Agriculture environment monitoring: rapid soil fertility evaluation by means of near infrared spectroscopy

A A Munawar*, Y Yunus, D Devianti, P Satriyo

Department of Agricultural Engineering, Syiah Kuala University, Banda Aceh, Indonesia
Agricultural Mechanization Research Centre, Syiah Kuala University, Banda Aceh, Indonesia

Corresponding author: aamunawar@unsyiah.ac.id

Abstract. In agricultural practices, the development of environmental controlling programs is essential in maintaining soil fertility. It should be monitored constantly in real time situation by predicting soil properties such as nutrients and mineral contents. The main purpose of this presented paper is employ the near infrared spectroscopy (NIRS) technique in determining nitrogen (N) and magnesium (Mg) content of agricultural soil rapidly and simultaneously. Near infrared spectra data of soil samples were obtained by acquiring absorbance data in wavelength range from 1000 to 2500 nm using a portable sensing device NIR with a photodiode array detector. Spectra data were enhanced using mean normalization approach to improve prediction accuracy. Further, prediction models were developed by means of partial least square regression (PLSR) followed by cross validation. The results showed that both soil fertility properties (N and Mg) can be predicted by NIRS with maximum correlation coefficient (r) were 0.93 and 0.91 for N and Mg respectively. It may conclude that NIRS technique was able to be used in agriculture environment monitoring especialy to predict soil fertility properties.

1. Introduction
Population growth was normally followed by an increase in food needs, which in general food needed came from agricultural land. Nationally, the average population of Indonesia only has 0.25 ha of arable land of family. The land area of course is not enough to meet the food needs of the family, causing farmers to work on land that is destined for non-cultivation areas, or lands that are on a slope above 10%. The reality shows that in general farmers do not yet have the knowledge or awareness in managing topographic land, whereas intensive land is used for agricultural practices which results in erosion-related agricultural land. Erosion is the release of soil aggregates by the kinetic energy of the rain, and the contaminated soil is transported by runoff water and ends in the lower land [1].

The intensification of agricultural cultivation on topographic land certainly results in the loss of topsoil which is a layer that contains macro and micro nutrients. These elements are nutrients that are needed for plant development and productivity [2]. Obviously plants can grow ideally on soils that have physical and chemical properties that are suitable for plant growth.

Physically, healthy soil conditions can be seen visually such as texture, structure and humidity. Soil chemistry is related to the amount of nutrients needed by plants, with the amount needed will be different for each growth phase. Soil fertility is largely determined by the presence of nutrients in the soil, both macro nutrients, secondary nutrients and micro nutrients. Macro nutrients include nitrogen (N), phosphorus (P), potassium (K), and C, H, O. Medium secondary nutrients include calcium (Ca), magnesium (Mg), and sulfur [3], [4]. Nutrient losses due to erosion on agricultural land are known by conducting...
laboratory testing which in practice, this analysis requires a relatively long time, involving chemicals that can cause pollution and relatively large costs. Nutrient requirements must be immediately known so that plants can grow ideally. Nutrients in the soil must be in sufficient quantities and balanced composition, because if one of the elements is reduced it can cause plant growth to be unnatural. Each nutrient has a certain task and not a single element that can replace it perfectly [5].

At this time, considerable attention and effort continues to be developed by researchers to find alternative methods that are fast, effective, environmentally friendly and non-destructive to predict levels of nutrient concentrations and heavy metals in the soil [5], [6]. One technology that is currently developing rapidly and has the potential to be used to detect / predict the content of this heavy metal, is near infrared reflectance spectroscopy (NIRS) technology. The advantages of this method include not damaging the material, relatively easy sample preparation does not require chemicals and can predict soil nutrient contents simultaneously [7].

NIRS technology works based on the principle that every biological material, including soil has certain electro-optical characteristics where the spectrum produced can be analyzed to obtain information about the organic content of the soil including to predict the content of heavy metal concentrations as the main cause of pollution in the soil [8], [9]. This information is obtained by analyzing the data spectrum of this biological material through special mathematical methods known as chemometrics. This method includes a multivariate analysis method to find the relationship between chemical or physical data from biological materials with electro-optical data from the material being studied.

