12

Vis/Near- and Mid- Infrared Spectroscopy for Predicting Soil N and C at a Farm Scale

Haiqing Yang[1] and Abdul M. Mouazen[2]
[1]*College of Information Engineering, Zhejiang University of Technology, Hangzhou*
[2]*School of Applied Sciences, Cranfield University, Bedfordshire*
[1]*P.R. China*
[2]*United Kingdom*

1. Introduction

a. In spectroscopic analysis, visible (Vis), near infrared (NIR) and mid infrared (MIR) ranges are often used as they include plenty of information on physical, chemical and biological properties of objects. Commonly, wavelengths ranges are from 350 to 760 nm for Vis, 760-2500 nm for NIR, and 2500 to 25000 nm for MIR (often used in its wavenumber form 4000 to 400 cm^{-1}). Frequencies in the Vis are due to electronic transition while those in the NIR are generally overtones and combination bands from the fundamental vibrations occur in the MIR, mainly O-H, N-H, and C-H bonds (Viscarra Rossel, *et al.*, 2006). When NIR and MIR radiations are focused onto a sample, the molecules in the sample will increase their vibration energy by absorbing energy at specific frequencies depending on the molecular geometry, bond strengths and atomic masses. The resulting Vis, NIR and MIR lights are thus modified, creating a spectrum or 'signatures' of the targeted object with peaks at the absorbing frequencies.
b. The combined contributions from the various soil components can result in a very complex spectrum, difficult to analyze visually, but multivariate calibration models can be built to derive useful qualitative and quantitative relationships or models between the spectral signatures and many soil properties. Spectrometry is the combination of spectroscopy and chemometric (multivariate statistical) methods. It should be noted that the Vis-NIR-MIR spectrometry technique can predict multiple soil properties simultaneously.
c. Recently, there is an increasing interest in the development of time- and cost-effective methods for the measurement of soil nitrogen (N) and carbon (C), due to the growing concerns about the effect of excessive use of N fertilizer in the environment and the increase of atmospheric C content, which could be limited through soil C sequestration. In order to manage N and C in soils in an efficient manner detailed information about these properties is needed. Previous reports confirm the presence of within-field variability of soil properties including N and C, which requires analysis of large number of soil samples (Mouazen, *et al.*, 2007). Due to the fact that standard procedures for the measurement of soil N and C are time-consuming and expensive (Sinfield, *et al.*,

2010), attention is being given to possible alternatives such as Vis-NIR and/or MIR spectroscopy. Numerous analyses of soil N and C have been conducted during the past decades using this technique, for examples, to predict the soil C and N mineralization rates (Fystro, 2002; Mutuo, *et al*., 2006), to derive spectral characteristics for classifying conventional and conservation agricultural practices (Haché, *et al*., 2007), to assess soil changes due to site disturbance during forest harvesting (Ludwig, *et al*., 2002), to evaluate the recovery of microbial functions during soil restoration (Schimann, *et al*., 2007), to determine carbon inventories (Reeves III, *et al*., 2002), to determine (*in situ*) organic matter composition of coatings at crack surfaces and linings of earthworm burrow walls (Reeves III, *et al*., 2002) and others (Chang, *et al*., 2001; Chang & Laird, 2002; Yang, *et al*., 2011a).

d. Applying Vis-NIR spectroscopy to predict soil properties needs no special sample preparation. However, MIR spectra are traditionally obtained by a FT-IR spectrometer with samples pressed in KBr pellets, which requires labour and specific skills. Fortunately, newly-developed ATR (attenuated total reflection) and DRIFT (diffuse reflectance infrared Fourier Transform) accessories are becoming the predominant FT-IR sample analysis tool. This is because sample handling is greatly simplified and sample preparation is eliminated. Hence, Vis-NIR-MIR spectrometry without sample preparation will bring about new wave of soil research.
e. This study investigated the potential of calibrating Vis-NIR, ATR-FTIR and DRIFT spectra to soil N and C concentrations with an aim of comparing the performance of two spectrometers, namely, a Vis-NIR spectrometer vs. a FT-IR spectrometer with ATR and DRIFT accessories. The models developed for N and C were then compared to those developed with the combination of Vis-NIR and ATR-FTIR spectra (Vis-NIR-ATR) and the combination of Vis-NIR and DRIFT spectra (Vis-NIR-DRIFT) for investigating whether the combination of Vis-NIR spectrometer and FT-IR spectrometer could improve the prediction accuracy of soil N and C. For each spectrometer, spectral data were subjected to various spectral transformation approaches before model calibration, aiming at model optimization.

2. Material and methods

2.1 Samples

Samples archived in the Soil Labs at Cranfield University were originally collected from the top 0-20 cm of the soil layer from five fields in Silsoe experimental farm at Bedfordshire, United Kingdom. Figure 1 illustrates the location of these fields, namely, Avenue Field (#A), Orchard (#B), Ivy Ground (#C), Showground (#D) and Copse Field (#E). According to the soil descriptions presented on www.landis.org.uk, managed by the National Soil Resources Institute (MSRI), Cranfield University, these fields belong to two major soil World Reference Base (WRB) classifications, namely, Cambisol and Luvisol. Of them, Ivy Ground (Sample codes: C15, C21-C39), Orhcard (B01-B25) and Copse Field (E01- E23) are Cambisols and Showground (D01-D35) is Luvisol, while Avenue Filed (A01-A12, A14-A20) comprises of both soils. The parent material underlying these fields is mainly siliceous stones.

A total of 122 bulk soil samples used in this study are with various proportions of sand, slit and clay (Table 1) and hence belong to three soil textures, e.g. sandy loam, clay loam and clay, according to the United States Department of Agriculture (USDA) triangular diagram

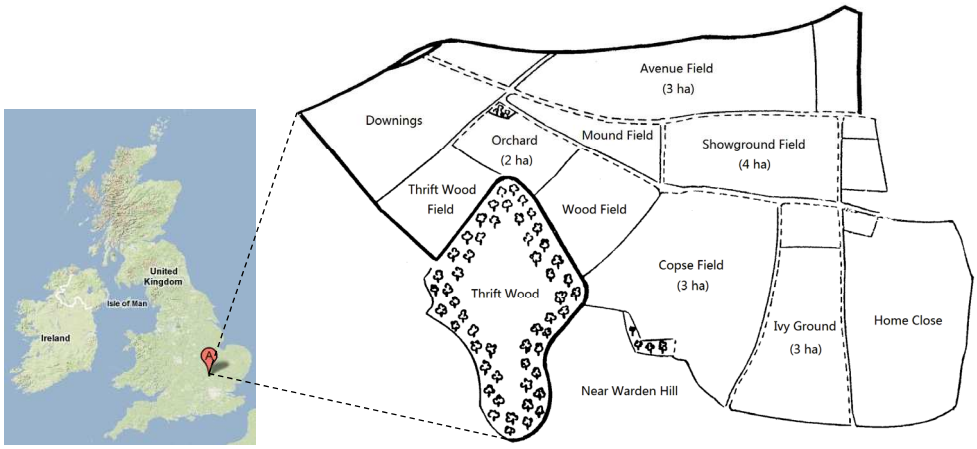


Fig. 1. Location of the five fields targeted in the study (www.landis.org.uk)

Code	Field Name	Vegetation	Soil type	Sand (%)	Silt (%)	Clay (%)	Soil texture[a]
A	Avenue Ground	wheat	C+L[b]	61.9	20.1	18.1	Sandy loam
B	Orchard	wheat	Cambisol	40.1	27.4	32.5	Clay loam
C	Ivy Ground	Soybean	Cambisol	21.1	27.2	51.7	Clay
D	Showground	wheat	Luvisol	65.0	20.9	14.1	Sandy loam
E	Copse Field	wheat	Cambisol	14.6	27.8	57.6	Clay

[a] according to USDA triangular diagram relating particle size distribution to soil texture.
[b] mixture of Cambisol and Luvisol types.

Table 1. Description of the five targeted farm fields

of soil texture classification. Soil samples were air-dried and crushed at first. Plant residues and stones were then removed. After that, the samples were sieved to pass a 2 mm mesh and air-dried again at 40℃ for 48h. A small amount of soil was used for chemical analysis, whereas the majorities were left for spectrophotometer measurement.

2.2 Reference methods

Particle size distribution was determined by a combination of wet sieving and hydrometer tests using the USDA soil texture classification system. Reference values of N and C were analyzed through a sequence of processes. First, a 50 mg sample was used for the measurement of TN and TC by a TrusSpecCNS spectrometer (LECO Corporation, St. Joseph, MI, USA) using the Dumas combustion method. Next, another 50 mg from each soil sample was mixed with 5% HCl and then oven-dried at 90℃ for 4 h in order to remove IC. Then, OC in IC-free samples were measured by the same Dumas combustion method. Finally, IC was calculated by the difference between TC and OC.

2.3 Vis-NIR spectra acquisition

The soil samples were equilibrated to room temperature (20-25℃) and carefully mixed before spectral measurement. A sub-sample of ~5g was loaded into a static ring cup and measured using a LabSpec 2500 spectrophotometer (Analytical Spectral Devices Inc. Boulder, CO, USA) equipped with a fibre-optic probe. The light source was a quartz-halogen bulb of 3000K°. The light source and reflectance fibre were gathered with a certain angle of 35°. One Si photodiode array in the range of 350-1000 nm and two Peltier cooled InGaAs detectors in the ranges of 1000-1800 nm and 1800-2500 nm were used. All spectra were recorded in diffuse reflection mode over the wavelength range of 400-2500 nm at 1 nm data spacing interval, which resulted in 2101 wavelengths per spectrum. The reflectance spectra were transformed into absorbance spectra using Log(1/R), as absorbance is directly proportional to the concentration of an absorber according to Beer-Lambert Law. The actual spectra resolution was 3 nm at 700 nm and 10 nm at 1400 and 2100 nm. Before sample spectral acquisition, twenty five reference scans were taken on a ceramic standard supplied with the spectrophotometer. Ten photometric scans were conducted for each sample, followed by another ten scans of the refilled sample cup. The twenty scans were then averaged in one spectrum for each sample.

2.4 MIR spectra acquisition

MIR spectra were collected by an ALPHA Fourier transform infrared (FT-IR) spectrometer (Bruker Optics, Billerica, MA, USA) with wavelength range of 7500–375 cm^{-1}, equipped with two exchangeable QuickSnap™ sampling modules. This instrument acquired spectra with two sampling accessories, namely, ATR and DRIFT. ATR is an easy-to-use FT-IR sampling method that is ideal for both solids and liquids and does not require any sample preparation. The Eco ATR is a single reflection ATR sampling module equipped with a versatile high throughput ZnSe ATR crystal for the analysis of powders, solids, pastes and liquids. The DRIFT module is an economical analysis option for a broad variety of solid samples: powders, inorganic material, gem stones, papers, textiles and others. The DRIFT module is designed for easy sampling and high light-throughput.

