Wavelengths selection of near infrared diffuse reflection spectroscopy analysis for soil organic matter

T. Pan,* H. Chen, J. Xie, H. Yin and X. Chen

Key Laboratory of Optoelectronic Information and Sensing Technologies of Guangdong Higher Educational Institutes , Department of Optoelectronic Engineering, Jinan University, Guangzhou 510632, P.R. China.

**E-mail: tpan@jnu.edu.cn*

Keywords: near infrared spectral analysis, soil, organic matter, MWPLS, ECMWMLR

Introduction

The rapid and simple analytical method for soil organic matter by near-infrared spectral (NIRS) analysis is very important. Wavelength selection of NIRS analysis is a key factor in improving model prediction effectiveness, reducing model complexity and designing special NIR spectroscopy instruments with high SNR (signal to noise ratio). In this paper, the equidistant combination moving window multiple linear regression (ECMWMLR) method for wavelength selection of NIRS analysis for soil organic matter is described.

Continuous spectral bands and discrete wavelength combinations for NIR spectroscopic analysis of soil organic matter were selected by moving window partial least square (MWPLS)[1] and ECMWMLR respectively. The selected spectral bands and wavelength combinations were preprocessed by Savitzky-Golay (SG)[2-6] smoothing and further, to establish PLS and MLR models respectively.

Experiment and methods

A total of 91 farmland soil samples were collected. The organic matter contents of the samples were measured by routine chemical methods. The organic matter chemical values ranged from 0.676 to 2.830 (%), and the mean and the standard deviation values were 1.622 and 0.520 (%) respectively. The NIR spectra were measured from 400 to 2500 nm by a FOSS XDS rapid liquid grating spectrometer, and a diffuse reflection accessory.

For MWPLS, the number of adopted wavelengths (N_M), beginning wavelength (B) and PLS factor (F) were set from 2 to 4200, from 400 to 2499 nm and from 1 to 30 respectively. The PLS

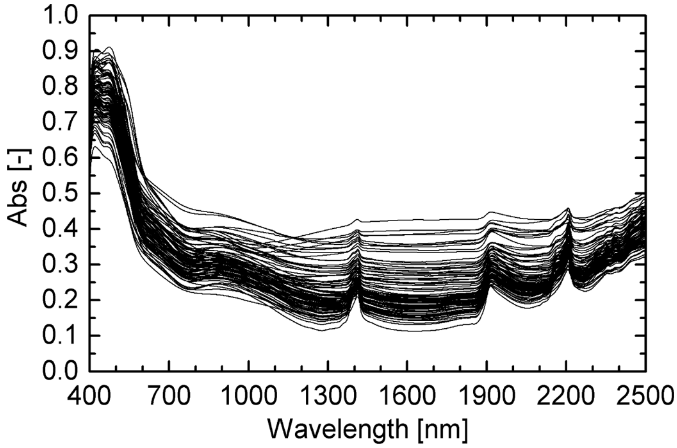

Figure 1. Near-infrared spectra of 91 soil samples.

model corresponding to each parameter combination (N_M, B, F) was established, and the optimal spectral regions and the optimal PLS factors were selected by the prediction statistics.

The ECMWMLR method extracted equidistant data in a continuous spectral array to establish an MLR model. The moving window method was applied to the selection of equidistant wavelengths. For ECMWMLR, the number of adopted wavelengths (N_E), gap of adopted wavelengths (G) and beginning wavelength (B) were set from 1 to 50, from 1 to 100 and from 400 to 2500 nm respectively. The MLR model corresponding to each parameter combination (N_E, G, B) was established, and the optimal wavelength combination was selected on the basis of the prediction statistics.

The model evaluation indicators include root mean squared error of prediction (*RMSEP*), correlation coefficient of prediction (R_P) and the relative root mean squared error of prediction (*RRMSEP%*). The *RMSEP* was used as the goal of model optimization and parameter design.

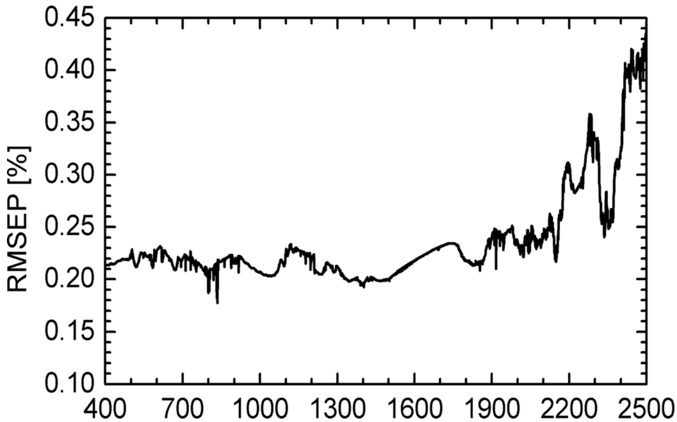

Figure 2. *RMSEP* of optimal models corresponding to each beginning wavelength by MWPLS.