The potential and track record of our research experience as a research team has made it possible to conduct research on the application of new alternative methods that are rapid, effective and simultaneous for determining nutrient losses due to erosion on topographic agricultural land in the area of study in Aceh Besar by predicting levels of nutrient loss on agricultural land by cultivation of annual crops by monoculture, intercropping on topography farming land. This research was carried out by combining NIRS technology and chemometrics methods and the results will be verified through laboratory standard testing to determine the accuracy and reliability of the applied technology in predicting nutrient losses due to erosion on topographic agricultural land.

NIRS technology that has developed since the last few decades has been and is still being studied and applied globally in various countries. Numerous studies and applications of NIRS technology have been reported widely in the field of agriculture such as fruit and horticulture products [10]–[14], animal feed and meat [15]–[17], agricultural crops, and also in environmental studies [18]–[21]. Thus, the main purpose of this presented research is employ the near infrared spectroscopy (NIRS) technique in determining nitrogen (N) and magnesium (Mg) content of agricultural soil rapidly and simultaneously. Prediction models were developed using partial least square regression (PLSR) approach.

2. Material and Methods
This present study was performed at Soil and Water laboratory, Agricultural Engineering Department, Syiah Kuala University, Banda Aceh. Soil and water samples due to erosion were obtained from making erosion plots with intercropping and monoculture seasonal management. The stages, procedures and methods carried out in this study in more detail are explained as follows:

2.1 Environmental preparations
Materials to be used in this study are soil and water runoff samples resulting from rain erosion in the measurement plot by adjusting the cropping patterns of annual crops in a monoculture and intercropping manner. Erosion plots were made using the small plot method, using land of 22 meters in length and 4 meters in width. Measurement of erosion is carried out every rain event, erosion that is accommodated at the outlet of the plot, dried air for one day then weighted [9]. All samples were taken directly to the laboratory and dried at room temperature for three days for uniform temperature and humidity before spectrum acquisition. Soil samples were previously cleaned from rocks, gravel and plant residues.
Further, to minimize the effect of spectral noise due to the influence of soil particle size, soil samples were mashed with mortar and filtered with a filter size of 20 mesh or 0.84 mm.

2.2 Spectra data acquisition
The near infrared spectra data for soil samples are acquired and obtained using NIR instruments (PSD-FTNIR i16) with workflow configurations built using Thermo Integration® software. Workflows are developed to acquire the diffuse reflectance spectrum, scan 64 times and then average the results, then record the scan results in three file formats for further data analysis. Integrating sphere is chosen as a spectrum acquisition method where background calibration is conducted every hour [22]. The wavelength range used is 1000-2500 nm with intervals of around 2 nm. Sampling of the soil sample spectrum is done by placing the soil sample (± 50 g) in the sample cup holder and arranged so that the sample bowl rotates 360 degrees during the scanning process of sample spectrum acquisition.

2.3 Spectra enhancement
Before used for data analysis and building prediction models, the near infrared spectrum for all soil samples were carried out for spectra correction and enhancement. This aims of this process is to eliminate various kinds of noise in the soil spectra data so that the results of nutrient contents predictions are more accurate and robust [3], [23]. The methods used in this spectra correction and enhancement are: Gaussian filter smoothing (GF) and area normalization (AN).

2.4 Environmental monitoring model
Nutrient content in the soil samples were predicted based on the NIR spectrum as an environmental monitoring models. They were generated through the process called as calibration in order to develop prediction models. The models were built by regressing the NIR spectrum (variable X) with actual measured soil nutrient contents (variable Y) from the measurement results in the laboratory. The regression method used is partial least square regression (PLSR).

3. Result and discussion
3.1 Soil spectra feature and profile
Soil spectra profiles in infrared spectrometry contain absorption at each wavelength of radiation absorbed. This can be represented by rows of numbers representing absorption at different wavelengths. This spectrum profile may contain information that reflects the nature of the sample or can be related to the physical or chemical nature of the system. Reflectant measurements of soil samples were carried out at wavelengths of 1000-2500 nm as shown in Figure 1.