2.4.1 Principles of ATR-FTIR

An attenuated total reflectance accessory operates by measuring the changes that occur in a totally internally reflected infrared beam when the beam comes into contact with a sample (Fig.2). An infrared beam is directed onto an optically dense crystal with a high refractive index at a certain angle. This internal reflectance creates an evanescent wave that extends beyond the surface of the crystal into the sample held in contact with the crystal. It can be easier to think of this evanescent wave as a bubble of infrared that sits on the surface of the crystal. This evanescent wave protrudes only a few microns ($0.5\mu - 5\mu$) beyond the crystal surface and into the sample. Consequently, there must be good contact between the sample and the crystal surface. In regions of the infrared spectrum where the sample absorbs energy, the evanescent wave will be attenuated or altered. The attenuated energy from each evanescent wave is passed back to the IR beam, which then exists at the opposite end of the crystal and is passed to the detector in the IR spectrometer. The system then generates an infrared spectrum (www.perkinelmer.com).

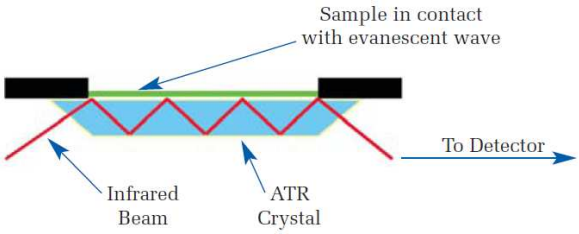


Fig. 2. A multiple reflection ATR system (www.perkinelmer.com)

When measuring solids by ATR, it is essential to ensure good optical contact between the sample and the crystal. The accessories have devices that clamp the sample to the crystal surface and apply pressure. This works well with elastomers and other deformable materials, but many solids give very weak spectra because the contact is confined to small areas. The effects of poor contact are the greatest at shorter wavelengths where the depth of penetration is the lowest. The issue of solid sample/crystal contact has been overcome to a great extent by the introduction of ATR accessories with very small crystals, typically about 2 mm across. The most frequently-used small crystal ATR material is diamond because it has the best durability and chemical inertness. These small area ATR crystal top-plates generally provide only a single reflection but this is sufficient, given the low noise levels of PerkinElmer's modern FT-IR spectrometers. Much higher pressure with limited force can now be generated onto these small areas. As a result, spectra can be obtained from a wide variety of solid materials including minerals.

After the crystal area has been cleaned and the background collected, the soil material is placed onto the small crystal area. Then the pressure arm should be positioned over the crystal/sample area. Force is applied to the sample, pushing it onto the diamond surface. It is good practice to apply pressure until the strongest spectral bands have an intensity which extends beyond 70%T, namely, from a baseline at 100%T down to 70%T. Then, the data are collected in the normal manner. Unlike transmission measurements, ATR sampling does not produce totally absorbing spectral bands because the effective path-length is controlled by the crystal properties thereby minimizing sample re-preparation time. After the spectrum has been collected, the crystal area must be cleaned before placing the next sample on the crystal. A 100%T line with no spectral features should be seen if the crystal is clean, if spectral features are seen, the crystal should be cleaned again using a solvent soaked tissue.

In the case of a solid sample, it is pressed into direct contact with the crystal. Because the evanescent wave into the solid sample is improved with a more intimate contact, solid samples are usually firmly clamped against the ATR crystal, so that trapped air is not the medium through which the evanescent wave travels, as that would distort the results.

2.4.2 Principles of DRIFT

Diffuse reflectance occurs when light impinges on the surface of a material and is partially reflected and transmitted. Light that passes into the material may be absorbed or reflected out again. Hence, the radiation that reflects from an absorbing material is composed of surface-reflected and bulk re-emitted components, which summed are the diffuse

reflectance of the sample (www.uksaf.org). DRIFT analysis of powders is conducted by focusing infrared light onto the powder (sometimes diluted in a non absorbing matrix, e.g. KBr) and the scattered light is collected and relayed to the IR detector.

In practice, DRIFT is most conveniently and rapidly used for soil analysis in diffuse reflection mode, where the incoming radiation is focused onto the soil sample surface, often in the form of a dry powder or <2 mm micro-aggregates, and the reflected radiation is passed back into the spectrophotometer (Fig.3, www.clw.csiro.au). In this study, infrared spectra were recorded in diffuse reflection mode with an Alpha spectrometer with DRIFT accessory. Bulk soil samples were scanned 20 times in the range from 4000 to 400 cm^{-1}. DRIFT spectra were corrected against atmospheric CO_2 and water vapour. Finally, the infrared reflectance spectra were transformed into absorbance spectra using Log(1/R).

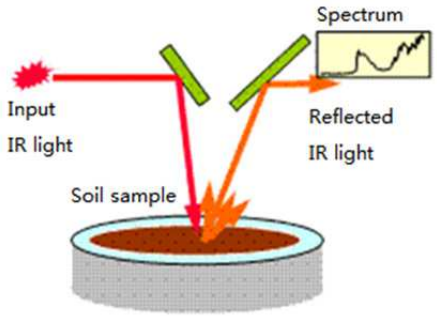


Fig. 3. A description of the method of acquiring a DRIFT spectrum (www.clw.csiro.au)

2.5 Spectral processing and development of calibration models for soil N and C

2.5.1 Principal components analysis (PCA)

PCA is a data compression process (i.e. a bilinear modelling process), which can be used to reduce a complex multidimensional data (e.g. spectra) into a smaller number of principal components (PCs) which reflect the underling structure of the original dataset. The first principal component typically explains most of the variation in the dataset with further principal components being orthogonal to the preceding PC and explaining less variation in the dataset. By plotting the PCs in two or three dimensional data space, interrelationships between the samples and variables can be examined (www.clw.csiro.au).

2.5.2 Partial least-squares regression (PLSR) analysis

For PLSR, spectral information is arranged as N × M matrix which consists of N spectra with absorbance values for M wavelengths, and the calibration data is expressed as a single vector with measured values for these spectra. The PLSR algorithm decomposes the M-dimensional spectra space into few factors termed latent variables (LVs), which represent the best projections of the calibration vector onto the N × M matrix. One of the advantages of PLSR compared to other chemometric methods like PCA is the possibility to interpret the first few LVs, because they show the correlations between the property values and the

spectral features. Furthermore, PLSR takes as well variations of the absorbance as variations of the calibration data into account. PLSR is a rapid analysis, can handle co-linear data, and can provide useful qualitative information.

2.5.3 Procedure of spectral processing and model calibration

Before the absorbance spectra were calibrated to predict soil properties, PCA was conducted to detect sample outliers in raw data set of Vis-NIR spectra, ATR spectra and DRIFT spectra. The identified sample outlier/s was/were excluded from further investigation. The remaining spectra were then subjected to various spectral pre-processing algorithms to reduce or eliminate noise, offset and bias in raw spectra. The investigated spectral pre-processing techniques included Savitzky-Golay smoothing, standard normal variate (SNV), multiplicative scatter correction (MSC), baseline offset correction (BOC), centre & scale, 1st- and 2nd- detrendings, and 1st- and 2nd- derivatives. Several spectral normalizations were also included. They were conducted according to maximum, range, mean and quantile values. Details of these algorithms are available in www.camo.com. PLSR algorithm was used to decompose both raw and transformed spectra matrix into 10 LVs. All PLSR models were validated with full cross-validation approach in which each spectrum was in turn excluded from the calibration sample set and was predicted by the PLSR model calibrated for the remaining spectra. By decomposing the spectra into 10 LVs, it was assumed that the PLSR model would be over-fitted because signal noise of the spectral measurements could also be correlated with the property vector. The optimal number of LVs was determined by minimizing the predicted residual error sum of squares (PRESS). For better understanding the importance of different wavelength ranges in the prediction of soil N and C, PLSR models were also developed for the combinational Vis-NIR-ATR and Vis-NIR-DRIFT spectra. Spectral transformation and model calibration were conducted using the UnscramblerX10.1® (CAMO, Oslo, Norway).

2.6 Model assessment criteria

The validation accuracy of PLSR models is given by the root mean squared error (RMSE):

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{N}}\sum\nolimits_{\mathrm{N}} (\mathrm{X_i} - \mathrm{Y_i})^2}$$

where X_i is the predicted value, Y_i the measured (reference) value and N the number of soil samples. To compare model performance, we recorded the residual predictive deviation (RPD), which is the ratio of standard deviation of reference values to RMSE of the calibration set during cross-validation. The criteria adopted for RPD classification (Mouazen, *et al.*, 2006) was that an RPD value below 1.5 indicates very poor model/predictions and that such as value could not be useful; an RPD value between 1.5 and 2.0 indicates a possibility to distinguish between high and low values, while a value between 2.0 and 2.5 makes approximate quantitative predictions possible. For RPD values between 2.5 and 3.0 and above 3.0, the prediction is classified as good and excellent, respectively. Meanwhile, we compared the coefficient of determination (R^2) in cross-validation of calibration models. Generally, a good model would have high values of R^2 and RPD for cross-validation.

3. Results and discussion

3.1 Laboratory analyses

Means and distributions of the reference values of total nitrogen (TN), total carbon (TC), organic carbon (OC), inorganic carbon (IC), the ratio of TC to TN, and the ratio of OC to TN in samples are summarized in Table 2. The averaged content (±s.d.) for TN, TC and OC are 0.2(±0.06)%, 2.11(±0.57)% and 1.98(±0.54)%, respectively. The IC concentration for the studied samples is very low with an average value of 0.12%, although its coefficient of variance (c.v.) is 1.08. This leads to large skewness of IC content distribution in samples. The ratio of OC to TN is as nearly constant as 10 with very low c.v. value of ~0.05. The inter-correlation coefficients among these properties are summarized in Table 3. TN, TC and OC were strongly correlated to each other (r=0.97~0.99), while they have weak correlation with IC (r=0.13~0.35). The TC/TN or OC/TN was poorly correlated to TN, TC and OC, however, they had good correlation with IC (r=0.53 or -0.49).

