Rapid and Simultaneous Prediction of Soil Quality Attributes using Near Infrared Technology

Agus Arip Munawar*, Devianti, Purwana Satriyo, Syahrul, Yuswar Yunus*

Abstract— The main purpose of this present study is to apply the near infrared (NIR) technology as a rapid and robust method in predicting soil quality parameters in form of potassium (K), Magnesium (Mg) and calcium (Ca) simultaneously. Diffuse reflectance spectra data were acquired for a total of 40 bulk soil samples (60 g per each bulk) in near infrared (NIR) wavelength range from 1000 to 2500 nm. On the other hand, actual reference K, Mg and Ca were measured using standard laboratory procedures. Prediction models, used to predict those three quality parameters were established using principal component regression (PCR) and partial least square regression (PLSR) method. Moreover, prediction accuracy and robustness were evaluated based on correlation coefficient (r) and residual predictive deviation (RPD) index respectively. The result showed that K, Mg and Ca of soil samples can be predicted simultaneously using NIR technology with maximum r coefficient and RPD index were 0.97 and 5.14 for K, 0.98 and 8.34 for Mg, 0.98 and 8.90 for Ca respectively, which categorized as excellent model performance. Thus, it may conclude that NIR technology can be used and applied as rapid and simultaneous method to predict quality parameters of soil samples satisfactory.

Keywords— NIRS, soil, prediction, nutrients, quality, fast, method.

1. INTORDUCITON

Tin general, the major function of soil is to provide fundamental natural resources for survival of all living creatures which depends on the balances of soil structure and nutrient compositions [1]. The maintenance of soil quality is critical for ensuring the sustainability of the environment and it requires real time monitoring in order to take further actions. Soil quality parameters are included macro and micro nutrients, carbon organic, soil textures and also minerals content [2], [3]. As a result, plants can grow optimally in a healthy soil, heavy metals free and fertile. Soil chemical properties related to the amount of nutrients and other soil quality parameters required by plants. the amount needed will vary each growth phase. To determine soil quality parameters as referenced above, several methods were already employed. However, most of these methods were based on solvent laboratory analysis and followed by complicated procedures [4]–[6]. Generally, it is difficult to determine nutrient and minerals contents on soil in real time and without sample preparation [7]. Normally, it requires standard laboratory procedures in which took some time with complicated sample preparation, involved chemical materials and destructive in nature [8]. Meanwhile, soil nutrient contents must be determined rapidly in order to take an action required and ensure optimum plant growth. Therefore, alternative fast and robust method is required to determine soil quality parameters. During last few decades, the near infrared (NIR) technology has been widely studied and applied as an effective technique for rapid analysis of soil and other biological quality properties [9], [10].

- *Agus* Arip Munawar*, Devianti, Purwana Satriyo, Syahrul, Yuswar Yunus*
- Department of Agricultural Engineering, Syiah Kuala University, Banda Aceh – Indonesia
- *Corresponding author: yuswaryunus@unsyiah.ac.id

Compared with standard chemistry analysis, NIR technology can be used as an alternative method since this analysis is rapid, cost effective, non-destructive, requires minimal sample preparation and can be used directly on-field [7], [11], [12]. More importantly, NIR method allows a simultaneous quantitative assessment of several quality attributes properties in a single measurement. This technique mainly measures overtones and combinations of fundamental vibrational bands for O-H, N-H and C-H bonds from the mid-infrared region [13]. The near infrared technology is a technique which uses infrared radiation in wavelength range from 700 to 2500 nm of the electromagnetic spectrum to analyze the chemical composition of organic matter [7], [14]. It provides information through spectra signatures and patterns, regarding with the intrinsic organic bonds of the molecules and thus the primary chemical constituents of the object can be determined [11]. Numerous studies have been conducted and reported related to the application of NIR technology in determining several quality parameters of biological objects such as fruits and horticultural products [15], [16], coffee and cocoa [17], [18], meat and dairy products [19], [20], soil and root [21]–[23]. Overall findings stated that the NIR technology was able to predict several quality parameters of biological objects rapidly. In term of accuracy and robustness, there are varied achievement among them from which categorized as sufficient to excellent prediction performances. Further studied still need to be carried out in order to find suitable methods related to NIR analysis. These methods are included NIR spectra correction method, calibration method and others. Therefore, the main objective of this present study is to apply the NIR technology as rapid and robust method in assessing soil quality attributes in form of potassium (K), magnesium (Mg) and calcium (Ca) through calibration models development. This study also aimed to find the optimum calibration method which can provide more accurate and robust prediction performances.

