Predicting Soil Salt Content Over Partially Vegetated Surfaces Using Non-Negative Matrix Factorization

Ya Liu, Xian-Zhang Pan, Rong-Jie Shi, Yan-Li Li, Chang-Kun Wang, and Zhi-Ting Li

***Abstract*—Remote sensing has been widely applied to map soil salinity in the last few decades. However, a notable decrease in the accuracy of soil salt content (SSC) predictions occurred when the soil surfaces were partially vegetated. To minimize the influence of partial vegetation cover on spectral reflectance, we applied a spectral separation method, non-negative matrix factorization (NMF), to extract soil spectral information from a controlled field experiment with three varying factors [vegetation coverage, soil moisture content (SMC), and SSC]. The method was applied without prior knowledge of, or restrictions on the mixed and source spectra. Soil samples and spectral reflectance collected on three periods were used to determine the effectiveness of NMF-extracted soil spectra with partial least squares regression (PLSR). The results indicated that SSC can be predicted by bare soil spectra. NMF effectively separated soil spectra from the observed spectra, and the SSC was successfully predicted from the extracted soil spectra within a wide range of vegetation cover (0%–64.7%) within defined moisture levels ($<0.15\ \mathrm{g\,g^{-1}}$ by weight). The approach proposed in this study will improve the prediction accuracy of SSC for partially vegetated surfaces and will expand the application of remote sensing.**

***Index Terms*—Non-negative matrix factorization (NMF), partially vegetated surfaces, remote sensing, soil moisture content (SMC), soil salt content (SSC).**

ABBREVIATIONS

SSC	soil salt content.
NMF	non-negative matrix factorization.
PLSR	partial least squares regression.
NDVI	normalized difference vegetation index.
RSU	residual spectral unmixing.
SMC	soil moisture content.
BSS	blind source separation.
SA	spectral angle.
EPO	external parameter orthogonalization.

Manuscript received April 20, 2015; revised June 30, 2015; accepted September 08, 2015. Date of publication October 15, 2015; date of current version January 18, 2016. This work was supported in part by the National Natural Science Foundation of China under Grant 41071140; in part by the Strategic Priority Research Program of the Chinese Academy of Sciences under Grant XDB15040300; and in part by China Soil Scientific Database under Grant XXH12504-1-02. *(Corresponding author: Xian-Zhang Pan.)*

Y. Liu, R.-J. Shi, Y.-L. Li, and Z.-T. Li are with the Key Laboratory of Soil Environment and Pollution Remediation, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China, also with the University of Chinese Academy of Science, Beijing 100049, China.

X.-Z. Pan and C.-K. Wang are with the Key Laboratory of Soil Environment and Pollution Remediation, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China (e-mail: panxz@issas.ac.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2015.2478490

I. INTRODUCTION

SOIL salinization is one of the most common forms of land degradation and environmental hazard [1]. In recent years, the primary and secondary soil salinization of the world continues to increase, resulting in turning fertile fields into degraded land even deserts. For sustainable development of ecological environment, monitoring the extent, magnitude, and spatial distribution of salt-affected soil are essential.

Image spectroscopy is a powerful tool to estimate the spatial distribution of soil salinity [2]–[5]. However, the application of spectroscopy techniques in monitoring soil properties is relatively restricted, particularly in areas where the soil surface is covered with vegetation [6], [7]. When using spectroscopy, vegetation seriously decreased the accuracy of soil organic carbon and clay contents prediction in partially vegetated fields [6], [7], and vegetation cover was also identified as one of the most important constraints on the use of remote sensing data to map soil salinity [1]. A number of strategies have been attempted, with different degrees of success, to account for vegetation effects and improve spectroscopic calibrations in order to predict soil properties in vegetated areas. The approaches dealt with soil salinity prediction in the vegetated area using remote sensing technique could be summarized as three kinds, they were as follows.

The first one was the traditional solution which was to mask out the areas with high vegetation cover, using vegetation indices with certain threshold values [8]. Therefore, imaging spectroscopy was often restricted to bare soil surfaces due to inadequate spectral information on soil salinity in vegetated area.

The second one was using vegetation indices. In a few studies, soil salinity was estimated with vegetation reflectance, and many of these studies preferred to use vegetation indices. The NDVI was sensitive to salinity, particularly in croplands [9]. Other vegetation indices such as the photochemical reflectance index (PRI) [10], the red edge position (REP) index, the chlorophyll-normalized difference index (Chl NDI) [10], [11], the modified NDVI [4], and the soil-adjusted vegetation index (SAVI) [12] were highly correlated with SSCs. However, the general applicability and transferability of these vegetation indices were too limited. Zhang *et al.* found that most vegetation indices, except SAVI, had weak relationships with soil salinity (with an average R^2 of 0.28) and some (e.g., PRI and REP) were not sensitive to all species of vegetation [12]. Zhang *et al.* [13] and Douaoui *et al.* [14] found that NDVI was a poor predictor of soil salinity [13], [14]. Therefore, even

though some vegetation indices worked well in some cases, the indices may not always be good indicators because they were not specifically constructed for direct determination of soil salinity. Accordingly, a new method to estimate soil salinity in partially vegetated areas must be developed.