Soil property	Reference value[a]				
	Mean	Median	Range	s.d.[b]	c.v.[c]
TN(%)	0.20	0.19	0.09-0.31	0.06	0.30
TC(%)	2.11	2.00	0.95-3.41	0.57	0.27
OC(%)	1.98	1.84	0.85-3.02	0.54	0.27
IC(%)	0.12	0.09	0.00-0.64	0.13	1.08
TC/TN	10.57	10.47	9.56-13.12	0.59	0.06
OC/TN	9.95	9.92	8.88-12.47	0.52	0.05

[a] one sample outlier (A02) detected by PCA was not included.
[b] standard deviation
[c] coefficient of variance(=s.d./Mean)

Table 2. Laboratory reference statistics for soil total nitrogen (TN), total carbon (TC), organic carbon (OC), inorganic carbon (IC), TC/TN and OC/TN

All	TN(%)	TC(%)	OC(%)	IC(%)	TC/TN	OC/TN
TN(%)	1	0.97	0.99	0.23	-0.25	-0.24
TC(%)	0.97	1	0.97	0.35	-0.07	-0.18
OC(%)	0.99	0.97	1	0.13	-0.19	-0.07
IC(%)	0.23	0.35	0.13	1	0.53	-0.49
TC/TN	-0.25	-0.07	-0.19	0.53	1	0.41
OC/TN	-0.24	-0.18	-0.07	-0.49	0.41	1
PC-1(Vis-NIR)	0.77	0.74	0.76	0.13	-0.22	-0.16
PC-2(Vis-NIR)	-0.48	-0.48	-0.46	-0.23	0.08	0.22
PC-1(ATR-FTIR)	0.90	0.87	0.88	0.19	-0.29	-0.28
PC-2(ATR-FTIR)	0.15	0.16	0.21	-0.17	0.05	0.30
PC-1(DRIFT)	0.56	0.55	0.52	0.26	-0.12	-0.24
PC-2(DRIFT)	-0.71	-0.68	-0.71	-0.05	0.24	0.16

Table 3. Correlation matrix between soil properties and the first two principal component scores of the Vis-NIR spectra, ATR-FTIR spectra and DRIFT spectra of 121 samples

3.2 Vis-NIR spectral analysis and model calibrations

Different wavelength bands respond to different chemical compositions or molecular groups in soil. However, this response is strongly influenced by soil texture classes. Figure 4 shows the representative Vis-NIR absorption spectra of samples from each soil texture class, e.g. sandy loam (field #D), clay loam (field #B) and clay (fields #C and #E), with high and low TC content. The shift in overall baseline in the Vis-NIR spectra is likely caused by the overall difference in the particle size distribution (Madari, *et al.*, 2006). The clay soil has a finer texture compared to the others, which results in a higher baseline. In general, smaller particle size results in higher reflectance or lower absorbance, but for our case, the higher absorption coefficients for the clay fraction apparently dominate the particle size effect resulting in higher absorbance. Within the clay texture class, the samples with higher TC content tend to exhibit stronger absorption in Vis-NIR spectra than those with lower TC contents (Fig.4). This observation seems true for sandy loam soils but not for clay loam soils, which might be attributed to the effect of soil colour. In Fig.4, the sample from clay loam class with low TC content of 1.59% shows higher absorbance than that with high TC content of 2.44%. This is probably due to particle size effect. The Vis-NIR spectra are characteristic of absorption bands associated with colour (400-760 nm), the bending (1413 nm) and stretching (1916 nm) of the O-H bonds of free water and lattice minerals at around 2210 nm (Madari, *et al.*, 2006).

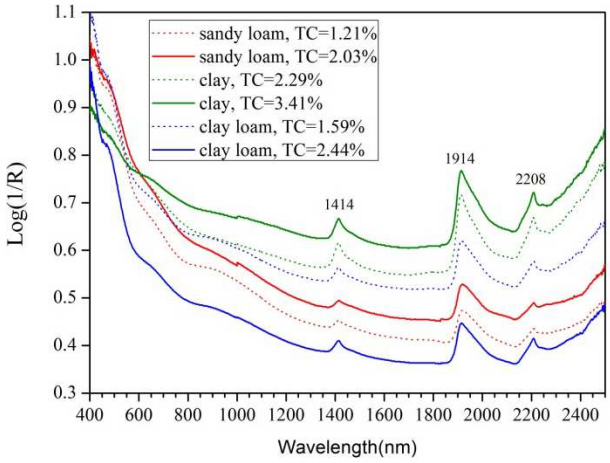


Fig. 4. Vis-NIR absorption spectra of samples from each soil texture class with high and low TC content.

3.2.1 PCA analysis for Vis-NIR spectra

Figure 5 shows all raw Vis-NIR spectra, PCA scores plot for the spectra, and residual X-variance for PC-1 and PC-2. PC-1 and PC-2 explains 96% and 3% of total variance, respectively. The PC-1 may explain variation related to SOM of the samples, as the PC-1 was better correlated to TN, TC and OC (r=0.74~0.76) than to PC-2 (r =-0.46~-0.48) (Table 3). Samples originated from different fields can be divided into two clusters: one for Luvisol soils (Showground Field, #D) and another for Cambisol soils (Orchard, #B; Ivy Ground, #C;

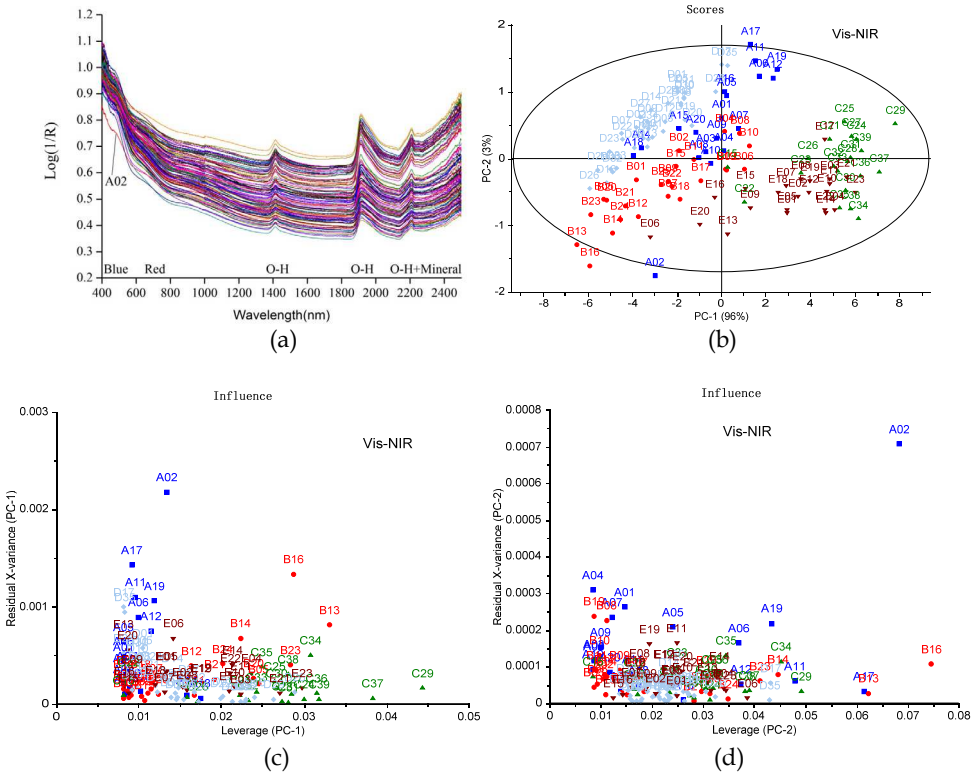


Fig. 5. Vis-NIR absorption spectra of all samples (a), principal components analysis scores plot for the Vis-NIR spectra (b), and residual X-variance for PC-1 (c) and PC-2 (d).

Copse Field, #E). Samples from Avenue Field (#A) exhibit partially mixing with the two clusters (Fig.5b). This might be attributed to the nature of this field as it is a mixture of Cambisol and Luvisol types. Although being of the same soil type of Cambisol, samples from field #B are clearly separated from those from fields #C and #E (Fig.5b). This is mainly due to different soil textures, namely, clay loam vs. clay (Table 1). Several samples located outside of the Hotelling T^2 ellipse become the candidates of samples outliers (Fig.5b). However, apart from sample A02, other outliers locate close to their member samples. In addition, sample A02 exhibits large residual X-variance for PC-1 (Fig.5c) and PC-2 (Fig.5d). The raw spectrum of sample A02 also shows distinct color features (lowest absorbance in visible range) from other samples (Fig.5a). Thus, sample A02 was considered as an outlier and excluded from further investigation.

3.2.2 Vis-NIR calibration models

Table 4 summarizes the cross-validation results of the PLSR models developed with raw and various transformed Vis-NIR spectra against each soil property. For raw spectra, PLSR models produced good or excellent prediction accuracy with R^2 of 0.86~0.90 and RPD of 2.73~3.33 for soil TN, TC and OC. Coupled with appropriate spectral pre-processing

Spectral pretreatment method	PLSR models calibrated for Vis-NIR spectra											
	Total N			Total C			Organic C			Inorganic C		
	LVs	R^2	RPD	LVs	R^2	RPD	LVs	R^2	RPD	LVs	R^2	RPD
None	5	0.90	3.33	5	0.86	2.73	5	0.90	3.16	**5**	**0.42**	**1.34**
SNV	**3**	**0.91**	**3.75**	3	0.87	2.74	**4**	**0.91**	**3.44**	3	0.37	1.29
MSC	4	0.90	3.33	4	0.87	2.74	4	0.90	3.20	3	0.39	1.31
BOC	4	0.90	3.33	4	0.86	2.70	5	0.91	3.27	5	0.37	1.29
Center & Scale	5	0.91	3.53	**5**	**0.87**	**2.78**	5	0.90	3.21	6	0.42	1.34
Normalization												
Maximum-	4	0.89	3.33	4	0.86	2.69	4	0.89	3.05	5	0.38	1.29
Range-	5	0.90	3.33	5	0.87	2.74	5	0.90	3.23	4	0.34	1.26
Mean-	4	0.89	3.16	4	0.85	2.61	4	0.89	3.00	4	0.39	1.31
Quantile-	5	0.91	3.53	5	0.87	2.78	5	0.91	3.25	6	0.42	1.34
De-trending												
1st-	4	0.90	3.53	4	0.87	2.75	4	0.90	3.20	4	0.38	1.30
2nd-	3	0.89	3.16	3	0.85	2.57	4	0.89	3.05	3	0.39	1.31
Derivative												
1st-	1	0.85	2.73	1	0.80	2.26	1	0.85	2.63	3	0.17	1.12
2nd-	4	0.61	1.71	1	0.56	1.51	4	0.63	1.65	1	0.14	1.11

Table 4. Cross validation result of PLSR models calibrated for raw and various transformed Vis-NIR spectra with 121 samples

algorithms, model performance was improved for a certain degree. For examples, PLSR model developed for TN after SNV-transformed spectra resulted in R^2 of 0.91 and RPD of 3.75. By the same pre-processing technique, prediction of soil OC was improved with R^2 of 0.91 and RPD of 3.44. The best calibration model for TC was obtained when the spectra were transformed by Center-&-Scale technique. It is worth noting that these optimized PLSR models need less latent variables (3-5) than those for raw spectra. In general, the fewer the latent variables used, the better is the model developed, as the calibration is more apt to be applicable to new samples (Madari, *et al.*, 2006). For IC, PLSR calibration with raw and various transformed spectra failed to produce useful models with $R^2 \leq 0.42$ and RPD≤1.34. Figure 6 shows the correlation between the measured and PLSR-predicted values of each soil property. The linear fitting slopes for TN, TC and OC are close to 1, which suggests that PLSR models developed with Vis-NIR absorption spectra for predicting soil TN, TC and OC are successful.

