Effect of Soil Moisture and Particle Size on Soil Total Phosphorus Estimation by Near-Infrared Spectroscopy

Lei Zhang[1, 2]*, Rongbiao Zhang[2]

[1]School of Electrical and Control Engineering, Henan University of Urban Construction,
Pingdingshan, 467044, China
[2]School of Electrical and Information Engineering, Jiangsu University,
Zhenjiang, 212013, China

Abstract

Near-infrared spectroscopy (NIRS) can detect soil total phosphorus in agricultural environments. Considering the soil moisture and particle size on total phosphorus prediction, we applied NIRS to the detection of soil samples with different soil moistures and particle sizes. Thus the effect of soil moisture and particle size was analyzed quantitatively and qualitatively. The procedures to remove the effect of soil moisture and particle size on total phosphorus prediction were also described. First, the near-infrared reflectance spectra of soil samples with different soil moistures and particle sizes were obtained and the absorbance values were determined. Next, the original spectra were corrected by using moisture absorbance index (MAI) and hybrid correction to counteract the effects of soil moisture and particle size, respectively. Absorbance of soil samples showed high correlation with soil moisture at wavelengths of 1,450 nm and 1,940 nm. MAI is a tool for normalizing the original spectral data so as to correct for soil moisture. Hybrid correction is based on the superposition of NIR spectra, and a particle size different from that of the original soil samples is generated. This is an effective means of correcting for the effect of soil particle size. Finally, using the corrected absorbance values at eight wavelengths (655, 722, 1,055, 1,255, 1,467, 1,678, 1,890, and 2,246 nm), the soil total phosphorus prediction model was built based on LS-SVM. Compared with the model used for original spectral data, the new model exhibited higher accuracy and stability. Results showed that MAI and hybrid correction are effective for correcting for soil moisture and soil particle size during the prediction of soil total phosphorus.

Keywords: soil moisture, soil particle size, near-infrared spectroscopy, soil total phosphorus, least-squares support vector machine

*e-mail: 847670717@qq.com

Introduction

Near-infrared spectroscopy (NIRS) is a fast, highly efficient, non-invasive, online detection method. In recent years, NIRS is increasingly applied to the real-time estimation of soil nutrients [1-12]. Phosphorus is an essential nutrient for animals and plants, and numerous studies have been carried out over the use of NIRS in the detection of soil total phosphorus [13-18]. The NIRS-based detection method has already achieved high accuracy, but the disturbance caused by soil moisture and particle size still exists. Minasny performed orthogonalization of external parameters to correct for the effect of soil moisture on the prediction of soil organic matter and discussed the problem of parameter optimization [19]. Bogrekci analyzed the effect of soil moisture and particle size on the predicted soil phosphorus, and built several soil phosphorus prediction models based on multiple linear regression and partial least squares regression [20-21]. James applied PLSR to the prediction of soil organic matter and pH value. It was found that soil moisture had little impact on soil organic matter, while the effect of soil pH value was greater [22].

Fewer studies concern the combined effect of soil moisture and particle size on total phosphorus prediction. This study proposed the algorithms for removing the effect of soil moisture and particle size and built a high-accuracy and high-stability soil total phosphorus prediction model based on LS-SVM.

Materials and Methods

Soil samples were collected from riverside farmland in Zhenjiang City (32°6'66" north latitude, 119°12'4" east longitude). Situated in the plain of the middle and lower reaches of the Yangtze River, the study area belongs to a northern subtropical monsoon climate. Yellow brown earth, typical of this region, is an alluvial soil. The soil samples were collected by the cross method to a depth of 0-20 cm. The stones and vegetation were removed during sampling and a rectangular hole measuring 20 cm in depth was dug. Along the wall of the hole, a 5 cm layer of soil was collected into a bag. A total of 320 samples were collected using this method. The soil samples were dried to obtain soil moisture measurements. Soil total phosphorus was determined by the sodium hydroxide alkali fusion-Mo-Sb colorimetric method.

We used an ARCoptix FT-NIR Rocket for NIRS with a range of from 350 nm to 2,600 nm. The spectrometer had a 35W halogen lamp and fiber optic probe, which were posed at an angle of 8° from each other and installed on the spectrometer frame. The distance from the fiber optic probe to the soil was adjustable. The fiber optic probe detected the reflectance spectra of the soil, which was analyzed by the spectrometer and transmitted to the computer via USB interface. The sampling interval was 1 nm, and each soil sample was measured five times and an average was taken.

