

A Portable *in-situ* Near-infrared LEDs-based Soil Nitrogen Sensor Using Artificial Neural Network

Nur Aisyah Syafinaz Suarin¹, Kim Seng Chia^{2,*}, Siti Fatimah Zaharah Mohamad Fuzi³

¹Faculty of Electrical and Electronic Engineering, Universiti Malaysia Pahang, Pekan, Pahang, Malaysia

²Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, Malaysia

³Faculty of Applied Science and Technology, Universiti Tun Hussein Onn Malaysia Pagoh Campus, Muar, Johor

²Sensor Devices Research Laboratory, MiNT-SRC, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Johor, Malaysia

Received 7 March 2018; accepted 6 June 2018, available online 6 August 2018

Abstract: Monitoring soil Nitrogen content for palm oil cultivation is paramount to produce high-quality palm oil. This study aims to investigate the feasibility of a proposed portable near-infrared (NIR) light emitting diodes (LEDs)-based soil Nitrogen sensor in predicting the soil Nitrogen content using artificial neural network (ANN). First, soil samples that collected from a local oil palm plantation were scanned using the developed sensor and then followed by a conventional method, i.e. Kjeldahl analysis to measure the actual soil Nitrogen content. ANN was used for Chemometric analysis to develop a predictive model to *in-situ* predict the soil Nitrogen content using the near infrared light. The performance of ANN was validated using leave one out cross-validation. Results indicate that ANN with one hundred hidden neurons achieved the best accuracy with a root mean square error of cross-validation of 0.031%. This finding suggests that the proposed portable sensor coupled with ANN is promising to satisfactorily predict soil Nitrogen content.

Keywords: Near-infrared, Sensor, LED, Artificial Neural Network, Nitrogen, Soil,

1. Introduction

The success of oil palm cultivation is determined by soil conditions and water availability which includes physical, biological, and chemical properties of soil [1]. Information that can be obtained by analyzing soil content including organic matter, mineral composition, humidity level, acidity level, and nutrients. Palm oil trees need nitrogen, phosphorus, and potassium in large quantities [1]. Monitoring soil nitrogen content for palm oil cultivation is paramount to produce high-quality palm oil. This is because nitrogen is the most important nutrient element that will determine the level of soil fertility for oil palm cultivation. Fertilizers and decomposed materials are nitrogen sources in soil. It is worthy to highlight that different plants required different level of Nitrogen content to grow [2].

Nutrient uptake by oil palm is varied as the palm developed from seedling to an adult palm [3]. Nitrogen is important to palm oil plant for tissue development. Deficiency in nitrogen is usually associated with topsoil erosion [1]. It will first appears as a discoloration of young oil palm fronds, which lose their healthy dark green color and turn into yellow i.e. chlorosis [4]. Consequently, the plant might be unable to conduct biological process efficiently, which may lead to a reduction of bunch size as well as number of effective fruit bunches. Since the healthiness of the oil palm plant

would affect the quality of oil, it is vital to ensure the soil Nitrogen level is sufficient. On the other hand, it is worth to highlight that the excessive soil Nitrogen content might cause pollution to the environment e.g. river pollution [1]. The excessive soil Nitrogen content is uneasy to be observed and it could affect the environment in a long run. Imbalance nutrient can negatively affect plant [5]. Maintaining a balanced nutrition has an important role in determining plant resistance or susceptibility to disease. Therefore, monitoring the soil Nitrogen level is crucial to achieve a good fertilizing schedule. In fact, adding fertilizers at an unnecessary time might lead to the excessive soil Nitrogen content which could cause water pollution and make plants more vulnerable to pest if rapid growth occurs. Additionally, nitrogen pollution fuels the growth of harmful algal which could devastate aquatic ecosystems.

Soil Nitrogen content testing is an excellent method for estimating the fertility status of a soil, and it provides valuable information for developing a good fertility management program. The conventional method of Kjeldahl analysis that used in monitoring and surveillance of soil Nitrogen content often carried out with extra work field for sampling and processing [6]. Normally, the time taken for the whole process from collecting soil samples, transporting the samples to the lab, until getting the result would consume 1-2 weeks. Apparently, this method is time-consuming for regular monitoring. Besides, as the

*Corresponding author: kschia@uthm.edu.my
2018 UTHM Publisher. All right reserved.
penerbit.uthm.edu.my/ojs/index.php/ijie

procedure required the use of laboratory tools and reagent, extra care is needed to handle for Kjeldahl analysis [1]. Cost for soil testing in the laboratory depends on the type of fertility test per sample. In order to obtain information regarding soil nutrients, a lot of soil samples from different locations in the same field were required to be tested. This approach is time-consuming, costly, and produce chemical wastes that are harmful to the environment. Due to the soil Nitrogen content level cannot be visually determined, there is a need to investigate an alternative that is simple, fast, and environmental friendly to measure soil Nitrogen content.