3.2.3 B-coefficients analysis of PLSR models for Vis-NIR spectra

B-coefficients curves of the PLSR models calibrated for the best transformed Vis-NIR spectra against each soil property are shown in Fig.7. The B-coefficients for TN, TC and OC exhibit strong similarity among them. This is mainly due to their high correlation obtained for

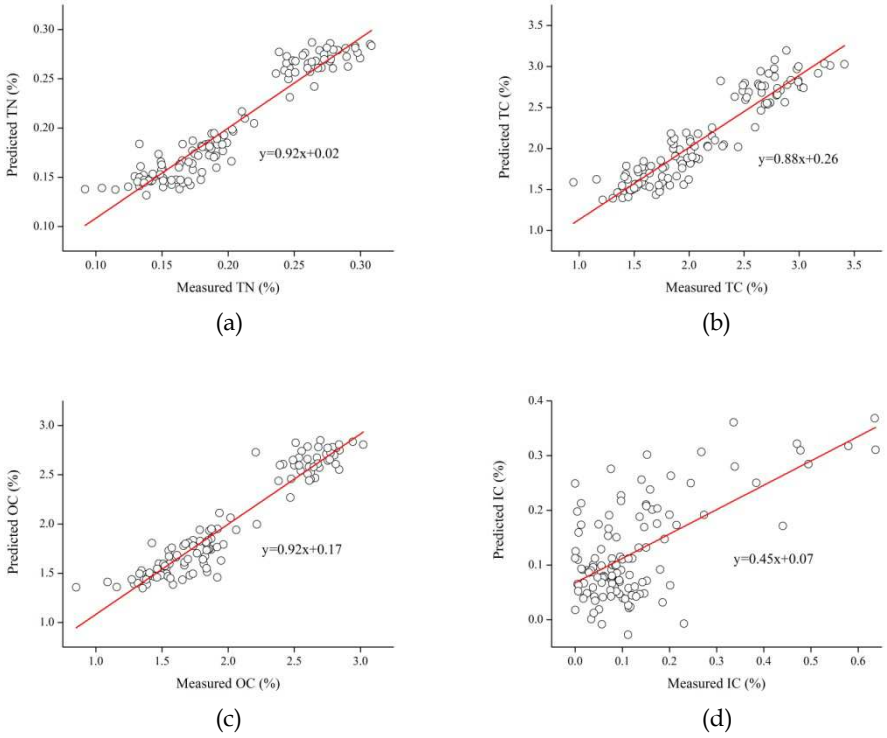


Fig. 6. Measured vs. predicted values of soil TN (a), TC (b), OC (c) and IC (d) based on Vis-NIR absorption spectra.

reference values (r=0.97-0.99, Table 3). It is worth noting that the visible range (400-760 nm) associated with soil colour shows huge influence on model accuracy, which is in line with other reports (Stenberg, *et al.*, 2010; Viscarra Rossel, *et al.*, 2006). The absorption feature in the visible and short-wave NIR (400-1000 nm) might be due to the Fe oxides in soil, mainly haematite and goethite (Viscarra Rossel & Behrens, 2010). The influential wavelengths located between 1000 and 2500 nm can be attributed to water, clay minerals and organic matter (Viscarra Rossel & Behrens, 2010). Using samples collected from Belgium and Northern France, Mouazen *et al*. (2006) compared the performance of two commercially available spectrophotometers with different wavelength ranges for the measurement of selected soil attributes including TC and TN. They found that the best accuracy was obtained when using a full wavelength range of 451-2459 nm, as compared to a short wavelength range of 401-1770 nm. Using samples collected from two depths along 11 km section of floodplain, Vohland and Emmerling (2011) reported that genetic algorithm (GA) allocated significant wavelengths most frequently to the range of 1970-2490 nm for soil OC, which is not particularly in line with those bands shown in Fig.7. One reason might be the use of farm-scale local set in our case in which the visible range associated with soil colours has the most influence on model calibration, whereas the NIR range has a smaller influence. This contradictory results against those published by others might be explained by the same

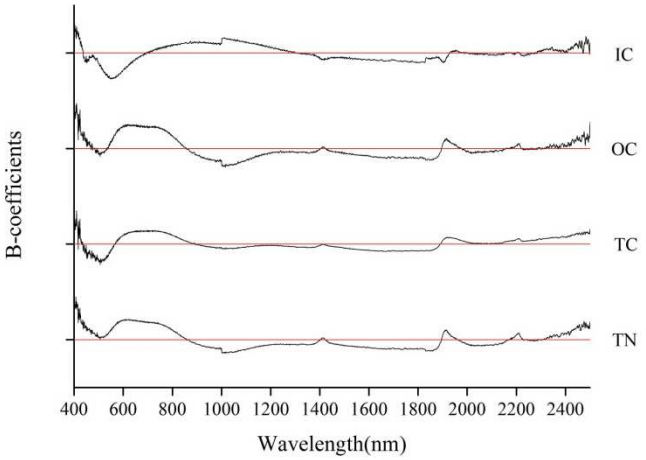


Fig. 7. B-coefficients curves obtained from PLSR analysis with the SNV-transformed Vis-NIR spectra for TN and OC, Center-&-Scale-transformed Vis-NIR spectra for TC, and raw Vis-NIR spectra for IC.

mineralogy of the data set originated from the same parent material in the current study, as compared to those using larger scale data sets (e.g. Mouazen, *et al.*, 2006; Vohland & Emmerling, 2011). Although the prediction accuracy of IC did not satisfy with lowest quantification standard (RPD>2.0), the characteristic bands of carbonate in 2300 and 2500 nm (Gaffey, 1986; Reeves III, *et al.*, 2002; Viscarra Rossel & Behrens, 2010) can be clearly identified in its B-coefficients curve.

3.3 DRIFT spectral analysis and model calibrations

The procedure of analyzing DRIFT spectra was the same as that presented for Vis-NIR spectral analysis.

3.3.1 DRIFT spectral response

Figure 8 shows the representative DRIFT absorption spectra of samples from each soil texture class, e.g. sandy loam (field #D), clay loam (field #B) and clay (fields #C and #E), with high and low TC content. Obviously, the absolute magnitude and range of the DRIFT absorptions are much greater than those found for corresponding Vis-NIR spectrum (Fig.4). The DRIFT spectrum generally shows more distinctive spectral features than the Vis-NIR spectrum. The peaks between 3698 and 3620 cm^{-1} and various peaks below 1100 cm^{-1} are characteristic for kaolinite (Madari, *et al.*, 2006). Soil organic matter also has characteristic absorption bands in the Mid-IR range (Table 5), however, due to its low concentration in the soil samples, and their overlapping with mineral peaks, most of these could not readily be identified by simple visual analysis of spectra. For example, the small bands around 2940-2935 cm^{-1} and 2886-2877 cm^{-1} indicate the presence of organic aliphatic C-H stretching (Table 5). This band is more evident in the spectra of soil samples having higher concentrations of organic carbon (Madari, *et al.*, 2006). Other organic peaks overlap with the mineral peaks. Viscarra Rossel *et al.* (2006) has compared the usefulness of visible, NIR and Mid-IR diffuse

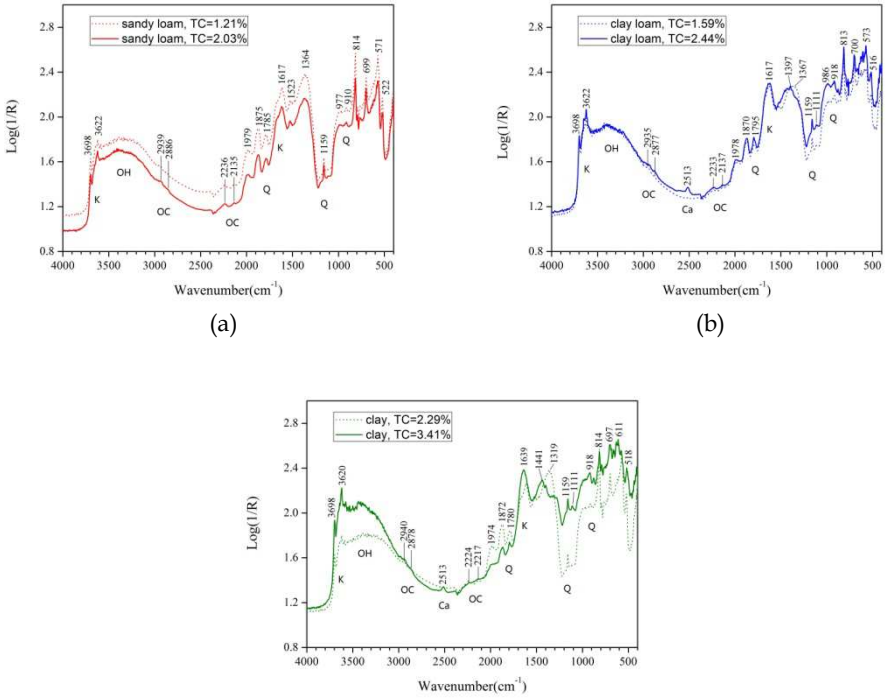


Fig. 8. DRIFT absorption spectra of samples from each soil texture class with high and low TC content. Some dominant soil components and absorption peaks are shown for quartz (Q), organic compounds (OC), calcite (Ca), kaolinite (K), and the (OH) features of free water and lattice minerals (Viscarra Rossel, *et al.*, 2006).

reflectance spectroscopy by examining PLSR factor loadings weights. They showed that frequencies in the Mid-IR range corresponding to the absorption of organic compounds, like organic acids, or acidic functional groups, like alkyl, amide and aromatic groups, were indicators of correlation between organic carbon concentrations in the bulk soil samples and the Mid-IR spectra.

