Study on the Prediction Model Based on a Portable Soil TN Detector

Xiaofei An[1,2(✉)], Guangwei Wu[1,2], Jianjun Dong[1,2], Jianhua Guo[1,2], and Zhijun Meng[1,2]

[1] Beijing Research Center for Intelligent Agricultural Equipment,
Beijing 100097, China
anxf@nercita.org.cn
[2] National Research Center of Intelligent Equipment for Agriculture,
Beijing 100097, China

Abstract. As the development of precision agriculture, it is necessary to obtain soil total nitrogen (TN) content and other element parameters. With the NIRS technology, a soil detector for soil total nitrogen content was developed. It included two part: optical part and control part. The detector took each lamp-house connected with the incidence of Y type optical fiber in turn by a manual rotation, The different wavelength lamp-house signal was transferred to the surface of soil by the input fibre. The reflected signal would be converted by photoelectric sensor, the optical signal was converted to electrical signal. After the power circuit, amplier circuit, and AD convert circuit, the electrical signal was processed by MCU. Finally, the result of soil total nitrogen content could be displayed on LCD. With the forty-eight apple orchard soil samples of Beijing surburb, the predicted models were established by seven different methods (MLR, PLSR, SVM, BPNN, GA + BPNN, GA + SVM and PSO + SVM). The model established by genetic algorithm (GA) optimizing BP neural network was optimal, with R_C of 0.94, R_V of 0.78, RMSEC and RMSEP of 0.037 and 0.067. The results showed that the soil total nitrogen content detector had a stable performance. The established model had perfect accuracy and strong robustness.

Keywords: NIR · Soil parameter · Modeling method · Soil sensor

1 Introduction

Soil macro and micro parameters such as TN, P, K, and OM could demonstrate the fertilization directly. It was the basis to carry out precision agriculture. At present, a lot of scholars [2–12] have carried out the farmland soil characteristics from spatial variability. The research content mainly concentrated on the soil nutrient (N, P, K, organic matter, et al.), soil moisture, elements such as electricity conductivity, pH value. Soil fertility, the physical and chemical parameters could be analyzed by spectral analysis methods, which could short analysis time, reduce cost and improve the analysis efficiency in the laboratory. With nearly 30 years' development, it has made remarkable progress and achievements under laboratory conditions by near infrared spectral method for the soil spectral data preprocessing.

In 1999, Shibusawa has studied on the parameters of soil (nitrogen, phosphorus and potassium, organic matter, electricity conductivity, et al.) by near infrared spectral analysis technology. The sensitive wavebands have been selected and the parameters corresponding prediction models were also established. Huan-junliu, et al. selected 675 soil samples of Heilongjiang province to build VIS NIR waveband and linear model of soil organic matter that the prediction accuracy was as high as 0.936(R^2). Gao Hongzhi, JiWenJun et al. have used the maximum continuous projection algorithm to eliminate the effect between multivariate linear, extracted the modeling characteristics of soil total nitrogen and organic matter wavelength, the use of random forest, SVM, and ANN, partial least squares method (PLS) high precision models were obtained. Dong-jian et al. has used 900-1700 nm wavelength range of diffuse reflection spectrum to identify and eliminate abnormal samples. The accuracy of the model became excellent with the continuous projection algorithm (SPA, Successive Projections Algorithm) to the optimal modeling variable wavelength selection, and then through the three linear modeling methods affection the result of the organic matter content prediction was analyzed. PLS method and RBF neural network prediction model were also established. Yan shanshan introduced the continuum removal method to extract the sensitive wavebands of SOM, analysis when the change of soil organic matter content, the change rule of the spectrogram to extract the sensitive wavelength of 600 nm, 900 nm and 2210 nm as the center wavelength, the BP neural network to establish the model of SOM. With the prediction model data, root mean square error MSE was 0.286, the correlation coefficient was 0.979 and the model was superior to the whole band.

