2011 3rd International Conference on Environmental
Science and Information Application Technology (ESIAT 2011)

Spectroscopic Calibration for Soil N and C Measurement at a Farm Scale

Haiqing Yang[*]

College of Information Engineering, Zhejiang University of Technology, Hangzhou 310032, P.R.China

Abstract

Calibration modeling is necessary to implement spectroscopic measurement of soil properties. The aim of this study is to compare the performance of calibration models developed for an agricultural farmland using two spectrophotometers with different scanning ranges, e.g. 400-2100nm and 400-2500nm. Soil samples were collected from an experimental farm at Silsoe, Bedfordshire, United Kingdom. The targeted soil properties under investigation were total nitrogen (TN) and organic carbon (OC). Spectra were divided into a calibration set (75%) and an independent validation set (25%). A partial least squares regression (PLSR) with leave-one-out cross validation was carried out to build calibration models based on the two spectral ranges. Validation result shows that the PLSR models developed on the range of 400-2500nm outperform those on 400-2100nm. It suggests that wavelengths located in the spectral band of 2100-2500nm contribute to model calibration of soil N and C. It is concluded that spectroscopic calibration using 400-2500nm could produce higher prediction accuracy for soil N and C measurement at a farm scale than short spectral ranges.

Selection and/or peer-review under responsibility of Conference ESIAT2011 Organization Committee.

1. Introduction

Increasing attentions are paid to the development of fast measurement approaches of soil carbon (C), due to the growing concerns about the increase in atmospheric C content, which could be limited through soil C sequestration. Fast estimate of soil nitrogen also arouses intensive attentions from academia for precision management of nitrogen fertilizer inputs to farmland. Due to spatial variability and non-linear temporal dynamics, accurate estimation of soil N and C requires measurement on numerous samples. This

[*] Corresponding author. Tel.: +86-571-56330001; fax: +86-571-56330001.
E-mail:yanghq@zjut.edu.cn.

would be impossible as standard procedures for soil N and C measurement are time-consuming and expensive[1]. Therefore, alternatives such as near-infrared diffuse reflectance spectroscopy (NIR-DRS) become welcome to scientists and engineers. NIR-DRS technique mainly measures overtones and combination bands of fundamental vibrations of O-H, N-H, and C-H bonds in the mid-infrared region.

The factors affecting the accuracy of soil property measurement originate from soil heterogeneity [2], moisture content [3], soil texture [4,5], soil color [6], model size [7], sample pretreatment [8,9], data preprocessing [10,11], and calibration procedures [12]. The effect of different spectral ranges for the prediction of soil properties has been examined [13,14]. Islam *et al*.[13] compared the ability of reflectance spectroscopy in the UV (250-400 nm), VIS (400-700 nm) and NIR (700-2500nm) to predict several soil properties including organic carbon (OC). Viscarra Rossel *et al*. [14] analyzed the capability of using VIS (400-700 nm), NIR (700-2500 nm), MIR (2500-25,000 nm) and combined VIS-NIR-MIR for simultaneous assessment of soil OC and other properties. They found the cross-validated prediction of soil OC based on the VIS and NIR range of similar accuracy (R2-adj of 0.60), but less accurate than the MIR and combined VIS-NIR-MIR models (R2-adj > 0.72).

The objective of this study is to apply two commercial spectrophotometers with different scanning ranges to measure soil total nitrogen (TN) and organic carbon (OC) at a farm scale with an emphasis on comparing the performance of calibration models developed using different spectral ranges.

2. Materials and methods

2.1. Soil samples

Soil samples were collected from the top layer of an experimental farm at Silsoe, Bedfordshire, United Kingdom. A total of 122 samples with variable proportions of sand, slit and clay were examined. Samples were air-dried at 40℃ for 24h, crushed and sieved to pass a 2-mm mesh. Plant residues and stones were removed. The sieved samples were then air-dried again at 40℃ for 48h. About 50 mg in each sample was used for the measurement of total nitrogen (TN) and organic carbon (OC) by a TrusSpecCNS spectrometer (LECO Corporation, St. Joseph, MI, USA) using the Dumas combustion method.

