Agricultural soil fertility properties in the near infrared spectrum

A A Munawar1*, D Devianti1, P Satriyo1, Zainabun2

1Department of Agricultural Engineering, PUSMEPTAN, Universitas Syiah Kuala, Banda Aceh, Indonesia.
2Department of Soil Science, Universitas Syiah Kuala, Banda Aceh, Indonesia

*Email: aamunawar@unsyiah.ac.id

Abstract. Soil spectrum in the near infrared (NIR) wavelength region can be used to reveal fertility properties which is related to plant cultivations. The main purpose of this presented paper is to study the soil spectrum in the NIR region and its related to the fertility properties in form of heavy metals like Fe and Cu. Soil samples were obtain from several land-use including agriculture, mining and ground field. Near infrared spectrum of soil samples were acquired in wavelength range from 1000 to 2500 nm. Prediction models used to determine Fe and Cu were built by means of partial least squares regression (PLSR) followed by leverage cross validation. Prediction performance was evaluated using coefficient of determination (r²) and ratio of prediction to deviation (RPD). The results showed that both Fe and Cu can be revealed simultaneously using the NIR spectrum with maximum r² and RPD indexes were 0.93 and 3.86 for Fe and 0.71 and 1.88 for Cu prediction respectively. Based on the achieved results, it may conclude that soil fertility properties can be revealed simultaneously and rapidly using near infrared spectral data.

1. Introduction

In smart agricultural and farming practices, soil fertility is one of the most important key factor affecting yield and agricultural products quality. The chemical balance of soil and elements of living organisms is the main requirement for plant growth, this is the main reason why soil treatments such as liming are often performed to keep the plant alive [1].

There are several micro elements that are needed by plants even though in relatively high concentrations they are not toxic, namely: B, Cl, Cu, Fe, Mo, Ni. Yet, those elements have a negative impact on plant growth. While other elements that have no effect when in low concentrations such as cadmium, chromium, mercury, and lead, but if in higher concentrations these elements can be deadly [2,3]. It is important to know that heavy metals in soil and plants are very dangerous, especially in the soil it will not decompose without special treatment so that it can accumulate continuously. Several studies that have been carried out have discussed the impact of heavy metals on plants and the concentrations of heavy metals for plants have been determined [4–6].

Soil fertility are greatly affect the ability of the soil to fulfill its function. There are three functions of soil that are closely related to soil quality, namely as a medium for plant growth, regulating and dividing water and can act as environmental filters. The total heavy metal content in the soil is highly dependent on the clay content of the soil. About 96% of the clay fraction contains heavy metals such as...
Co, Cu, Ni, Pb and Zn. Soil organic carbon content is the main characteristic for soil quality which can affect various kinds of organic compounds and soil physical properties [7–9].

The presence of heavy metals is related to the levels of organic matter in the soil. The presence of organic matter in the soil will cause chelation of metal cations. The processes that occur in the soil are mostly carried out by the relatively few soil constituents, namely clay and humus. The application of various organic materials such as manure, compost, and municipal waste/waste indirectly contributes to the accumulation of heavy metals in the soil. Examples of these heavy metals are As, Cd, Cr, Cu, Pb, Hg, Ni, Se, Mo, Zn, Ti, Sb, and so on. Certain animal manures such as poultry, cattle, pigs and fertilizers produced in agriculture are generally applied to crops and pastures as solids or liquids [10,11].

To determine soil fertility properties, several methods have been widely applied. However, most of them are based on chemical analysis and measurements which are time consuming, complicated sample preparations and destructive in nature [4,12]. Therefore, alternative methods are required to determine soil fertility which is non-destructive, environmental friendly, and rapid. Near infrared spectroscopy is one among those methods that is popular and widely employed in many fields including agriculture [13]. Thus, the main purpose of this present paper is to study the soil spectrum in the NIR region and its related to the fertility properties in form of heavy metals like Fe and Cu.

2. Materials and methods

2.1. Soil samples

A total of 25 bulk of soil samples from different agricultural land-use were used in this study. At each location, soil samples were taken with a depth approximately 25-35 cm with a sample size of 1-2/ha. Then the soil was collected and stored for three days to balance the standardization of moisture levels, small stones and plant residues contained in the soil sample were cleaned first. The soil sample will be ground with a mortar and sieved to a size of 20 mesh (0.84 mm) to minimize the impact of particle size on the soil [14].