![Figure 1. Soil spectra profile and feature in 10 cm and 20 cm depth.](image-url)
Reflectant data is indirectly used to measure the amount of energy absorbed by the sample. Spectrum measurement results can be transformed into absorbance data (absorption). Absorbance data is obtained by transforming the reflectant value into a form of absorbance and its derivative (D1). This transformation was carried out because the chemical composition of a material has a linear relationship with the NIR absorbance data. In general, the typical diffuse reflectance spectrum in the infrared spectrum range with a wavelength of 1000 - 2500 nm for soil samples in this study looks like in those figure. From this spectrum it appears that the spectrum features indicate the presence of organic matter and the nutrient content and chemical properties of the soil due to molecular bond interactions OH, CH, CO and NH area.

The diffuse reflectance spectrum form for this soil sample occurs due to changes in energy vibrations in the form of overtone, bending and stretching. The carbonate element dominates chemical properties in soil samples. It can be seen that there are relevant wavelengths or band assignments for chemical bonds that reflect soil properties and nutrients such as CH bonds in bands 1070-1090 nm, 1105-1210 nm, 2325-2380 nm, OH in bands 1430 and 1910 nm, CHO on bands 1680-1720 nm, 2300 - 2490 nm [2], [24]. The electro-optical properties of this soil can then be used as a reference in predicting nutrient content and mineral elements in the soil by applying near infrared technology.

Absorbance data that have high concentrations of chemical compounds are represented by the peaks on the graph. This will make it easier to interpret the data, because high concentrations are represented by high peaks. The occurrence of wave crests on the NIR spectrum graph is due to the absorption of wavelengths by certain chemical contents [25].

In this study, the samples used were soil samples from the erosion of rain events in the measurement plot by adjusting the cropping patterns of annual crops in a monoculture and intercropping manner. Soil samples were taken with a total of 40 samples with the assumption that there are differences in the levels of distributed soil nutrients, so that an adequate range of data is obtained. The statistical description of nutrient data on the soil is shown in Table 1.

| Table 1. Descriptive statistics of actual measured soil nutrients |
|---------------------------------------------------------------|
| **Descriptive statistics** | **N** | **Mg** |
| Mean                     | 0.15  | 6.43  |
| Max                      | 0.52  | 18.07 |
| Min                      | 0.02  | 0.31  |
| Range                    | 0.50  | 17.76 |
| Std. Deviation           | 0.14  | 4.75  |
| Variance                 | 0.02  | 22.60 |
| Skewness                 | 1.25  | 0.43  |
| Kurtosis                 | 0.47  | -0.52 |
| Median                   | 0.09  | 6.54  |

Prediction models for N and Mg determination were established by means of partial least square regression (PLSR) approach. In general, this method is more widely used because with this method data in the spectrum are reduced to prevent overfitting problems without losing one or some very useful information. Calibration carried out with PLSR, takes into account reference data in the formation of the model. This approach is a multivariate calibration method for estimating the content of a material using a linear combination to estimate the independent variables from the original variables [26], [27]. In the PLSR method, variables that show a very high correlation with dependent variables will be given a large weight because these variables will be more effective in estimating. The NIRS spectrum calibration stage using the PLSR method also uses some data processing. Each data processing carried out has their respective functions.
3.2 Soil nitrogen prediction

Soil nitrogen levels in respective samples was predicted using infrared spectrum data that has been obtained. In the initial stage, the prediction model is built by regressing raw spectrum data (without spectrum improvement) with actual nitrogen content (N) data measured in the laboratory. The regression method used is partial least squares regression (PLSR). In general, soil N levels are predicted with unfavorable results using raw absorbance spectrum with correlation coefficient ($r$) = 0.73, as well as ratio prediction to deviation (RPD) index is 1.48. Scatter plot between actual and predicted N content of soil samples is shown in Figure 2.