Zhang et.al developed an automatic fluidic system for the rapid detection of soil nutrients [7]. Micro pump and valves controlled by a laptop computer delivered a sequence of test solutions to a test chamber where ion-selective electrodes are mounted. Ion-selective electrodes (ISE) were employed as the soil nutrient detection. This device showed an acceptable result for the soil nutrient detection. However, the cost of this system is high. This is because a computer is needed to process the data, a mechanical system is needed control valve, and the user is needed to be well trained. This could be a barrier to be a solution that benefits individual farmers who have limited resource and time to familiar the complex procedure.

In recent decades, spectroscopic technology has widely been used in various fields which involve in chemistry such as agriculture, medical and food industry. Near-infrared (NIR) spectroscopy is a spectroscopic method which uses the near infrared radiation energy to interact with the sample of interest [9]. Electromagnetic spectrum consists of several spectrums based on wavelength. The wavelength of NIR spectrum is range from 750nm to 1000nm. NIR energy typically can penetrate much farther into sample compare with mid-infrared radiation [10]. This is due to the energy of near-infrared is higher and the wavelength is shorter than mid-infrared. Near-infrared spectroscopy is a reliable method, fast, and non-destructive measurement technique. NIR technology has been widely studied in various areas e.g. food technology by predicting total volatile basic nitrogen (TVB-N) in pork meat [10], predicting diesel fuel properties [11], the internal quality of pineapples [12], and glucose level [13]. However, it is worthy to highlight that near-infrared energy is only able to be absorbed and to vibrate molecules that contain O-H, C-H, and N-H bonds.

The NIR spectroscopic technique is a green technology. However, a commercial NIR spectroscopy is expensive and bulky in size. Thus, an alternative that

applied spectroscopic method yet portable and low cost is needed to unleash the potential of this technology further. A recent investigation indicates that the NIR light emitting diodes (LEDs) based sensor is capable of predicting the glucose level using much affordable setup compared with a conventional NIR spectroscopy [13]. Thus, the objectives of this study are to develop a portable soil Nitrogen sensor using NIR LEDs and artificial neural network, to establish the relationship between soil Nitrogen content and the transmitted NIR light, and to evaluate the predictive accuracy of the proposed sensor using leave one out cross-validation (LOOCV).

2. Materials and methods

In this paper, the flow of the development proposed sensor is presented. There are three main phases in this study that initiate with the development of proposed sensor (Section 2.1), followed by the data acquisition phase (Section 2.2), and then ended with the data processing and modelling (Section 2.3). The result of artificial neural network (ANN) with the different hidden neurons number have been compared and discussed based on the root mean square error (RMSE), the correlation coefficient (R-value) and the root mean square error cross-validation (RMSECV). Leave one out cross-validation (LOOCV) method has been used to validate the performance of the predictive models due to the low number of data used in this study. The best predictive model is selected from the analysis and applied in the developed sensor to predict independent samples of soil Nitrogen content.

2.1 Development of proposed sensor - NIR LEDs-based Soil Nitrogen sensor

The spectroscopic technique is applied in this project to develop the proposed sensor due to its ability to measure organic content in the sample non-invasively. The proposed sensor was developed by using near infrared LEDs which emit electromagnetic radiation with wavelength range of 700 nm to 1000 nm. According to Beer-Lambert's Law, the absorbance of light is directly proportional to the thickness of the media through which the light is being transmitted. The design of proposed sensor adapts Beer-Lambert's Law [14] based on the transmission of near-infrared light through a soild sample, and then a light sensing element was used to measure the light intensity in linear position as illustrated in Figure 1.