2.2. Near infrared spectrum of soil

Spectra data of soil samples in form of diffuse reflectance spectrum were obtained in near infrared wavelength region from 1,000 to 2,500 nm with co added of 32 number of scans [8,15,16]. Collected spectrum were enhanced using mean normalization (MN) and the prediction performance was then compared.

2.3. Spectra calibration

Nutmegs soil spectrum were calibrated to obtain prediction models used to determine Fe and Cu on soil samples. The partial least squares regression (PLSR) approach was used to develop the models followed by leverage cross validation to avoid overfitting [17,18].

2.4. Model performance evaluation

Obtained models were evaluated based on their prediction accuracy both for raw spectrum and corrected ones. Several statistical indicators were introduced to evaluate the models namely coefficient of determination ($R^2$), correlation coefficient ($r$), root mean square error (RMSE) or also known as root mean square deviation (RMSD) in calibration (RMSEC), in cross validation (RMSECV), and ratio of prediction to deviation (RPD) index as shown in equation (1) to equation (4) below [17,19–21]:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

Where

$$SS_{res} = \sum_i (y_i - f_i)^2 = \sum_i e_i^2$$

and

$$SS_{tot} = \sum_i (y_i - \bar{y})^2$$

(2)
It is obvious that good model performance is to be expected has high $r^2$, $r$ and RPD whilst RMSE is low. The number of latent variables or factor also considered to justify the models [17,22]. Fewer number of latent variables is preferable.

3. Results and discussion

3.1. Soil spectrum

The near infrared spectrum of soil samples is presented in Figure 1. From the results of the acquisition of raw spectrum from soil samples using NIRs technology, it can be observed that the shape of the spectrum of the sample of the paddy field soil with the spectrum of the field soil does not have a very significant difference. This is because the content of organic, inorganic and other substances contained in the sample of paddy field soil and sample of field soil is the same. However, the difference is the absorbance or transmission of the NIR rays on the biological material in the sample at a certain wavelength, resulting in differences in peaks at each wavelength due to stronger wave vibrations.

![Figure 1. Soil spectrum in near infrared wavelength region from 1,000 to 2,500 nm.](image)

There are several other things that also influence the difference in the soil spectrum, including differences in soil texture, causing different grains of soil particles. This affects the presence of air gaps and air cavities, causing scatter effects in measurement using NIRs technology. Basically the denser the particles contained in the soil sample used, the smaller the scatter effects in the scanning process, so the distance between the spectrum will be closer.

Various kinds of other negative impacts that can be caused by heavy metals in soil, namely poisoning in biological processes include various processes catalysed by microorganisms. Critical and allowed amount of heavy metals in several medium is presented in Table 1. Microbial commodities can be used as indicators of changes in soil quality due to heavy metal contamination or changes in soil quality, this can occur in excessive use of agrochemicals such as fertilizers and pesticides. In addition, this also results from industrial activities, especially mining activities.
Table 1. Critical and allowed heavy metals in several medium (ppm)

| Heavy metals | Soil  | Water  | Plants | Rice |
|--------------|-------|--------|--------|------|
| Pb           | 100   | 0.03   | 50     | 1.0  |
| Cd           | 0.50  | 0.05-0.10 | 5-30 | 0.5  |
| Co           | 10    | 04-0.6 | 15-30  | -    |
| Cr           | 2.5   | 0.5-1.0 | 5-30  | -    |
| Ni           | 20    | 0.2-0.5 | 5-30  | -    |
| Cu           | 60-125| 2-3    | 20-100 | 10   |
| Mn           | 1,500 | -      | -      | -    |
| Zn           | 70    | 5-10   | 100-400| 40   |

3.2. Fe and Cu calibration models

The main part in the NIRS application is to establish prediction models used to determine Fe and Cu through calibration. The method used for the calibration is partial least square regression from which required spectral data as independent variables and actual measurement of Fe and Cu. Descriptive statistics of the actual Fe and Cu on soil samples is presented in Table 2.