Experimental Methods

Treatment of Soil Samples

The effect of soil moisture on total phosphorus prediction was discussed and the algorithm for removing the effect of soil moisture was proposed. Then the effect of soil particle size was analyzed and corrected by hybrid correction method. Finally the two methods were combined for the prediction of soil total phosphorus. Different soil samples were used for finding the optimal correction algorithms. Therefore, soil samples were divided into three groups and pretreated by different methods.

The first group had 80 soil samples, which were used for analyzing the effect of soil moisture on total phosphorus prediction. All samples were ground and passed through a 20-mesh sieve (0.8 mm). The reflectance and absorbance were measured for each soil sample at eight wavelengths (655, 722, 1,055, 1,255, 1,467, 1,678, 1,890, and 2,246 nm), respectively.

The second group also had 80 soil samples for studying the effect of soil particle size on total phosphorus prediction. All samples were dried at 60ºC for 24 h in an oven to remove soil moisture.

The soil samples in the third group (N = 80) were used for finding the correction algorithms. They were dried, ground and passed through a 20-mesh sieve.

SVM-based Prediction

Support vector machine (SVM) is a data mining technique that can solve regression problems and recognize non-linear and high-dimensionality patterns with a small sample size. Least square SVM (LS-SVM) is an improved version of SVM and relies on kernel parameter σ^2, and penalty factor γ. γ is mainly for controlling the degree of penalty for wrongly classified samples. By choosing a different penalty factor, one can balance between misclassification and complexity of algorithm. The choice of kernel parameter γ in the RBF kernel has a large impact on model accuracy. If γ is too large, it will result in over-learning; but if γ is too small, it will result in under-learning.

Model performance can be measured by degree of correlation, stability, and prediction capacity. Here, coefficient of correlation R was used. The larger the R value, the higher the correlation. Stability and prediction capacity were measured by root mean square error of calibration (RMSEC) (formula 1), root mean square error of prediction (RMSEP) (formula 2), and residual prediction deviation (RPD) (formula 3). The values of RMSEC and RMSEP are expected to be small, and the closer the two values, the higher the stability and the strength of the prediction capacity. RPD is divided into three types. When RPD≥2.0, the model is fit for the prediction of soil total phosphorus based on NIRS; when 1.4<RPD<2.0, other modeling techniques can be combined to increase the reliability of the current model; when RPD≤1.4, the model is not reliable at all.

$$RMSEC = \sqrt{\frac{1}{l_c - k - 1} \sum_{i=1}^{l_c} (y_i - \hat{y}_i)^2} \quad (1)$$

…where l_c is the number of calibration samples, is the number of free variables, y_i is the real value of sample *i*, and $\hat{y}_i$ is the predicted value of sample.

$$RMSEP = \sqrt{\frac{1}{l_p - 1} \sum_{i=1}^{l_p} (y_i - \hat{y}_i)^2} \quad (2)$$

…where l_p is the number of calibration samples, k is the number of free variables, y_i is the real value of sample, and $\hat{y}_i$ is the predicted value of sample i.

$$RPD = SD/RMSEP \quad (3)$$

…where SD is standard deviation of the calibration sample set.

Results and Discussion

Pretreatment of Spectral Data

Table 1 shows the statistics of soil samples. The samples in Group 2 were dried and the soil moisture was negligibly small. Table 2 shows the classification of soil samples and the correction precision factor. All spectral measurements are substituted into formulas (4)-(6) for five-point quadratic Savitzky Golay smoothing. Thus the first-order differential of reflectance and absorbance are obtained:

Table 1. Statistics of soil samples.

Sample group number	The number of soil samples	Phosphorus Concentration (mg/kg)	Moisture Content (%)
1	80	20-1,000	4-12
2	80	20-1,000	-
3	80	20-1,000	4-12

Table 2. Accuracy of classification and correction factors.