Since the 1990's, some soil scientists and agricultural information technology scholars have researched on soil sample spectrum analysis of research status in real time based on the NIR spectroscopy. And several kinds of prototypes have been developed. [10–20] (Adamchuk, 1999; Shibusawa, 2000; Mouazen, 2005; Maleki, 2007; Christy, 2008; XingZhen, 2008; Gao Hongzhi, 2011). Sudduth and Hummel have developed a rapid detection sensor which was used to identify the soil texture and soil organic matter in the 1990's based on NIRS. The accuracy of the detector could reach 90 %. With the development of NIRS technology, the NIR lamphouse could be used in the field condition which made it easy to collect data in the outside. Several large companies have also product specificity lamp-house for the portable soil sensor and on-the-go soil sensor. Several foreign scholars have also obtained excellent result of soil moisture, soil organic matter and so on. Different scholars from studying the spectral characteristics of the soil, combined with chemometrics methods, established the soil moisture, soil organic matter, soil total nitrogen and soil parameters prediction model. foreign scholars also started to research on the effect of factor under the field condition to improve the model accuracy.

Hence, we decided to develop a kind of portable soil TN detector based on NIRS technology. And then it would make the basis for the field condition to detect soil TN.

2   Materials and Methods

2.1   Soil TN Detector Design

(1) Selection of sensitive wavelengths

It was necessary to determine the used NIR wavelength lamp-house for the soil TN sensor, firstly. Several scholars had studied on the determination of NIR wavelength lamp-house. Zheng Lihua et al. had tried four kinds of methods to estimate soil TN content, such as BPNN, MLR, SVM and wavelet analysis. With the different methods, she also proposed the selected sensitive wavelength, including 24 suggestion wavelengths. The selected wavebands were 844 nm,859 nm, 923 nm,931 nm, 972 nm, 984 nm, 1028 nm,1064 nm,1092 nm,1124 nm,1187 nm,1208 nm, 1215 nm,1286 nm, 1311 nm,1389 nm,1394 nm,1536 nm,1559 nm,1673 nm,1684 nm,1895 nm,1833 nm, 1991 nm, 2150 nm and 2234 nm.

According to the mentioned results, An et al. suggested a new group of wavelengths to build the BPNN estimation model, 1550 nm, 1300 nm, 1200 nm, 1100 nm, 1050 nm, 940 nm. The R^2 (calibration and validation) were 0.85 and 0.77,respectively. Under the laboratory condition and field condition, the models have different affect factor, as a result, it would be necessary to determine new model parameters after finishing the development.

(2) Structure design of soil TN sensor

With the NIRS technology, a soil detector for soil total nitrogen content was developed. It was consisted of optical part and control part. The detector took each lamp-house connected with the incidence of Y type optical fiber in turn by a manual rotation, The different wavelength lamp-house signal was transferred to the surface of soil by the input fibre. Once the signal was converted by the photoelectric sensor, the optical signal was converted to electrical signal. After the power circuit, amplier circuit, and AD convert circuit, the electrical signal was processed by MCU. Finally, the result of
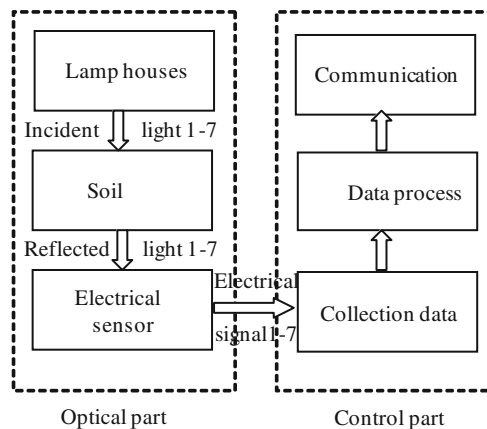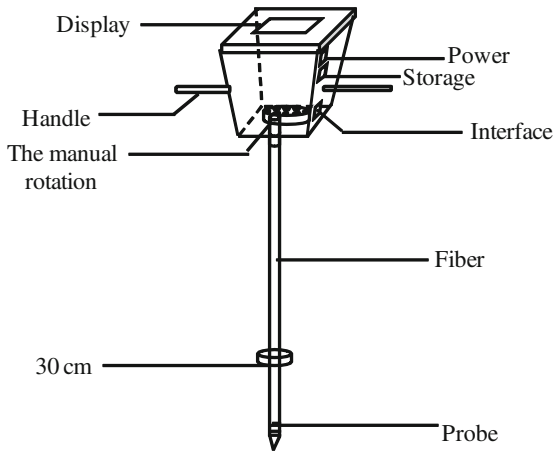

Fig. 1. Structure

Fig. 2. Soil TN content detector structure

soil total nitrogen content could be displayed on LCD. Figure 1 shows the structure, and Fig. 2 shows Soil TN content detector structure.