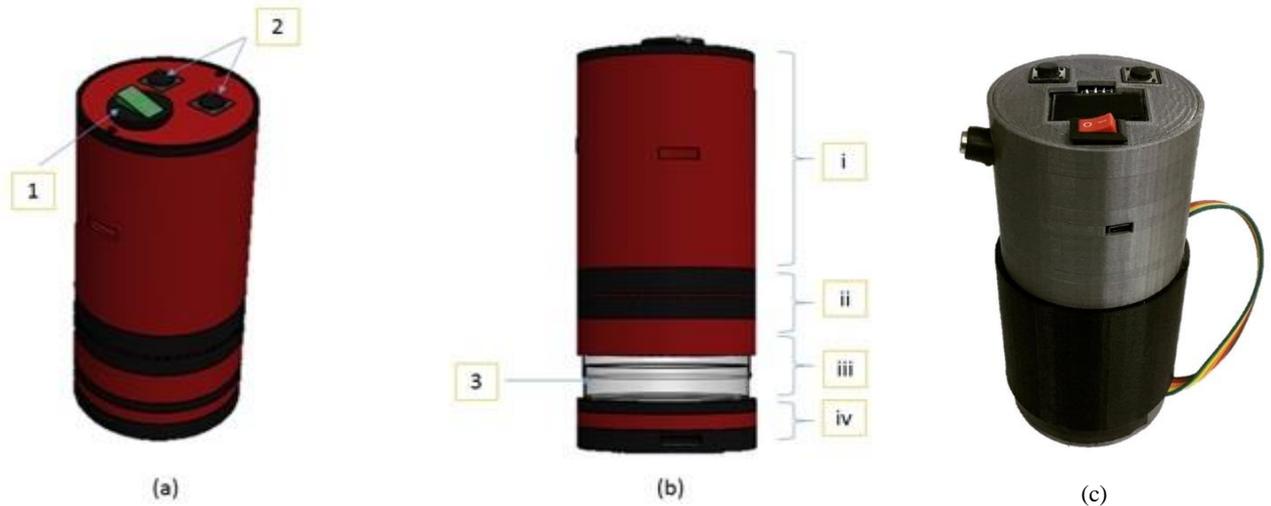


Fig. 1: The design of the developed portable soil Nitrogen sensor: (a) perspective view, (b) front view, and (c) proposed sensor after fabrication process: (1) switch, (2) push buttons, (3), petri dish, (i) main circuit compartment, (ii) NIR LEDs compartment, (iii) soil sample compartment, (iv) light sensing element circuit compartment.

Figure 1 (a) and (b) show the design of the proposed hardware. A cylinder shape is chosen as a basic shape to design and fabricate the hardware. The sequence of the compartment is decided based on their function and priority. Figure 1 (c) shows the image of the hardware that produced by the three dimensional (3D) printing technology. The design of the hardware consists of several parts which required a custom made fabrication so that the sensor is able to properly function.

The developed hardware consists of three (3) different circuits to suit the design of the hardware. They are the main controller circuit, the NIR LEDs circuit, and light sensing element circuit. The printed circuit board (PCB) layout for circuit design is shown in Figure 2. The light sensing element circuit and NIR LEDs circuit were controlled by the main circuit which consists of a micro-controller.

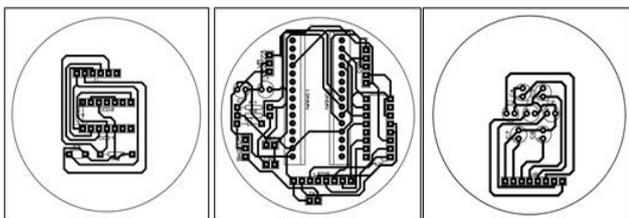


Fig. 2: PCB Layouts: (i) CCD circuit, (ii) Main Circuit, and (iii) NIR LEDs Circuit.

2.2 Soil sampling

A total of five (5) soil samples used in this project that were collected from the area of two acres oil palm field coordinate N 1 50 3.984 E 103 8'37.5936. The soil sample was extracted by using an augmented soil tool with a length of 15 cm and a diameter of 10 cm. This soil tool was used to extract the targeted soil samples that

were located 10 cm depth from the earth surface. The highest Nitrogen content could be detected in depth 0-30 cm [16]. Thus, we decided to extract soil that was 10 cm depth from the earth surface. Nitrogen content level depends on the location of the soil. For instance, the Nitrogen content might be lower if the soil is extracted from the area which have a lot of plants planted in it. The samples were stored in a plastic container and kept in a box before the procedure of scanning measurement and reference measurement was conducted. The experiment was conducted at Biochemical Laboratory, Faculty of Science, Technology and Human Development, UTHM.