3.3.2 PCA analysis for DRIFT spectra

Figure.9a shows all Mid-IR spectra obtained by FT-IR spectrometer with DRIFT accessory. Figure 9 also shows the PCA scores plot for the DRIFT spectra, and residual X-variance for PC-1 and PC-2. PC-1 and PC-2 explains 83% and 14% of total variance contained in the spectra, respectively. The PC-1 axis may explain differences attributed to SOM content of the samples as the PC-1 was better correlated to TN, TC and OC (r=0.87~0.90) than to PC-2 (r=0.15~0.21) (Table 3). Samples originated from different fields can be separated into two clusters according to soil types: one for Luvisol soils from Showground Field (#D) and another for Cambisol soils from Orchard (#B), Ivy Ground (#C) and Copse Field (#E),

Mid-IR band (cm^{-1})	Assignments	Vis-NIR wavelength (nm)
3380	O-H stretching of phenolic OH	
3400-3300	O-H stretching (H bonded OH groups),	
3300	N-H stretching	1500,1000,751
3030	Aromatic C-H strecthing	1650,1100,825
2940-2900	Aliphatic C-H stretching	
2930,2850	Alkyl asymmetric-symmetric C-H stretching	1706,1754,1138, 1170,853,877
2600	O-H stretching of H-bonded -COOH	
1725-1720	C=O stretching of –COOH and ketones	1930,1449
1660-1630	C=O stretching of amide groups (amide I band), quinine C=O and/or C=O of H-bonded conjugated ketones	2033,1524
1620-1600	Aromatic C=C stretching and/or asymmetric –COO stretching	
1610	N-H stretching of Amine	2060
1590-1517	COO- symmetric stretching, N-H deformation+C=N stretching (amide II band)	
1525	Aromatic C=C strecthing	
1460-1450	Aliphatic C-H	2275,1706
1400-1390	OH deformation and C-O stretching of phenolic OH, C-H deformation of CH2 and CH3 groups, COO- asymmetric stretching	
1350	Symmetric COO- stretching and/or –CH bending of aliphatics	
1270	C-OH stretching of phenolic OH	1961
1280-1200	C-O stretching and OH deformation of COOH, C-O stretching of aryl ethers	
1225	C-O stretching and OH deformation of COOH	
1170-950	C-O stretching of polysaccharides or polysaccharide-like substances	2137
1170	C-OH stretching of aliphatic OH, C-C stretching of aliphatic groups	
1050	C-O stretching of carbohydrates	2381
830	Aromatic CH out of plane bending	
775	Aromatic CH out of plane bending	

Table 5. Absorption bands of C and N in organic bonds in the Mid-IR range (Madari, *et al.*, 2006) and corresponding wavelengths in the Vis-NIR range (Stenberg, *et al.*, 2010)

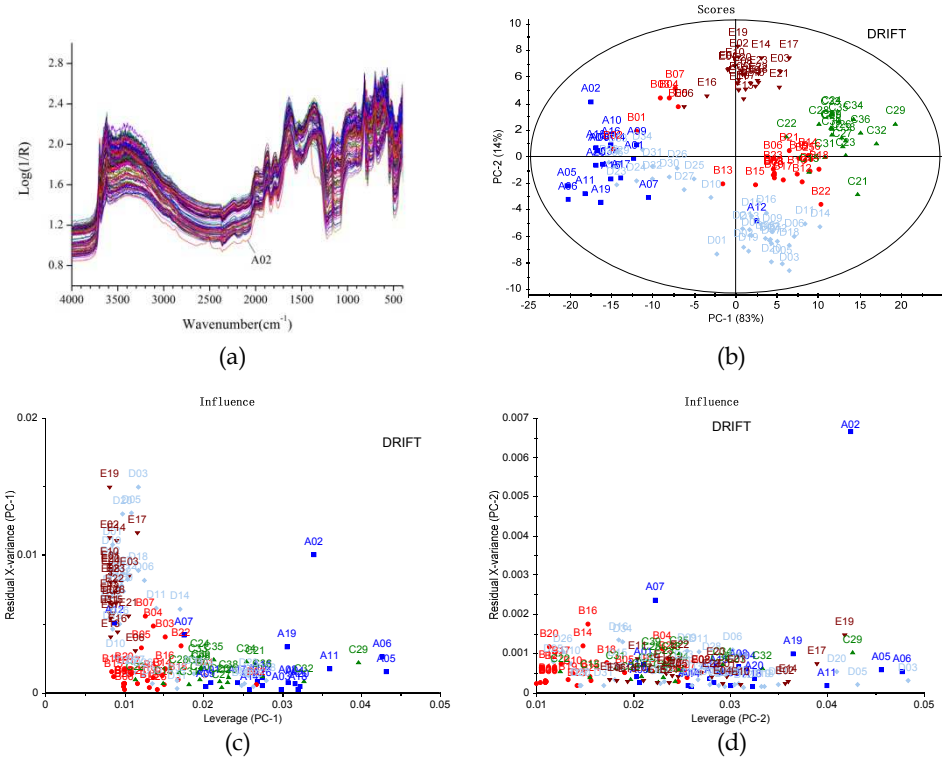


Fig. 9. DRIFT absorption spectra of all samples (a), principal components analysis scores plot for the DRIFT spectra (b), residual X-variance for PC-1 (c) and PC-2 (d).

although several samples from the field #B were mixed with former cluster. Samples from the field #A completely mixed together with the former cluster, although some of them belong to latter cluster. It seems that the soil type has no effect on samples separation into different classes but the texture diversity within the sample population used for calibration has clear effect (Madari, *et al*., 2006). Samples from the fields #C and #E are totally separated although they are of same clay texture. This is mainly due to their distinct SOM-related soil properties concentrations as different vegetations have been growing in both fields (soybean in the field #C with TN of 0.28±0.02%, TC of 2.94±0.23% and OC of 2.73±0.15%; wheat in the field #E with TN of 0.25±0.02%, TC of 2.62±0.15% and OC of 2.56±0.16%). No sample locates outside of the Hotelling T^2 ellipse. Although sample A02 exhibits low residual X-variance for PC-1 (Fig.9c) which accounts for the most amount of variance in the DRIFT spectra, it displays large residual X-variance for PC-2 (Fig.9d). Also, in the raw spectrum, sample A02 shows quite different from the others (Fig.9a). Thus, sample A02 was considered as a sample outlier and excluded from further investigation.

3.3.3 DRIFT calibration models

Table 6 summarizes the cross-validation results of the PLSR models developed with raw and various transformed DRIFT spectra against each soil property, e.g. TN, TC, OC and IC.

Spectral pretreatment method	PLSR models calibrated for DRIFT spectra											
	Total N			Total C			Organic C			Inorganic C		
	LVs	R^2	RPD	LVs	R^2	RPD	LVs	R^2	RPD	LVs	R^2	RPD
None	5	0.95	4.62	5	0.93	3.93	5	0.94	4.19	5	0.70	1.86
SNV	4	0.94	4.29	5	0.94	3.93	4	0.94	4.06	3	0.68	1.82
MSC	4	0.94	4.29	4	0.92	3.63	4	0.94	4.15	3	0.68	1.81
BOC	5	0.95	4.62	**6**	**0.94**	**4.10**	5	0.95	4.25	4	0.71	1.88
Center & Scale	5	0.95	4.62	5	0.93	3.88	5	0.94	4.25	**4**	**0.71**	**1.89**
Normalization												
Maximum-	6	0.94	4.29	6	0.93	3.80	6	0.94	4.15	5	0.68	1.80
Range-	6	0.94	4.62	6	0.93	3.85	6	0.94	4.22	5	0.69	1.84
Mean-	5	0.94	4.29	6	0.93	3.85	5	0.94	4.12	5	0.67	1.78
Quantile-	**5**	**0.95**	**5.00**	5	0.94	4.01	**5**	**0.95**	**4.43**	4	0.70	1.88
De-trending												
1st-	4	0.94	4.62	4	0.94	4.04	4	0.95	4.35	4	0.70	1.86
2nd-	5	0.94	4.62	5	0.93	3.83	5	0.95	4.32	3	0.67	1.79
Derivative												
1st-	2	0.83	2.61	2	0.79	2.18	2	0.79	2.21	5	0.30	1.22
2nd-	2	0.69	1.94	2	0.66	1.70	2	0.65	1.71	1	0.10	1.01

Table 6. Cross validation result of PLSR models calibrated for raw and various transformed DRIFT spectra with 121 samples

For raw spectra, PLSR models produced excellent prediction accuracy with R^2 of 0.93~0.95 and RPD of 3.93~4.62 for TN, TC and OC. These models were calibrated with 5 latent variables. Coupled with proper spectral pre-processing techniques, model prediction was improved. For examples, the best model for soil TN produced prediction accuracy with R^2 of 0.95 and RPD of 5.00 using spectral pre-processing of quantile normalization. It was also effective for OC model improvement with R^2 of 0.95 and RPD of 4.43. For TC, the best model was built using spectral transformation of BOC, although this model needed 6 latent variables. Figure 10 shows the correlation between the measured and PLSR-predicted values of each soil property. The linear fitting lines for TN, TC and OC are closer to 1:1 as compared to the corresponding plots for Vis-NIR spectra (Fig.6). DFIRT spectra outperformed Vis-NIR spectroscopy for the prediction of these three properties (Table 4 & 6). Even for IC, almost all PLSR models developed with raw and various transformed DRIFT spectra outperformed their counterparts for Vis-NIR spectra. Accuracy obtained for IC can be used to distinguish high and low content with RPD>1.5 (Table 6). These results suggest that PLSR models developed with DRIFT absorption spectra for predicting these soil properties are more accurate.

