The application of NIRS method for non-destructive measurement of fat and carbohydrates in pumpkin seeds

Ifmalinda*, Andasuryani, Santosa, I Putri
Agricultural Engineering Department, Andalas University, Limau Manis, Padang, Sumatera Barat 25163, Indonesia

*Email: ifmalinda@ae.unand.ac.id

Abstract. Pumpkin is one of the agricultural products that can be used as an alternative food ingredient. Chemical content such as carbohydrates and fats can be obtained by the near infrared spectroscopy (NIRS) method by using NIR light that is penetrated from the material so that the reflectance and absorbance spectra will be obtained. This spectral data is the value of the reflectance intensity and is the raw data that still contains noise. Furthermore, it is processed with the NIRS data processing method to reduce the influence of wave inference and noise on the spectral data in order to obtain more accurate results. The purpose of this study was to obtain the best calibration mode to estimate the fat and carbohydrate content of pumpkin seeds using the NIRS method. Pre-treatment of the spectrum data was carried out with Gap Derivative and Derivative Savitzky-Golay. NIRS spectra data were processed using the multivariate Partial Least Squares (PLS) calibration method. The results showed the best calibration model for carbohydrate and fat content using Derivative Gap data processing with values of $r = 0.95$, $R^2 = 0.98$, SEC = 1.25, and RMSEC = 1.23. and using Latent Variable (LV) factor 3, while for fat content $r = 0.99$, $R^2 = 0.89$, SEC = 0.17 and RMSEC = 0.17. and using Latent Variable (LV) factor 4.

1. Introduction
Pumpkin is an agricultural food product that can be used as an alternative material. One of the advantages of pumpkin is that it has a longer shelf life compared to other agricultural products. Pumpkin fruit that is quite old and without defects when harvested can be stored at room temperature for approximately six months without undergoing much change [1, 2]. In addition to the fruit, pumpkin seeds can also be used. Pumpkin seeds contain various nutrients that are beneficial to the body, such as carbohydrates, fats, proteins, vitamin C and other ingredients.

Measurement of the chemical content of a material is usually carried out destructively (conventional methods) or by damaging the material carried out in the laboratory. This method is expensive, takes a long time to prepare samples and creates chemical waste. Near Infrared Spectroscopy (NIRS) is a non-destructive method used to quickly analyze and obtain information on the chemical content of a material without using chemicals. Prediction of chemical content values such as carbohydrates and fats can be obtained using the Near Infrared Spectroscopy (NIRS) method using NIRS light that is penetrated from a material so that the reflectance and absorbance spectra will be obtained. This spectral data is the value of the reflectance intensity and is the raw data that still contains noise. Furthermore, it is processed with the NIRS data processing method to reduce the
influence of wave inference and noise (noise) on the spectral data in order to obtain more accurate results [1].

The NIRS data calibration technique uses partial least squares (PLS). PLS is a step carried out to build a model that relates the response spectra of each sample at each wavelength to the chemical concentration known from laboratory analysis, where the PLS method uses a linear combination to estimate the independent variable and the original variable [3]. [4] states that linear combinations are chosen to take the assumption that variables with very high correlation and independent variables have the same weight because these variables will be more effective in estimating, so it is expected to produce a few linear combinations of independent variables needed to explain most of the variation.

Many studies related to NIRS have been carried out including near infrared spectroscopy on medical foods [5], near infrared spectroscopy in intact mango fruit [6], near infrared spectroscopy on feed nutritive [7]. The purpose of this study was to obtain the best calibration mode to estimate the fat and carbohydrate content of pumpkin seeds using NIRS method.

2. Materials and methods

2.1. Pumpkin seed preparation

The materials used in this study were pumpkin seeds of the bokor (cerme) variety with the criteria for harvesting dark yellow pumpkins that were 4 months old from growing flowers [8]. The pumpkin seed variety used in this study was the bokor variety. Prior to measurements, the pumpkin seeds were separated from the pumpkin fruit and then cleaned and removed the mucus. Furthermore, it is washed and air-dried or drained. Then the pumpkin seeds were dried using a rack-type oven dryer at a temperature of 60°C until the moisture content was <12%. The number of samples prepared was 50 samples, then the NIRS spectrum was taken and then continued with the measurement of moisture content.

