

Fast Detection of Inorganic Phosphorus Fractions and their Phosphorus Contents in Soil Based on Near-Infrared Spectroscopy

Lei Zhang*, Rongbiao Zhang

School of Electric Information Engineering, Jiangsu University, Zhenjiang, 212013, China.
 847670717@qq.com

For fast detection of inorganic phosphorus fractions and their phosphorus contents in soil, a method employing near-infrared spectroscopy (NIRS) combined with partial least squares (PLS) and least squares support vector machine (LS-SVM) was proposed. Fifty soil samples for each of iron phosphate, magnesium phosphate, calcium phosphate and aluminum phosphate, with application rates of phosphate being respectively 0, 60, 200 and 500 mgkg⁻¹, were taken to acquire the near infrared reflectance spectroscopy of the soil samples. PLS model regression coefficients were used to obtain the effective wavelengths (EWs) of the four phosphates as the input of LS-SVM to create a prediction model for detection of these phosphates and their phosphorus contents in the soil samples. The result showed that LS-SVM model using EWs as input had an advantage over PLS model. The correlation coefficient R^2 of EWs-LS-SVM model was respectively 0.90, 0.86, 0.88 and 0.85, and the mean square error of prediction (RMSEP) was respectively 12, 10, 15 and 18. For classification of various phosphates, a number of wavelengths equal to 12 yielded the optimal result with a 2% minimum classification error. The result showed that the method of NIRS combined with PLS and LS-SVM offered high accuracy and allowed for fast detection of inorganic phosphorus fractions and their phosphorus contents in soil.

1. Introduction

For a long time available phosphorus content in soil has been used as the basis to determine abundance or deficiency of phosphorus in soil, with little consideration on fractionation of phosphorus in soil. Phosphorus in soil includes organic phosphorus and inorganic phosphorus, of which inorganic phosphorus accounts for 60%-80% of total phosphorus and is a major source of phosphorus needed by plants (Gomez et al., 2008; Reeves and Smith, 2009; Pan et al., 2009; Reeves, 2010). Most inorganic phosphorus in soil takes the form of phosphate, so detection of phosphates and their phosphorus contents in soil is an important way to know about the conditions of phosphorus in soil, which provides a theoretical basis for correctly understanding and evaluating the effect of inorganic phosphorus fractions in soil on phosphorus nutrition in plants. In recent years, using near-infrared spectroscopy to obtain soil nutrient information becomes a focus of study both at home and abroad (Bogrekci and Lee, 2006; Bao et al., 2007; Barthes et al., 2008; Brichlemyer and Brown, 2010; James and Reeves, 2010; Minasny et al., 2011). It is necessary to carry out research on a fast detection method for inorganic phosphorus fractions and their phosphorus contents in soil.

Domestic and foreign scholars have carried out fruitful research on the application of near-infrared spectroscopy in soil neighborhood. Delel Abdi et al. (2012) utilized near-infrared spectroscopy to predict phosphorus in soil and its relevant attributes, who concluded that the prediction of PH value, magnesium, calcium and manganese in soil was successful, while prediction of total phosphorus, total nitrogen, iron, zinc and aluminum was feasible. Guillaume Debaene et al. (2014) utilized near-infrared spectroscopy to predict organic carbon, N, PH value, P₂O₅, K₂O, Ca, Mg, K and Na contents in soil, who concluded that it was appropriate to use PLS and SVM for modeling. Liu Xuemei et al. (2012) utilized near-infrared spectroscopy to measure available nitrogen and potassium contents in soil and used PLS and LS-SVM methods respectively to create a correction model.

Currently the study on NIRS prediction of phosphorus fertilizer in soil mainly focuses on the prediction and analysis of total fractions, while NIRS has not been used for detection of inorganic phosphorus fractions and

their phosphorus contents in soil. This paper takes the agricultural soil along the Yangtze River region in Zhenjiang, Jiangsu as the object of study, in which inorganic phosphorus exists mainly in the form of iron phosphate, magnesium phosphate, calcium phosphate and aluminum phosphate. NIRS technology was utilized to obtain reflectance spectrum data of different phosphates, and PLS model regression coefficients were used to obtain the effective wavelengths (EWs) of the four phosphates as the input of LS-SVM to create a prediction model for detection of these phosphates and their phosphorus contents in soil.