The third one was spectral unmixing. Recently, soil spectral information was extracted from mixed spectra with spectral unmixing techniques, to improve the accuracy of soil properties estimations. Ghosh *et al.* [15] used linear spectral unmixing analysis to map salt-affected soils with Hyperion hyperspectral remote sensing data (EO-1), and classified not quantified slight, moderate and highly salt-affected soils. Bartholomeus *et al.* [6] first proposed the new method of RSU to filter out the influence of maize from the mixed spectral signal, and they used the residual spectra to quantify soil organic carbon content [6]. The RSU method effectively eliminated the influence of vegetation from the mixed spectral reflectance, but it was necessary to know in advance the portions of soil and vegetation. The soil reflectance was then calculated using (1), and was used to predict soil properties as usual. In this case, both the efficiency and feasibility were restricted, since the fraction of vegetation and soil should be known in advance

$$R_{soil} = R_{mix} - \frac{(f_{maize} \times R_{maize})}{f_{soil}} \tag{1}$$

where R_{mix} is the reflectance of the mixed spectra, R_{maize} is the reflectance of the maize endmember, f_{maize} and f_{soil} are the fraction of maize and soil, respectively.

Ouerghemmi *et al.* [7] applied BSS techniques to extract soil spectra from mixed hyperspectral spectra, and predicted clay contents using the extracted soil spectra. The biggest advantage of BSS was no need of any prior knowledge about the source spectra and how they mixed [7]. However, Ouerghemmi *et al.* [7] used independent component analysis (ICA) technique to implement BSS. The biggest disadvantages of ICA were that the source signals were required to be independent of each other, and some of the results were in negative domains, which required further processing, thus limited the wide application of ICA, because most source signals were correlative in practical application.

As another technique of the BSS family, NMF which was not limited to the independent source signals overcame the disadvantages of ICA, and extracted source spectra with all positive values. For its wide applicability, NMF was previously applied widely in image analyses [16]–[18], and was also proposed to compress spectral data [19]. We explore the NMF technique to extract the soil spectrum from mixed hyperspectral data containing both vegetation and soil.

However, in addition to vegetation, many other factors influence soil reflectance, such as soil moisture, soil color, and surface roughness. Of these factors, soil moisture is the most prominent one because it cannot be avoided and is variable in the field. Soil moisture affected the entire visible and near infrared reflectance (vis–NIR) domain from 400 to 2400 nm [20], [21] and seriously reduced the prediction accuracy of some soil properties [20], [22], [23]. Specific to soil salinity estimation, the effect of soil moisture was greater, as SSC changed greatly with the changing SMC [24].

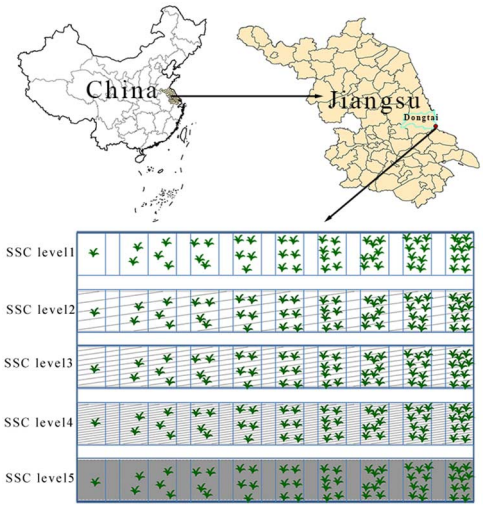

Fig. 1. Geographic location of the experimental site.

Thus, soil moisture effects would be considered in the SSC estimation in this study.

In this paper, according to the approach proposed by Ouerghemmi *et al.* [7] which consisted in a "double-extraction" technique: 1) extraction of soil spectra from mixed spectra and 2) extraction of soil property contents from the extracted soil spectra [7], we used NMF to extract the soil spectra from the mixed spectra contained both vegetation and soil spectra. Then to predict SSC, PLSR models established by the extracted soil spectra were used. Therefore, the objectives of this study were 1) to further identify the effects of vegetation on spectral estimates of SSC; 2) to apply NMF to alleviate vegetation effects to improve the accuracy of estimates of SSC; and 3) to testify the effects of soil moisture in the SSC prediction after NMF process.