2.2. NIRS spectrum acquisition

The tools used in this study include one unit of NIR Spectrometer brand NIFLEX N-500 with a wavelength range of 1000 nm to 2500 nm or 10000 - 4000/cm which is equipped with a laptop unit which is used to process spectral data, and LED lights and Spectroscopy Software for data spectrum analysis. NIR measurements were carried out at the Food and Agricultural Product Processing Engineering Laboratory (TPPHP), Department of Agricultural Engineering, Bogor Agricultural University. The number of samples to be taken for the spectrum is 50 grams. The pumpkin seed samples were arranged into a petri dish with a weight of 50 grams each (9.7 cm in diameter and 1.5 cm in height) then placed on a rotary plate on the FT-NIR spectrometer and scanned 3 times. The scanning results from the three repetition were transformed into absorbance spectra and then averaged. The reflectance of pumpkin seeds was measured at a wavelength of 1000-2500 nm.

2.3. Chemical analysis

2.3.1 Fat level. 3 grams of sample (W0) was put into filter paper which had been closed at both ends with fat-free cotton and put into a fat sleeve whose fixed weight had been measured (W1) and connected to a Soxhlet tube. The fat shell was inserted into the extractor chamber of the Soxhlet tube and rinsed with fat solvent (n-Hexan). Then reflux for 3 hours. The fat solvent in the fat flask is distilled until all the fat solvent evaporates. Furthermore, the pumpkin fat was dried in an oven at a temperature of 105°C and the pumpkin was dried in a desiccator until the weight was constant (W2). Calculation of fat content can use the equation [9]:

\[
\text{Fat Level (\%) = \frac{(W2 - W1)}{W0} \times 100 \%} \quad \text{………………………………………………………………(1)}
\]

Information :

- \( W0 \) = sample weight (g)
- \( W1 \) = pumpkin weight without fat (g)
- \( W2 \) = pumpkin weight with fat (g)
2.3.2. Carbohydrate level. Carbohydrate level in pumpkin seed flour was determined using the by different method (SNI 01-2892-1992). Carbohydrate levels can be calculated using the following equation:

\[
\text{Carbohydrate Level (\%)} = 100\% - \%(\text{protein+fat+moisture content+ash content})
\]

2.3.3. Research data analysis. The results of NIR spectroscopy measurements obtained spectral data in the form of reflectance data, to obtain a linear correlation between NIR absorption values and chemical data, the reflectance was transformed into absorbance spectra using Equation (3).

\[
\log (1/R)
\]

Before analyzing the data to build a prediction model, pre-processing of the NIRS spectrum data on pumpkin seeds was carried out to minimize and eliminate disturbances or noise on the spectrum so that the prediction model produced was more accurate and stable [10]. In this study, the NIR spectra data processing method was carried out with (1) GapDerivative (second derivative), (2) Savitzky-Golay Derivative (first derivative). Data processing with the first and second derivatives serves to decipher overlapping chemical components, reveal information on chemical components hidden in the spectra and suppress unwanted components so that the desired components will appear [11].

Analysis of NIRS spectra data using the multivariate Partial Least Squares (PLS) calibration method. Calibration was carried out to determine the correlation between chemical content of pumpkin seeds obtained from destructive determination and NIRS reflectance. The Partial Least Squares (PLS) method is a method used to estimate a series of dependent variables (responses) from a large number of independent variables (predictors), having a linear or nonlinear systematic structure, with or without high collinearity data [12]. According to [2] the Partial Least Squares method can estimate a large number of independent variables and dependent variables.

2.3.4. Evaluation of calibration results. The results of the calibration model can be evaluated by looking back at the value of the correlation coefficient (r), coefficient of determination (R2), root mean square error (RMSE), standard error of calibration set (SEC), and residual predictive deviation (RPD). Standard Error Calibration (SEC), indicates the accuracy of the calibration equation obtained from chemical and NIRS data. A low SEC value indicates a low predictive error of calibration. A good model has high r and R2 values, low RMSEC [13], [14].