2. Experiments

2.1 Soil samples

Field data for this study was acquired from the agricultural land along the Yangtze River region in the city of Zhenjiang. Due to its location in the Middle and Lower Yangtze Valley Plain, most of the soil is yellow-brown soil formed by the impact of the Yangtze River. Here inorganic phosphorus in soil exists mainly in the form of iron phosphate, magnesium phosphate, calcium phosphate and aluminum phosphate. Sampling depth was 1-20cm after removal of topsoil, and sampling weight was 1.0-1.5kg. To minimize man-made error, the soil samples taken were spread out in a room and air-dried naturally. After air-dried, the soil samples were ground in a grinder to remove all nutrients and organic matters, then passed through a 2mm sieve, numbered and sealed for storage. After quartering, two quarters of the soil samples each totaling 200 were taken, the first quarter for routine chemical measurement and analysis, and the second quarter for near-infrared spectral scan and analysis.

To study the effect of different phosphorus compounds on soil reflectance spectrum, four phosphates, namely iron phosphate, magnesium phosphate, calcium phosphate and aluminum phosphate, were taken to undergo soil sample treatment as shown in Table 1 and obtain 50 study samples for each of the four different phosphates. HCL solution of 0.1mol/L and deionized water were used to remove phosphorus in the existing soil samples. After filtering, pH value and phosphorus concentration in the soil samples were analyzed (Mouazen et al., 2005, 2006).

Table 1: Soil sample treatment

Phosphorus compound	Particle size (mm)	Moisture content	PH value	Application rate of phosphorus compound (mgkg ⁻¹)	
Iron phosphate	2	8%	6	None	0
Magnesium phosphate				Minor	60.0
Calcium phosphate				Moderate	200.0
Aluminum phosphate				Massive	500.0

Table 2 shows the application rates of these four phosphate fertilizers and the average actual total phosphorus concentrations. The highest and lowest phosphorus concentrations in the soil samples were respectively 98 mgkg⁻¹ and 11 mgkg⁻¹. Although phosphorus in the soil samples had been removed, the phosphorus concentration still ranged from 11 to 18 mgkg⁻¹.

Table 2: Phosphate fertilizer application rate and average actual total phosphorus concentration

Compound	Phosphate fertilizer application rate (mgkg ⁻¹)	Average actual total phosphorus concentration (mgkg ⁻¹)
Iron phosphate	0	18
	60.0	22
	200.0	38
	500.0	60
Magnesium phosphate	0	12
	60.0	22
	200.0	36
Calcium phosphate	500.0	98
	0	13
	60.0	21
Aluminum phosphate	200.0	34
	500.0	78
	0	11
Aluminum phosphate	60.0	18
	200.0	48
	500.0	90

2.2 Instrument and equipment

Fourier transform near-infrared spectrometer of Swiss Arcspectro FT Rocket type with a measurement range of 700-2600 nm was used for spectroscopic measurement. The soil was measured for spectral sampling at an interval of 1 nm, scanned for 15 times at a resolution of 3 nm and probe viewing angle of 25°, using a 12V/45W tungsten halogen lamp as light source.

2.3 Spectral acquisition and data preprocessing

The soil samples were placed into a sample holder, with the spectrometer 12cm above the soil surface. Measurement was done using diffuse reflection approach and a light incident angle of 45°. Since scattering caused by instrument noise and sample particle size can affect the analysis and extraction of spectral information, the spectrum of a soil sample was acquired for three times to obtain the average value as the final spectral data for that soil sample. To reduce the effects from high frequency random noise, baseline drift, inhomogeneous light scattering in sample, etc., spectral data was preprocessed with multiple scattering correction (MSC) and first order difference.

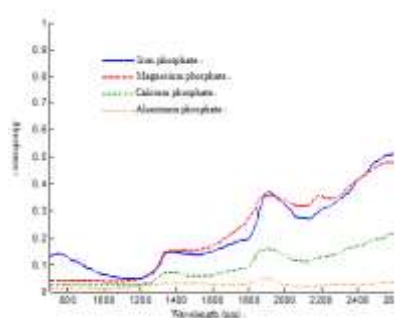


Figure 2: Average absorbance of four phosphates

The absorbance spectrum of the four phosphates is shown in Figure 2. From the figure above a conclusion can be made that there are three significant peaks for iron phosphate in the near-infrared spectrum, 754 nm, 1334 nm and 1911 nm, while from 2120 nm to 2600 nm, the absorbance value increases gradually. There are two significant peaks for magnesium phosphate, 1336 nm and 1855 nm, while from 2265 nm to 2600 nm, the absorbance value increases gradually. There are two significant peaks for calcium phosphate, 1369 nm and 1868 nm, while from 2135 nm to 2600 nm, the absorbance value increases gradually. There is one significant peak for aluminum phosphate, 1908 nm, while from 2065 nm to 2600 nm, the absorbance value does not change significantly.