2.2. Spectrophotometers and spectral acquisition

About 5g in each sample was loaded into a static ring cup and measured separately by a mobile fiber-type AgroSpec Vis-NIR spectrophotometer (Tec5 Co., Germany) with spectral range of 350-2200nm and a LabSpec 2500 spectrophotometer (ASD Inc. Boulder, CO, USA) with spectral range of 350-2500nm. Both instruments scan samples at 1-nm interval. A 100% white reference was used before spectral scanning. Ten scans were recorded for each sample, followed by another ten scans of the re-filled sample cup. The twenty scans were then combined and averaged as one sample.

2.3. Data processing techniques

The spectra were treated with the Unscrambler® (CAMO Inc., Oslo, Norway). Spectral pretreatment methods include moving average smoothing (MA), multiplicative scatter correction (MSC), standard normal variate (SNV), baseline offset correction (BOC), 1st and 2nd de-trending, 1st and 2nd derivatives and three types of spectral normalization.

To eliminate the noise in spectra, only ranges of 400-2100nm and 400-2500nm were adopted in the study. After spectral transformation, the spectra were divided into a calibration set (70%) and an independent validation set (30%). The calibration spectra were subjected to a partial least squares regression (PLSR) with leave-one-out cross validation. The optimal number of latent variables (LVs) was determined by minimizing the predicted residual error sum of squares (PRESS). The performance of the

calibration models was assessed using the coefficient of determination (R^2) and root-mean-square error in calibration (RMSEC) and cross validation (RMSECV). The coefficient of determination (R^2), root-mean-square error of the prediction (RMSEP), Bias and residual predictive deviation (RPD) were used for evaluating prediction performance of the established PLSR models for the independent validation set. We adopted the criteria of classifying RPD values [15] as follows: an RPD value below 1.5 indicates that the calibration is not usable; an RPD value between 1.5 and 2.0 indicates a possibility to distinguish between high and low values; an RPD value between 2.0 and 2.5 makes approximate quantitative predictions possible. For RPD value between 2.5 and 3.0 and above 3.0, the prediction is classified as good and excellent, respectively. Generally, a good model would have high values of R^2 and RPD, and low values of RMSEC, RMSECV, RMSEP and Bias.

3. Results and discussion

3.1. Reference measurement of soil properties

Table 1 shows the sample statistics of calibration set and independent validation set. The two NIR-spectrally active soil properties of TN and OC have similar coefficient of variation (CV) of 0.26-0.29 and high inter-correlation of 0.99.

Table 1 Sample statistics for calibration and independent validation sets

Soil properties	Calibration set				Independent validation set				Inter-correlation	
	Mean	Range	SD[b]	CV[c]	Mean	Range	SD	CV	TN/%	OC/%
TN (%)	0.20	0.09 - 0.31	0.057	0.29	0.20	0.12-0.30	0.055	0.28	1.00	0.99
TC (%)	1.97	0.85 - 3.02	0.544	0.28	1.99	1.24-2.84	0.526	0.26		1.00

[b] standard deviation;
[c] coefficient of variation

3.2. Spectral characteristics

Figure 1 shows the spectra measured by two spectrophotometers. It is observed that spectra remain similar shapes in the common range of 400-2100nm. The extension part in the range of 2100-2500nm provided by ASD spectrophotometer exhibits new information unavailable in Tec5 instrument, which may allow for contributing more information to build accurate calibration models
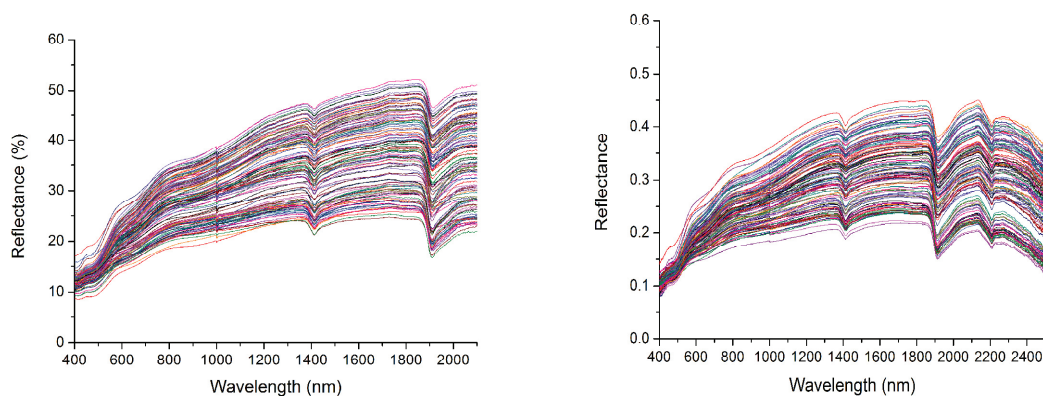