Table 2. Descriptive statistics of actual Fe and Cu (mg/kg)

|               | Fe (mg/kg) | Cu (mg/kg) |
|---------------|------------|------------|
| n             | 40         | 40         |
| Mean          | 356.325    | 3.821      |
| Max           | 781.700    | 8.150      |
| Min           | 114.700    | 0.950      |
| Range         | 667.000    | 7.200      |
| Std. Deviation| 213.087    | 2.196      |
| RMS           | 413.810    | 4.394      |
| Skewness      | 0.675      | 0.498      |
| Kurtosis      | -0.998     | -0.984     |
| Median        | 263.720    | 3.350      |
| Q1            | 185.763    | 1.718      |
| Q3            | 503.608    | 5.515      |

A total of 40 data samples were then subjected onto PLSR approach to construct the models. At first, raw untreated spectra data or original spectrum were used to develop the models for both Fe and Cu prediction. The prediction performance for Fe and Cu using raw spectrum is presented in Figure 2. It shows that the coefficient of determination ($r^2$) for Fe and Cu prediction are 0.72 and 0.61 respectively. It seems that Fe can be predicted with sufficient accuracy from which the ratio of prediction to deviation (RPD) index was 1.92 whereas for Cu, the RPD is 1.51 which categorized as coarse prediction performance and need improvement. Based on literatures [20,21,23], near infrared model’s performance can be judged using the RPD index as follows: RPD between 1.5 – 1.9 means that coarse quantitative prediction is possible, but still need some improvement in calibration. A value between 2 and 2.5 indicates that prediction model is sufficient. Meanwhile, an RPD value between 2.5 and 3 or above corresponds to good and excellent prediction accuracy respectively.
Moreover, for Fe determination, the models required a total 5 latent variables or PLS factors to achieve mentioned results whilst for Cu, the number of PLS factors required to achieve this performance is 9 factors.

![Figure 2](image2.png)

**Figure 2.** Prediction performance of PLSR based on raw spectrum for Fe and Cu determination.

As commonly performed in NIRS practices, the raw spectrum can be corrected to minimize and eliminate irrelevant background information appeared in the spectra. In this study, we attempted to correct the spectrum using mean normalization (MN) and first derivative (D1). The prediction performance after MN spectra correction was improved as shown in Figure 3.

![Figure 3](image3.png)

**Figure 3.** Prediction performance of PLSR based on MN spectrum for Fe and Cu determination.

Prediction accuracy was increased when the models are constructed using MN spectrum both for Fe and Cu prediction. The coefficient of determination increased to 0.93 for Fe with the same required PLS
factor 5. Consequently, the RPD index was significantly improve to 3.86 as root means square error is decreased. According to RPD classification, the prediction performance for Fe using MN spectrum is categorized as excellent prediction performance. Similar finds also noted for Cu prediction whereas the coefficient of determination was also increased to 0.71 from previous achieved one using raw spectrum 0.61. Thus, the RPD index was also improved as categorized as sufficient model performance. The required PLS factor for Cu prediction was also improved to 8 from previous achieved by the raw spectrum. It is obvious that spectra correction has a significant impact to the overall prediction performance of NIRS model for both soil fertility properties.

4. Conclusions

The main purpose of this presented paper is to study the soil spectrum in the NIR wavelength region ranging from 1000 to 2500 nm, and its related to the fertility properties in form of heavy metals like Fe and Cu. The results showed that both Fe and Cu can be revealed simultaneously using the NIR spectrum with maximum $r^2$ and RPD indexes were 0.93 and 3.86 for Fe and 0.71 and 1.88 for Cu prediction respectively. Based on the achieved results, it may conclude that soil fertility properties can be revealed simultaneously and rapidly using near infrared spectral data. Moreover, it is obvious that spectra correction has a significant impact to the overall prediction performance of NIRS model for both soil fertility properties.

Acknowledgements

We would to express our sincere gratitude and acknowledge to the DRPM Ristek BRIN for funding this study through hibah penelitian dasar 2021.

References

[1] Toohey R C, Boll J, Brooks E S and Jones J R 2018 Effects of land use on soil properties and hydrological processes at the point, plot, and catchment scale in volcanic soils near Turrialba, Costa Rica Geoderma 315 138–48

[2] Ning Y, Li J, Cai W and Shao X 2012 Simultaneous determination of heavy metal ions in water using near-infrared spectroscopy with preconcentration by nano-hydroxyapatite Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 96 289–94

[3] Ren H Y, Zhuang D F, Singh A N, Pan J J, Qi D S and Shi R H 2009 Estimation of As and Cu Contamination in Agricultural Soils Around a Mining Area by Reflectance Spectroscopy: A Case Study Pedosphere 19 719–26