2 MATERIALS AND METHODS

2.1 Soil samples

A bulk of soil samples were collected from 8 different sites in *Banda Aceh* and *Aceh Besar* district area in Aceh Province, Indonesia. Those soil samples cover different land use: rice field, ground/planted field and bare land. Soils samples were taken to the lab, dried and stored for 3 days before spectra acquisition and further data analysis. A total of 40 bulk of soil samples (60 g per each bulk) were used to acquire and measure near infrared spectra data.

2.2 Near infrared spectra data acquisition

Near infrared spectra data of all soil samples were acquired in form of diffuse reflectance spectral data in wavelength range from 1000 to 2500 nm. To ensure acquisition stability, background spectra correction was performed every hour automatically. Diffuse spectra data were collected and recorded in NIR wavelength region as referenced above with the increment of 0.2 nm resolution. Soil samples were scanned 32 times and the averaged data were recorded and stored in SPA and CSV extension files.

2.3 Prediction models

Soil quality attributes in form of K, Mg and Ca were predicted by developing prediction models through calibration. Two different regression approaches namely principal component regression (PCR) and partial least square regression (PLSR) were used and the prediction performances were compared.
In order to achieve optimum and robust prediction models, leverage cross validation was subjected during calibration models development [7], [24]. Prediction model performance were quantified using several statistical indicators namely: coefficient determination (R^2), coefficient correlation (r), the root mean square error (RMSE) and residual predictive deviation (RPD) index which corresponds to prediction accuracy and robustness indicators respectively [25], [26].

3 RESULTS AND DISCUSSIONS

3.1 Infrared spectra feature of soil

A captured single spectrum in near infrared region of soil samples was shown in Fig. 1. Soil spectral features in the NIR wavebands are highly correlated to the vibration energy of chemical bonds and functional groups like O, H, C, N, and other chemical structures. These bonds represent the energy changes from which two or more vibration patterns exist in these bonds including stretch vibration and bend vibration. Soil nutrients and minerals are consisted of several chemical structures vibrated along near infrared wavebands and thus can be predicted using such particular data analysis. Several soil quality attributes like minerals were spread-out along NIR wavebands especially at the end of NIR region around 2200 nm, and little portion of minerals also observed in early NIR wavelength, at around 1160 nm. As can be seen in this referenced figure, the highest peak from the NIR wavebands of soil samples were observed in 1410 nm and 1920 nm which are associated with moisture molecules (O-H). This also in agreement with Cen and He (2007) stated that water absorption bands of organic materials can be found in wavelength around 1400 nm and 1900 nm respectively. Figure axis labels are often a source of confusion.

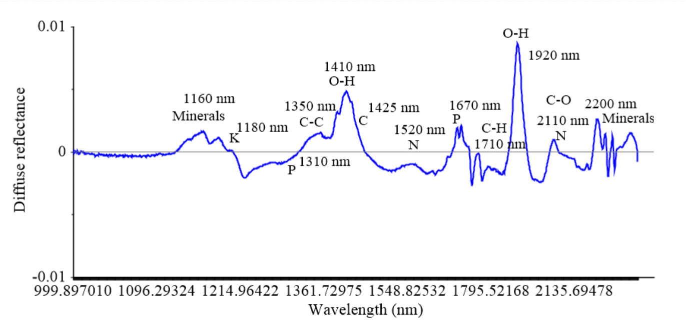

Fig. 1. Spectra features of soil samples in NIR wavelength region.

Moreover, soil macro nutrients like nitrogen (N), phosphorus (P) and potassium (K) were absorbed light radiation in wavelength range around 1520 nm and 2110 nm for N, 1310 nm and 1670 nm for P and 1180 nm for K. Similar findings also noted by other publications stated that soil macro nutrients can be predicted in early to middle NIR wavelength region since those nutrients reached highest light absorption. On the other hand, soil organic carbon was absorbed more light in wavelength range from 1350 to 1425 nm.

3.2 Soil quality attributes prediction

The main part of infrared technology application is to develop models used to predict several quality attributes of respected materials simultaneously. Those models were established using spectra data and reference data obtained from the laboratory analysis. In this study, we attempted to develop prediction models used to determine K, Mg and Ca contents of soil samples by employing regression approaches. Firstly, the principal component regression (PCR) was used to establish prediction models.