(3) Optical part

In order to develop the portable detector, appropriate lamp-houses and transmission channel should be selected correctly. In this study, Lighting Emitting Diode (LED) was selected as active lamp-house. According to the result of 2.1 section, six wavelength of LED were selected. All the selected lamphouses had the advantagement of low cost, narrow bandwidth and portable. They were at the wavelength of 1550, 1300, 1200, 1100, 1050, and 940 nm. In order to analyze the effect of soil moisture, another special LED was also selected, which was at the wavelength of 1450 nm. When the soil TN sensor began to work, all the LEDs would be open by turn, and the optical signal could be transmitted by the input fiber. When the LED optical signal arrival at the surface of soil, it would be divide into two part. One part would be absorbed by the soil and the other part would be reflected by the soil. According to the different content of soil properties, the reflectance value would be different. For the soil moisture content, as the soil moisture content increased, the reflectance value would reduce. When the locating dowel was pulled up, the rotary table was back to original place by the clockwork spring.

(4) Control part

The soil TN detector was controlled by the MCU. Figure 3 shows the MCU process chart. Once the optical signal was converted into electrical signal, it was processed by the IU convert circuit, amplier circuit, AD convert circuit and LCD display circuit by turn. All the weak electrical signal would become standard signal for the MCU. In the MCU, it could be display stored. Both the reflectance value and absorbance value would be stored at the same time. With the help of soil TN predicted model, the soil TN content could be displayed on LCD.
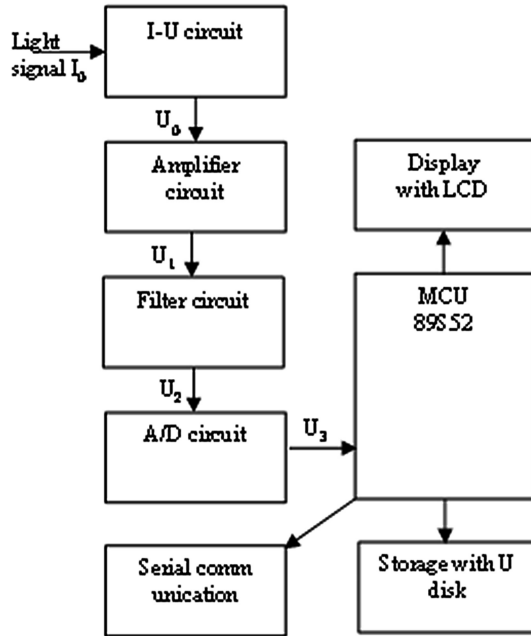
Fig. 3. MCU process chart

2.2   Experimental Methods

The soil coverage rate was 82 % in Beijing suburb, the area was about 137.8 hectare. Among them, the brown soil area was 89.1 hectare, about 65 % of Beijing soil area. Soil type was typical brown earth soil which has a strong representation. They were collected from Beijing suburb with the ranges of N 40.14395738° ∼ 40.1434362° and E 116.2154066° ∼ 116.2150152°. The depth was 1-20 cm under the soil surface. In order to reduce the effect of operation error, all the soil samples were taken as uniform number, crushing and drying. And then they were divided into two parts, one part was measured by the traditional chemical methods; the other part was used to detect by the developed portable sensor.

All the data were measured for twice by the soil TN detector. For the first time, the absorbance of standard whiteboard was measured. And for the second time, the probe output signal was also obtained. With the help of equation of (1) and (2), the absorbance of the soil sample could be obtained. According to the Eqs. (1) and (2), the absorbance would be obtained.

$$r_i = V_i' / V_i \times 100\% \tag{1}$$

$$A_i = \lg(1/r_i) \tag{2}$$

Where: i is 1550 nm, 1450 nm, 1300 nm, 1200 nm, 1100 nm, 1050 nm, and 940 nm; V_i is the output value at the wavelength of i by putting probe on the standard

whiteboard; $V_i^{'}$ is the output at the wavelength of i by pushing probe into the soil sample; r_i is the reflectance value at the wavelength of i; A_i is absorbance of the wavelength of i.

