were calibrated with chemical data using principal component regression (PCR) and partial least square (PLS) methods. The best estimation for vanillin content using absorbance spectra with OSC pre-treatment at 3 PLS factors with the accuracy parameters of r = 0.91, SEC = 1.31%, SEP = 1.34%, CV = 45.60%, RPD = 2.30 with a consistency of 97.96%. The best prediction of vanillin content using PCR method was obtained with SNV pre-treated spectra at 6 PCR factors with the accuracy parameters of r = 0.89, SEC = 1.47%, SEP = 1.49%, CV = 43.21%, RPD = 2.09 and 101.40% consistency. The best estimation for water content using the PLS method is using the SNV pre-treatment spectra with a PLS factor of 5 with the accuracy parameters of r = 0.99, SEC = 2.11%, SEP = 2.13%, CV = 6.12%, RPD = 7.98 with consistency = 98.86%. While using the PCR method, the best estimation for moisture content was also obtained by SNV pre-treated spectra at 4 PCR factors with the accuracy values of r = 0.99, SEC = 2.33%, SEP = 2.25%, CV = 6.25%, RPD = 7.58 and consistency = 103.90%. The PLS calibration provide a better accuracy than the PCR. The NIR spectroscopy associated with the selected pre-treatments and factor numbers of PLS and PCR can be used for determination of chemical content of vanilla pods.

1. Introduction

Indonesia is the second largest vanilla producer in the world after Madagascar with production reaching 20% to 30%. Vanilla exports in 2020 reached 363 thousand tons with the value of 60.2 million USD. The main export destination country is the United States which amounts to 254 thousand tons or 70% of total exports [1]. Vanilla production in Indonesia is mostly carried out in the provinces of Central Java, East Java, Bali, NTT, NTB, Sulawesi, West Sumatra and North Sumatra. Based on the SNI 01-0010-2002, there are several standard indicators of vanilla pod quality, including vanillin and water content [2]. The quality of vanilla should be maintained so that vanilla from Indonesia can still compete in the international market. The accurate determination of vanillin and moisture content so far is by

using chemical laboratory analysis, but this method takes a long time and expensive. So, a more efficient method of determining the moisture and vanillin content of vanilla pods is necessary for ensuring the chemical quality of Indonesia' vanilla.

Near infrared spectroscopy (NIRS) is one of the non-destructive methods that extensively studied for many purposes in industries such as chemical quality of products, products adulteration, and products classification. It provides several advantages ; fast and easy to do, it does not require chemicals, and able to perform internal quality tests on agricultural products non-destructively. This method has also widely studied in agriculture field to determine physicochemical content of agricultural products such as catechin in gambir [3], fat and free fatty acid in jatropha [4], trigonelline, chlorogenic acid, caffeine content of coffees [5][6], and classification of coffee origin [7]. Research on the estimation of vanillin and moisture content in Thailand' vanilla using NIRS has been carried out [8]. It is reported that the vanillin and moisture content in vanilla can be predicted accurately by NIRS and PLS method. This study aims to assess NIR spectroscopy to predict vanillin and moisture content of Indonesia' vanilla pods non-destructively using PCR and PLS calibration method.

2. Materials and Method

2.1 Materials and Apparatus

The samples used were commercial vanilla pods obtained from the islands of Sulawesi and Sumatra. The number of samples used was 24 samples (@55 grams). The materials used to determine the chemical content of vanilla pods were N_aOH, ethanol, and filter paper. The instrument used to measure the reflectance of vanilla is one unit of NIRFlex N-500 spectrometer (Buchi, Switzerland) equipped with a fiber optic solid and a computer. The other apparatus used include an absorbance spectrophotometer, and a petri dish as a vanilla container, mortal, volumetric flask, funnel, erlenmeyer, oven, scales, and cups.

2.2 Measurement of Reflectance of Vanilla Pods

The sample of amount 55 gram was arranged in a neat and tight position placed on petri dish, and the reflectance of samples were measured using a gun (fiber optic solid) NIRFlex N-500 spectrometer at the wavelength of 1000 – 2500 nm or 4000 cm^{-1} – 10000 cm^{-1}. The NIR instrument was firstly calibrated with the tool's built-in software. The number of samples were 24 samples, each sample were measured in nine different measurement points so total spectrum obtained was 216 spectra.