Table 1.
Prediction performance of K, Mg and Ca using principal component regression (PCR) method

Quality parameters	Statistical Indicators				
	R^2	r	RMSE	RPD	LVs
K	0.76	0.84	0.19	2.71	7
Mg	0.82	0.88	1.24	3.83	6
Ca	0.84	0.90	1.44	4.45	6

Ca: calcium, K: potassium, LVs: number of latent variables, Mg: magnesium, PCR: principal component regression, r: correlation coefficient, R^2: coefficient of determination, RMSE: root mean square error, RPD: residual predictive deviation index.

As shown in Table 1, all three quality parameters (K, Mg and Ca) can be predicted with maximum correlation coefficient 0.90. In term of accuracy and robustness, Mg and Ca can be determined very well with correlation coefficient (r) and residual predictive deviation (RPD) index of 0.88 and 3.83 for Mg, and r = 0,90, RPD = 4.45 for Ca prediction which was categorized as excellent prediction model performance. On the other hand, good model performance was achieved when the model used to predict K content. Even though generated lowest r coefficient and RPD index, K content can be predicted effectively and sufficiently using the PCR regression approach. In this study, we also attempted to develop prediction models using partial least square regression (PLSR) approach. Just like PCR method, the PLSR also regressed the NIR spectral data of soil samples and actual quality parameters obtained from laboratory analysis. Prediction performance for K, Mg and Ca using the PLSR approach is presented in Table 2 from which we can see that all those three quality parameters (K, Mg and Ca) of soil samples can be predicted very well and the overall results were better than obtained by the PCR method.

Table 2.
Prediction performance of K, Mg and Ca using partial least square regression (PLSR) method

Quality parameters	Statistical Indicators				
	R^2	r	RMSE	RPD	LVs
K	0.96	0.97	0.10	5.14	7
Mg	0.97	0.98	0.57	8.34	6
Ca	0.97	0.98	0.72	8.90	6

Ca: calcium, K: potassium, LVs: number of latent variables, Mg: magnesium, PCR: principal component regression, r: correlation coefficient, R^2: coefficient of determination, RMSE: root mean square error, RPD: residual predictive deviation index.

The correlation coefficients for K, Mg and Ca were significantly improved. The maximum correlation coefficient achieved was 0.98 with maximum robustness index of RPD was 8.90. All prediction models obtained by using PLSR method were categorized as excellent prediction models. Judging from the prediction performance, it seems that PLSR regression approach generated more accurate and robust prediction results than that in PCR one. Scatter plot derived from PLSR approach for all those three quality parameters of soil samples was presented in Figure 2. It shows that PLSR improved prediction accuracy by lowering prediction error (RMSE), and thus resulting high correlation coefficient and robustness index. Similar findings also reported by other researchers that PLSR generally provided and generated more accurate and robust prediction results compared to PCR approach. It is obvious that the PLSR regression approach enhanced and improved prediction performance for all three quality attributes prediction. The main reason behind this improvement is that in PLSR, the algorithm seeks to find best correlation between X variable (spectra data of all soil samples) and Y variable (K, Mg and Ca content) by mapping them onto latent variable spaces. Meanwhile, on the PCR, the algorithm only subjected X data onto the principal component map.

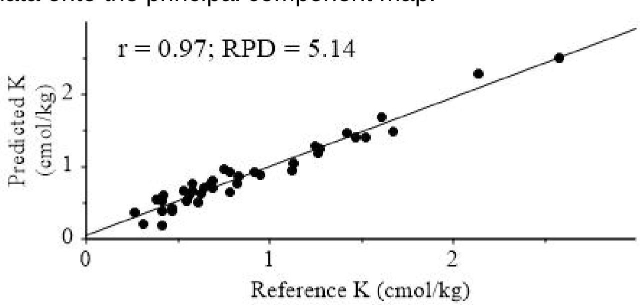

Fig. 2a. Scatter plot between actual reference and predicted K of soil samples using PLSR approach.