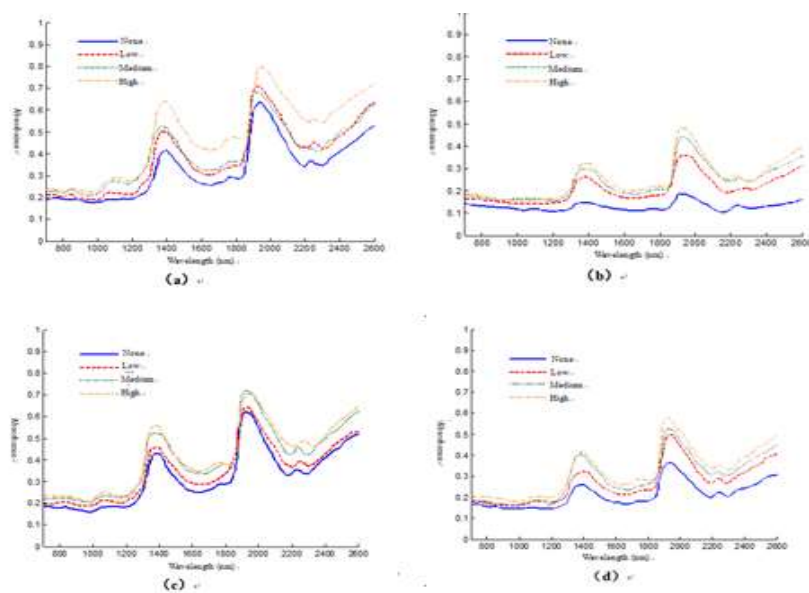


Figure 3: Average absorbance of soil samples with different concentrations: (a) iron phosphate, (b) magnesium phosphate, (c) calcium phosphate, (d) aluminum phosphate

Figure 3a shows that from zero to high iron phosphate concentration, the average absorbance of soil samples has two peaks, 1396 nm and 1933 nm, which is due to the effect of soil moisture on the spectrum. As soil moisture increases, soil reflectance decreases while absorbance increases. In the 700 nm- 1054 nm zone, phosphorus concentration and absorbance in iron phosphate exhibit linear relationship; the lower absorbance is, the lower phosphorus concentration will be. Figure 3b shows that from zero to high magnesium phosphate concentration, average absorbance of soil samples has two peaks, 1374 nm and 1945 nm, and the water absorption zone is also obvious. In the 700 nm- 1129 nm zone, phosphorus concentration and absorbance in magnesium phosphate exhibit linear relationship. Figure 3c shows that from zero to high calcium phosphate concentration, average absorbance of soil samples has two peaks, 1378 nm and 1982 nm, and the water absorption zone is also obvious. In the 700 nm- 1160 nm and 2356 nm- 2600 nm zones, phosphorus concentration and absorbance in calcium phosphate exhibit linear relationship. Figure 3d shows that from zero to high aluminum phosphate concentration, average absorbance of soil samples has two peaks, 1374 nm and 1945 nm, and the water absorption zone are also obvious. In the 700 nm- 1129 nm and 2380 nm- 2600 nm zones, phosphorus concentration and absorbance in aluminum phosphate exhibit linear relationship.

3. Creating LS-SVM prediction model

PLS model regression coefficients were used as the input of LS-SVM to obtain the EWs of iron phosphate, magnesium phosphate, calcium phosphate and aluminum phosphate. Effective wavelengths were selected according to the PLS model regression curve preprocessed with first order difference, mainly based on the assumption that there should exist a maximum value of absolute regression coefficient for wavelengths within a certain band range. For example, the wavelength between the peak and trough of the regression curve is an effective wavelength. The result is shown in Table 3.