Serial number	Moisture classification (%)	MAI value	Accuracy rate (%)	λ_j
1	Low level (0-2.0)	<6.3	95.0	1
2	Low-medium level (2.0-4.0)	6.3-7.0	85.5	0.92
3	Middle level (4.0-8.0)	7.0-8.2	88.0	0.87
4	High-medium level (8.0-12.0)	8.2-9.0	92.0	0.78
5	High level (>12.0)	>9.0	100	0.66

$$R_i = \frac{\sum_{j=-r}^{r} R_i^* \times W_j}{\sum_{j=-r}^{r} W_j} \quad (4)$$

$$A_i = \log \frac{1}{R_i} \quad (5)$$

$$Y_i = \frac{1}{6h} (11A_i + 18A_{i-1} - 9A_{i-2} + 2A_{i-3}) \quad (6)$$

…where R_i is an element in the vector of spectral reflectance before and after smoothing, respectively; W_j is the weight factor in smoothing based on moving a window (window length 2r+1); A_i is the absorbance of soil at the wavelength of nm; Y_i is the first-order differential of soil absorbance; and h is the step length of differential operation.
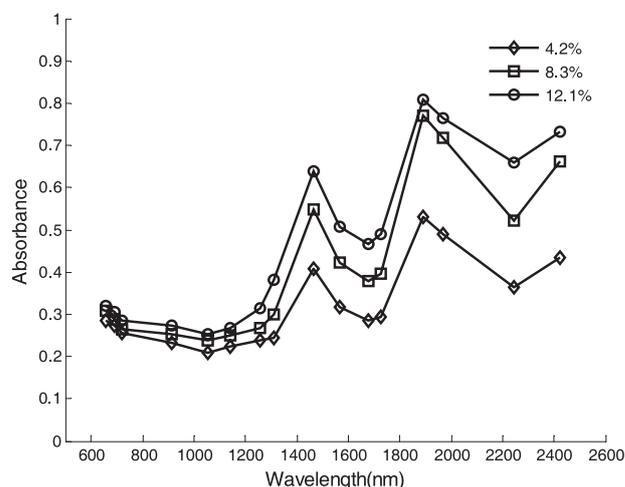
Removing the Effect of Soil Moisture

Fig. 1. Different absorbance curves with different soil moisture contents.
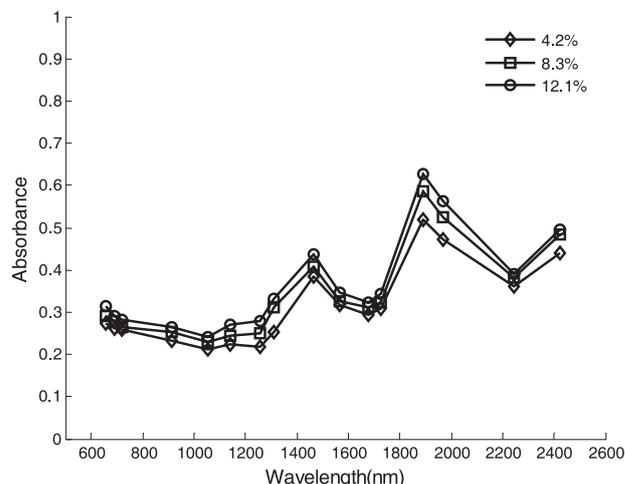

Fig. 2. Comparing the original and modified absorbance.

The soil total phosphorus prediction model based on NIRS is built using dried soil samples. Absorbance values at eight wavelengths are chosen as the prediction factors. Then using the original absorbance, the prediction model is built based on LS-SVM. R_C, R_V, RMSEC, RMSEP, and RPD of the model are 0.62, 0.50, 0.048, 0.185, and 0.8, respectively, indicating poor accuracy because of soil moisture. Soil moisture has obvious absorption peaks at 1,450 and 1,940 nm, respectively. Studies have indicated the feasibility of using the linear method for predicting soil moisture content. Therefore, the absorbance at these two wavelengths is used to obtain soil moisture, which is more accurate than using the absorbance at only one wavelength alone. To remove the effect of soil moisture on spectral data, MAI is used and calculated as follows:

$$MAI = \frac{A_{1940} - A_{1450}}{A_{1940} + A_{1450}} \times 10 \qquad (7)$$

In spectral correction, classification is first performed based on soil moisture. Then for each interval of soil moisture, different MAI is used. Five intervals of soil moisture are classified: for soil moisture of 0-2.0%, MAI is below 5.3; for 2.0-4.0%, 4.0-8.0%, and 8.0-12.0%, MAI is 5.3-6.0, 6.0-7.2, and 7.2-8.0, respectively. When soil moisture is above 12%, MAI is larger than 8.0. Our experiment shows that the accuracy of MAI-based classification of soil samples is above 86%. After classification, correction factor for absorbance λ_j is proposed for different intervals of soil moisture. The correction factor is calculated using Formula (8). Table 2 shows the MAI values for different soil moistures and the correction factors.