![Figure 2. Scatter plot between actual and predicted N content of soil samples using raw spectrum NIR.](image)

Nitrogen content prediction in the soil samples using raw spectrum still belongs to the group of rough predictions (sufficient performance). Therefore, in order to increase the value of the RPD, calibration is re-modelled using enhanced spectra data: Gaussian filter (GF) and area normalization (AN) correction methods. The results of the prediction of N levels after being corrected and enhanced using the GF correction method showed different results obtained as presented in Figure 3.

![Figure 3. Scatter plot between actual and predicted N content of soil samples using GF spectrum.](image)

The prediction results for soil N determination using the GF and AN correction methods produce a better prediction performance than raw spectra data. The accuracy and robustness index achieved by
the GF correction method; $r = 0.73$, RPD = 1.48. And finally the AN correction method; $r = 0.72$, RPD = 1.47. All correction methods produce RPD values lower than 2 which according to the literature belong to the group of rough predictions (sufficient performance) as presented in Figure 4 using area normalization (AN) spectra data.

![Figure 4](image_url)

**Figure 4.** Scatter plot between actual and predicted N content of soil samples using AN spectrum.

Based on obtained results, spectra corrections did not make any different to the overall performance. This indicates that for certain spectrum corrections, important information on the spectrum is considered noise and is eliminated when the correction method algorithm is performed. Evaluation performance for N and Mg prediction in soil samples using GF and AN near infrared spectra data is presented in Table 2 and Table 3.

**Table 2.** Prediction performance for N and Mg determination using GF spectra data.

| Soil fertilities  | SD   | $R^2$ | $r$  | RMSEC | LV | RPD |
|-------------------|------|-------|------|-------|----|-----|
| Nitrogen (N)      | 0.14 | 0.62  | 0.78 | 0.08  | 8  | 1.64|
| Magnesium (Mg)   | 4.75 | 0.48  | 0.69 | 3.36  | 8  | 1.41|

LV: number of latent variables, $r$: correlation coefficient, $R^2$: coefficient of determination, RMSEC: root mean square error in calibration, RPD: ratio prediction to deviation, SD: standard deviation of actual measurement.

**Table 3.** Prediction performance for N and Mg determination using AN spectra data

| Soil fertilities  | SD   | $R^2$ | $r$  | RMSEC | LV | RPD |
|-------------------|------|-------|------|-------|----|-----|
| Nitrogen (N)      | 0.14 | 0.60  | 0.77 | 0.08  | 20 | 1.61|
| Magnesium (Mg)   | 4.75 | 0.49  | 0.70 | 3.33  | 20 | 1.42|

LV: number of latent variables, $r$: correlation coefficient, $R^2$: coefficient of determination, RMSEC: root mean square error in calibration, RPD: ratio prediction to deviation, SD: standard deviation of actual measurement.

Soil magnesium content (Mg) prediction using raw spectrum aims to determine the error rate of the statistical parameters of the resulting raw spectrum. Furthermore, the results of Mg prediction using the raw spectrum are corrected using the Gaussian filter (GF) and area normalization (AN) correction methods. The improvement aims to compare the results of the raw spectrum prediction and the prediction results using the correction method in order to obtain better and more accurate data. Based on obtained result, scatter plot shows Mg prediction data using the partial least squares (PLSR) method.
with the results of the correlation coefficient is 0.95 and the resulting RPD value index is 3.41 which is classified as a very good prediction as presented in Figure 5.

**Figure 5.** Scatter plot between actual and predicted Mg content of soil samples using raw spectrum.

In general, the Gaussian filter correction method is able to improve the ground infrared spectrum by minimizing disturbances (noises) due to light mixing that occurs during the process of spectrum data acquisition. Light mixing can be caused by several factors, including: changes in curvature and changes in sensor and detector position due to temperature fluctuations in the electronic elements of the instrument, the presence of gaps in the sample particles that allow empty space so that light hits the air or other medium, and other additive matters that in relation to the biological material conditions of the sample being tested as presented in Figure 6.