3.3.4 B-coefficients analysis of PLSR models for DRIFT spectra

B-coefficients curves obtained after PLSR analyses with the best transformed DRIFT spectra for each soil property are combined in Fig.11. The B-coefficients for TN, TC and OC exhibit

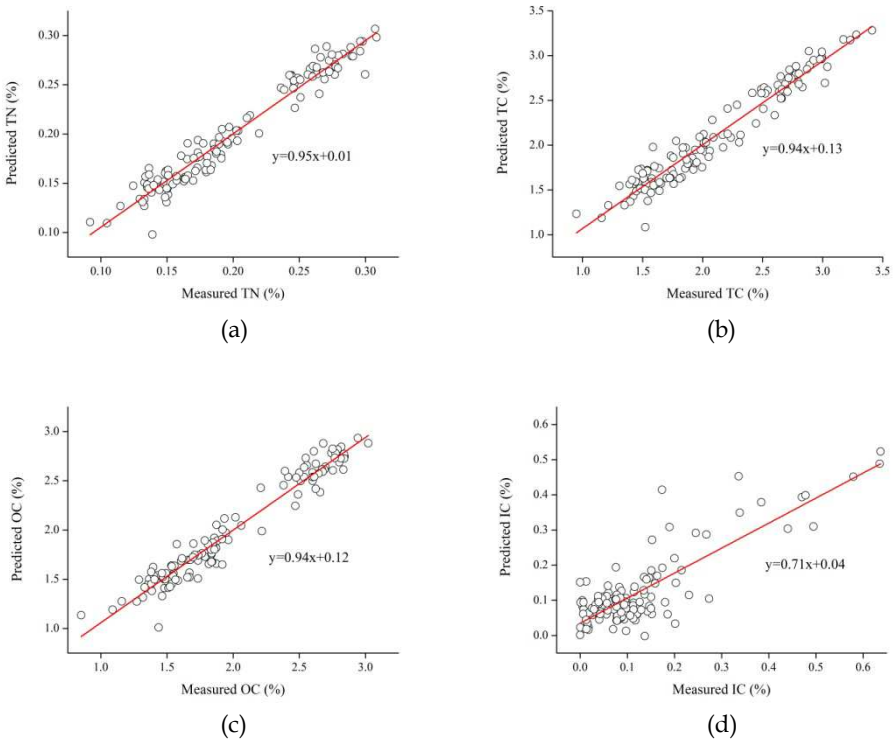


Fig. 10. Measured vs. predicted values of soil TN (a), TC (b), OC (c) and IC (d) based on DRIFT absorption spectra.

strong similarity among them. This is mainly due to their high correlation obtained with reference values (r=0.97-0.99, Table 3). Overall, the most influential frequencies for predicting these soil properties locate in the wavenumber range from 2100 to 1200 cm^{-1}. As indicated in Table 5, the peaks at around 1620-1558 cm^{-1} are for aromatic C=C stretching and/or asymmetric –COO stretching; the peaks at 1460 cm^{-1} and 1229 cm^{-1} are corresponding to Aliphatic C-H stretching and C-O stretching/OH deformation of COOH. Other significant wavebands can be found at round 2930 cm^{-1}, corresponding to Aliphatic C-H stretching. Interestingly, although the DRIFT-calibrated models are not accurate enough for IC quantification, the corresponding B-coefficients curve exhibits several distinctive frequencies for model calibration. The most influential frequency locates at around 1474 cm^{-1} with two subordinate ones at 1795 cm^{-1} and 2513 cm^{-1}, which are characteristic of existence of calcium carbonate (Reeves III, *et al.*, 2002). Viscarra Rossel and Behrens (2010) found the carbonate of clay minerals in the NIR band of 2336 nm to be associated with the third overtones of CO_3^{-2}in MIR of 1415 cm^{-1}.

3.4 ATR-FTIR spectral analysis and model calibrations

The procedure of analyzing ATR-FTIR spectra was the same as that presented for Vis-NIR and DRIFT spectral analysis.

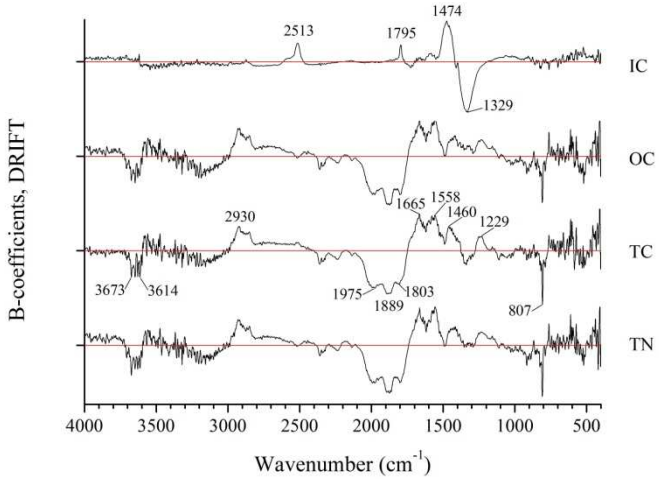


Fig. 11. B-coefficients curves obtained from PLSR analyses with the transformed DRIFT spectra by quantile normalization for TN and OC, baseline offset correction for TC, and Center & Scale for IC.

3.4.1 ATR-FTIR spectral response

Figure12 shows the representative ATR-FTIR spectra of samples from each soil texture class with high and low TC content. Obviously, the absolute magnitude and range of the ATR-FTIR light intensity are much lower than those for corresponding DRIFT spectra. Besides, the ATR-FTIR spectra of the samples present quite different shape together with absence of many characteristic peaks between 3000 and 1700 cm^{-1}, compared to the corresponding DRIFT spectra of these samples. The peaks between 3696 and 3620 cm^{-1} and various peaks

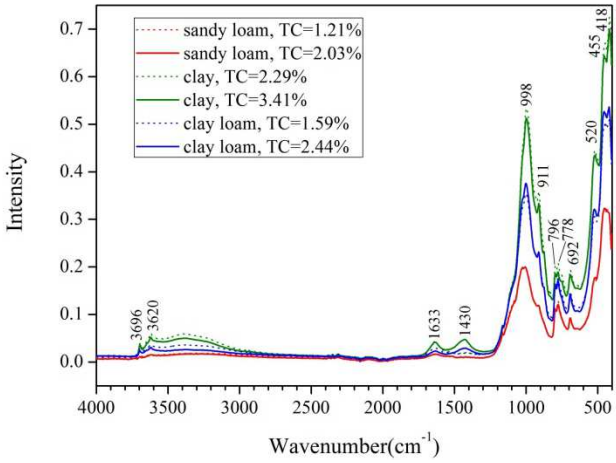


Fig. 12. ATR-FTIR spectra of samples from each soil texture class with high and low TC content.

below 1100 cm^{-1} might be characteristic for kaolinite (Madari, *et al.*, 2006). However, most characteristic wavebands of soil organic matter presented in Table 5, due to their low concentrations in the soil samples and their overlapping with mineral peaks, could not readily be identified by simple visual observation. For example, the bands around 2940-2935 cm^{-1} and 2886-2877 cm^{-1}, indicating the presence of organic aliphatic C-H stretching (shown in DRIFT spectra, Fig.8), are nearly vanished. Although the ATR-FTIR spectra present severe deviations from corresponding DRIFT spectra, the light intensity in ATR-FTIR spectra exhibits strong correlation with particle size distribution. For example, samples of clay texture with fine particles present strongest light intensity, whereas samples with sandy loam texture with coarse particle size have lowest light intensity. The clay loam texture samples with middle size particles correspond to intermediate light absorption. For each soil texture class, the spectral difference due to high and low TC content is not apparent. All these spectra appear to be free from baseline offset.

3.4.2 PCA analysis for ATR-FTIR spectra

Figure 13 shows all Mid-IR spectra obtained by FI-IR spectrometer with ATR accessory together with the PCA scores plot for the spectra and residual X-variance for PC-1 and PC-2. PC-1 and PC-2 explained 98% and 1% of total variance of the spectra, respectively. The PC-2

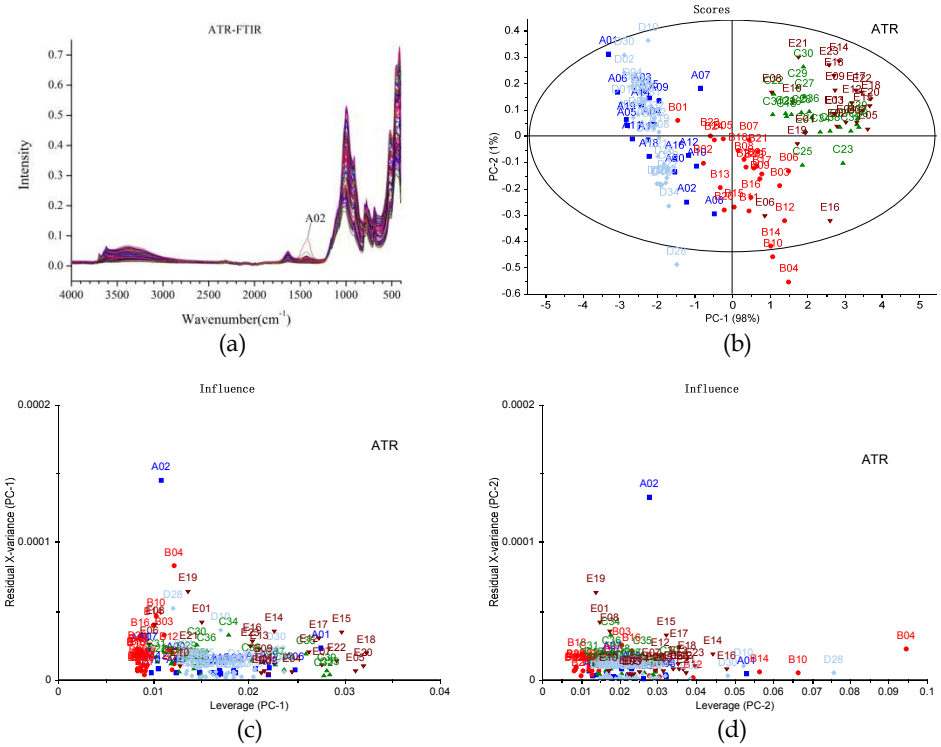


Fig. 13. All ATR-FTIR spectra (a), principal components analysis scores plot for the ATR-FTIR spectra (b), and residual X-variance for PC-1 (c) and PC-2 (d).

shows variation related to SOM of the samples as the PC-2 was better correlated to TN, TC and OC (r=-0.68~0.71) than to PC-1 (r=0.52~0.56) (Table 3). Samples originated from different fields can be divided into one cluster for Luvisol type soils (field #D) and another for Cambisol type soils (fields #B, #C, and #E). Although samples from the field #A are the mixture of Cambisol and Luvisol types, most of them locate within the former cluster. Samples from the fields #C and #E are hardly separated, although they are of the same soil texture of clay (Table 1). Compared to the samples D28, B04 and B10 located outside of the Hotelling T^2 ellipse (Fig.13b), sample A02 exhibits large residual X-variance for PC-1 (Fig.13c) and PC-2 (Fig.13d). Thus, sample A02 was considered as a sample outlier and excluded from further investigation.