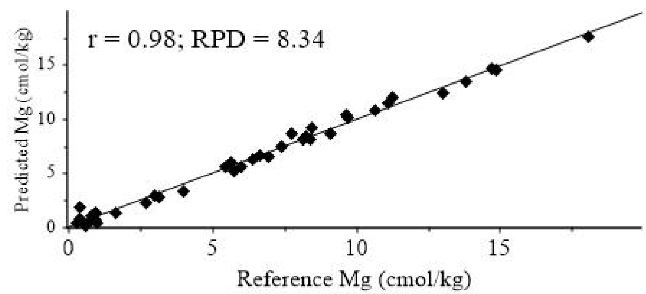

Fig. 2b. Scatter plot between actual reference and predicted Mg of soil samples using PLSR approach.

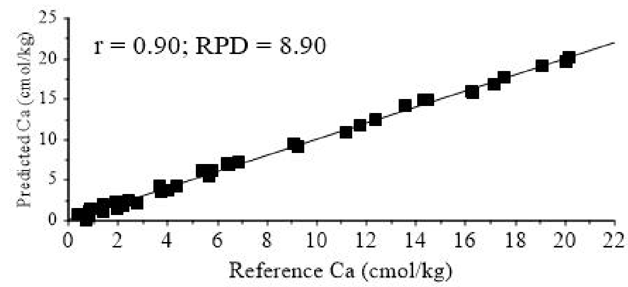

Fig. 2c. Scatter plot between actual reference and predicted Ca of soil samples using PLSR approach.

In term of the number latent variables (LVs) required to construct prediction models, both PCR and PLSR approaches required same number for LVs (7 for K prediction and 6 for both Mg and Ca prediction). The number LVs often considered and taken into account to evaluate the effectiveness of prediction models in NIRS application. It is obvious that lower number of LVs is preferable to avoid over-fitting and over optimistic prediction models. Based on literatures, it is recommended that maximum number of latent variables is 9 to 11 [7, 19].

4 Conclusion

Based on obtained results, it may conclude that near infrared reflectance spectroscopy (NIRS) can be applied as a fast and simultaneous method for major quality parameters prediction of soil samples (K, Mg and Ca). Prediction models established by the PLSR approach generated more accurate and robust prediction result for all those three quality parameters with maximum correlation coefficients between actual reference and predicted quality attributes: 0.97 for K and 0.98 for both Mg and Ca prediction respectively.

Acknowledgment

The authors wish to thank to LPPM Syiah Kuala University for research support and funding through *Hibah Penelitian Professor* scheme.

References

[1] Devianti, Sufardi, Zulfahrizal, and A. A. Munawar, “Rapid and Simultaneous Detection of Hazardous Heavy Metals Contamination in Agricultural Soil Using Infrared Reflectance Spectroscopy,” IOP Conf. Ser. Mater. Sci. Eng., vol. 506, p. 012008, 2019.
[2] M. Knadel et al., “Comparing predictive ability of laser-

induced breakdown spectroscopy to visible near-infrared spectroscopy for soil property determination," Biosyst. Eng., vol. 156, pp. 157–172, 2017.

[3] D. B. Madhavan et al., "Rapid prediction of particulate, humus and resistant fractions of soil organic carbon in reforested lands using infrared spectroscopy," J. Environ. Manage., vol. 193, pp. 290–299, 2017.

[4] J. Liu, Y. Zhang, H. Wang, and Y. Du, "Study on the prediction of soil heavy metal elements content based on visible near-infrared spectroscopy," Spectrochim. Acta - Part A Mol. Biomol. Spectrosc., vol. 199, pp. 43–49, 2018.

[5] P. Zhao, F. Yu, R. Wang, Y. Ma, and Y. Wu, "Sodium alginate/graphene oxide hydrogel beads as permeable reactive barrier material for the remediation of ciprofloxacin-contaminated groundwater," Chemosphere, vol. 200, pp. 612–620, 2018.

[6] M. Zovko et al., "A geostatistical Vis-NIR spectroscopy index to assess the incipient soil salinization in the Neretva River valley, Croatia," Geoderma, vol. 332, no. April, pp. 60–72, 2018.

[7] A. Munawar, D. V. Hörsten, D. Mörlein, E. Pawelzik, and J. K. Wegener, "Rapid and non-destructive prediction of mango sweetness and acidity using near infrared spectroscopy," in Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft fur Informatik (GI), 2013, vol. P-211.

[8] 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.

[9] N. Zhang et al., "Determination of total iron-reactive phenolics, anthocyanins and tannins in wine grapes of skins and seeds based on near-infrared hyperspectral imaging," Food Chem., vol. 237, pp. 811–817, 2017.