**Figure 6.** Scatter plot between actual and predicted Mg content of soil samples using GF spectrum.

Furthermore, we also attempted to perform spectra correction by applying the area normalization (AN) spectrum method. This is done to ensure the choice of the correct spectrum correction method for further use in the prediction of soil nutrient levels. The results of the correction of the spectrum of soil samples using the AN spectrum correction method. Each data processing performed has its own function. Data processing with normalization functions to transform the spectrum into a scope and all data at the same scale.
The typical AN method of the electrum is not much different from the raw spectrum. The difference between the two is that the AN method band spectrum is slightly thicker than the raw spectrum. Thick spectrum bands indicate noise on the spectrum as shown in Figure 7.

Figure 7. Scatter plot between actual and predicted Mg content of soil samples using AN spectrum.

The performance evaluation results of predicted Mg levels in the soil using GF and AN enhancement methods produce various prediction models. Prediction results obtained from the enhanced spectra are better and some are declining in value from the raw spectrum. As for the GF correction method; \( r = 0.95, \text{ RPD} = 3.51 \) while the AN correction method; \( r = 0.92, \text{ RPD} = 2.66 \). The GF spectra data methods can be classified as very good predictions while AN correction methods are classified as good performance. In this case the correction of AN spectrum does not make the prediction results better than the raw spectrum. This indicates that in this spectrum correction, important information on the spectrum is considered noise and is removed when the correction method algorithm is performed. Therefore the area normalization (AN) spectrum method is less suitable in predicting Mg levels in the soil. A good spectrum method used at soil Mg levels is the Gaussian Filter (GF) correction method because it can increase the value of the raw spectrum RPD.

4. Conclusion
The feasible ability of near infrared technology for environmental monitoring is being investigates in this study. Soil fertility properties were predicted simultaneously in form of nitrogen (N) and magnesium content (Mg) using three different spectra data namely raw spectrum, Gaussian filter and area normalization. Based on obtained results, it may conclude that soil fertility properties can be predicted and determined using NIRS technology and thus, it is feasible to apply this technique in environmental monitoring practices.

References
[1] X. Zhang, W. Sun, Y. Cen, L. Zhang, and N. Wang, “Predicting cadmium concentration in soils using laboratory and field reflectance spectroscopy,” Sci. Total Environ., vol. 650, pp. 321–334, 2019.
[2] C. Tao, Y. Wang, W. Cui, B. Zou, Z. Zou, and Y. Tu, “A transferable spectroscopic diagnosis model for predicting arsenic contamination in soil,” Sci. Total Environ., vol. 669, pp. 964–972, 2019.
[3] A. A. Munawar, Devianti, P. Satriyo, Syahrul, and Y. Yunus, “Rapid and simultaneous prediction of soil quality attributes using near infrared technology,” Int. J. Sci. Technol. Res.,
vol. 8, no. 9, pp. 725–728, Sep. 2019.

[4] Syahrul, Y. Yunus, P. Satriyo, and A. A. Munawar, “Applying infrared reflectance spectroscopy to predict water quality in Aceh river,” Int. J. Sci. Technol. Res., vol. 8, no. 10, pp. 969–972, Oct. 2019.

[5] Y. Yunus, Devianti, P. Satriyo, and A. A. Munawar, “Rapid Prediction of Soil Quality Indices Using Near Infrared Spectroscopy,” in IOP Conference Series: Earth and Environmental Science, 2019, vol. 365, no. 1.

[6] I. Ramli, S. Rusdiana, H. Basri, A. A. Munawar, and V. A. Zelia, “Predicted Rainfall and discharge Using Vector Autoregressive Models in Water Resources Management in the High Hill Takengon,” in IOP Conference Series: Earth and Environmental Science, 2019, vol. 273, no. 1.

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

[8] Darusman, Zulfahrizal, Y. Yunus, and A. A. Munawar, “Soil quality assessment by near infrared spectroscopy: Predicting pH and soil organic carbon,” Int. J. Sci. Technol. Res., vol. 8, no. 10, pp. 2512–2516, 2019.

[9] D. Devianti, S. Sufardi, Z. Zulfahrizal, and A. A. Munawar, “Rapid and Simultaneous Detection of Hazardous Heavy Metals Contamination in Agricultural Soil Using Infrared Reflectance Spectroscopy,” in IOP Conference Series: Materials Science and Engineering, 2019, vol. 506, no. 1.