3. Results and discussion

3.1. Pumpkin seed NIRS reflectance spectra wave

The NIRS wavelength range is 1000-2500 nm with 1501 spectra data. The results of the original raw NIRS spectrum analysis without processing can be seen in Figure 1. The results of the original spectra explain the absorption of pumpkin seed samples to NIRS light at certain wavelengths. The shape of the spectral graph has peaks and valleys which indicate the difference in chemical content contained in pumpkin seeds. Particle size, wavelength and chemical content can affect the NIRS spectra that will be produced [15]. According to [16] the peaks that appear on the graph are caused by vibrations that occur when chemical bonds interact with NIRS rays.
The original spectra of pumpkin seeds still have noise and have distances between spectra caused by the presence of cavities in the empty sample pile in the petri dish so that the absorption of the spectrum is not perfect and is also caused by interference in the form of an acoustic signal that arises due to an electronic circuit that is not part of the unwanted signal [7]. To improve the results of the original spectra and produce accurate data and a good calibration model, it is necessary to pretreatment with the Savitzky-Golay Derivative method and GapDerivative Pretreatment. The results of the spectra that have been processed with Savitzky-Golay Derivative Pretreatment and GapDerivative Pretreatment can be seen in Figures 2 and 3.

Figure 1. Spectrum of Raw NIRS Original Pumpkin Seeds

Figure 2. Spectrum of Pumpkin Seeds with Savitzky-Golay Derivative Pretreatment
According to [2] the first derivative method (Derivative Savitzky-Golay) can reveal hidden information in the spectra, this method is usually used for prediction of low amounts of content. The first-derivative method also reduces baseline effects, background information, and reduces overlapping spectra.

![Figure 3. Pumpkin Seed Spectrum with GapDerivative Pretreatment](image)

3.2. Carbohydrate and fat level of pumpkin seeds
The chemical content of pumpkin seeds in the form of carbohydrate and fat content can be seen in Table 1. The minimum and maximum values obtained in the chemical data were used as calibration data sets in building the calibration model.

| Chemical Content | Mean  | Min   | Max   | Standard Deviation |
|------------------|-------|-------|-------|--------------------|
| Carbohydrate     | 25.06 | 17.36 | 32.75 | 3.77               |
| Fat              | 33.32 | 29.97 | 36.25 | 1.38               |

3.3. Loading pumpkin seed spectra plot
The NIRS waves emitted by the organic matter spectrometer will partly be absorbed to vibrate the C-H, N-H, and O-H bonds and partly will be reflected. The part of the reflected wave will be captured by the detector as a reflectant [18]. The loading plot obtained to identify the fat content of pumpkin seeds can be seen in Figure 4. The spectra resulting from the measurement of the NIRS wavelength have peaks and valleys that indicate the absorption of the chemical content contained in pumpkin seeds. The wavelength range of pumpkin seed fat content is 1433, 1666, 1749 and 1875 nm. The results of other studies have fat content in Gayo coffee 1639-2325 nm [19], Java Preanger coffee 1477 and 1680 nm
Bondowoso coffee 1410, 1700 and 1892 nm [9]. Solok Arabica coffee fat is located at wavelengths of 1379 - 1388 nm, 1720 - 1733 nm, and 2300 - 2346 nm [21].

The loading plot obtained to identify the carbohydrate content of pumpkin seeds can be seen in Figure 5. The spectra resulting from the measurement of the NIRS wavelength have peaks and valleys indicating the absorption of carbohydrates in pumpkin seeds with a wavelength range of 1464, 1614, and 1988 nm. These results are also in accordance with other studies of carbohydrate content in Java Preanger coffee 1477 and 1680 nm [20], Bondowoso coffee 1410, 1700 and 1892 nm [9].
3.4. Calibration with PLS method

The evaluation of the calibration model aims to obtain the best calibration model. Analysis of NIRS spectra data was carried out using the multivariate Partial Least Squares (PLS) calibration method. The PLS method forms a model from the existing variables to form a series of responses using least squares regression in matrix form \([\ldots]\). Statistical parameters to evaluate the results of the calibration model can be seen from the correlation coefficient \((r)\), coefficient of determination \((R^2)\), standard error of calibration set \((SEC)\), standard error of validation set \((SEP)\), \([\ldots]\).

The plot of calibration data for pumpkin seed fat content test using PLS method data processing with non-pretreatment can be seen in Figure 6.