2.3   The Evaluation of Soil TN Predicted Model

In order to evaluate the portable soil TN detector accuracy and performance, three parameters were chose to evaluate them, include RMSEC, RMSEP and RPD value. Equations (3), (4) and (5) shows the calculation formulas.

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_{ci})^2}{n_c}} \tag{3}$$

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y_{pi})^2}{n_p}} \tag{4}$$

$$RPD = \frac{\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}}{\sqrt{\sum_{i=1}^{n}[(y_i - y_{pi})^2 - \frac{\sum_{i=1}^{n}(y_i - y_{pi})^2}{n_c}]^2}} \tag{5}$$

Where: i was soil sample serial; y_i was soil TN measured by FOSS Kjeltec^{TM}2300; y_{ci} is calibration group soil TN predicted value; y_{pi} is validation group soil TN predicted value; $\overline{y}$ is soil total nitrogen content average valuemeasured by FOSS Kjeltec^{TM}2300; n_c calibration group number; n_p is validation group number.

3   Result and Discussion

3.1   Soil Spectral Data Pretreatment

Table 1 showed the statistics data of soil parameters. The content range of that was 0.007 % ∼ 0.286 %, average content of soil TN was 0.16 % and the standard deviation was 0.007. From the distribution coverage, soil total nitrogen variation range was across crop growth cycle, it was suitable for modeling analysis.

The data obtained from the detector has been processed by the average filtering software and hardware of the first-order RC low-pass filter. According to the formula (1) and (2), soil reflectance and absorbance values at the six different wavelengths were obtained. Relevant relations showed that the correlation coefficient between absorbance and soil TN content were 0.19, 0.31, 0.10, 0.21, 0.10 and 0.11, respectively.

This paper proposed a new soil spectral data pretreatment method, the differential absorbance was selected as a new independent variable factor. Differential absorbance data was obtained by formula (6). Figure 4 showed the result of absorbance and soil TN content with the new data pretreatment method. The correlation coefficient between

Table 1. Statistics data of soil parameters

Number	Parameter	Soil sample	Max value	Min value	Average value	Median value	Standard deviation
1	SM (%)	48	20.9	10.5	16.4	17.2	0.004
2	TN (%)	48	0.286	0.007	0.16	0.17	0.007
3	pH	48	8.34	7.71	7.99	8.00	0.02
4	OM (%)	48	1.31	0.69	0.95	0.93	0.02
5	P (mg/L)	48	454.2	45.90	118.15	107.20	9.59
6	K (mg/L)	48	351.90	81.73	165.20	152.95	9.37

differential absorbance ($A_{940,1100}$, $A_{1050,1550}$, $A_{1100,1200}$, $A_{1200,1550}$, $A_{1300,1550}$, $A_{1550,1450}$) and TN content with the new data pretreatment method were 0.26, 0.35, 0.28, 0.35, 0.28 and 0.21, respectively. Although the correlation coefficient was still low, it was improved by 36.8 %, 12.9 %, 180 %, 28.6 %, 180 % and 90.9 %, comparing with the original absorbance.

$$A_{i,j} = \frac{A_i - A_j}{A_i + A_j} \tag{6}$$

Where: i and j are 940,1050,1100,1200,1300,1450 and 1550 nm; $A_{i,j}$ is differential absorbance value; A_i is original absorbance value.
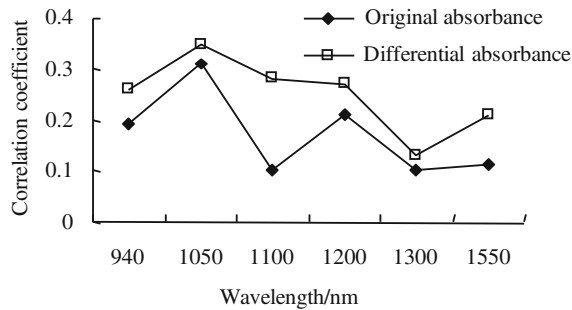

