Soil Quality Assessment By Near Infrared Spectroscopy: Predicting Ph And Soil Organic Carbon

Darusman, Zulfahrizal, Yuswar Yunus, Agus Arip Munawar

Abstract: The main objective of this present study is to apply near infrared reflectance spectroscopy (NIRS) combined with multivariate analysis in predicting soil quality attributes rapidly and simultaneously. Those quality attributes are: pH and soil carbon organic (SOC). Near infrared spectra data, in form of absorbance spectrum were acquired for a total of 10 bulk soil samples amounted 55 g per bulk. Spectra data were acquired and recorded in wavenumbers range from 4000 to 10 000 cm^{-1} with number of scans is set to 64 scans respectively. On the other hand, actual soil quality attributes (pH and SOC) were measured by means of standard laboratory methods. Prediction models were developed and established using principal component regression (PCR) and partial least square regression (PLS) followed by leave one out cross validation (LOOCV) method. The results showed that all quality attributes can be predicted using NIRS in tandem with PCR and PLS with maximum correlation coefficient between predicted and measured parameters are: 0.97 for both pH and SOC quality parameters. Moreover, robustness index for pH and SOC prediction are 4.60 and 4.67 respectively, which referred as excellent prediction performance. It may conclude that NIRS can be used to assess soil quality attributes rapidly and simultaneously.

Keywords: NIRS, soil, pH, SOC, agriculture.

1. INTRODUCTION

As we know that soil is one of the most important media and source for plants to grow and develop. In principle, plants can grow optimally when soil is in good and healthy condition. Soil is characterized by its physical and chemical properties from which defines soil fertility, textures and inner qualities [1]. Soil quality attributes is highly related with the micro and macro nutrients on soil such as N, P, K, pH, carbon organic, minerals and free from heavy metal contaminations [2]–[4]. Each phase of growing plants required different soil nutrients from which affect total yield and whole quality parameters of plants and agricultural products [5]. In order to provide sufficient nutrients on soil, fertilization practices are common to be employed and applied in many nations worldwide, especially in developing countries [6], [7]. To date, it fertilization and irrigation management practices relies on the basic need of plants for the related nutrients required. However, the amount of nutrient on agricultural soil is hard to know rapidly in such instant time. Generally, to determine soil quality parameters such as micro and macro nutrients, several methods were already employed and applied widely. Nonetheless, most of those practical methods are based on solvent extractions from which require specific instrument for each quality attributes to be measured [8]–[10]. Moreover, these kind of measurements are normally time consuming, laborious, complicated sample preparations and involved chemical materials [11].

- *Department of Soil Science, Syiah Kuala University, Banda Aceh – Indonesia*
- *Department of Agricultural Engineering, Syiah Kuala University, Banda Aceh – Indonesia Corresponding author: aamunawar@unsyiah.ac.id*

Recently, in agriculture era 4.0, advanced technology had been widely studied, particularly related to the precision farming practices. One of the most promising technology to be utilized in agriculture is near infrared reflectance spectroscopy (NIRS). This technology works based on the phenomena of the interaction between light radiation and biological objects [11]–[13]. Near infrared cover a wavenumber range from 4000 to 10 000 cm-1 or 1000 to 2500 nm. It is slightly above visible radiation and short wave near infrared from 400 to 900 nm [14], [15]. In principle, when the light goes through the biological objects, there three main reactions occurred namely reflectance, absorbance and transmittance. These reactions had been studied to reveal chemical properties inside those biological materials [16], [17]. Respective information to be revealed are mainly related to the inner quality attributes of those materials. The main advantage of NIRS method are: simple sample preparation, pollution free, without involving chemical materials, non-destructive, no chemical waste and thus environmental friendly to be used. NIRS have been widely studied and applied in many sectors, including agriculture [12], [18], [19]. Numerous studies and publications are reported on the application of NIRS as non-destructive and fast method in determining several quality attributes of foods and agricultural products such as horticulture [13], [20], [21], cocoa [22], [22], [23], coffee [24]–[27], milk and dairy products [28]–[30], meat [31]–[33], animal feed [34], [35] and food authentications [36]–[38]. The increasing numbers of research and publication on NIRS application indicates that NIRS technology are feasible and potential to be employed in precision agriculture practices [39]. Our own recent studies on NIRS applications also showed that spectral based technology like near infrared is provided promising accuracy on quality attributes prediction of several agricultural products. As NIRS itself can not reveal related information on the buried spectra data, multivariate analysis is therefore required in combination with near infrared spectra data to determine related quality attribute information. Multivariate analysis is a specific method on statistics that used to analyze multivariate data through calibration models development. Therefore, the main objective of this present study is to employ near infrared reflectance spectroscopy (NIRS) in tandem with multivariate analysis as a rapid and simultaneous method in predicting and determining soil quality attributes in form of pH and soil organic carbon (SOC) on agricultural soil samples with different purposes.