II. Materials and Methods

A. Experimental Site Description

In order to objectively testify the ability of NMF, a wide SSC and vegetation coverage ranges should be required. Therefore, we artificially added salt to the soil and planted barley in various seeding densities. Additionally, we regularly collected spectra, took photos, and sampled soil of each plot, during the growth process. Thus, we could obtain more variation in SSC and vegetation coverage. This study was conducted on the Huanghai Raw Seed Growing Farm ($32^\circ \sim 38'-40'$ N, $120^\circ \sim 52'-54'$ E) in Dongtai city, Jiangsu Province, China (Fig. 1), with the Yellow Sea to the east. The farm is approximately 5 km to the coastline of China Yellow Sea, and the topography is flat with an average elevation of 1.0–1.5 m. The climate is subtropical monsoon with four different seasons and with a mean annual temperature of 14.7 °C and an annual precipitation of 1042 mm. The field implemented our experiment was 220 m^2 (22 m $\times$ 10 m) in area, and located in the western part of the farm. The soil is an Aquent of the U.S. Soil Taxonomy classification and is a silt loam soil according to the U.S.D.A. texture classification. The parent material is marine sediment [25]. The SSC was about 0.58 g kg^{-1}.

B. Planting, Sampling, and Spectra Collecting

The experimental site was divided into 50 plots (5 lines × 10 columns corresponding to 5 SSC levels × 10 sowing densities), each with size of 1.5 m × 1 m. Since the SSC in the study field was relatively uniform, we added 0, 1.17, 1.95, 2.73, and 3.12 kg NaCl to each plot of 5 lines, respectively, to form a wider range of SSC levels on November 7, 2013. Three days later after thoroughly mixing of salt with soil, we sowed barley at 10 different densities (21, 36, 51, 66, 81, 96, 111, 126, 141, and 156 g seeds per plot) on each line of plots (Fig. 1). The barley grew under local farm management practices. Then, we synchronously observed the vegetation coverage and reflectance spectra of each plot, and sampled soils on December 6, 2013, December 30, 2013, and January 21, 2014 before the soil surface was fully covered with barley. Since the SSC and vegetation coverage for each plot could change temporally, a wider range of vegetation coverage, and variational SSC and SMC were obtained.

Before sampling soil, digital photos of the each plot were taken using a digital camera (EOS 10D, Canon Inc., Tokyo, Japan) positioned 1 m above the top of the vegetation for determining vegetation coverage later. Spectral reflectance of each plot was measured simultaneously with an ASD spectroradiometer (Fieldspec 3 Hi-Res, PANalytical, B.V, Boulder, CO, USA, formerly Analytical Spectral Devices) that covered the 350–2500 nm wavelength regions at 1 nm intervals. The aperture angle of the fore optic was 25°, and the spectral measurements were taken from the nadir at a height of 1.3 m and were centered in the middle of each plot, resulting in a 57-cm diameter field of view. Spectral measurements were performed under clear skies between 12:00 A.M. and 2:00 P.M. A white panel (1 × 1 m; Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences) was used to calibrate and optimize the instrument before each measurement. Ten spectra of each plot were taken.

After photos and spectra collection, in each plot, five subsamples were collected to a depth of 5 cm using an auger and were then mixed.

C. Soil and Spectral Data Preparation

A portion of each sample was air-dried, ground to pass through a 2-mm mesh and then analyzed for electrical conductivity (EC). The 1:5 soil and water suspensions were used to measure the EC with a conductivity meter conventionally. The EC was transformed into SSC (g kg^{-1}) using (2) [26]

$$y = 3.257x - 0.357 \tag{2}$$

where y is the SSC (g kg^{-1}) and x is the EC 1:5 value (ms cm^{-1}).

The other portion was used to determine the moisture content by oven drying. The SMC (g g^{-1}) was calculated as the ratio of water content (weight minus oven-dried weight) to oven-dried weight of each sample.

The vegetation coverage was extracted using the G–R thresholding method [27]. The green channel minus the red channel of a photo, then a threshold was set, pixels with G–R value higher than the threshold were sorted as vegetation and the rest as soil. Vegetation coverage can be obtained by calculating the proportion of vegetation.