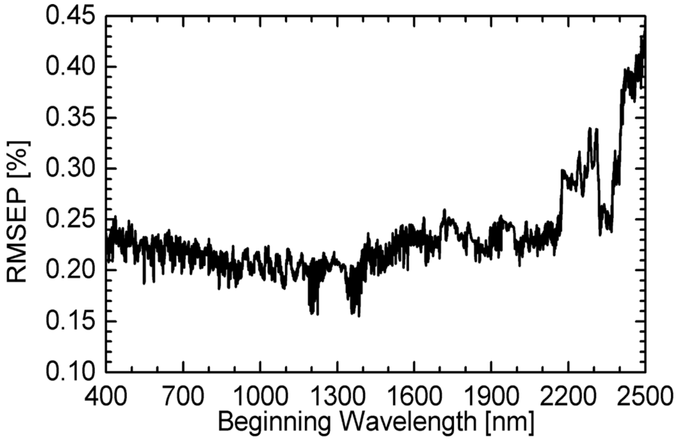

Figure 3. *RMSEP* of optimal models corresponding to each beginning wavelength by ECMWMLR.

Results and discussion

The NIR spectra of 91 samples are shown in Figure 1.

Based on prediction statistics of an optimal single wavelength model, all samples were divided into a calibration set (64 samples) and a prediction set (27 samples).

The *RMSEP* of the optimal models corresponding to each beginning wavelength by MWPLS and by ECMWMLR were shown in Figure 2 and Figure 3, respectively.

By using ECMWMLR and MWPLS methods, the wavelength combinations and spectral bands for NIRS analysis of soil organic matter were selected. The selected wavelength combinations and spectral bands were preprocessed by SG smoothing, and further treatment, to establish

Table 1. Selected spectral regions, wavelength combinations and their prediction effects.

	MWPLS			ECMWMLR		
Wavelength (nm)	835–846.5	1402–1631.5	1391.5–1651	Begin	1385	1223.5
				End	1622.5	1315.5
N_M	24	460	520	N_E	26	24
				G	19	8
OD	3	0	0	-	0	0
DP	5	2	4	-	2	6
NSP	9	77	65	-	7	15
F	21	20	16	-	-	-
RMSEP(%)	0.174	0.182	0.180	-	0.153	0.157
R_P	0.940	0.936	0.938	-	0.958	0.951
RRMSEP(%)	11.1	11.6	11.5	-	9.7	10.0

MLR and PLS models. Selected wavelength combinations, spectral regions and their prediction statistics are summarized in Table 1.

The results show that the optimal prediction statistics obtained by ECMWMLR were slightly better than those of the MWPLS model.

Conclusion

The optimal models for NIRS analysis of soil organic matter were selected by MWPLS and ECMWMLR. For the wavelength combination of the optimal model by ECMWMLR, N_E was 26, G was 19, B was 1385 nm, the smoothing mode was the original spectral smoothing, 2, 3 degree polynomial, and 7 smoothing points. The prediction correlation coefficient R_P was 0.958, *RMSEP* was 0.153 % and *RRMSEP* reached 9.7%. For the optimal model obtained by MWPLS, the optimal wavelength band was 835–846.5 nm, N_M was 24, smoothing mode was 3rd order derivative smoothing, 5, 6 degree polynomial, 9 smoothing points, F was 21, the prediction correlation coefficient R_P was 0.940, *RMSEP* was 0.174 %, and the *RRMSEP* reached 11.1%. The results show that the optimal prediction effectiveness obtained by ECMWMLR was slightly better than was obtained by MWPLS. The work showed that ECMWMLR can be effectively applied to the wavelength selection for NIR analysis of soils.

Acknowledgements

This work was supported by NSF of China (10771087), the NSF of Guangdong (7005948), the Science and Technology Project of Guangdong Province (2007B020714001, 2007A020905001), the Science and Technology Project of Guangzhou (2007Z3-E0281). *Corresponding author: Tao Pan (tpan@jnu.edu.cn)

References

1. J. Jiang, R.J. Berry, H.W. Siesler and Y. Ozaki, *Anal. Chem.* **74,** 3555 (2002).
2. A. Savitzky and M.J.E. Golay, *Anal. Chem.* **36,** 1627 (1964).
3. K. Nakanishi, A. Hashimoto and T. Pan, *et al.*, *Appl. Spectrosc.* **57,** 1510 (2003).
4. T. Pan, A. Hashimoto, M. Kanou, *et al.*, *Bioprocess Biosyst. Eng.* **26,** 133 (2003).
5. J. Chen, T. Pan and X. Chen, *Optics Precision Eng.* **14,** 1 (2006).
6. P. Cao, T. Pan and X. Chen, *Optics Precision Eng.* **15,** 1952 (2007).