2. MATERIALS AND METHODS

2.1 Soil samples

In this study, soil samples were collected from three different land use, rice field, bare land and orchards. Soil samples were obtained in Aceh Besar district area in Aceh Province, Indonesia. Samples were taken to the lab, grinded and sorted for unwanted materials, such as rocks, insects and others. Soil samples were then stored in room temperature of 26^{o}C and relative humidity 86% for two days to equilibrate.

2.2 Spectra data acquisition

Near infrared spectra data in form of absorbance spectrum were acquired for soil samples using a self-developed infrared instrument (PSD FTNIRS i16) with support of number of scans set to 64. Spectra data were recorded in wavenumber range from 4000 to 10 000 cm^{-1} and optical gain of 8x. Spectra data were then saved in network drive in two different extension files.

2.3 Actual pH and SOC measurements

After spectra acquisition was completed, all soil samples were taken in the same day to measure soil quality attributes in form of pH and SOC. Actual referenced pH was determined by means of pH conductivity meter, while actual SOC was measured using elemental analyzer. Soil organic carbon on soil samples were expressed as % SOC. All Actual soil quality attributes were measured in triplicate and averaged.

2.4 Multivariate data analysis

Prediction models were developed in order to determine pH and SOC of agricultural soil samples. Principal component regression (PCR) and partial least square regression (PLS) were applied as multivariate analysis method. Leave one out cross validation (LOOCV) was then also used to quantify and evaluate prediction models. It is obvious that good, accurate and robust prediction models must generate and provide high correlation coefficient (r), low error (RMSE), and high residual predictive deviation (RPD).

3. RESULTS AND DISCUSSIONS

3.1 absorbance spectrum of soil samples

Absorbance spectral data in wavenumbers range from 4000 to 10 000 cm^{-1} for soil samples are presented in Fig.1. The presence of peaks and valleys represent the vibration of molecular bonds of C-C, O-H, N-H, C-H and organic minerals of soil samples due to electromagnetic radiation.

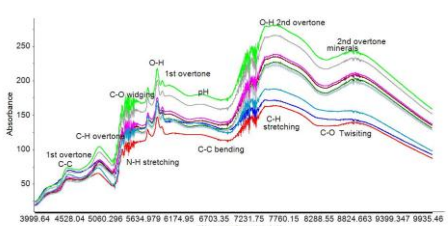

Fig. 1. Absorbance spectrum of soil samples.

Spectra data consists of different reactions of organic materials. As shown in Fig. 1, soil moisture contents can be predicted in wavenumbers 5839 cm^{-1} due to first overtone occurred, and also can predicted in 7664 cm^{-1} due 2nd overtone, since moisture content is constructed by O-H molecule. Moreover, bending and stretching also happened along the near infrared region due to C-H, C-C and N-H vibrations. Soil minerals can also be predicted using NIRS due to 1st overtone founded in 8847 cm^{-1}.

3.2 Predicted pH and SOC

The main core in NIRS application is to develop prediction models with support of multivariate analysis, used to predict soil quality attributes in form of pH and soil organic carbon (SOC). Those prediction models were established using two different regression approaches namely using principal component regression (PCR) and partial least square regression (PLS).Descriptive statistics data for actual referenced pH and SOC of soil samples is shown in Table 1. As shown in this table, the average measured pH and SOC of soil samples were 6.998 and 0.867 respectively.

Table 1. Descriptive statistics of actual pH and SOC

Statistical parameters	Soil quality attributes	
	pH	SOC
Mean	6.998	0.867
Max	7.36	1.85
Min	6.59	0.2
Range	0.77	1.65
Std. Deviation	0.232	0.567
Variance	0.054	0.322
RMS	7.001	1.020
Skewness	-0.282	0.431
Kurtosis	-0.174	-1.063
Median	7.005	0.725
Q1	6.895	0.4575
Q3	7.135	1.29

At first, soil quality attributes prediction models were developed using PCR regression method. The number of latent variables (LVs) required to establish both models were 8 and prediction performance of pH and SOC by means of PCR approach is presented in Table 2.