2.3 Chemical Analysis

Chemical analysis was carried out to obtain data on vanillin and moisture content data from vanilla pods. Furthermore, the data is used as reference data in building the NIR calibration model. The measurement of water content is based on SNI 01-2891-1992 using the oven method with a temperature of 105 C for 3 hours [9] . Meanwhile, the vanillin content was measured by comparing the absorption of the vanillin sample solution with a standard vanilla solution using absorbance spectrophotometer at a wavelength of 348 nm [10].

2.4. NIR Data Processing and Analysis

The collected NIR and chemical dataset were divided into calibration and cross-validation datasets i.e. 2/3 part is used for calibration and 1/3 part for validation. The calibration between NIR and chemical data were carried out using PLS partial least square (PLS) and principal component regression (PCR) methods. Before calibration process, some pretreatment spectra such as normalization, De-trending, SNV, orthogonal signal correction (OSC), and multiplicative scatter correction (MSC) were applied to increase accuracy of NIR prediction [11,12]. All data processing was carried out using the Unscrambler X 10.4 application and MS office excel. Parameters used for determining the accuracy of the calibration were coefficient correlation (r), standard error (SE), coefficient of variation (CV), and residual predictive deviation (RPD) [13].

3. Results and Discussion

3.1 Chemical Data of Vanilla Pods

Table 1 shows the chemical data from destructive measurements of vanilla pods, namely moisture content and vanillin content. Average moisture content of vanilla pods is 35.16% with standard deviation of 16.99%. Meanwhile, vanillin content ranged from 0.38-9.65% with an average of 3.18% and standard deviation of 3.07%. . The standard deviation value obtained were relatively high both for water and vanillin content due to vanilla pods used in this study from two different region (Sumatera and Sulawesi).

Table 1 Chemical components in vanilla pods

Content Chemical	Average (%)	Standard Deviation (%)	Maximum (%)	Minimum (%)
Moisture Content (%)	35 .16	16.99	57.83	16.62
Vanillin Concentration (%)	3.18	3.07	9.65	0.38

3.2 Absorbance spectra of Vanilla pods

Figure 1 shows the NIR absorbance spectrum for vanilla pods (low, medium and high absorbance level of vanilla pods) at the wavelength of 1000-2500 nm. The absorbance values ranged from 0.19 to 1.97. The peaks of the absorbance spectrum occurred at the wavelengths of 2193 nm, 1940 nm, 1446 nm, and at 1200 nm. NIR absorption that appears at wavelength of 1200 nm occurs due to CH3 bonding; a wavelength of 1446 nm indicates the presence of CH aromatic compounds; the wavelength of 1940 nm indicates that there is a C=O chemical bond and an OH chemical bond with the product H_2O; and at a wavelength of 2193 nm indicated HC = CH bond [14]. This is in accordance with the chemical composition of the vanillin pods, namely the vanillin content of $C_8H_8O_3$ which is dominated by CH3, OH, C=O bonds. There is a match between the wavelength at the peak of spectra and the chemical compound of the vanilla pod.