![Figure 6. Plot of Calibration Data for Non Pretreatment Pumpkin Seed Fat Level Test](image)

The results of the calibration of pumpkin seed fat content with non-pretreatment resulted in the following values: \(r = 0.72\), \(R^2 = 0.51\), \(SEC = 0.96\), and \(RMSEC = 0.95\), and used Latent Variable (LV) factor 7. The calibration model showed a value that was still not good enough. The number of Latent Variables used aims to reduce spectral interference, a high PLS factor value can reduce the ability to predict due to spectral interference \([\ldots]\). The range of \(r\) values is from 0 to 1 \([\ldots]\).

The plot of calibration data for pumpkin seed fat content test using PLS method data processing with Derivative Savizky-Golay pretreatment can be seen in Figure 7.
The prediction results of the calibration value of pumpkin seed fat content with the Derivative Savitzky-Golay pretreatment were able to improve the prediction model to be better than without pretreatment. This is indicated by the values: $r = 0.95$, $R^2 = 0.89$, SEC = 0.45, and RMSEC = 0.44.

and using Latent Variable (LV) factor 5. Savitzky-Golay Derivative Pretreatment with PLS factor 5 is considered good for building a calibration model with a correlation value close to 1. According to [26] stating that $r$ values above 0.90 are considered good. The PLS factor used in the Savitzky-Golay derivative is also lower than without pretreatment. Overfitting occurs if the number of PLS factors is too large [27]. The plot of calibration data for pumpkin seed fat content test using PLS method data processing with Derivative Gap pretreatment can be seen in Figure 8.
The prediction results of the calibration value of pumpkin seed fat content with DerivativeGap pretreatment were able to improve the prediction model to be better than without pretreatment. This is indicated by the values: $r = 0.99$, $R^2 = 0.98$, SEC = 0.18, and RMSEC = 0.17 and uses the Latent Variable (LV) factor 4. The calibration model shows a stable model with a value of $r$ close to 1. This indicates that the calibration model acceptable because it meets the requirements with a high coefficient of determination and low error [4]. According to [28], low SEC and SEP values result in a good calibration model and conversely a large value indicates a calibration set model that does not represent the validation set. The plot of calibration data for the pumpkin seed carbohydrate content test using PLS method data processing with non-pretreatment can be seen in Figure 9.

![Predicted vs. Reference Plot](image)

**Figure 9.** Plot of Calibration Data for Carbohydrate Test of Non Pretreatment Pumpkin Seeds

The prediction results of the carbohydrate content calibration value of pumpkin seeds with non-pretreatment still show a low prediction model, this is indicated by the values: $r = 0.65$, $R^2 = 0.43$, SEC = 3.16, and RMSEC = 3.12. and using Latent Variable (LV) factor 7. The correlation coefficient indicates a positive correlation between the actual data variables and a positive correlation between the actual data variables [29]. The plot of calibration data for the Carbohydrate content test of pumpkin seeds using the PLS method data processing with the Derivative Savizky-Golay pretreatment can be seen in Figure 10.
The prediction results of the calibration value of pumpkin seed carbohydrates with the Derivative Savitzky-Golay pretreatment were able to improve the prediction model, this was shown by the values: $r = 0.93$, $R^2 = 0.86$, $SEC = 1.55$, and $RMSEC = 1.53$, and used the Latent Variable (LV) factor. The results of the calibration model are better because it only involves 5 Latent Variables to produce these predictions. [30] states that the prediction model with the least number of latent variables is better and more efficient than the prediction model with the higher number of LV or PC. The plot of calibration data for the pumpkin seed carbohydrate content test using PLS method data processing with DerivativeGap pretreatment can be seen in Figure 11.

**Figure 10.** Plot of Carbohydrate Content Calibration of Pumpkin Seeds with Savitzky-Golay Derivative Pretreatment

**Figure 11.** Plot of Carbohydrate Content Calibration of Pumpkin Seeds with DerivativeGap Pretreatment
The prediction results of the calibration value of pumpkin seed carbohydrates with DerivativeGap pretreatment resulted in a better prediction model than without pretreatment. This is indicated by the value: $r = 0.95$, $R^2 = 0.91$, SEC = 1.25, and RMSEC = 1.23, and uses the Latent Variable (LV) factor 3. The prediction model produced is very good, it can be seen from the value of $r$ close to 1, in accordance with statement [31, 32].