Table 2. Prediction performance of pH and SOC using PCR regression approach

Quality attributes	R^2	r	RMSE	RPD
pH	0.61	0.78	0.14	1.64
SOC	0.78	0.88	0.25	2.24

R: correlation coefficient, R^2: coefficient of determination, RMSE: root mean square error, RPD: residual predictive deviation, SOC: soil organic carbon.

Soil quality attributes, pH and SOC can be predicted by using NIRS combined with PCR approach with correlation coefficient maximum for pH is 0.78 and prediction error is 0.14. On the other hand, maximum correlation coefficient for SOC prediction using PCR is 0.88 and error is 0.25. In term of prediction robustness, pH can be predicted with sufficient performance since the RPD resulted is 1.68 whilst for SOC prediction, the RPD is 2.24 from categorized as good model performance. Based on literatures, 1 < RPD < 1.5 indicates coarse prediction performance and to be improved. The RPD between 1.5 < RPD < 2.0, indicates as sufficient performance while 2 < RPD < 2.5 refereed to good model performance. Moreover, excellent and robust prediction model performance is reached when the RPD index of predicted value is above 2.5. Scatter plot derived from actual versus predicted pH and SOC of soil samples using PCR regression approach is presented in Fig. 2 and Fig. 3.

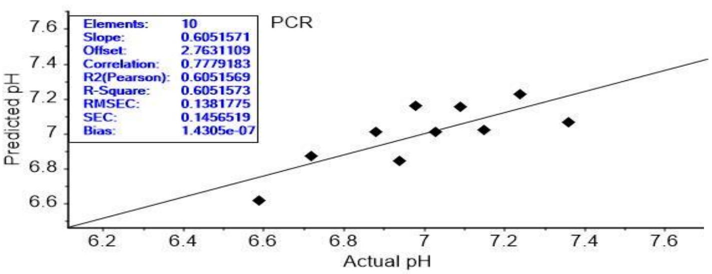

Fig. 2. *pH prediction performance using PCR regression approach.*

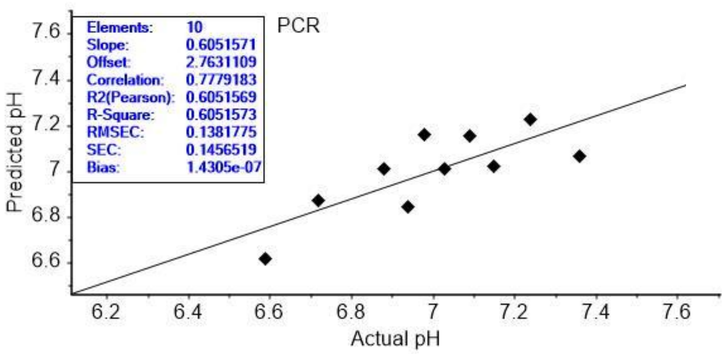

Fig. 3. *SOC prediction performance using PCR regression approach.*

approach. The correlation coefficient for pH prediction was improved to 0.97 and RPD index increased to 2.87. PLS approach performance in predicting pH and SOC is presented in Table 3.

Table 3. Prediction performance of pH and SOC using PLS regression approach

Quality attributes	R^2	r	RMSE	RPD
pH	0.94	0.97	0.05	4.60
SOC	0.94	0.97	0.12	4.67

R: correlation coefficient, R^2: coefficient of determination, RMSE: root mean square error, RPD: residual predictive deviation, SOC: soil organic carbon.

On the other hand, SOC prediction was also slightly improved, when prediction model is developed by means of partial least square regression approach. Scatter plot derived from actual versus predicted pH and SOC of soil samples using PLS regression approach are presented in Fig. 4 and Fig. 5 below.