Fig. 4. Data pretreatment result

3.2   Establishment of Soil TN Predicted Models

Both the calibration and validation set were established by the 48 soil samples. The original absorbance A_i and the differential absorbance $A_{i,j}$ were as input factor. The soil TN predicted model was established by the MLR, PLSR, BPNN method, SVM method, GA + BPNN, improved support vector machine (GA + SVM) method and the improved support vector machine method (PSO + SVM) respectively. Table 2 showed the results of different soil TN predicted models.

Table 2. Comparison of different soil TN predicted models

Num	Method	Variable	Calibration set	R$_C$	Validation set	R$_V$	RMSEC	RMSEP	RPD
1	MLR	A_i	32	0.50	16	0.51	0.174	0.165	0.367
2	PLSR	A_i	32	0.83	16	0.65	0.076	0.106	1.446
3	BPNN	A_i	36	0.90	12	0.71	0.045	0.026	1.042
4	GA + BPNN	A_i	36	0.89	12	0.79	0.051	0.064	2.554
5	SVM	A_i	36	0.81	12	0.68	0.065	0.135	1.179
6	GA + SVM	A_i	36	0.90	12	0.77	0.044	0.090	1.941
7	PSO + SVM	A_i	36	0.96	12	0.70	0.135	0.070	2.514
8	MLR	$A_{i,j}$	32	0.58	16	0.44	0.161	0.137	0.521
9	PLSR	$A_{i,j}$	32	0.89	16	0.65	0.069	0.059	2.916
10	BPNN	$A_{i,j}$	36	0.86	12	0.74	0.054	0.040	1.064
11	GA + BPNN	$A_{i,j}$	36	0.94	12	0.78	0.037	0.067	2.577
12	SVM	$A_{i,j}$	36	0.69	12	0.72	0.097	0.106	1.638
13	GA + SVM	$A_{i,j}$	36	0.99	12	0.61	0.006	0.063	3.183
14	PSO + SVM	$A_{i,j}$	36	0.99	12	0.67	0.006	0.053	4.191

3.3   Discussions

The model with differential absorbance was obvious better than that with original absorbance for all the seven different methods from Table 2. For the RMSEC, it reduced from 0.085 to 0.061 and improved about 27.4 %. For the RMSEP, it reduced from 0.094 to 0.075, improved about 3.1 %. The model accuracy become more better.

For the linear method, PLSR model had the Rc of 0.83, Rv of 0.65, which were both higher than that of MLR model. RMESC and RMSEP were also lower than that of MLS model. For the nonlinear methods, BPNN, SVM, GA + BPNN, GA + SVM and PSO + SVM had the Rc of 0.69, 0.86, 0.94, 0.86 and 0.94, respectively. It had the increased trend gradually. The correlation coefficients of validation were 0.72, 0.74, 0.78, 0.74 and 0.78, respectively. Among them, the GA + BPNN method had the highest value. All the RMSEC and RMSEP satisfied the requirement except SVM method. According to the RPD value, GA + BPNN, GA + SVM and PSO + SVM model had excellent performance.

As a result, the GA + BPNN model had high Rc and Rv, and the RMSEC, RMSEP, RPD value satisfied the requirement. It could be as the soil TN forecasting model.

4   Conclusions

The soil TN forecasting model was also established by seven different methods (MLR, PLSR, SVM, BPNN, GA+BPNN, GA+SVM, PSO+SVM). The model was evaluated by Rc, Rv, RMSEC, RMSEP and RPD.

(1) A kind of new portable soil TN detector was developed. LED was selected as the lamp house, and optimization amplification and filter circuit were designed.

(2) A new soil spectral data pretreatment method suitable for portable detector was proposed. According to the method, the correlation coefficient between spectral absorbance and soil TN was improved, obviously.

(3) The soil TN forecasting model was established by seven different methods (MLR, PLSR, SVM, BPNN, GA + BPNN, GA + SVM and PSO + SVM). The GA + BPNN model had high Rc of 0.94 and Rv of 0.78, and the RMSEC, RMSEP, RPD satisfied the requirement. Especially, the RPD value reached 2.577. It could be as the soil TN forecasting model.