4. Conclusions
Carbohydrate and fat content can be estimated non-destructively using the NIRS method. The best calibration model for fat and carbohydrate content was obtained by processing data using DerivativeGap pretreatment with values of $r = 0.99$, $R^2 = 0.98$, SEC = 0.18, and RMSEC = 0.17 and using Latent Variable (LV) factor 4, while for carbohydrate content $r = 0.95$, $R^2 = 0.91$, SEC = 1.25, and RMSEC = 1.23 and uses a Latent Variable (LV) factor of 3.

References
[1] Jannah, R. 2018. Prediksi Kadar Proksimat Tiga Varietas Biji Kopi Arabika Dengan Near Infrared Spectroscopy. Tesis. Institut Pertanian Bogor. Bogor.
[2] Kusumiyati., Munawar,A., Suhandi, D. 2020. Prediksi Vitamin C, Total Asam Tetrirasi, dan Total Padatan Terlarut Pada Buah Mangga Menggunakan Near-Infrared Reflectance Spectroscopy. Jurnal Teknologi Pertanian Vol. 21 No. 3. 145-154.
[3] Cahyo, L. D. 2016. Pengembangan Metode Pengukiran Tingkat Kematangan Buah Jambu Kristal Menggunakan NIR (Near Infrared). Tesis. Institut Pertanian Bogor. Bogor.
[4] Miller, J.N., and Miller. J.C. 2010. Statistics and Chemometrics for Analytical Chemistry Sixth Edition. Pearson Education Limited, England.
[5] Y. Deng, Y. Wang, G. Zhong, and X. Yu. —Simultaneous Quantitative Analysis of Protein, Carbohydrate and Fat in Nutritionally Complete Formulas of Medical Foods by Near-Infrared Spectroscopy, Infrared Phys. Technol., vol. 93, no. July, pp. 124–129, 2018, doi: 10.1016/j.infrared.2018.07.027.
[6] Munawar, A.A., Kusumiyati, Hafidh, R. Hayati, and D. Wahyuni, —The Application of Near Infrared Technology as a Rapid and Non-Destructive Method to Determine Vitamin C Content of Intact Mango Fruit, INMATEH - Agric. Eng., vol. 58, no. 2, pp. 1–12, 2019, doi: 10.35633/INMATEH- 58-31.
[7] Samadi, S. Wajizah, A. A. Munawar. 2018. Rapid and Simultaneous Determination of Feed Nutritive Values by Means of Near Infrared Spectroscopy. Tropical Animal Science Journal. 41(2). 121-127
[8] Hendrasty, H.K. 2003. Pembuatan Tepung Labu Kuning dan Pengolahannya. Kanisius : Yogyakarta.
[9] Athfiyah, S. N. 2017. Penentuan Kandungan Kimia Utama Biji Kopi Arabika Bondowoso Secara Nondestruktif dengan Near Infrared Spectroscopy. Skripsi. Institut Pertanian Bogor. Bogor.
[10] Rosita, R. 2016. Penentuan Kandungan Kimia Biji Kopi Arabika Gayo secara Non Destruktif dengan Near Infrared Spectroscopy. [Tesis]. Bogor. Sekolah Pasca Sarjana Institut Pertanian Bogor.
[11] Lengkey LCCh, Budiasta IW, Seminar KB, Purwoko BS. 2013. Determination of Chemical Properties in Jatropha Curcas L. eed IP-3P by Partial Least Squares Regression and Near-Infrared Reflectance Spectroscopy. J. Agriculture Innovations and Research 2(1):2319-1473
[12] Kurinawan, F. 2017. Karakterisasi dan Klasifikasi Biji Kopi Java Arabika Berdasarkan Indikasi Geografis Menggunakan Metode NIR Spectroscopy dan Analisis Diskriminan. Tesis. Institut Pertanian Bogor. Bogor.
[13] Jha, S. N., A. R. P. Kinsly, S. Chopra. 2006. Non-Destructive Determination of Firmness and Yellowness of Mango During Growth and Storage Using Visual Spectroscopy. Jurnal Biosystem Engineering. 94:397-402.
[14] Zulfahrizal., A.A. Munawar dan H. Meilina. 2016. Estimasi Kandungan Lemak Pada Biji Kakao Utuh Secara Cepat dan Non-Destruktif dengan Menggunakan Teknologi NIRS. Jurnal Otomasi, Kontrol dan Instrumentasi. Vol. 8(1). Hal 17.