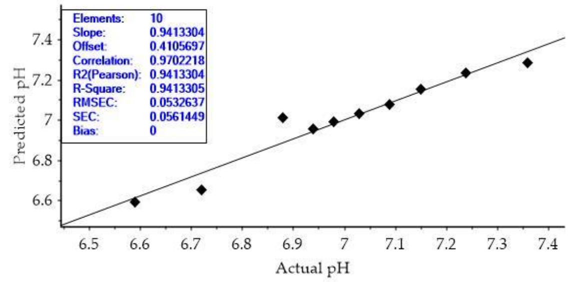

Fig. 4. *pH prediction performance using PLS regression approach.*

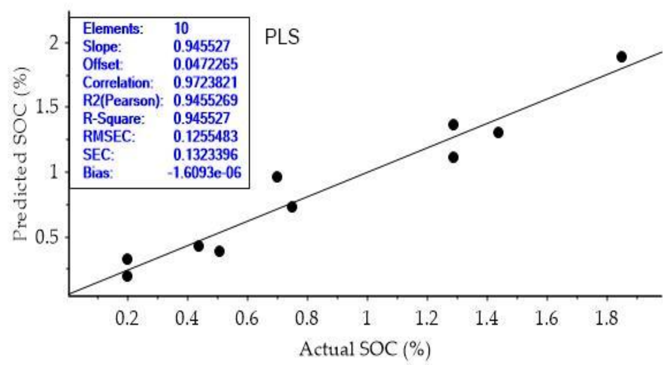

Fig. 5. *SOC prediction performance using PCR regression approach.*

Based on required latent variables point of view, the number of latent variables (LVs) required to establish and construct prediction models are 6 and all data generated good results with the number of LVs is less than 10. The number LVs often considered and taken into account to validate and evaluate the efficiency and 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. PLS found to be better in accuracy and robustness compared to PCR regression approach. Both methods work based on the multivariate data and subjected spectra data onto new map of latent variables. The only different between PLS and PCR is that in PLS, both independent and dependent variables were included in those map. The independent variables here acted is spectra data obtained from the spectra acquisitions, while dependent variable data is actual referenced soil quality attributes of soil samples namely pH and SOC in this case. The most optimum and relevant wavenumbers to predict pH and SOC is presented in Fig 6. As shown on this figure, the relevant wavenumbers are 5187 cm^{-1}, 5704 cm^{-1}, 5839 cm^{-1}, 7636 cm^{-1}, and 8651 cm^{-1} respectively.

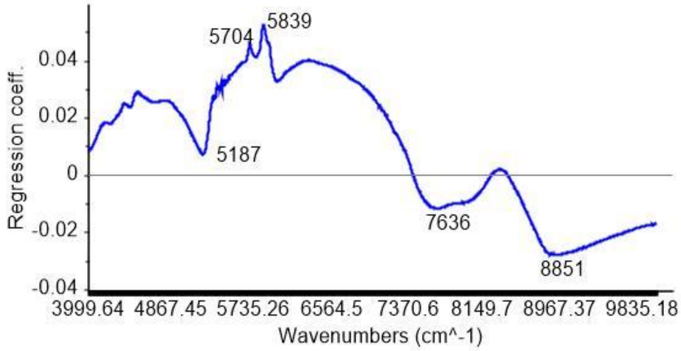

Fig. 6. Relevant wavenumbers to predict pH and SOC based on regression coefficient curve.

When the model is used to predict pH and SOC independently using LOOCV validation method, the most optimum LVs used for those prediction is 8 as presented in Fig. 7.

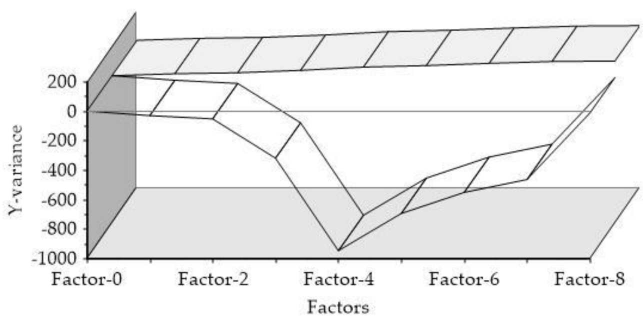

Fig. 7. Optimum number of LVs based on explained variance in LOOCV validation method.