3.4.3 ATR-FTIR calibration models

Table 7 summarizes the cross-validation results of PLSR models developed with the raw and various transformed ATR-FTIR spectra against each soil property. For raw spectra, PLSR models produced excellent prediction accuracy for TN, TC and OC with R^2 of 0.89~0.92 and RPD of 3.02~3.75. These models were developed with 6 latent variables. None of spectral pre-processing techniques adopted can effectively improve the prediction accuracy of these models. This might be due to absence of the baseline offset of spectra (Fig.12). For IC, all models can be used to distinguish high and low concentration with RPD

Spectral pretreatment method	PLSR models calibrated for ATR-FTIR spectra											
	Total N			Total C			Organic C			Inorganic C		
	LVs	R^2	RPD	LVs	R^2	RPD	LVs	R^2	RPD	LVs	R^2	RPD
None	**6**	**0.92**	**3.75**	**6**	**0.89**	**3.02**	**6**	**0.89**	**3.10**	6	0.71	1.91
SNV	5	0.90	3.33	5	0.88	2.89	3	0.87	2.74	4	0.69	1.82
MSC	5	0.90	3.33	4	0.87	2.75	3	0.87	2.74	4	0.69	1.82
BOC	7	0.91	3.53	7	0.88	2.95	7	0.89	3.07	4	0.70	1.87
Center & Scale	6	0.91	3.53	6	0.89	2.97	6	0.89	2.97	4	0.70	1.88
Normalization												
Maximum-	5	0.90	3.33	5	0.87	2.79	5	0.87	2.78	5	0.70	1.85
Range-	5	0.90	3.53	4	0.86	2.71	6	0.88	2.93	5	0.70	1.85
Mean-	6	0.91	3.53	4	0.86	2.69	5	0.88	2.87	5	0.70	1.87
Quantile-	7	0.91	3.75	6	0.88	2.89	6	0.89	3.00	4	0.70	1.87
De-trending												
1st-	6	0.90	3.53	6	0.87	2.79	7	0.89	3.09	**5**	**0.72**	**1.93**
2nd-	6	0.91	3.53	6	0.87	2.79	5	0.88	2.92	5	0.72	1.91
Derivative												
1st-	4	0.88	3.16	4	0.85	2.60	4	0.86	2.69	5	0.71	1.91
2nd-	3	0.78	2.31	3	0.75	1.99	2	0.74	1.98	7	0.60	1.61

Table 7. Cross validation result of PLSR models developed with raw and various transformed ATR-FTIR spectra with 121 samples

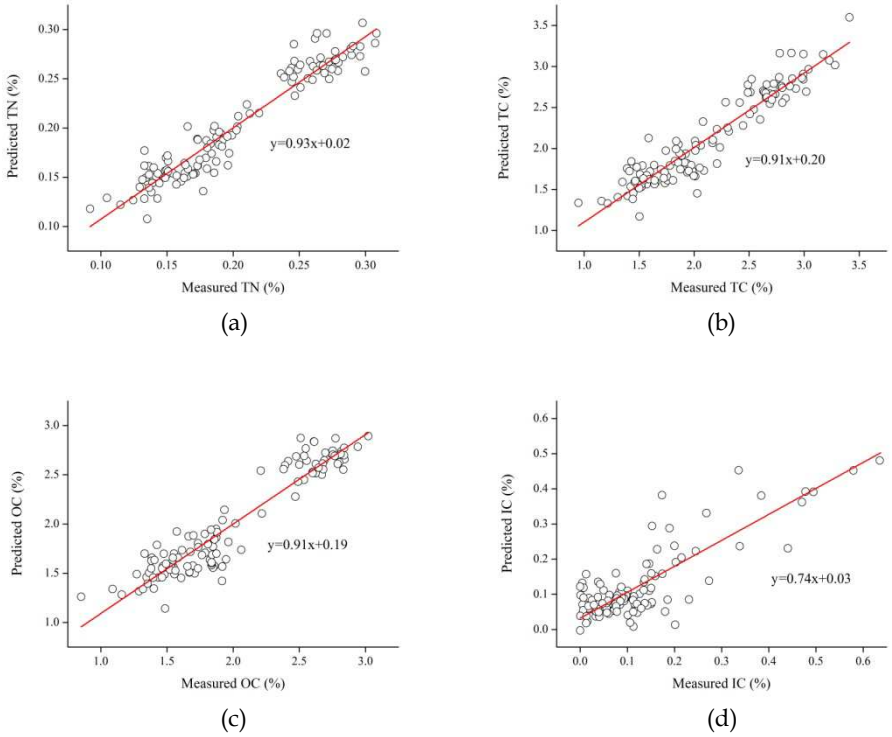


Fig. 14. Measured vs. predicted values of soil TN (a), TC (b), OC (c) and IC (d) based on the ATR-FTIR spectra.

of 1.61~1.93. Figure 14 shows the correlation between the measured and predicted values of each soil property. The linear fitting lines for TN, TC and OC are closer to the 1:1 lines as compared to the corresponding plots for Vis-NIR spectra (Fig.6a-c), but not so well as those for DRIFT spectra (Fig.10a-c). However, the linear fitting for IC is better than that for Vis-NIR spectra (Fig.6d) and that for DRIFT spectra (Fig.10d).

3.4.4 B-coefficients analysis of PLSR models for ATR-FTIR spectra

B-coefficients curves of the PLSR models developed with the best transformed ATR-FTIR spectra for each soil property are compared in Fig.15. The B-coefficients for TN, TC and OC exhibit strong similarity among them. This is mainly due to the high correlation obtained with the reference values (r=0.97-0.99, Table 3). Overall, the most significant wavebands for predicting these soil properties locate in the wavenumber range between 1700 and 400 cm^{-1}. As indicated in Table 5, the peaks at 1612-1551 cm^{-1} are for aromatic C=C stretching and/or asymmetric –COO stretching and the peaks at around 1050 cm^{-1} correspond to Polysaccharides. Other significant wavebands can be found at round 2925 cm^{-1}, which correspond to Aliphatic C-H stretching. Interestingly, although the ATR-calibrated models are not accurate enough for IC quantification, its B-coefficients curve for IC exhibits several

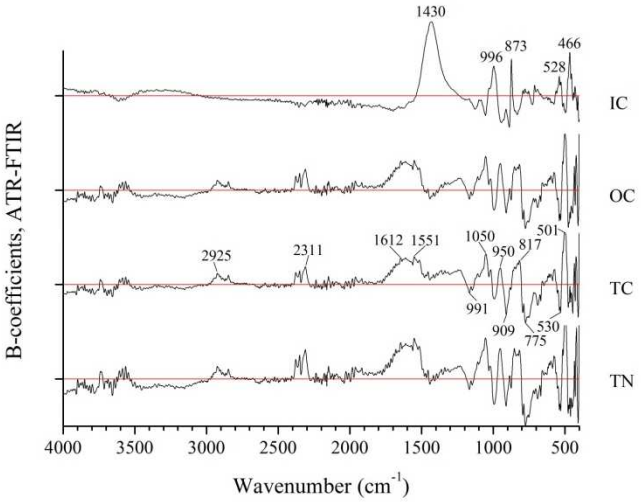


Fig. 15. B-coefficients curves obtained from PLSR analysis with the raw spectra for TN, TC and OC, and the 1st-detrending-transformed spectra for IC.

distinctive bands, similar to those obtained with DRIFT spectroscopy. The most significant band locates at around 1430 cm^{-1}, which corresponds to calcium carbonate ($CaCO_3$). However, compared to the B-coefficients curve for IC based on DRIFT spectra (Fig.11), two characteristic bands of $CaCO_3$ at 1795 cm^{-1} and 2513 cm^{-1} disappear completely, which is mainly due to the weaker light response of ATR-FTIR spectra compared to DRIFT spectra.

3.5 Model calibrations for combinational Vis-NIR-MIR spectra

The raw Vis-NIR and Mid-IR spectra were combined to develop PLSR calibration models of soil TN, TC, OC and IC concentrations. Table 8 shows the cross-validation results of the PLSR models. For the Vis-NIR-ATR spectra, PLSR models produced excellent prediction accuracy with R^2 of 0.89~0.91 and RPD of 3.00~3.75 for TN, TC and OC. By comparison, PLSR models calibrated for the Vis-NIR-DRIFT spectra achieved higher prediction accuracy than those with Vis-NIR-ATR spectra, with R^2 of 0.93~0.95 and RPD of 3.83~4.62 for TN, TC and OC. For IC, PLSR model for Vis-NIR-ATR spectra was developed with 10 latent variables and validated

Combinational spectra	PLSR models calibrated for raw combinational Vis-NIR-MIR spectra											
	Total N			Total C			Organic C			Inorganic C		
	LVs	R^2	RPD	LVs	R^2	RPD	LVs	R^2	RPD	LVs	R^2	RPD
Vis-NIR-ATR	5	0.91	3.75	5	0.89	3.00	6	0.91	3.42	10	0.69	1.83
Vis-NIR-DRIFT	6	0.95	4.62	6	0.93	3.83	6	0.95	4.32	5	0.70	1.88

Table 8. Cross validation result of PLSR models developed with raw Vis-NIR-MIR spectra with 121 samples

with R^2 of 0.69 and RPD of 1.83, whereas PLSR model for Vis-NIR-DRIFT spectra calibrated with 5 latent variables performed better than Vis-NIR– ATR with higher prediction accuracy (R^2 of 0.70 and RPD of 1.94). These models for combinational spectra performed slightly better than those for Vis-NIR spectra (Table 4). However, these models did not produce better performance than those for DRIFT spectra (Table 6) and ATR spectra (Table 7) with only exception of Vis-NIR-ATR models for OC with R^2 of 0.91 and RPD of 3.42.

3.6 Comparison of PLSR model performance among Vis-NIR, ATR-FTIR, DRIFT and combinational spectra

As shown in Tables 4, 6 and 7, model performance is not only a function of wavelength ranges used during PLS regression analysis, but also a function of spectral pre-processing techniques. Overall, for TN, TC and OC, PLSR models calibrated for DRIFT spectra outperformed those for Vis-NIR spectra and ATR-FTIR spectra. For IC, both ATR-FTIR and DRIFT models outperformed Vis-NIR models no matter what spectral pre-processing techniques were applied. However, if coupled with appropriate spectral pre-processing techniques, Vis-NIR models for TN and OC can produce competitive prediction performance (R2>0.90 and RPD>3.0) with less number of latent variables (3 or 4) as compared to best ATR-PLSR models calibrated with 6 latent variables. For TC, ATR-FTIR models performed slightly better than Vis-NIR models. The lower accuracy for the calibration of IC compared to TN, TC and OC may be attributed to errors in the reference method for IC determination, since IC is calculated by difference between TC and OC.