2.3 Proposed measurement

Spectra acquisition is the process where the Nitrogen contents of soil samples were scanned by the developed proposed sensor. There were four (4) steps: First, putting the soil sample from the container into a petri dish. Second, putting the petri dish on the weighing scale to ensure that the weight of soil in the petri dish is 25g. Third, the soil sample was compressed in the petri dish to ensure the thickness of the soil in the petri dish did not exceed 5 mm. The thickness of soil in petri dish was decided based on the design of the proposed sensor as shown in Figure 1. The distance from the near infrared LEDs and the sensing element were able to give a stable reading when the thickness was at 5 mm. When the thickness was more than 5mm, the unwanted signals were increased. On the other hand, when the thickness was lower than 5 mm, the soil did not fully fill in the petri dish. In other words, the reading might not be the NIR light that penetrated the soil. After the NIR signals acquisition was completed, the soil samples were transferred from the petri dish into a test tube to conduct the conventional soil Nitrogen measurement by using Kjeldahl analysis experiment.

2.4 Reference measurement - Kjeldahl Analysis

Kjeldahl analysis experiment was chosen as the reference measurement for this research due to the availability of the instrument in our university. Besides, Kjeldahl analysis is a reliable conventional Nitrogen measurement approach and is always been used to measure the Nitrogen content of various samples. Kjeldahl analysis consists of three (3) steps, i.e. digestion, distillation, and titration process. Sample preparation required to dilute 5g of soil sample in a 35 ml of distilled water in a test tube. Next, two tablets of anti-foam and 12 ml of sulphuric acid were added in the test tube before start the digestion process. For distillation process, 35% of sulphuric acid was prepared. This required 525g of sulphuric acid to be diluted in 1.5 liter distilled water. Titration process required the preparation for boric acid solution, sulphuric acid, and scrubber solution. Boric acid solution is made from Bromocresol Green, ethanol, Methyl Red and Boric acid. A lot of chemical reagents were used in Kjeldahl analysis which needed precaution during the sample preparation and the correct steps for the waste disposal. The reference measurement is important to obtain the actual Nitrogen content that would be used as the target for building a prediction model from the seven input parameters of NIR LEDs. The Kjeldahl analysis is one of the reference methods that can be used to test the validity of proposed sensor, however, there are other technologies can also be used to prove the reliability of the sensing element.

2.5 Pre-processing data

First, the acquired transmission intensity was converted into absorbance value, A, according to Lambert's Law as stated in Equation (1). Next, five samples were divided into five sets consist of four samples as training set and one sample for validation. Each set has a different sample for validation as shown in Table 1. In future, more soil samples can be used to perform better prediction model for the proposed sensor.

$$A = -\log (T) \tag{1}$$

Table 1: Data arrangement for leave-one-out cross validation

Number of set for LOOCV	Soil sample for validation
1 st set (Soil 2, 3, 4, 5)	Soil 1
2 nd set (Soil 1, 3, 4, 5)	Soil 2
3 rd set (Soil 1, 2, 4, 5)	Soil 3
4 th set (Soil 1, 2, 3, 5)	Soil 4
5 th set (Soil 1, 2, 3, 4)	Soil 5

2.6 Data processing - Chemometric analysis by using Artificial Neural Network

The performance of the prediction modeling was determined by the root mean square error (RMSE) and coefficient correlation (R-value). Equation 2 shows the formula to calculate R-value and Equation 3 shows the formula to calculate RMSE. The root mean square error cross validation (RMSECV) equation is shown in Equation 4. The *n* is the number of spectra, *y* is the actual value, *y_{cal}* denotes the calculated value, *y_{mean}* denotes the mean value, and *y_{pred}* is the predicted value of the corresponding parameters. Different hidden neurons were used to test five (5) sets of training and validation data in order to obtain the best result of predictive models. R value was used to determine the correlation between predicted and actual value for prediction analysis, RMSE value was calculated and recorded in order to know the root mean square error between actual and predicted value. Lastly, RMSECV value is important when cross validation analysis is performed in data processing to validate the result.