Infrared spectral data recorded and acquired from the instrument sometimes may contain unwanted background information and noises due to light dispersion and scattering effect. These noises may reduce prediction accuracy and robustness in which interfered desired relevant soil quality attributes information. Noises and other interfered effects need to be removed for a better prediction performance. Thus, it is very necessary to pre-process infrared spectral data prior to further multivariate data analysis. In this study, we did not performed those spectra corrections, because the achieved performances are good and excellent enough to predict both pH and SOC of soil samples rapidly and simultaneously.

4. CONCLUSION

Based on prediction performance evaluation, it may conclude that NIRS technology combined with multivariate analysis was able to be applied in predicting soil quality attributes in form of pH and SOC rapidly and simultaneously. The maximum correlation coefficient for pH and SOC predictions were 0.97 using PLS approach. Moreover, the robustness index for pH is 4.40 and for SOC prediction is 4.67 which were categorized as excellent prediction performance.

REFERENCES

[1] A. Morellos et al., “Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy,” Biosyst. Eng., vol. 152, pp. 104–116, 2016.
[2] S. Chatterjee, N. Dey, and S. Sen, “Soil moisture quantity prediction using optimized neural supported model for sustainable agricultural applications,” Sustain. Comput. Informatics Syst., 2018.
[3] A. M. Mouazen, B. Kuang, J. De Baerdemaeker, and H. Ramon, “Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy,” Geoderma, vol. 158, no. 1–2, pp. 23–31, 2010.
[4] J. Moros, M. J. Martínez-Sánchez, C. Pérez-Sirvent, S. Garrigues, and M. de la Guardia, “Testing of the Region of Murcia soils by near infrared diffuse reflectance spectroscopy and chemometrics,” Talanta, vol. 78, no. 2, pp. 388–398, 2009.
[5] 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.
[6] C. Wang, T. Zhang, and X. Pan, “Potential of visible and near-infrared reflectance spectroscopy for the determination of rare earth elements in soil,” Geoderma, vol. 306, no. June, pp. 120–126, 2017.
[7] C. Vogel et al., “Combining Diffusive Gradients in Thin films (DGT) and Spectroscopic Techniques for the Determination of Phosphorus Species in Soils,” Anal. Chim. Acta, 2019.
[8] M. A. N. Coutinho, F. de O. Alari, M. M. C. Ferreira, and L. R. do Amaral, “Influence of soil sample preparation on the quantification of NPK content via spectroscopy,” Geoderma, vol. 338, no. July 2018, pp. 401–409, 2019.
[9] S. Nawar and A. M. Mouazen, “Predictive performance of mobile vis-near infrared spectroscopy for key soil properties at different geographical scales by using spiking and data mining techniques,” Catena, vol. 151, pp. 118–129, 2017.
[10] D. Aosai, Y. Yamamoto, T. Mizuno, T. Ishigami, and H. Matsuyama, “Size and composition analyses of colloids in deep granitic groundwater using microfiltration/ultrafiltration while maintaining in situ hydrochemical conditions,” Colloids Surfaces A Physicochem. Eng. Asp., vol. 461, no. 1, pp. 279–286, 2014.
[11] C. Pasquini, “Near infrared spectroscopy: A mature analytical technique with new perspectives – A review,” Anal. Chim. Acta, vol. 1026, pp. 8–36, 2018.
[12] 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.
[13] 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, 2016.
[14] M. Ruiz-Altisent et al., “Sensors for product