Researchers have reported that the particle size distribution within the soil sample population and also within each sample of the calibration set affects the accuracy of calibration for TC and OC both in Mid-IR and Vis-NIR (Madari, *et al.*, 2006; Mouazen, *et al.*, 2005, Yang, *et al.*, 2011b). However, Vis-NIR proved to be more sensitive to particle size effects than the Mid-IR range (Madari, *et al.*, 2006). Vis-NIR spectroscopy performed very well for a very homogenous sample population, even slightly better than Mid-IR, but with increasing heterogeneity among and within the soil samples the accuracy decreased drastically. By contrast, the particle size distribution had less effect in the Mid-IR range. For the very homogeneous sample population, the accuracy was slightly lower than Vis-NIR, but with the increase in the heterogeneity of the sample population the accuracy did not diminish drastically and was higher than using with Vis-NIR (Madari, *et al.*, 2006). Thus Mid-IR spectroscopy coupled with appropriate chemometrics can be considered to be more robust than Vis-NIR.

3.7 Fundamentals of predicting N in soil

Soil N content is often highly correlated with C (Martin, *et al.*, 2002). For example, Chang, *et al.* (2001) reported r of 0.95 between TC and TN. In this study, the mean (±s.d.) values of TC/TN and OC/TN are 10.6(±0.59) and 9.95(±0.52), respectively (Table 2). It is an interesting point to explore whether there is an independent spectral basis for the determination of N in soil by infrared (IR) spectroscopy or whether N is predicted through high correlation with C. In the work by Chang and Laird (2002), in which C and N were added to a soil resulting in a wide range of C-to-N ratios, N was proved to be predicted in soil independently of C. Although the N absorbers are present in the soil spectra, their absorbance is not as strong as that of C bonds, as the mass of C in soil is generally an order of magnitude higher than that of N (about 10 times in our case). Thus, Martin *et al.* (2002) explained that N is predicted best

on its correlation with C if a high C-to-N correlation exists. In the current study, the excellent prediction of TN might be due to its high correlation with TC or OC (r=0.97-0.99) (Table 3), which can be proved by the similar distribution of their B-coefficients curves for Vis-NIR spectra (Fig.7), DRIFT spectra (Fig.11), and ATR-FTIR spectra (Fig.15).

4. Conclusions

The Mid-IR spectroscopy, including ATR and DRIFT, and Vis-NIR spectroscopy were implemented for the prediction of soil TN, TC, OC and IC. Results proved that both Vis-NIR and Mid-IR when combined with chemometric methods have great potential to quantify soil N and C at the field scale. It was also shown that DRIFT is more robust than Vis-NIR or ATR in terms of prediction accuracy. Although the Mid-IR spectra holds more information and usually easier to interpret as compared to Vis-NIR spectra with overtones and combinations features, until recently the MIR instruments are less portable and born to easier damage of optical materials. In contrast, the Vis-NIR has some advantages related to portability, mobile (on-line) measurement, remote sensing and others. This study suggests that Vis-NIR spectroscopy, if coupled with proper spectral pre-processing techniques, has the potential for successful prediction of soil N and C, although the combination and overtone peaks in the Vis-NIR spectral range are usually weak.

5. Acknowledgment

We gratefully thank the Soil Labs in Cranfield University for the provision of spectrophotometer. Sincere thanks are given to Dr. Boyan Kuang for assisting soil sampling and laboratory reference measurement, and Dr. Robert J. A. Jones for assisting soil description. Financially supported by Natural Science Foundation of Zhejiang Province, P.R.China (Project No.Y1090885) and by the State Scholarship Fund of China (Grant No.[2009]3004).

6. References

Chang, C.-W. & Laird, D. A. (2002). Near-Infrared Reflectance Spectroscopic Analysis of Soil C and N. *Soil Science*, Vol.167, No.2, pp.110-116, ISSN 0038-075X

Chang, C.-W.; Laird, D. A.; Mausbach, M. J. & Hurburgh, C. R. (2001). Near-Infrared Reflectance Spectroscopy–Principal Components Regression Analyses of Soil Properties. *Soil Sci. Soc. Am. J.*, Vol.65, No.2, (2001/3), pp.480-490, ISSN 0361-5995

Fystro, G. (2002). The prediction of C and N content and their potential mineralisation in heterogeneous soil samples using Vis-NIR spectroscopy and comparative methods. *Plant and Soil*, Vol.246, pp.139-149, ISSN 0032-079X

Gaffey, S. J. (1986). Spectral reflectance of carbonate minerals in the visible and near infrared (0.35-2.55 microns):calcite, aragonite, and dolomite. *American Mineralogist*, Vol.71, pp.151-162, ISSN 0003-004X

Haché, C.; Shibusawa, S.; Sasao, A.; Suhama, T. & Sah, B. P. (2007). Field-derived spectral characteristics to classify conventional and conservation agricultural practices. *Computers and Electronics in Agriculture*, Vol.57, No.1, pp.47-61, ISSN 0168-1699

Ludwig, B.; Khanna, P. K.; Bauhus, J. & Hopmans, P. (2002). Near infrared spectroscopy of forest soils to determine chemical and biological properties related to soil

sustainability. *Forest Ecology and Management*, Vol.171, No.1-2, pp.121-132, ISSN 0378-1127

Madari, B. E.; Reeves Iii, J. B.; Machado, P. L. O. A.; Guimarães, C. M.; Torres, E. & McCarty, G. W. (2006). Mid- and near-infrared spectroscopic assessment of soil compositional parameters and structural indices in two Ferralsols. *Geoderma*, Vol.136, No.1-2, pp.245-259, ISSN 0016-7061

Martin, P. D.; Malley, D. F.; Manning, G. & Fuller, L. (2002). Determination of soil organic carbon and nitrogen at the field level using near-infrared spectroscopy. *Canadian Journal of Soil Science*, Vol.82, No.4, pp.413-422, ISSN 0008-4271

Mouazen, A. M.; Baerdemaeker, J. D. & Ramon, H. (2006). Effect of wavelength range on the measurement accuracy of some selected soil constituents using visual-near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, Vol.14, No.3, pp.189-199, ISSN 0967-0335

Mouazen, A. M.; Karoui, R.; De Baerdemaeker, J. & Ramon, H. (2005). Classification of soil texture classes by using soil visual near infrared spectroscopy and factorial discriminant analysis techniques. *Journal of Near Infrared Spectroscopy*, Vol.13, No.4, pp.231-240, ISSN 0967-0335

Mouazen, A. M.; Maleki, M. R.; De Baerdemaeker, J. & Ramon, H. (2007). On-line measurement of some selected soil properties using a VIS-NIR sensor. *Soil and Tillage Research*, Vol.93, No.1, pp.13-27, ISSN 0167-1987

Mutuo, P. K.; Shepherd, K. D.; Albrecht, A. & Cadisch, G. (2006). Prediction of carbon mineralization rates from different soil physical fractions using diffuse reflectance spectroscopy. *Soil Biology and Biochemistry*, Vol.38, No.7, pp.1658-1664, ISSN 0038-0717

Reeves III, J.; McCarty, G. & Mimmo, T. (2002). The potential of diffuse reflectance spectroscopy for the determination of carbon inventories in soils. *Environmental Pollution*, Vol.116, No.Supplement 1, pp.S277-S284, ISSN 0269-7491

Schimann, H.; Joffre, R.; Roggy, J.-C.; Lensi, R. & Domenach, A.-M. (2007). Evaluation of the recovery of microbial functions during soil restoration using near-infrared spectroscopy. *Applied Soil Ecology*, Vol.37, No.3, pp.223-232, ISSN 0929-1393

Sinfield, J. V.; Fagerman, D. & Colic, O. (2010). Evaluation of sensing technologies for on-the-go detection of macro-nutrients in cultivated soils. *Computers and Electronics in Agriculture*, Vol.70, No.1, pp.1-18, ISSN 0168-1699

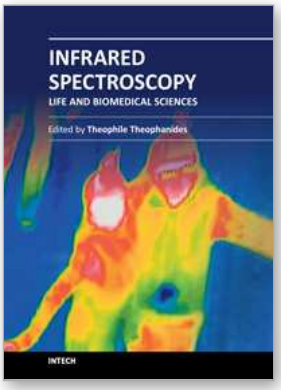
Viscarra Rossel, R. A. & Behrens, T. (2010). Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, Vol.158, No.1-2, pp.46-54, ISSN 0016-7061

Viscarra Rossel, R. A.; Walvoort, D. J. J.; McBratney, A. B.; Janik, L. J. & Skjemstad, J. O. (2006). Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, Vol.131, No.1-2, pp.59-75, ISSN 0016-7061

Vohland, M. & Emmerling, C. (2011). Determination of total soil organic C and hot water-extractable C from VIS-NIR soil reflectance with partial least squares regression and spectral feature selection techniques. *European Journal of Soil Science*, Vol.62, pp.598-606, ISSN 1351-0754

Yang, H.; Kuang, B. & Mouazen, A. M. (2011a). Prediction of soil TN and TC at a farm-scale using VIS-NIR spectroscopy. *Advanced Materials Research Journal*, Vol.225-226, pp.1258-1261, ISSN 1022-6680

Yang, H.; Kuang, B. & Mouazen, A. M. (2011b). Selection of Preprocessing Parameters for PCA of Soil Classification Affected by Particle Sizes Based on Vis-NIR Spectroscopy. *Key Engineering Materials Journal*, Vol.467-469, pp.725-730, ISSN 1013-9826



Infrared Spectroscopy - Life and Biomedical Sciences
Edited by Prof. Theophanides Theophile

ISBN 978-953-51-0538-1
Hard cover, 368 pages
Publisher InTech
Published online 25, April, 2012
Published in print edition April, 2012

This informative and state-of-the art book on Infrared Spectroscopy in Life sciences designed for researchers, academics as well as for those working in industry, agriculture and in pharmaceutical companies features 20 chapters of applications of MIRS and NIRS in brain activity and clinical research. It shows excellent FT-IR spectra of breast tissues, atheromatic plaques, human bones and projects assessment of haemodynamic activation in the cerebral cortex, brain oxygenation studies and many interesting insights from a medical perspective.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:

Haiqing Yang and Abdul M. Mouazen (2012). Vis/Near- and Mid- Infrared Spectroscopy for Predicting Soil N and C at a Farm Scale, Infrared Spectroscopy - Life and Biomedical Sciences, Prof. Theophanides Theophile (Ed.), ISBN: 978-953-51-0538-1, InTech, Available from: http://www.intechopen.com/books/infrared-spectroscopy-life-and-biomedical-sciences/vis-near-and-mid-infrared-spectroscopy-for-predicting-soil-n-and-c-at-a-farm-scale

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821