$$R = \sqrt{(1 - \frac{\sum(y_{cal} - y)^2}{\sum(y_{cal} - y_{mean})^2})} \tag{2}$$

$$RMSE = \sqrt{(1/n \sum(y_{cal} - y)^2)} \tag{3}$$

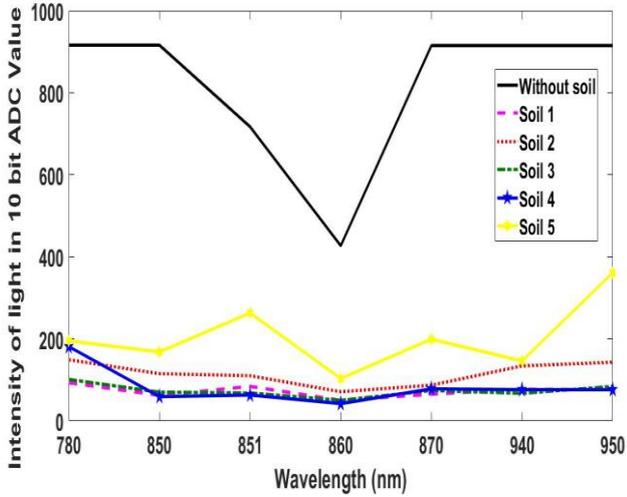
$$RMSECV = \sqrt{(1/n \sum(y_{pred} - y)^2)} \tag{4}$$

3. Results and Discussion

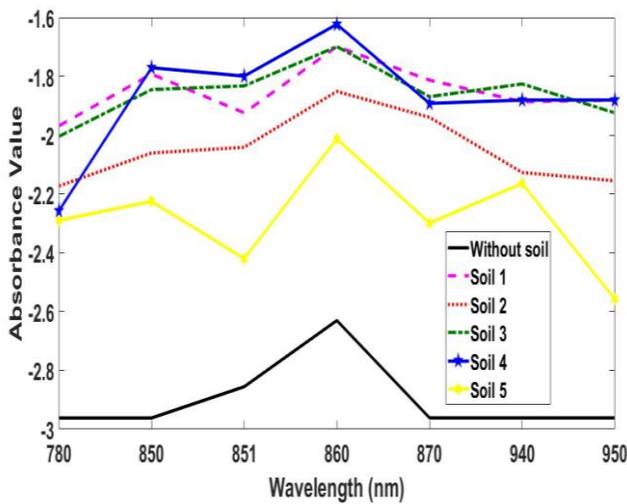
Figure 3 (a) shows the acquired data after the near infrared energy transmitted through each soil samples from each NIR LEDs. Line graph labelling without soil indicates that the scanning measurement for empty petri dish which is 100% transmission. Soil 1, soil 2, soil 3, soil 4, and soil 5 were the scanning measurements for 5 soil samples with different soil Nitrogen values which were validated using Kjeldahl analysis experiment after measurement by using proposed sensor. The different intensity of light that acquired from an empty petri dish compared to the petri dishes that contained soil samples indicates that infrared energy was absorbed by the soil sample. The near infrared energy will be absorbed by the chemical content of the soil sample to release photon energy which causes vibration of molecules and is suitable for measuring the components of interest that consists of either carbon-hydrogen (C-H), oxygen-hydrogen (O-H) or Nitrogen-Hydrogen (N-H) [14]. The higher the content of N-H bond in the soil sample, the higher the absorbance of near infrared energy that vibrates the molecules. Consequently, the lower the remaining near infrared energy that passed through the soil sample.

On the other hand, Figure 3 (b) shows the absorbance for the five soil samples. Higher transmission energy was measured for empty petri dish using the developed soil Nitrogen sensor. The absorbance value was obtained

using Equation (1). Absorbance is inversely proportional with the transmittance i.e. when the Nitrogen content is higher, the transmittance of the light is lower. This is illustrated in Figure 3 (a) in which the transmittance energy without soil is the highest, while the absorbance values without soil is the lowest in the Figure 3 (b).



(a)



(b)

Fig. 3: The transmission intensity of NIR energies: (a) in 10 bit ADC Value, and (b) in the absorbance value.

As the absorption of near infrared energy by the soil higher, the absorbance value is bigger. The difference between each line illustrates that each soil sample consists of different content of Nitrogen soil content which absorbed different amount of energy from near infrared light.