$$\lambda_i = \frac{A_i}{A_s} \qquad (8)$$

…where λ_j is correction factor, A_s is average absorbance at soil moisture of 0-2.0%, and A_i is the average absorbance of five intervals of soil moisture. As analyzed above, the correction factor can convert the spectral data of high-moisture content soils into that of dried soils.

The corrected absorbance A_{ri} is given by formula (9):

$$A_{ri} = A_i \times \lambda_j \qquad (9)$$

…where A_{ri} is corrected absorbance, A_i is original absorbance, λ_j is correction factor, i and ri are wavelengths ranging from 650 nm to 2,600 nm, and j is the classification interval to which the soil sample belongs (with value of 1-5).

For total phosphorus prediction, absorbance values at eigth wavelengths are used as prediction factors. Fig. 1 shows the absorbance values at different soil moisture when total phosphorus is 0.125 %. The average absorbance is 0.60, 0.55, and 0.40 at 1,450 wavelength, respectively; at 1,940 nm, the average absorbance is 0.80, 0.76, and 0.50, respectively. Fig. 2 is the comparison of curves

of original and corrected absorbance. At 1,450 nm, the average absorbance is 0.42, 0.40, and 0.38, respectively; at 1,940 nm, the average absorbance is 0.60, 0.58, and 0.50, respectively. As compared with Fig. 1, the coefficient of correlation with total phosphorus increases, while that with soil moisture decreases, indicating the removal of soil moisture.

Removing the Effect of Soil Particle Size

Soil samples in the second group were dried in an oven at 60ºC for 24h. Then the soil samples were ground and passed through different sieves to obtain particle sizes of 0.25, 0.45, and 0.8mm. The effect of soil particle size was studied by using absorbance values at eight wavelengths for each sample as the prediction factors. Fig. 3 shows the absorbance curves for different particle sizes when total phosphorus is 0.115%. When the particle size is 0.8, 0.45, and 0.25 mm, the average absorbance at 1,450 nm is 0.65, 0.50, and 0.45, and that at 1,940 nm is 0.85, 0.74, and 0.65, respectively. As the particle size decreases, the absorbance decreases as well. As seen from the spectral curves in Fig. 2, the effect of soil moisture can be removed by drying the soil samples. Therefore, the spectra corresponding to different particle sizes are obtained for hybrid correction.

Particle size is divided into four levels: <0.25 mm, 0.25-0.45 mm, 0.45-0.80 mm, and 0.80-2.0 mm. The samples with 2.0 mm are eliminated. Fifty samples are used for corrections and divided into four groups, with 10, 10, 20, and 10 samples in each group. The particle size of each group is 0.25, 0.45, 0.8, and 2.0mm, respectively (Table 3).

Table 3 shows the hybrid correction and the total phosphorus prediction model. For this prediction model based on LS-SVM, the coefficient of correlation after hybrid correction is 0.88 and that of calibration is 0.82. After single correction, the coefficient of correlation is
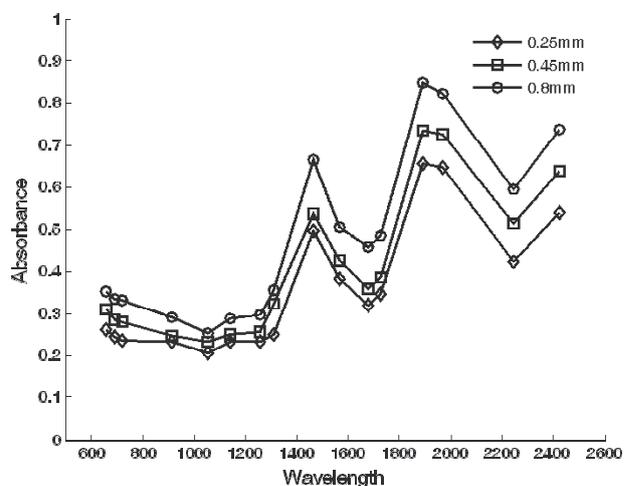

Fig. 3. Absorbance curves of soil samples with different soil particle sizes.