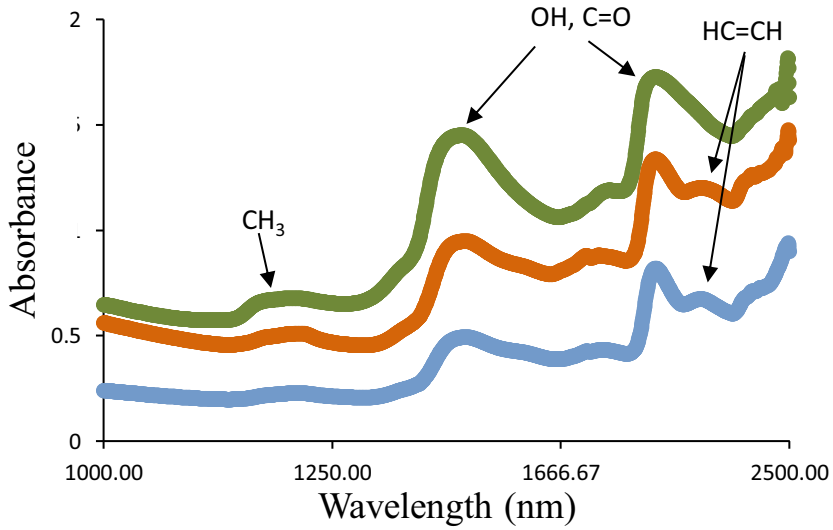


Figure 1 Absorbance spectra of vanilla pods

3.3 Prediction of chemical content of vanilla pods.

The results of calibration and validation of the estimation of vanillin and moisture content of vanilla pods based on PLS and PCR calibration methods and several NIR spectra pre-treatments is shown in Table 2.

In PLS calibration, the best prediction of vanillin content of vanilla pods was obtained using absorbance spectra with OSC pre-treatment and 3 PLS factors. The accuracy parameters of prediction are high (r=0.91, SEC=1.31%, CV=45.51%, RPD=1.30). It is selected because r value is high, the lowest SEP values, the lowest difference between SEC and SEP, the lowest CV and the highest RPD value. Despite MSC and SNV have a higher r values and a lower SEC than OSC, but the SEP and CV are high and the lower values of RPD. Moreover, OSC pre-treatment is also more efficient than MSC and SNV since it has the lowest numbers of PLS factor. The best prediction of moisture content was obtained using SNV pre-treatment and 5 PLS factors (r=0.99, SEC=2.11%, CV=6.12%, and RPD=7.98) indicating that all prediction accuracy parameters are excellent.

.

Table 2 Results of calibration and validation of the estimation of vanillin content and moisture content of vanilla pods

Chemical Content	Calibration method	Pre-processing Data	Factor	R	SEC (%)	SEP (%)	CV (%)	RPD	Consistency (%)
Vanillin content	PLS	Original	9	0.91	1, 36	1.48	50.60	2.08	91.74
		Normalize	8	0.91	1.34	1.52	51.72	2.03	88.24
		OSC	**3**	**0.91**	**1.31**	**1.34**	**45.61**	**2.30**	**97.96**
		MSC	8	0.92	1.28	1.49	50.77	2.07	86.35
		SNV	8	0.92	1.28	1.44	49.02	2.14	89.12
		de-trending	8	0.90	1.37	1.49	50.80	2.07	92.23
	PCR	Original	12	0.89	1.49	1.54	45.20	2.00	97.09
		Normalize	4	0.85	1.68	1.53	45.04	2.01	109.43
		MSC	8	0.89	1.47	1.50	43.99	2.06	98.51
		SNV	**6**	**0.89**	**1.47**	**1.49**	**43.21**	**2.09**	**101.40**
		OSC	3	0, 89	1.49	1.52	44.54	2.03	98.05
Moisture Content	PLS	Original	6	0.99	2.34	2.47	7.09	6.89	94.54
		Normalize	6	0.99	2.22	2.34	6.71	7.29	95.09
		MSC	5	0.99	2.13	2.16	6.19	7.89	98.56
		SNV	**5**	**0.99**	**2.11**	**2.13**	**6.12**	**7.98**	**98.86**
		de - *trending*	4	0.99	2.48	2.64	7.56	6.46	94.18
	PCR	Original	4	0.99	2.55	2.73	7.61	6.23	93.42
		MSC	4	0.99	2.35	2.27	6.33	7.49	103.58
		SNV	**4**	**0.99**	**2.33**	**2.25**	**6.25**	**7.58**	**103.90**
		De*trending*	7	0.99	2.54	2.55	7.10	6.68	99.66

In PCR calibration, the best prediction for vanillin content was obtained by SNV pre-treatment with 8 PCR factors (r=0.89, SEC=1.47%, CV=43.21% and RPD=2.09). The r value obtained from this model is above 0.80, a good correlation between the NIR spectra and the reference [11]. The SEC and SEP values obtained are 1.47% and 1.50%, the difference in values between the two is small, indicating that the prediction is good. The RPD value is 2.09%, that between 2 – 2.5 indicate a fairly good prediction [12]. For moisture content, the best prediction is also obtained using SNV pre-treatment but with 4 PCR factors (r=0.99, SEC=2.33%, CV=6.25%, and RPD=7.58).