[10] R. Hayati, A. A. Munawar, and F. Fachruddin, “Enhanced near infrared spectral data to improve prediction accuracy in determining quality parameters of intact mango,” Data Br., p. 105571, Apr. 2020.

[11] A. 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, pp. 208–215, Jul. 2016.

[12] A. A. Munawar, Kusumiyati, and D. Wahyuni, “Near infrared spectroscopic data for rapid and simultaneous prediction of quality attributes in intact mango fruits,” Data Br., vol. 27, Dec. 2019.

[13] A. A. Munawar, 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.

[14] A. 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, pp. 219–222.

[15] C. D. Iskandar, Zainuddin, and A. A. Munawar, “Beef Freezing Optimization by Means of Planck Model Through Simulation,” in IOP Conference Series: Earth and Environmental Science, 2019, vol. 365, no. 1.

[16] Samadi, S. Wajizah, and A. A. Munawar, “Fast and simultaneous prediction of animal feed nutritive values using near infrared reflectance spectroscopy,” in IOP Conference Series: Earth and Environmental Science, 2018, vol. 122, no. 1.

[17] Samadi, S. Wajizah, and A. A. Munawar, “Near infrared spectroscopy (NIRS) data analysis for a rapid and simultaneous prediction of feed nutritive parameters,” Data Br., vol. 29, Apr. 2020.

[18] A. A. Munawar, H. Syah, and Yusmanizar, “Fast and robust quality assessment of honeys using near infrared spectroscopy,” in IOP Conference Series: Earth and Environmental Science, 2019, vol. 365, no. 1.

[19] Y. Saputri, Yusriana, and A. A. Munawar, “Infrared spectroscopic features of turmeric powder,” in IOP Conference Series: Earth and Environmental Science, 2019, vol. 365, no. 1.
[20] R. Agustina and A. A. Munawar, “Electro-Optic Properties of Dried Pliek U Powder: Local Ingredients from Aceh,” in *IOP Conference Series: Earth and Environmental Science*, 2019, vol. 365, no. 1.

[21] Yusmanizar et al., “Fast and Non-Destructive Prediction of Moisture Content and Chologenic Acid of Intact Coffee Beans Using Near Infrared Reflectance Spectroscopy,” in *IOP Conference Series: Materials Science and Engineering*, 2019, vol. 506, no. 1.

[22] Sudarjat, Kusumiyati, Hasanuddin, and A. A. Munawar, “Rapid and non-destructive detection of insect infestations on intact mango by means of near infrared spectroscopy,” in *IOP Conference Series: Earth and Environmental Science*, 2019, vol. 365, no. 1.

[23] Agussabti, Rahmadiansyah, P. Satriyo, and A. A. Munawar, “Data analysis on near infrared spectroscopy as a part of technology adoption for cocoa farmer in Aceh Province, Indonesia,” *Data Br.*, vol. 29, Apr. 2020.

[24] E. S. Mohamed, A. M. Saleh, A. B. Belal, and A. A. Gad, “Application of near-infrared reflectance for quantitative assessment of soil properties,” *Egypt. J. Remote Sens. Sp. Sci.*, vol. 21, no. 1, pp. 1–14, 2018.

[25] H. Cheng et al., “Estimating heavy metal concentrations in suburban soils with reflectance spectroscopy,” *Geoderma*, vol. 336, no. August 2018, pp. 59–67, 2019.

[26] S. Chakraborty et al., “Rapid assessment of regional soil arsenic pollution risk via diffuse reflectance spectroscopy,” *Geoderma*, vol. 289, pp. 72–81, 2017.

[27] D. Miranda Salazar, H. L. Martínez Reyes, M. E. Martínez-Rosas, M. M. Miranda Velasco, and E. Arroyo Ortega, “Visible-near infrared spectroscopy to assess soil contaminated with cobalt,” *Procedia Eng.*, vol. 35, pp. 245–253, 2012.

**Acknowledgment**

The authors would like to thank directorate for research and community services DRPM Kemristek Dikti and LPPM Syiah Kuala University for funding support through research grant PD and PP.