Table 3. Estimation models of soil AP content with proposed mixed-calibration set method.

Serial number	Calibration group	R_c	RMSEC	Validation group	R_v	RMSEP	RPD
1	50	0.60	0.076	30	0.53	0.076	0.87
2	$10^a+10^b+20^c+10^d$	0.88	0.051	30	0.82	0.038	2.65

(a) 10 soil samples, particle size 0.25 mm
(b) 10 soil samples, particle size 0.45 mm
(c) 10 soil samples, particle size 0.8 mm
(d) 10 soil samples, particle size 2.0 mm

Table 4. Comparison of four kinds of elimination effects.

Serial number	Model factor	Calibration set	R_c	Validation set	R_v	RMSEC	RMSEP	RPD
1	(a)	40	0.62	40	0.55	0.048	0.185	0.80
2	(b)	40	0.80	40	0.73	0.036	0.155	1.35
3	(c)	40	0.86	40	0.78	0.030	0.028	2.86
4	(d)	40	0.90	40	0.82	0.027	0.021	3.85

0.60 and that of calibration is 0.53. This indicates higher accuracy and stability as compared with the model with single correlation. RDP of the model with hybrid correction is 2.65 [23]. Therefore, hybrid correction greatly reduces the effect of soil particle size.

Evaluating the Correction Algorithms

The correction algorithms are evaluated using the total phosphorus prediction model based on LS-SVM.
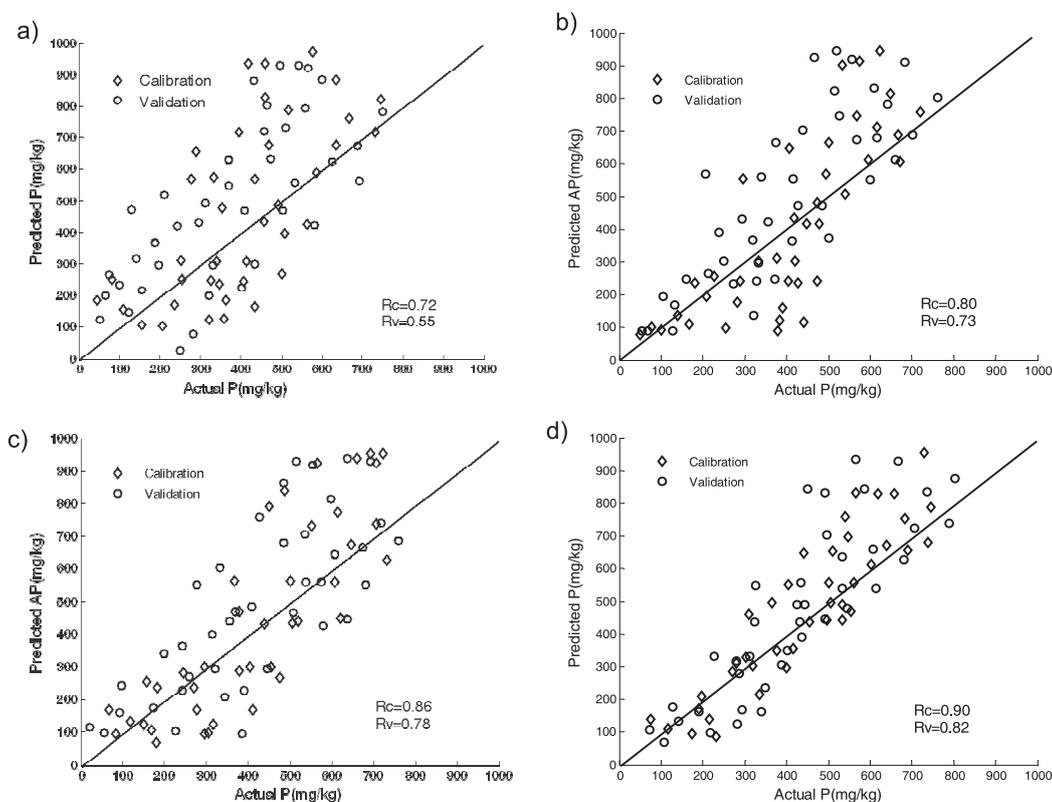

Fig. 4. Comparison of elimination effects: a) estimation of soil P content with original absorbance; b) estimation of soil P content with absorbance after eliminating the interference of soil moisture content; c) estimation of soil P content with absorbance after eliminating the interference of soil moisture content and soil particle size; d) estimation of soil P content with absorbance after drying and grinding soil samples.