[15] Sahfinti. 2019. Identifikasi Kandungan Lemak dan Karbohidrat pada Kupi Green Bean Arabika Lintong Menggunakan NIR Spectroscopy dan PLS. Skripsi. Universitas Sumatera Utara.

[16] Zulfahrizal. 2014. Pengembangan Metode Pengukuran Nondestruktif untuk Menentukan Mutu dan Fermentasi Biji Kakao Utuh Menggunakan Nir Spectroscopy. [Desertasi]. Bogor (ID): Institut Pertanian Bogor

[17] Munawar, A.A. 2014. Multivariate Analysis and Artificial Neural Network Approaches of Near Infrared Spectroscopic Data for Non-Destructive Quality Attributes Prediction of Mango. [Desertasi]. Goettingen: Georg-August University.

[18] He, Y., L. Xiaoli, and D. Xunfei. 2007. Discrimination of Varieties of Tea uUing Near Infrared Spectroscopy and PLS. Skripsi. Universitas Sumatera Utara.

[19] Aditama, H. F, Budiastra, I. W., dan Widodo, S. (2019) Pengembangan Model Jaringan Syaraf Tiruan untuk Penentuan Kandungan Kimia Biji. Kopi Arabika Gayo dengan NIRS.Warta IHP/Journal of Agro-Industry Vol.36 07 2019: 22-29

[20] Heruwati, I. 2011. Deteksi Gejala Chilling Injury Buah Belimbing (Averrhoa Carambola L) yang Disimpan pada Suhu Rendah dengan NIR Spectroscopy. Skripsi. Institut Pertanian Bogor. Bogor.

[21] Yuwita, F., Ifmalinda, Makky, M., 2019. Uji Nondestruktif Kandungan Kafein, Protein, dan Lemak Biji Kopi Solok Radjo menggunakan Near Infrared Spectroscopy (NIRS). Tesis. Univeristas Andalas.

[22] Agustina, Sri., Purwanto, A.Y., Budiastra, I.W., 2015. Prediksi Kandungan Kimia Mangga Arumanis Selama Penyimpanan dengan Spektroskop NIRS. Jurnal Keteknikan Pertanian Vol. 3 No. 1. 57-63

[23] 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.

[24] Jankovská, R. dan Šustová, K. (2003). Analysis of Low Milk by Near-Infrared Spectroscopy. Czech Journal of Food Sciences 21(4): 123-128.

[25] Mattjik AS dan Made IS. 2006. Perancangan Percobaan dengan Aplikasi SAS dan MINITAB. Bogor [ID]: IPB Press.

[26] Williams, P. dan Norris, K. (1990). Near-Infrared Technology in the Agricultural and Food Industries. American Association of Cereal Chemical, Inc. St. Paul. (USA): American Association of Cereal Chemist. Inc.

[27] Cozzolino, D., I. Murray, A Chreeb, and J.R. Scaife. 2005. Multivariate Determination of Free Fatty Acids and Moisture in Fish oils by Partial Least-Squares Regression and Near-Infrared Spectroscopy. Lebensmittel Wissenschaft und Technologie. 38: 821-828.

[28] Lammertyn J, Peirs A, Beardemarker JD, Nicolai ‘B. 2013. Ligth Penetration Properties NIR Radiation in Fruit with Respect to Non-Destructive Quality Assessment. Postharvest Biology & Technology. 18: 21-132.

[29] Taylor RER. 1990. Interpretation of the Correlation Coefficient : a Basic Review. Journal of Diagnostic Medical Sonography. 1: 35–39.

[30] Mouazen, A. M., Kuang, B., De Baerdemaeker, J., & Ramon, H. (2010). Comparison Among Principal Component, Partial Least Squares and Back Propagation Neural Network Analyses for Accuracy of Measurement of Selected Soil Properties with Visible and Near Infrared Spectroscopy. Geoderma, 158(1–2), 23–31. https://doi.org/10.1016/j.geoderma.2010.03.001.

[31] Nircal 5.5 manual. 2013. Nircal manual. Switzerland: Buchi Labortechnik AG, CH Flawil.

[32] Munawar A A, Yunus Y, Devianti and Satriyo P 2020 Calibration models database of near infrared spectroscopy to predict agricultural soil fertility properties Data Br. 30