[4] Mohamed E S, Saleh A M, Belal A B and Gad A A 2018 Application of near-infrared reflectance for quantitative assessment of soil properties Egypt. J. Remote Sens. Sp. Sci. 21 1–14

[5] Chakraborty S, Weindorf D C, Deb S, Li B, Paul S, Choudhury A and Ray D P 2017 Rapid assessment of regional soil arsenic pollution risk via diffuse reflectance spectroscopy Geoderma 289 72–81

[6] Tao C, Wang Y, Cui W, Zou B, Zou Z and Tu Y 2019 A transferable spectroscopic diagnosis model for predicting arsenic contamination in soil Sci. Total Environ. 669 964–72

[7] Darusman D, Juwita I R, Munawar A A, Zainabun Z and Zulfahirzal Z 2021 Rapid determination of mixed soil and biochar properties using a shortwave near infrared spectroscopy approach IOP Conf. Ser. Earth Environ. Sci. 667

[8] 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

[9] Ramli I, Rusdiana S, Basri H, Munawar A A and Zelia V A 2019 Predicted Rainfall and discharge Using Vector Autoregressive Models in Water Resources Management in the High Hill Takengon IOP Conference Series: Earth and Environmental Science vol 273 (Institute of Physics Publishing)
[10] Soriano-Disla J M, Janik L J and McLaughlin M J 2018 Assessment of cyanide contamination in soils with a handheld mid-infrared spectrometer *Talanta* **178** 400–9

[11] Liu J, Han J, Xie J, Wang H, Tong W and Ba Y 2020 Assessing heavy metal concentrations in earth-cumulic-orthic-anthrosols soils using Vis-NIR spectroscopy transform coupled with chemometrics *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **226** 117639

[12] Galvez-Sola L, Moral R, Perez-Murcia M D, Perez-Espinosa A, Bustamante M A, Martinez-Sabater E and Paredes C 2010 The potential of near infrared reflectance spectroscopy (NIRS) for the estimation of agroindustrial compost quality *Sci. Total Environ.* **408** 1414–21

[13] Munawar A A, von Hörsten D, Wegener J K, Pawelzik E and Mörlein D 2016 Rapid and non-destructive prediction of mango quality attributes using Fourier transform near infrared spectroscopy and chemometrics *Eng. Agric. Environ. Food* **9** 208–15

[14] Munawar A A, Yunus Y, Devianti D and Satriyo P 2021 Agriculture environment monitoring: rapid soil fertility evaluation by means of near infrared spectroscopy *IOP Conf. Ser. Earth Environ. Sci.* **644** 012036

[15] Devianti D, Sufardi S, Zulfahrizal Z and Munawar A A 2019 Rapid and Simultaneous Detection of Hazardous Heavy Metals Contamination in Agricultural Soil Using Infrared Reflectance Spectroscopy *IOP Conference Series: Materials Science and Engineering* vol 506 (Institute of Physics Publishing)

[16] Ichwana I, Nasution Z and Arip Munawar A 2020 Near-Infrared Spectroscopy As A Rapid And Simultaneous Assessment of Agricultural Groundwater Quality Parameters *INMATEH Agric. Eng.* **233**–40

[17] Pasquini C 2018 Near infrared spectroscopy: A mature analytical technique with new perspectives – A review *Anal. Chim. Acta* **1026** 8–36

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

[19] Cozzolino D 2014 An overview of the use of infrared spectroscopy and chemometrics in authenticity and traceability of cereals *FRIN* **60** 262–5

[20] Cen H and He Y 2007 Theory and application of near infrared reflectance spectroscopy in determination of food quality *Trends Food Sci. Technol.* **18** 72–83

[21] Nicolaï B M, Beullens K, Bobelyn E, Peirs A, Saeyes W, Theron K I and Lammertyn J 2007 Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review *Postharvest Biol. Technol.* **46** 99–118

[22] Deng Y, Wang Y, Zhong G and Yu X 2018 Simultaneous quantitative analysis of protein, carbohydrate and fat in nutritionally complete formulas of medical foods by near-infrared spectroscopy *Infrared Phys. Technol.* **93** 124–9

[23] Darusman D, Zulfahzial Z, Yunus Y and Munawar A A 2019 Soil Quality Assessment By Near Infrared Spectroscopy: Predicting Ph And Soil Organic Carbon *Int. J. Sci. Technol. Res.* **8**