Table 3: Four effective wavelengths of phosphate

Phosphate	Effective wavelength (nm)
Iron phosphate	867,934,1022, 1464,1934,2152,2208,2450
Magnesium phosphate	845,1023, 1419,1866,1938,2055,2230,2430
Calcium phosphate	985,1022, 1439, 1948,2123,2378,2442, 2516
Aluminum phosphate	908,1024, 1374,1467, 2231,2387,2465,2550

4. Experimental results and analysis

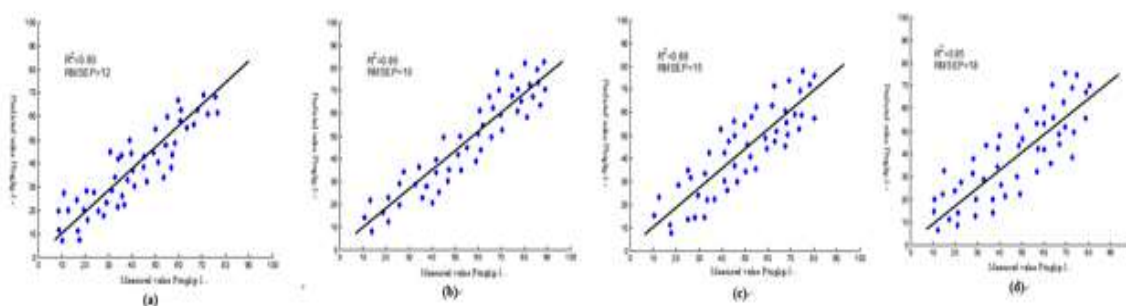


Figure 4: EWs-LS-SVM phosphorus concentration prediction model: (a) iron phosphate, (b) magnesium phosphate, (c) calcium phosphate, (d) aluminum phosphate

Figure 4 shows the prediction result for phosphorus concentration of various phosphates in the soil samples using EWs-LS-SVM model, where R^2 is 0.90, 0.86, 0.88 and 0.85, and RMSEP is 12, 10, 15 and 18, respectively. The result is better than the correlation coefficient R^2 and RMSEP for phosphorus concentration of various phosphates established using PLS prediction model, as shown in table 4.

Table 4: PLS results for the soil samples with four compounds (the number of factors was six)

Phosphate	Calibration value		Verification value	
	R ²	RMSEC (mgkg ⁻¹)	R ²	RMSEP (mgkg ⁻¹)
Iron phosphate	0.80	16	0.75	18
Magnesium phosphate	0.82	12	0.68	15
Calcium phosphate	0.72	18	0.65	20
Aluminum phosphate	0.70	20	0.50	21

The EWs-LS-SVM model performs better than the PLS model, probably because it uses PLS model regression coefficient to obtain effective wavelength, and also because LS-SVM takes into account the nonlinear information of spectral data, thus improving prediction precision.

Table 5: Classification results for the soil samples with four phosphates

Criteria	Number of selected wavelengths	Error (%)
Correlation coefficient > 0.6, stepwise discriminant analysis	45	14.5
Correlation coefficient > 0.65, stepwise discriminant analysis	15	2.0
Correlation coefficient > 0.7, stepwise discriminant analysis	3	45.5

Table 5 shows the classification result for the four phosphates of selected wavelengths in the soil samples under different given conditions. The threshold level of the correlation coefficient was increased ($r=0.7$), and stepwise discriminant procedure was used to reduce the number of selected wavelengths, which, however, resulted in higher classification errors (45.5%). The threshold level of the correlation coefficient was decreased ($r=0.6$), and stepwise discriminant procedure was used to decrease the number of selected wavelengths, which reduced classification error to 14.5%. A correlation coefficient equal to 0.65 resulted in a lower classification error (2.0%). Table 5 shows that when the number of independent variables (number of selected wavelengths) increases, classification error increases, and when the number of independent variables decreases, classification error increases again. When the number of wavelengths was equal to 15, a lower classification error (2.0%) resulted.