The calibration methods and types of NIR data pre-treatment affect the accuracy parameters in predicting vanillin and moisture content using NIRS. In general, the PLS calibration gives a better accuracy than PCR. The accuracy is also depended on amount and type of chemical content of material. The NIR spectra pre-treatment slightly increased NIR accuracies both in PLS and PCR calibration. The SNV was selected as the best pre-treatment both for prediction of vanillin and moisture content except for vanillin prediction using PLS. It might be due to the variation in NIR spectra was predominantly influenced by the chemical composition rather than light scattering.

The data plot of vanillin content predicted by the best NIR calibration using OSC pre-treatment and 3 PLS factors versus vanillin content reference is shown in Figure 2. This best prediction of NIRS produces the SEC and SEP values which are small, namely 1.31 % and 1.34% and the difference in value of 0.03%. The difference between SEC and SEP which is very small indicates that the calibration is good [15]. However, the CV value is too high. The ideal CV value for agricultural materials and considered reasonable is 20% to 25% [16]. It is due to the vanillin content is not well distributed along the samples tested, especially in 1-8% vanillin content.

The data plot of moisture content predicted by the best NIR calibration using SNV pre-treatment and 5 PLS factors versus moisture content reference is shown at Figure 3. The best prediction of NIRS gives a small SEC and SEP namely 2.11% and 2.13% and the difference in value of 0.02%. The difference between SEC and SEP which is very small indicates that the calibration is good and calibration set represent the validation set [14].

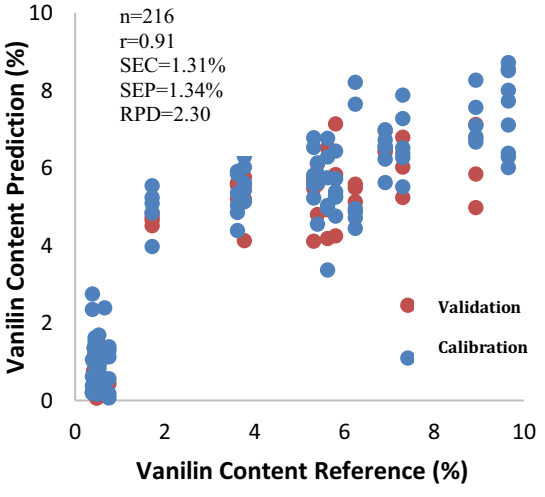


Figure 2 Plot of vanillin content predicted by PLS and SNV pre-treatment vs reference

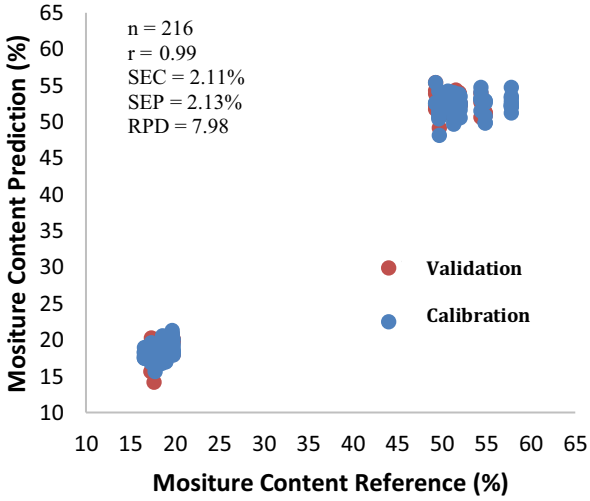


Figure 3 Plot of moisture content predicted by using using PLS and OSC pre-treatment vs referen