characterization and quality of specialty crops-A review," Comput. Electron. Agric., vol. 74, no. 2, pp. 176–194, 2010.
[15] 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.
[16] 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.
[17] H. Cen and Y. He, "Theory and application of near infrared reflectance spectroscopy in determination of food quality," Trends Food Sci. Technol., vol. 18, no. 2, pp. 72–83, 2007.
[18] T. Nordey, J. Joas, F. Davrieux, M. Chillet, and M. Léchaudel, "Robust NIRS models for non-destructive prediction of mango internal quality," Sci. Hortic. (Amsterdam)., vol. 216, pp. 51–57, 2017.
[19] F. Comino, V. Aranda, R. García-Ruiz, M. J. Ayora-Cañada, and A. Domínguez-Vidal, "Infrared spectroscopy as a tool for the assessment of soil biological quality in agricultural soils under contrasting management practices," Ecol. Indic., vol. 87, no. January, pp. 117–126, 2018.
[20] P. P. Subedi and K. B. Walsh, "Assessment of sugar and starch in intact banana and mango fruit by SWNIR spectroscopy," Postharvest Biol. Technol., vol. 62, no. 3, pp. 238–245, 2011.
[21] 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.
[22] M. A. Quelal-Vásconez, M. J. Lerma-García, É. Pérez-Esteve, A. Arnau-Bonachera, J. M. Barat, and P. Talens, "Fast detection of cocoa shell in cocoa powders by near infrared spectroscopy and multivariate analysis," Food Control, vol. 99, no. November 2018, pp. 68–72, 2019.
[23] M. A. Quelal-Vásconez, É. Pérez-Esteve, A. Arnau-Bonachera, J. M. Barat, and P. Talens, "Rapid fraud detection of cocoa powder with carob flour using near infrared spectroscopy," Food Control, vol. 92, pp. 183–189, 2018.
[24] D. F. Barbin, A. L. de S. M. Felicio, D. W. Sun, S. L. Nixdorf, and E. Y. Hirooka, "Application of infrared spectral techniques on quality and compositional attributes of coffee: An overview," Food Res. Int., vol. 61, pp. 23–32, 2014.
[25] M. B. S. Scholz et al., "Validation of near-infrared spectroscopy for the quantification of cafestol and kahweol in green coffee," Food Res. Int., vol. 61, pp. 176–182, 2014.
[26] C. Assis, H. Vinicius Pereira, V. Silva Amador, R. Augusti, L. Soares de Oliveira, and M. Martins de Sena, "Combining mid infrared spectroscopy and paper spray mass spectrometry in a data fusion model to predict the composition of coffee blends," Food Chem., vol. 281, no. December 2018, pp. 71–77, 2018.
[27] A. Adnan, D. von Hörsten, E. Pawelzik, and and D. Mörlein, "Rapid Prediction of Moisture Content in Intact Green Coffee Beans Using Near Infrared Spectroscopy," Foods, vol. 6, no. 6, p. 38, 2017.
[28] R. M. Correia et al., "Banknote analysis by portable near infrared spectroscopy," Forensic Chem., vol. 8, pp. 57–63, 2018.
[29] P. Gottardo, M. Penasa, N. Lopez-Villalobos, and M. De Marchi, "Variable selection procedures before partial least squares regression enhance the accuracy of milk fatty acid composition predicted by mid-infrared spectroscopy," J. Dairy Sci., vol. 99, no. 10, pp. 7782–7790, 2016.
[30] J. R. R. Dórea, G. J. M. Rosa, K. A. Weld, and L. E. Armentano, "Mining data from milk infrared spectroscopy to improve feed intake predictions in lactating dairy cows," J. Dairy Sci., vol. 101, no. 7, pp. 5878–5889, 2018.
[31] F. Righi et al., "The use of near infrared spectroscopy to predict faecal indigestible and digestible fibre fractions in lactating dairy cattle," Livest. Sci., vol. 206, no. April, pp. 105–108, 2017.
[32] M. Kamruzzaman, Y. Makino, and S. Oshita, "Non-invasive analytical technology for the detection of contamination, adulteration, and authenticity of meat, poultry, and fish: A review," Analytica Chimica Acta, vol. 853, no. 1. 2015.
[33] M. M. Reis and K. Rosenvold, "Early on-line classification of beef carcasses based on ultimate pH by near infrared spectroscopy," Meat Sci., vol. 96, no. 2, pp. 862–869, 2014.
[34] 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.
[35] Samadi, S. Wajizah, and A. A. Munawar, "Rapid and simultaneous determination of feed nutritive values by means of near infrared spectroscopy," Trop. Anim. Sci. J., vol. 41, no. 2, 2018.
[36] B. C. Geronimo et al., "Computer vision system and near-infrared spectroscopy for identification and classification of chicken with wooden breast, and physicochemical and technological characterization," Infrared Phys. Technol., vol. 96, no. September 2018, pp. 303–310, 2019.
[37] C. Charnier et al., "Fast ADM1 implementation for the optimization of feeding strategy using near infrared spectroscopy," Water Res., vol. 122, pp. 27–35, 2017.
[38] L. Moran, S. Andres, P. Allen, and A. P. Moloney, "Visible and near infrared spectroscopy as an authentication tool: Preliminary investigation of the prediction of the ageing time of beef steaks," Meat Sci., vol. 142, no. April, pp. 52–58, 2018.
[39] 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.