As shown in Fig. 4, the prediction model consists of (a) original spectral data, (b) spectral data corrected for soil moisture using MAI, (c) spectral data corrected for soil moisture using MAI and particle size using hybrid correction, and (d) spectral data of soil samples dried and ground. When using the original absorbance, the coefficient of correlation of correction and calibration is 0.72 and 0.55, respectively; RPD is 0.80. The original spectral data are not qualified as prediction factors. As an improvement, the first-order differential of the corrected absorbance is taken as the variable of the input layer, and total phosphorus as the variable of the output layer. The prediction model is then built based on LS-SVM. After correcting for soil moisture and soil particle size, the RPD of the prediction model is 2.86, and the coefficient of correlation of correction and calibration is 0.86 and 0.78, respectively.

In a laboratory environment, the total phosphorus prediction model is built based on the absorbance values of dried and sieved soil samples (as shown in Table 4 and Fig. 4). For models c, R_C, R_V, RMSEC, RMSEP, and RPD are 0.86, 0.78, 0.030, 0.028, and 2.86, respectively; for model d, the values are 0.90, 0.82, 0.027, 0.021, and 3.85, respectively. The two models have similar accuracy, and the effect of soil moisture and particle size is successfully removed using MAI and hybrid correction, respectively. But the difference between the two models indicates the residual impact of soil moisture and particle size and there is still room for improvement.

Discussion

Soil moisture and particle size can considerably impact NIRS-based estimation of total phosphorus, leading to reduced accuracy and stability. The effect of soil moisture and particle size is analyzed qualitatively and quantitatively in this paper and the correction methods are proposed. Absorbance of soil samples will increase with soil moisture, while reflectance will decrease. As particle size decreases, reflectance will increase and absorbance will decrease. Our experiment has proven the ability of MAI and hybrid correction in removing the effects of these two factors. But given the complexity of field conditions, more experiments are needed to improve the accuracy of soil total phosphorus prediction.

Conclusions

Soil spectra are deeply influenced by soil moisture and particle size. As soil moisture increases, reflectance will decrease and absorbance will increase, both in a non-linear manner. As particle size decreases, the reflectance will increase while absorbance will decrease. When particle size is below 0.25mm, the particle size has little impact on prediction accuracy.

Different MAI is given to soil samples with different soil moisture, and then the correction factor is determined for each interval of soil moisture. Our experiment shows that the coefficient of correlation between the corrected absorbance and total phosphorus is significantly improved. We propose the use of spectral data of dried soil samples with different particle sizes for hybrid correction, which reduces the effect of particle size on prediction.

The prediction models before and after corrections for soil moisture and particle size are built based on LS-SVM. R_C, R_V, RMSEC, RMSEP, and RPD of the model built with correction are 0.86, 0.78, 0.030, 0.028, and 2.86, respectively, which are greatly improved as compared with the model without correction. Thus the proposed correction algorithms can effectively remove the effect of soil moisture and particle size on total phosphorus estimation.

Acknowledgements

References

1. FIDÊNCIO P.H., POPPI R.J., DE ANDRADE J.C. Determination of organic matter in soils using radial basis function networks and near infrared spectroscopy.Anal. Chim.Acta. **453**, 125, **2002**.

2. BOGREKCI I., LEE W.S. Spectral soil signatures and sensing phosphorus. Biosyst. Eng. **92**, 527,**2005a**.

3. VISCARRA ROSSEL R.A.V., WALVOORT D.J.J., MCBRATNEY A.B., JANIK L.J., SKJEMSTAD J.O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma. **131**, 59, **2006**.

4. GOMEZ C., VISCARRA ROSSEL R.A., MCBRATNEY A.B. Soil organic carbon prediction by hyperspectral remote sensing and field vis – NIR spectroscopy: an Australian case study. Geoderma. **146**, 403, **2008**.