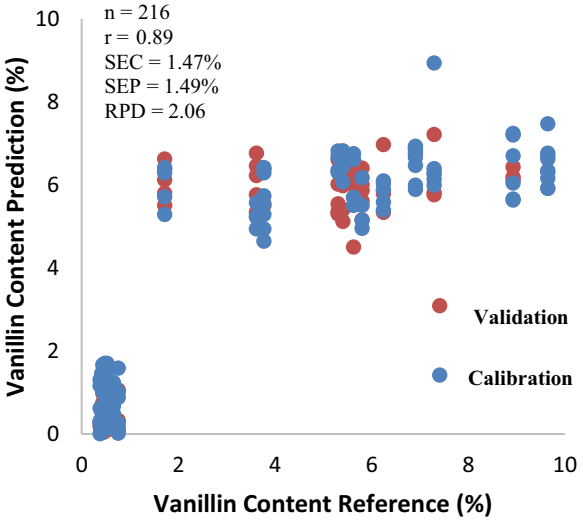


Figure 4 Plot of vanillin content predicted by using PCR and SNV pre-treatment vs reference

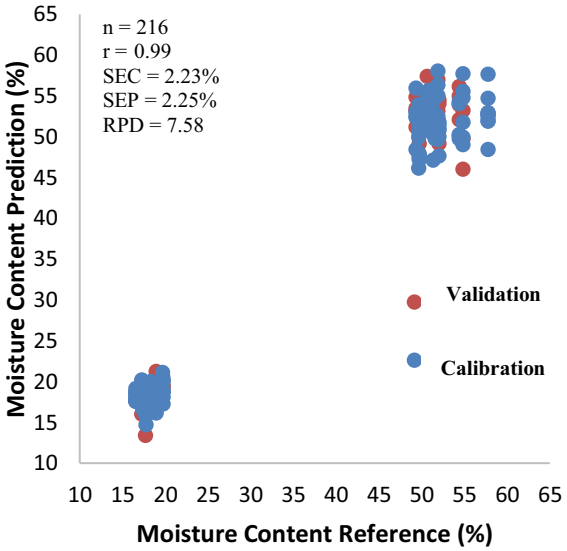


Figure 5 Plot of moisture content predicted by using PCR and SNV pre-treatment vs reference

It also has a very high correlation value of 0.99 and an RPD value of 7.98%. The RPD value of more than 3 indicates a good and very good prediction [11,12]. Compared to previous studies [8], the correlation coefficient of this study is higher than that of previous studies (r=0.96). The moisture content of samples is also not well distributed, and it is likely categorized into two groups of samples. There are no vanilla pods with moisture content of 22-45%.

The data plot of vanillin content predicted by the best calibration with PCR versus vanillin content reference is shown in Figure 4. The r value obtained from this calibration is 0.99. The value is above 0.80, a good correlation between the NIR spectra and the actual data (reference) [11]. The SEC and SEP values obtained are 2.54% and 2.55% and the difference in values between the SEC and SEP is small, indicating that the calibration is also good. The RPD value in this model is 6.80, more than 3 indicating a very good calibration [11]. The CV of 7.10% and consistency value of 99.66% indicated that the calibration can be used to predict vanillin content with acceptable accuracy.

The data plot of moisture content predicted by the best NIR calibration using SNV pre-treatment and PCR calibration versus moisture content reference is shown in Figure 5. The calibration is good and acceptable. The accuracy is higher than calibration with the original spectra. This shows that the use of pre-treatment before calibration can improve the accuracy of the calibration [12,17]. More chemical data especially in moisture content of 22-45% and vanillin content of 1-8% in order to develop the robust calibration for more accurate prediction of chemical composition of vanilla pods.

4. Conclusion

The best estimation for vanillin content using absorbance spectra with OSC pre-treatment at 3 PLS factors with the accuracy parameters of r = 0.91, SEC = 1.31%, SEP = 1.34%, CV = 45.60%, RPD = 2.30 with a consistency of 97.96%. The best prediction of vanillin content using PCR method was obtained with SNV pre-treated spectra at 6 PCR factors with the accuracy parameters of r = 0.89, SEC = 1.47%, SEP = 1.49%, CV = 43.21%, RPD = 2.09 and 101.40% consistency. The best estimation for water content using the PLS method is using the SNV pre-treatment spectra with a PLS factor of 5 with the accuracy parameters of r = 0.99, SEC = 2.11%, SEP = 2.13%, CV = 6.12%, RPD = 7.98 with consistency = 98.86%. While using the PCR method, the best estimation for moisture content was also obtained by SNV pre-treated spectra at 4 PCR factors with the accuracy parameters of r = 0.99, SEC = 2.33%, SEP = 2.25%, CV = 6.25%, RPD = 7.58 and consistency = 103.90%. The PLS calibration

provide a better accuracy than the PCR for estimating the chemical content of vanilla pods. The NIR spectroscopy associated with the selected pre-treatments and factor numbers can be used for determination of chemical content of vanilla pods.