The number of hidden neurons would affect the performance of the ANN model. Figure 5 shows the result of different hidden neurons for modeling with their corresponding RMSE for testing. ANN with 100 hidden

neurons achieved the best result among the rest with the lowest value of RMSE and the highest value of R-value. During the training, Levenberg-Marquardt algorithm was selected to perform the training. ANN with 30 hidden neurons showed the highest value of RMSE. As the value of RMSE smaller, the better the performance of predictive model. Percentage of root mean square error for cross validation (RMSECV) was used in this study to validate the performance of the predictive model due to the small number of the available soil samples.

There were five training sets used in this study with different combination of testing and training sample as shown in Table 1. The number of hidden neuron plays a vital role to ensure the accuracy of predictive model and complexity to processing the data. The training can be continue with more hidden neurons but the ANN with 100 hidden neurons had achieved a satisfied accuracy.

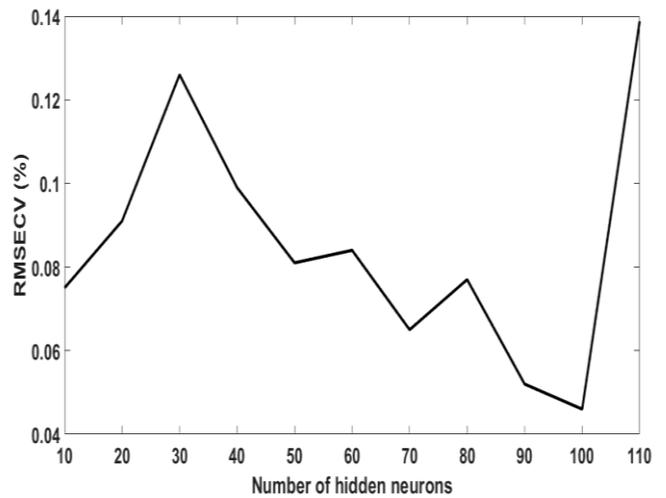


Fig. 4: The root mean squared errors of cross-validation (RMSECV) against the number of hidden neurons.

Table 2 tabulates the result of cross-validation performance of Artificial Neural Network (ANN) for different hidden neurons. Result shows that ANN model was able to produce good result by giving out low difference value of measured and prediction values, well-predicted for total five testing soils for five leave-one-out cross-validation models. By applying ANN model for 100 hidden neurons, the validation result was the best and was better compared with model of 90 hidden neurons and 110 hidden neurons. ANN with 110 hidden neurons produced over-fitted result. ANN with 90 hidden neurons, on the other hand, was under-fitted. R-Value was analysed to determine the correlation between actual value and predicted value of model. The higher the number of R-Value, the closer the different between actual and predicted value. Average root mean square error for cross validation (RMSECV) for 100 hidden neurons of ANN was 0.031 and the lowest. Thus, this shows that ANN model with 100 hidden neurons was able to produce well-predicted model for near-infrared

LED-based sensor to predict Nitrogen content in soil. Nevertheless, the accuracy of the ANN may be negatively affected by the environment conditions e.g. temperature and humidity. In fact, near infrared light is susceptible temperature, and thus the use of temperature control system should be considered to eliminate the influence of temperature [15].

Table 2: Leave-one-out cross validation (LOOCV) for ANN analysis with different hidden neurons (HNs)

Model (HNs)	Cross Validation	R-Value	RMSE	RMSECV (%)
ANN (90)	1	0.981	0.009	0.052
	2	0.962	0.005	
	3	0.887	0.017	
	4	0.792	0.036	
	5	0.812	0.060	
ANN (100)	1	0.997	0.005	0.046
	2	0.985	0.023	
	3	0.912	0.021	
	4	0.989	0.032	
	5	0.946	0.029	
ANN (110)	1	0.856	0.022	0.139
	2	0.879	0.035	
	3	0.889	0.014	
	4	0.901	0.042	
	5	0.817	0.037	

4. Conclusion

This study demonstrates that the proposed sensor is promising in rapidly predicting soil Nitrogen content using near infrared (NIR) spectroscopic technique and Artificial Neural Network (ANN) without dumping any chemical waste to environment. Results indicate that near infrared energy will be partially absorbed by soil sample according to the Nitrogen content in the soil sample. ANN with 100 hidden neurons achieved best performance with correlation coefficient of 0.966 and root mean square error of prediction (RMSEP) of 0.031 %. This shows that there is a strong relationship between NIR energy and soil Nitrogen content in soil.