5. Conclusions

- (1) It is feasible to use near-infrared spectroscopy (NIRS) for fast detection of inorganic phosphorus fractions in soil.
- (2) PLS model regression coefficients were used to obtain the effective wavelengths (EWs) of four phosphates.
- (3) The effective wavelengths (EWs) of the four phosphates were used as the input of LS-SVM to create a EWs-LS-SVM correction model for decoupling, measuring and quantifying the phosphorus concentrations in the phosphates. R² was 0.90, 0.86, 0.88 and 0.85, and RMSEP was 12, 10, 15 and 18, respectively, which were better than those established with the PLS prediction model, indicating that it is the only characteristic directly associated with spectral prediction within the near infrared zone. For classification of phosphates, a number of wavelengths equal to 15 yielded the optimal result with a 2% minimum classification error

References

- Abdi, D., Tremblay, G.F., Ziadi, N., Bélanger, G., et Parent, L.-É., 2012. Predicting Soil Phosphorus-related Properties Using Near Infrared Reflectance Spectroscopy. *Soil Science Society of America Journal*, 76, 2318-2326. DOI: 10.2136/sssaj2012.0155
- Bao, Y., He, Y., Fang, H., Pereira, A.G., 2007. Spectral characterization and N content prediction of soil with different particle size and moisture content. *Spectrosc. Spect. Anal.* 27 (1), 62–65. DOI: 10.3321/j.issn:1000-0593.2007.01.017

- Barthes, B.G., Brunet, D., Hien, E., Enjalric, F., Conche, S., Freschet, G.T., Annunzio, R., Toucet-Louri, J., 2008. Determining the distributions of soil carbon and nitrogen in particle size fractions using near-infrared reflectance spectrum of bulk soil samples. *Soil Biol. Biochem.* 40, 1533–1537. DOI: 10.1016/j.soilbio.2007.12.023
- Bogrekcı, I., Lee, W.S., 2006. Effects of soil moisture content on absorbance spectra of sandy soils in sensing phosphorus concentrations using UV–VIS–NIR spectroscopy. *Trans. ASABE* 49 (4), 1175–1180. DOI: 10.13031/2013.21717
- Brichlemyer, R., Brown, D., 2010. On-the-go VisNIR: potential and limitations for mapping soil clay and organic carbon. *Comput. Electron. Agric.* 70 (2), 209–216. DOI: 10.1016/j.compag.2009.10.006
- Gomez, C., Viscarra Rossel, R.A., McBratney, A.B., 2008. Soil organic carbon prediction by hyperspectral remote sensing and field vis – NIR spectroscopy: an Australian case study. *Geoderma*. 146 (3–4), 403–411. DOI: 10.1016/j.geoderma.2008.06.011
- Guillaume Debaene, Jacek Niedźwiecki, Alicja Pecio, Anna Żurek, 2014, Effect of the number of calibration samples on the prediction of several soil properties at the farm-scale, *Geoderma*, 114, 214–215. DOI: 10.1016/j.geoderma.2013.09.022
- James, B., Reeves III, J.B., 2010. 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, 3–14. DOI: 10.1016/j.geoderma.2009.04.005
- Liu, X.M., Liu, J.S., 2012, Based on the LS-SVM modeling method determination of soil available N and available K by using near-infrared spectroscopy, *Spectroscopy and Spectral Analysis*, 32(11), 3019–3023. DOI: 10.3964/j.issn.1000-0593(2012)11-3019-05
- Minasny, B., McBratney, A.B., Bellon-Maurel, V., Roger, J., Gobrecht, A., Ferrand, L., Joalland, S., 2011. Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon. *Geoderma*. 167, 118–124. DOI: 10.1016/j.geoderma.2011.09.008
- Mouazen, A.M., De Baerdemaeker, J., Ramon, H., 2005. Towards development of online soil moisture content sensor using a fibre-type NIR spectrophotometer. *Soil Tillage Res.* 80, 171–183. DOI: 10.1016/j.still.2004.03.022
- Mouazen, A.M., Karoui, R., De Baerdemaeker, J., Ramon, H., 2006. Characterization of soil water content using measured visible and near infrared spectra. *Soil Sci. Soc. Am. J.* 70 (4), 1295–1302. DOI: 10.2136/sssaj2005.0297
- Pan, L., Wang, J., Li, P., Sun, Q., Zhang, Y., Han, D., 2009. Region optimization of SSC model for *Pyrus Pyrifolia* by Genetic Algorithm. *Spectrosc. Spect. Anal.* 29 (05), 1246–1250. DOI: 10.3964/j.issn.1000-0593(2009)05-1246-05
- Reeves III, J.B., Smith, D.B., 2009. 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 (8), 1472–1481. DOI: 10.1016/j.apgeochem.2009.04.017
- Reeves III, J.B., 2010. 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 (1), 3–14. DOI: 10.1016/j.geoderma.2009.04.005