Fig.1 Spectra measured by a mobile fiber-type AgroSpec Vis-NIR spectrophotometer (Tec5 Co., Germany) with spectral range of 350-2200nm(left); and by a LabSpec 2500 spectrophotometer (ASD Inc. USA) with spectral range of 350-2500nm(right).

3.3. PLSR models on 400-2100nm

Table 2 lists the results of the PLSR models developed for the prediction of TN based on original and transformed spectra. Except for 2nd derivative spectra, PLSR models for TN provided good performance with R^2 of 0.81-0.86 and RMSECV of 0.021-0.025% for the cross-validation of calibration set, and R^2 of 0.84-0.89, RMSEP of 0.018-0.022% and RPD of 2.59-2.98 for the independent validation set.

Table 2 PLSR models for soil TN with spectral range of 400-2100nm

Spectral transformation	LVs	Calibration		Cross-validation		Independent validation			
		R^2	RMSEC	R^2	RMSECV	R^2	RMSEP	Bias	RPD
None	3	0.84	0.022	0.81	0.025	0.85	0.021	0.001	2.59
Max normalization	2	0.86	0.021	0.84	0.022	0.87	0.020	-0.003	2.77
Range normalization	3	0.87	0.020	0.85	0.022	0.88	0.019	-0.002	2.87
Mean normalization	**2**	**0.88**	**0.020**	**0.86**	**0.021**	**0.89**	**0.018**	**-0.002**	**2.98**
MA+SNV	3	0.87	0.020	0.86	0.021	0.88	0.019	-0.002	2.90
MA+MSC	3	0.86	0.021	0.85	0.022	0.87	0.019	-0.002	2.81
1st De-trending	2	0.84	0.023	0.83	0.024	0.88	0.019	-0.004	2.95
2nd De-trending	1	0.83	0.023	0.82	0.024	0.84	0.022	-0.004	2.57
BOC	3	0.85	0.022	0.83	0.023	0.85	0.021	-0.001	2.61
1^{st}derivative	2	0.84	0.023	0.82	0.024	0.88	0.019	-0.003	2.98
2nd derivative	3	0.87	0.021	0.70	0.031	0.71	0.029	-0.004	1.87

Table 3 lists the results of the PLSR models developed for the prediction of OC based on original and transformed spectra. Except for original and 2nd derivative spectra, PLSR models for OC provided good performance with R^2 of 0.82-0.87 and RMSECV of 0.199-0.236% for the cross-validation of calibration set, and R^2 of 0.84-0.88, RMSEP of 0.177-0.194% and RPD of 2.50-2.93 for the independent validation set.

Table 3 PLSR models for soil OC with spectral range of 400-2100nm

Spectral transformation	LVs	Calibration		Cross-validation		Independe validation			
		R^2	RMSEC	R^2	RMSECV	R^2	RMSEP	Bias	RPD
None	3	0.84	0.214	0.82	0.232	0.82	0.219	0.033	2.28
Max normalization	3	0.87	0.194	0.86	0.207	0.86	0.194	0.011	2.65
Range normalization	3	0.88	0.187	0.87	0.201	0.87	0.184	0.007	2.80
Mean normalization	2	0.88	0.188	0.87	0.201	0.87	0.187	0.015	2.74
MA+SNV	3	0.88	0.186	0.87	0.199	0.87	0.189	-0.090	2.74
MA+MSC	3	0.88	0.190	0.86	0.204	0.86	0.191	-0.002	2.70
1st De-trending	2	0.83	0.226	0.82	0.236	0.88	0.180	-0.022	2.89
2nd De-trending	1	0.82	0.227	0.82	0.235	0.85	0.198	-0.029	2.64
BOC	3	0.86	0.206	0.84	0.219	0.84	0.206	0.011	2.50
1^{st}derivative	**2**	**0.83**	**0.221**	**0.82**	**0.234**	**0.88**	**0.177**	**-0.019**	**2.93**
2nd derivative	4	0.90	0.174	0.70	0.300	0.80	0.232	-0.005	2.22