[10] M. Soto-Cámara, A. J. Gaitán-Jurado, and J. Domínguez, "Application of near infrared spectroscopy technology for the detection of fungicide treatment on durum wheat samples," Talanta, vol. 97, pp. 298–302, 2012.

[11] D. Cozzolino, "An overview of the use of infrared spectroscopy and chemometrics in authenticity and traceability of cereals," FRIN, vol. 60, pp. 262–265, 2014.

[12] B. M. Nicolaï et al., "Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review," Postharvest Biol. Technol., vol. 46, no. 2, pp. 99–118, 2007.

[13] E. Arendse, O. A. Fawole, L. S. Magwaza, and U. L. Opara, "Non-destructive prediction of internal and external quality attributes of fruit with thick rind: A review," J. Food Eng., vol. 217, pp. 11–23, 2018.

[14] C. Pasquini, "Near infrared spectroscopy: A mature analytical technique with new perspectives – A review," Anal. Chim. Acta, vol. 1026, pp. 8–36, 2018.

[15] V. Cortés et al., "Integration of simultaneous tactile sensing and visible and near-infrared reflectance spectroscopy in a robot gripper for mango quality assessment," Biosyst. Eng., vol. 162, pp. 112–123, 2017.

[16] K. Song, H. Cha, S. H. Park, and Y. I. Lee, "Determination of trace cobalt in fruit samples by resonance ionization mass spectrometry," Microchem. J., vol. 75, no. 2, pp. 87–96, 2003.

[17] N. Reis, A. S. Franca, and L. S. Oliveira, "Discrimination between roasted coffee, roasted corn and coffee husks by Diffuse Reflectance Infrared Fourier Transform Spectroscopy," LWT - Food Sci. Technol., vol. 50, no. 2, pp. 715–722, 2013.

[18] C. Li, H. Guo, B. Zong, P. He, F. Fan, and S. Gong, "Rapid and non-destructive discrimination of special-grade flat green tea using Near-infrared spectroscopy," Spectrochim. Acta - Part A Mol. Biomol. Spectrosc., vol. 206, pp. 254–262, 2019.

[19] N. Prieto, Ó. López-Campos, J. L. Aalhus, M. E. R. Dugan, M. Juárez, and B. Uttaro, "Use of near infrared spectroscopy for estimating meat chemical composition, quality traits and fatty acid content from cattle fed sunflower or flaxseed," Meat Sci., vol. 98, no. 2, pp. 279–288, 2014.

[20] H. T. Zhao, Y. Z. Feng, W. Chen, and G. F. Jia, "Application of invasive weed optimization and least square support vector machine for prediction of beef adulteration with spoiled beef based on visible near-infrared (Vis-NIR) hyperspectral imaging," Meat Sci., vol. 151, no. January, pp. 75–81, 2019.

[21] N. Meyer, H. Meyer, G. Welp, and W. Amelung, "Soil respiration and its temperature sensitivity (Q10): Rapid acquisition using mid-infrared spectroscopy," Geoderma, vol. 323, no. February, pp. 31–40, 2018.

[22] J. Liu, X. Luan, and F. Liu, "Adaptive JIT-Lasso modeling for online application of near infrared spectroscopy," Chemom. Intell. Lab. Syst., vol. 183, no. October, pp. 90–95, 2018.

[23] R. Matamala et al., "Predicting the decomposability of arctic tundra soil organic matter with mid infrared spectroscopy," Soil Biol. Biochem., vol. 129, no. July 2018, pp. 1–12, 2018.

[24] Samadi, S. Wajizah, and A. A. Munawar, "Fast and simultaneous prediction of animal feed nutritive values using near infrared reflectance spectroscopy," IOP Conf. Ser. Earth Environ. Sci., vol. 122, no. 1, 2018.

[25] A. Munawar, D. von Hörsten, J. K. Wegener, E. Pawelzik, and D. Mörlein, "Rapid and non-destructive prediction of mango quality attributes using Fourier transform near infrared spectroscopy and chemometrics," Eng. Agric. Environ. Food, vol. 9, no. 3, 2016.

[26] Yusmanizar et al., "Fast and Non-Destructive Prediction of Moisture Content and Chologenic Acid of Intact Coffee Beans Using Near Infrared Reflectance Spectroscopy," IOP Conf. Ser. Mater. Sci. Eng., vol. 506, p. 012033, 2019.