References

[1] Directorate General of Plantation. 2020. *Indonesian Plantation Statistics 2020 Vanilla.* Jakarta (ID): Secretariat of the Directorate General of Plantations.
[2] National Standardization Agency. 2002. SNI 01-0010-2002 . Jakarta.
[3] Y A. Purwanto, I W Budiastra, K Syamsu.2013. Application of near-infrared reflectance spectroscopy for determination of gambir (uncaria gambir roxb.) Moisture and catechin content. Acta Horticulturae 1011: DOI: 10.17660/Actahortic.2013.1011.37.
[4] L.C.E.Ch. Lengkey, D.P. Sari, I.W. Budiastra , K.B. Seminar, B.S. Purwoko.2013. Prediction of Chemical Contents of *Jatropha Curcas* L. Seeds IP-3P By Near Infrared Reflectance Method. ISHS Acta Horticulturae 1011: DOI: 10.17660/Actahortic.2013.1011.42.
[5] Budiastra I W, Sutrisno, Sukrisno. P C Ayu. 2020. Determination of trigonelline and chlorogenic acid (CGA) concentration in intact coffee beans by NIR spectroscopyy. *Agricultural Engineering International: CIGR Journal.* Vol 22 N0.1. 163-168 pp.
[6] PC Ayu, I W.Budiastra, A. Rindang. 2020. NIR spectroscopy application for determination caffeine content of Arabica green bean coffee. *IOP Conference Series: Earth and Environmental Science* 1 / 012049.
[7] F Kurniawan, I W Budiastra, Sutrisno, S Widyotomo. 2019. Classification of Arabica Java Coffee Beans Based on Their Origin using NIR Spectroscopy. *IOP Conference Series: Earth and Environmental Science* 309 / 012006.
[8] Wongsheree T, Rittiron R, Wongs-Aree C, Thongtheing T. 2016. Vanilla content evaluation in cured bean by near infrared technique. *Acta Hortic.* 1119(July 2020):297–304.
[9] National Standardization Agency. (1992). SNI 01-2891-1992: Food and Beverage Test Methods. Jakarta
[10] [AOAC] Association of Official Analytical Chemists. 1981. Vanillin in vanilla extract: rapid approximate method spectrophotometric method:19.011-19.012.
[11] Cen H, He Y. 2007. Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends Food Sci. Technol.* 18(2):72–83
[12] Purwanto A, H. Permatasari, I W Budiastra. 2015. Effects of Preprocessing Techniques in Developing a Calibration Model for Soluble Solid and Acidity in 'Gedong Gincu' Mango using NIR Spectroscopy. *International Journal of Engineering and Technology* Vol.7 No.5.1921-1927 .
[13] Farobie, O., Amrullah, A.; Bayu, A., Syaftika N., Anis, L.A., Hartulistiyoso E. *In-depth study of bio-oil and biochar production from macroalgae Sargassum sp. via slow pyrolysis*. RSC Advances. 2022. **12**(16), 9567–9578.
[14] Osborne BG, Fearn T, Hindle PH. 1993. *Practical NIR Spectroscopy with applications in food and beverage analysis*. 2nd ed. Singapore (SG): Longman Singapore Publishers (Pte) Ltd.
[15] Lammertyn J., Peirs A., De Baerdemaeker J., Nicolaï B. 2000. Light penetration properties of NIR radiation in fruit with respect to non-destructive quality assessment. Postharvest Biology Technology. 18:121-132. https://doi.org/10.1016/S0925- 5214(99)00071-X
[16] Mattjik AS, Made IS. 2006. Experimental Design with SAS and MINITAB Applications. Bogor (ID): IPB Press.
[17] 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 substances in green tea leaves. J. Agri. Food Chem. 475:5064-5067.