3.4. PLSR models on 400-2500nm

Table 4 and 5 list the result of PLSR models developed for the prediction of TN and OC based on original and transformed spectra. Except for derivative spectra, PLSR models for TN provided excellent prediction performance with R^2 of 0.87-0.92 and RPD of 3.00-3.63. Similarly, PLSR models for OC provided good and excellent performance with R^2 of 0.86-0.90 and RPD of 2.69-3.21 for the independent validation set.

Table 4 PLSR models for soil TN with spectral range of 400-2500nm

Spectral transformation	LVs	Calibration		Cross-validation		Independe validation			
		R^2	RMSEC	R^2	RMSECV	R^2	RMSEP	Bias	RPD
None	6	0.90	0.018	0.85	0.022	0.88	0.019	-0.008	3.15
Max normalization	5	0.91	0.017	0.88	0.020	0.91	0.016	-0.005	3.60
Range normalization	5	0.91	0.017	0.89	0.019	0.91	0.016	-0.005	3.49
Mean normalization	4	0.91	0.017	0.88	0.019	0.91	0.016	-0.005	3.58
MA+SNV	**3**	**0.90**	**0.018**	**0.89**	**0.019**	**0.92**	**0.015**	**-0.004**	**3.63**
MA+MSC	4	0.91	0.017	0.88	0.020	0.88	0.019	-0.006	3.00
1st De-trending	5	0.91	0.017	0.88	0.020	0.87	0.020	-0.009	3.06
2nd De-trending	3	0.87	0.020	0.85	0.022	0.89	0.018	-0.006	3.19
BOC	4	0.87	0.020	0.84	0.023	0.88	0.018	-0.005	3.08
1st derivative	2	0.88	0.020	0.82	0.024	0.82	0.023	-0.006	2.47
2nd derivative	1	0.67	0.032	0.54	0.039	0.54	0.037	-0.004	1.49

By comparison, these models for TN and OC developed on spectral range of 400-2500nm outperform those on spectral range of 400-2100nm (Table 2 and 3). The result is line with the report by Mouazen *et al*.[15], in which they compared the performance of two commercially-available spectrophotometers with different wavelength ranges for the measurement of selected soil attributes including total carbon (TC) and total nitrogen (TN). They found that the best accuracy is obtained when using a full wavelength range of 451-2459 nm, as compared to a short wavelength range of 401-1770 nm.

Table 5 PLSR models for soil OC with spectral range of 400-2500nm

Spectral transformation	LVs	Calibration		Cross-validation		Independe validation			
		R^2	RMSEC	R^2	RMSECV	R^2	RMSEP	Bias	RPD
None	4	0.89	0.177	0.88	0.191	0.86	0.188	0.009	2.76
Max normalization	3	0.92	0.154	0.91	0.167	0.89	0.163	0.009	3.18
Range normalization	5	0.93	0.145	0.91	0.165	0.89	0.165	0.004	3.13
Mean normalization	4	0.92	0.157	0.90	0.173	0.89	0.171	-0.012	3.03
MA+SNV	**3**	**0.92**	**0.152**	**0.91**	**0.163**	**0.90**	**0.161**	**-0.003**	**3.21**
MA+MSC	4	0.92	0.152	0.90	0.169	0.88	0.175	0.002	2.96
1st De-trending	2	0.88	0.188	0.87	0.197	0.86	0.193	-0.007	2.69
2nd De-trending	4	0.91	0.160	0.90	0.177	0.87	0.184	0.012	2.81
BOC	4	0.89	0.174	0.88	0.190	0.87	0.181	0.005	2.86
1st derivative(7)	2	0.89	0.174	0.86	0.207	0.83	0.208	-0.003	2.48
2nd derivative(7)	2	0.82	0.232	0.58	0.356	0.59	0.325	-0.004	1.59