Although the soil parameter predicting model were established, the robustness and universality still had disadvantage. In the next step, the effect of soil other properties should be considered in the model.

References

1. Sinfield, J.V., Fagerman, D., Colic, O.: Evaluation of sensing technologies for on-the-go detection of macro-nutrients in cultivated soils. Comput. Electron. Agric. **70**(1), 1–18 (2010)
2. Zhang, J., Tian, Y., et al.: Estimating model of soil total nitrogen content based on near-infrared spectroscopy analysis. Trans. CSAE **28**(12), 183–188 (2012)
3. Shenk, J.S., Workman Jr., J.J., Westerhaus, M.O.: Application of NIR spectroscopy to agricultural products. In: Handbook of Near-infrared Analysis, Practical Spectroscopy Series, pp. 383–431 (1992)
4. Li, M., Sasao, A., Shibusawa, S., Sakai, K.: Soil parameters estimation with NIR spectroscopy. J. Japan. Soc. Agric. Mach. **62**(3), 111–120 (2000)
5. Chang, C.W., Laird, D.A.: Near-infrared reflectance spectroscopic analysis of soil C and N. Soil Sci. **167**(2), 110–116 (2002)
6. Gao, H.Z., Luo, Q.P.: Near infrared spectral analysis and measuring system for primary nutrient of soil. Spectrosc. Spectral Anal. **31**(5), l245–1249 (2011)
7. Ji, W., Li, X., Li, C., et al.: Using different data mining algorithms to predict soil organicmatter based on visible near infrared spectroscopy. Spectrosc. Spectral Anal. **32**(9), 2393–2398 (2012)
8. He, D., Chen, X.: Research on soil organic matter content measurement by spectral analysis. J. Chin. Agric. Mech. **46**(2), 1–5 (2015)
9. Yan, S., Cheng, X., Song, H.: Study on the extraction of sensitive band of soil organic matter near infrared spectrum basd on continum removal. Shanxi Agric. Univ. J. **36**(1), 72–76 (2016)
10. Adamchuk, V.I., Lund, E.D., Sethuramasamyraja, B.: Direct measurement of soil chemical properties on-the-go using ion-selective electrodes. Comput. Electron. Agric. **48**(2), 272–294 (2005)
11. Reeves III, J.B., McCarty, G.W., Meisinger, J.J.: Near infrared reflectance spectroscopy for the analysis of agricultural soils. J. Near Infrared Spectrosc. **7**, 179–193 (1999)
12. Reeves III, J.B., McCarty, G.W.: Quantitative analysis of agricultural soils using near infrared reflectance spectroscopy and fibre-optic probe. J. Near Infrared Spectrosc. **9**, 25–34 (2001)

13. Bowers, S.A., Hanks, R.J.: Reflection of radiant energy from soils. Soil Sci. **100**(2), 130–138 (1965)
14. Bogrekci, I., Lee, W.S.: Improving phosphorus sensing by eliminating soil particle size effect in spectral measurement. Trans. ASABE **48**(5), 1971–1978 (2005)
15. 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**, 1175–1180 (2006)
16. Bernard, G.B., Didier, B., Edmond, H.: Deteriming 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 (2008)
17. Christy, C.D.: Real-time measurement of soil attributes using on-the-go near infrared reflectance spectroscopy. Comput. Electron. Agric. **61**, 10–19 (2008)
18. Xing, Z., Yang, Q., Yang, S., et al.: Design of seed pelleter and experimental study on pelleting technology. Agric. Mech. Res. **3**, 98–101 (2008)
19. Sudduth, K.A., Hummel, J.W.: Soil organic matter, CEC, and moisture sensing with a portable NIR spectrophotometer. Trans. ASAE **36**(6), 1571–1582 (1993)
20. Hummel, J.W., Sudduth, K.A., Hollinger, S.E.: Soil moisture and organic matter prediction of surface and subsurface soils using an NIR soil sensor. Comput. Electron. Agric. **32**(2), 149–165 (2001)