5. REEVES III J.B., SMITH D.B. The potential of mid- and near-infrared diffuse reflectance spectroscopy for determining major- and trace-element concentrations in soils from a geochemical survey of North America. Appl. Geochem. **24**, 1472, **2009**.

6. PAN L., WANG J., LI P., SUN Q., ZHANG Y., HAN D. Region optimization of SSC model for Pyrus Pyrifolia by Genetic Algorithm. Spectrosc. Spect. Anal. **29**, 1246, **2009**.

7. REEVES III J.B. Near- versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: where are we and what needs to be done? Geoderma. **158**, 3, **2010**.

8. STENBERG B. Effects of soil sample pretreatments and standardized rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon. Geoderma. **158**, 15, **2010**.

9. YU Y. Quantitative determination of parameters of substrate using NearInfrared Spectroscopy Technique. Spectrosc. Spect.Anal. **31**, 2928, **2011**.

10. AN X.F., LI M.Z., ZHENG L.H., LIU Y.M., SUN H. A portable soil nitrogen detector based on NIRS. Precision Agric. **15**, 3, **2014**.

11. RÁCZ A., HÉBERGER K., FODOR M. Quantitative determination and classification of energy drinks using near-infrared spectroscopy. Analytical and Bioanalytical Chemistry. **8**, 1, **2016**.

12. VINEELA CHALLAGULLA V., WALSH K.B., SUBEDI P. Microalgal fatty acid composition: rapid assessment using near-infrared spectroscopy. Journal of Applied Phycology. **28**, 85, **2016**.

13. PANKRATOVA K.G., SHCHELOKOV V.I., STUPAKOVA G.A., SYCHEV V.G. Study of the Suitability of NIR Spectroscopy for Monitoring the Contamination of Soils with Oil Products. Novel Methods for Monitoring and Managing Land and Water Resources in Siberia. **13**, 327, **2016**.

14. HE Y., HUANG M., GARCIA A., HERNANDEZ A., SONG H. Prediction of soil macronutrients content using near-infrared spectroscopy. Comput. Electron. Agric. **58**, 144, **2007**.

15. MORÓN A., COZZOLINO D. Measurement of phosphorus in soils by near infrared reflectance spectroscopy: Effect of reference method on calibration. Commun. Soil Sci. Plant Anal. **38**, 1965, **2007**.

16. MOUAZEN A.M., KUANG B., DE BAERDEMAEKER J., RAMON H. 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.**158**, 23, **2010**.

17. MESSIGA A.J., ZIADI N., MOREL C., PARENT L.É. Soil phosphorus availability in no-till versus conventional tillage following freezing and thawing cycles. Can. J. Soil Sci. **90**, 419, **2010**.

18. DALEL ABDI, GAËTAN F. TREMBLAY, NOURA ZIADI, GILLES BÉLANGER, LÉON-ÉTIENNE PARENT Predicting Soil Phosphorus-Related Properties Using Near-Infrared Reflectance Spectroscopy. Soil Sci. Soc. Am. J. **76**, 2318, **2012**.

19. MINASNY B., MCBRATNEY A.B., BELLON-MAUREL V., ROGER J., GOBRECHT A., FERRAND L., JOALLAND S. Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon. Geoderma. **167**, 118, **2011**.

20. BOGREKCI I., LEE W.S. Improving phosphorus sensing by eliminating soil particle size effect in spectral measurement. Trans. ASABE. **48**, 1971, **2005b**.

21. BOGREKCI I., LEE W.S. Effects of soil moisture content on absorbance spectra of sandy soils in sensing phosphorus concentrations using UV–VIS–NIR spectroscopy. Trans. ASABE. **49**, 1175, **2006**.

22. JAMES B., REEVES III J.B. Near-versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus onsite analysis: where are we and what needs to be done? Geoderma. **158**, 3, **2010**.

23. KAROUI R., MOUAZEN A.M., DUFOUR E., PILLONEL L., SCHALLER E., DE BAERDEMAEKER J., BOSSET J. Chemical characterisation of European Emmental cheeses by near infrared spectroscopy using chemometric tools. Int. Dairy J. **16**, 1211, **2006**.