4. Conclusions

In this study, the visible-near-infrared (VIS-NIR) spectroscopy was applied to build calibration models for predicting soil total nitrogen (TN) and organic carbon (OC) at a farm scale. Two spectrophotometers were used to obtain different ranges of spectral data. Model performance was compared between two spectral ranges, e.g., 400-2100nm and 400-2500nm. Calibration models were built using a partial least squares regression (PLSR) with leave-one-out cross validation. Validation result of the established models shows that PLSR models for soil TN and OC developed on the range of 400-2500nm outperform those on the range of 400-2100nm. It suggests that wavelengths located in the spectral band of 2100-2500nm contribute to model calibration of soil TN and OC. It is concluded that spectroscopic calibration using 400-2500nm could produce higher prediction accuracy for soil N and C measurement than short spectral ranges.

Acknowledgements

References

[1] Sinfield JV, Fagerman D, Colic O. Evaluation of sensing technologies for on-the-go detection of macro-nutrients in cultivated soils. *Computers and Electronics in Agriculture* 2010;**70**:1-18.

[2] Brunet D, Barthès BG, Chotte JL, Feller C. Determination of carbon and nitrogen contents in Alfisols, Oxisols and Ultisols from Africa and Brazil using NIRS analysis: Effects of sample grinding and set heterogeneity. *Geoderma* 2007;**139**:106-17.

[3] Mouazen AM, Karoui R, De Baerdemaeker J, Ramon H. Characterization of soil water content using measured visible and near infrared spectra. *Soil Science Society of America Journal* 2006;**70**:1295-302.

[4] Cozzolino D, Morón A. Potential of near-infrared reflectance spectroscopy and chemometrics to predict soil organic carbon fractions. *Soil and Tillage Research* 2006;**85**:78-85.

[5] Barthès BG, Brunet D, Hien E, Enjalric F, Conche S, Freschet GT, d'Annunzio R, Toucet-Louri J. Determining the distributions of soil carbon and nitrogen in particle size fractions using near-infrared reflectance spectrum of bulk soil samples. *Soil Biology and Biochemistry* 2008;**40**:1533-7.

[6] Mouazen AM, Karoui R, Deckers J, De Baerdemaeker J, Ramon H. Potential of visible and near-infrared spectroscopy to derive colour groups utilizing the Munsell soil colour charts. *Biosystems Engineering* 2007;**97**:131-43.

[7] Guerrero C, Zornoza R, Gómez I, Mataix-Beneyto J.Spiking of NIR regional models using samples from target sites: Effect of model size on prediction accuracy. *Geoderma* 2010;**158**:66-77.

[8] Yang H, Kuang B, Mouazen AM. Selection of preprocessing parameters for PCA of soil classification affected by particle sizes based on Vis-NIR spectroscopy. *Key Engineering Materials Journal* 2011;**467-469**:725-30.

[9] Terhoeven-Urselmans T, Schmidt H, Georg Joergensen R, Ludwig B. Usefulness of near-infrared spectroscopy to determine biological and chemical soil properties: Importance of sample pretreatment. *Soil Biology and Biochemistry* 2008; **40**:1178-88.

[10] Maleki MR, Mouazen AM, Ramon H, De Baerdemaeker J. Multiplicative scatter correction during on-line measurement with near infrared spectroscopy. *Biosystems Engineering* 2007;**96**:427-433.

[11] Yang H, Kuang B, Mouazen AM. Selection of spectral preprocessing methods for soil texture classification. *Advanced Materials Research Journal* 2011;**181-182**:416-21.

[12] Mouazen AM, 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* 2010;**158**:23-31.

[13] Islam K, Singh B, McBratney AB. Simultaneous estimation of several soil properties by ultra-violet, visible, and near-infrared reflectance spectroscopy. *Australian Journal of Soil Research* 2003;**41**:1101-14.

[14] Viscarra Rossel RA, Walvoort DJJ, McBratney AB, Janik LJ, Skjemstad JO. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* 2006;**131**:59-75.

[15] Mouazen AM, Baerdemaeker JD, Ramon H. Effect of wavelength range on the measurement accuracy of some selected soil constituents using visible-near-infrared spectroscopy. *Journal of Near Infrared Spectroscopy* 2006;**14**:189-99.