Acknowledgement

This study was sponsored by Final Year Project Fund provided by the Faculty of Electric and Electronic (FKEE) Universiti Tun Hussein Onn Malaysia. The authors would like to acknowledge Advanced Mechatronic Research Group (AdMiRe), Faculty of Electrical and Electronic Engineering, Faculty of Science, Technology and Human Development (FSTPi) UTHM for providing facilities for this study; En. Ishak bin Ayob, Assistant Science Officer for guiding in Kjeldahl analysis experiment, TechSource Malaysia for providing Matlab software, Sime Darby and Innovate Malaysia Design Competition 2017 for inputs; Tan Siew Chin and Tan Yit Peng for contribution in assist with this study.

References

- [1] H. Mohidin and M. M. Hanafi, "Bragantia Determination of optimum levels of nitrogen , phosphorus and potassium of oil palm seedlings in solution culture," pp. 1–11, 2016.
- [2] "Sustainability and Responsible Production of Malaysian Palm Oil – Sime Darby's Perspective IPOSC 2014, 19-20 August 2014 Tang Men Kon," no. August, pp. 19–20, 2014.
- [3] M. Read, "Nutrients for oil palm : A question of balance," pp. 12–14, 2016.
- [4] B. I. Rankine and T. H. Fairhurst, "Management of Phosphorus , Potassium and Magnesium in Mature Oil Palm," vol. 13, no. 1, pp. 10–15, 1999.
- [5] F. Guan, M. Xia, X. Tang, and S. Fan, "Catena Spatial variability of soil nitrogen , phosphorus and potassium contents in Moso bamboo forests in Yong an City , China," *Catena*, vol. 150, pp. 161–172, 2017.
- [6] W. Baethgen, M. Alley, "A manual colorimetric procedure for measuring ammonium nitrogen in soil and plant kjeldahl digests" vol. 20, pp. 961-969, 1989.
- [7] M. Zhang, M. Wang, L. Chen, S. S. Ang, and C. V Nguyen, "An Automatic Fluidic System for the Rapid Detection of Soil Nutrients," in Int. Conference on Automation and Logistic China, no. September, pp. 2742–2746, 2008.
- [8] Y. He, M. Huang, and A. Garc, "Prediction of soil macronutrients content using near-infrared spectroscopy," vol. 58, pp. 144–153, 2007.
- [9] W. Măntele and E. Deniz, "Spectrochimica Acta Part A : Molecular and Biomolecular Spectroscopy UV – VIS absorption spectroscopy : Lambert-Beer reloaded," *SAA*, vol. 173, pp. 965–968, 2017.
- [10] L. Huang, J. Zhao, Q. Chen, and Y. Zhang, "Nondestructive measurement of total volatile basic nitrogen (TVB-N) in pork meat by integrating near infrared spectroscopy, computer vision and electronic nose techniques," *Food Chem.*, vol. 145, pp. 228–236, 2014.
- [11] H. A. G. Al-kaf, K. S. Chia, and N. A. M. Alduais, "A comparison between single layer and multilayer artificial neural networks in predicting diesel fuel properties using near infrared spectrum," *Pet. Sci. Technol.*, vol. 36, no. 6, pp. 411–418, Mar. 2018.
- [12] M. N. H. Jam and K. S. Chia, "Investigating the relationship between the reflected near infrared light and the internal quality of pineapples using neural network," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 7, no. 4, 2017.

- [13] K. S. Chia, N. A. S. Suarin, and S. F. Z. M. Fuzi, "Prediction of glucose concentration using near infrared light and adaptive linear neuron," in *2017 IEEE 13th International Colloquium on Signal Processing & its Applications (CSPA)*, 2017, no. March, pp. 35–38.
- [14] M. Zude, M. Pflanz, L. Spinelli, C. Dosche, and A. Torricelli, "Non-destructive analysis of anthocyanins in cherries by means of Lambert – Beer and multivariate regression based on spectroscopy and scatter correction using time-resolved analysis," *J. Food Eng.*, vol. 103, no. 1, pp. 68–75, 2011.
- [15] A. A. Aminondin and K. S. Chia, "A temperature control system for near infrared spectroscopic analysis using proportional controller," *Int. J. Integr. Eng.*, vol. 9, no. 3, pp. 24–28, 2017.
- [16] Knight, S.M "Soil Mineral Nitrogen testing: Practice and intpretation," report project funded by The Home-Grown Cereals Authority (HGCA)