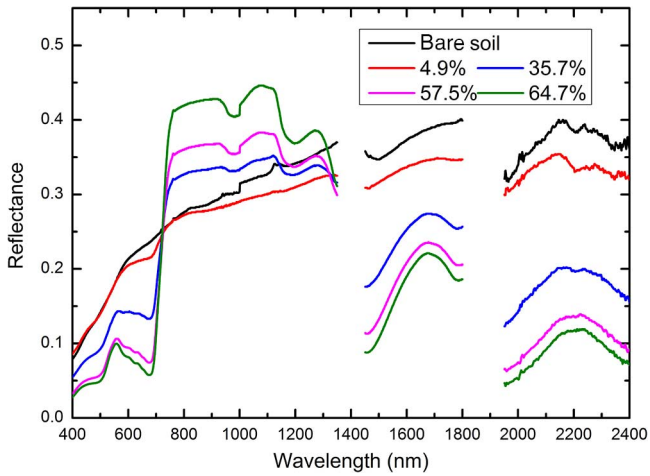

Fig. 2. Reflectance spectra of partially vegetated soil surfaces with different vegetation coverage.

Ten spectra taken in one plot were averaged and the averaged spectrum was identified as spectrum of each plot. In this study, the spectra were subjected to standard normal variate (SNV) preprocessing [28] to reduce the multiplicative interferences of scatter and particle size. Because the reflectance data in the extreme wavelength ranges from 350 to 399 nm and from 2401 to 2500 nm were significantly affected by noise, and the ranges from 1351 to 1450 nm and from 1801 to 1950 nm were affected by moisture in the air, the reflectance data from those bands were not included in our study. The reflectance spectra of partially vegetated soil surfaces with different coverage were shown in Fig. 2.

A total of 150 spectra (50 plots × 3 periods) and their corresponding values of SSC, SMC, and vegetation coverage were obtained. Due to mistakes in measurements, three spectra were not used in the later analyses. The total original spectral data were defined as Data set A, which then was split into two subsets, according to a method which had been widely applied to minimize the soil moisture effects in many studies [23], [29], [30]. Data set A_1 with lower SMC and Data set A_2 with higher SMC, separated by the threshold value of SMC of 0.15 g g^{-1}, as 0.15 was the median value of SMC in our study. The descriptive statistics of the SSC, SMC of the soil samples and the vegetation coverage of the three data sets were given in Table I.

D. Concept of BSS

BSS was a technique to recover source signals from their mixtures. The term "blind" has two meanings: 1) what the sources are is unknown and 2) how they are mixed is unknown. The mathematical model can be written as follows:

$$\mathbf{X} = \mathbf{AS} + \mathbf{R} \tag{3}$$

where $\mathbf{X}$ was the observed (mixed) signal matrix, $\mathbf{A}$ was the mixing matrix, $\mathbf{S}$ was the source signal matrix, and $\mathbf{R}$ was the residual matrix.

TABLE I
STATISTICS OF SMC, VEGETATION COVERAGE, AND SSC FOR THE THREE DATA SETS USED IN THIS STUDY

		Data set A (all samples)		Data set A_1 (samples of SMC<0.15 g g^{-1})		Data set A_2 (samples of SMC>0.15 g g^{-1})	
N		147		75		72	
Soil moisture content (g g^{-1})	Min	0.09	Min	0.09	Min	0.15	
	Max	0.20	Max	0.15	Max	0.20	
Vegetation coverage (%)	Min	0	Min	0	Min	0	
	Max	64.70	Max	57.90	Max	64.70	
Soil salt content (g kg^{-1})	Min	0.66	Min	0.66	Min	0.67	
	Max	30.32	Max	30.32	Max	22.83	

The aim of BSS was to solve (3), when both $\mathbf{A}$ and $\mathbf{S}$ were unknown. One of the most common techniques to solve BSS was ICA. With the premise that the data must meet some hypothesis, thus ICA can work. First, the source signals must be independent to each other; second, the number of observed signals must be more than the number of source signals; third, there were no equipment noise. In the ICA process, there were three parameters to be determined, according to users' specific intention: learning rate, the number of iterations, and optimization function. If these parameters were not chosen properly, the results would not be satisfactory.

E. Non-Negative Matrix Factorization

NMF was another technique of BSS, so it has all advantages of BSS, simultaneously overcomes some disadvantages of ICA. However, its application in spectral unmixing was not seen, we initially used NMF to separate soil and vegetation spectra in this study.

There were several NMF algorithms. In this study, we used the one initially proposed by Lee and Seung [31], with the constraint of non-negativity, allowed only additive and not subtractive combinations of original data.

The spectral reflectance database was an $n \times m$ matrix $\boldsymbol{X}$, each column of which contained n non-negative spectral reflectance values of one of the m spectra. The NMF constructed approximate factorizations of the form $X \approx WH$ [31]. The $\boldsymbol{W}$ and $\boldsymbol{H}$ matrices were purely non-negative, with the non-negativity constraints, and the dimensions of $\boldsymbol{W}$ and $\boldsymbol{H}$ were $n \times p$ and $p \times m$, respectively. The rank p was generally chosen by $p < (nm)/(n+m)$, and in our case, p was the number of spectral reflectance sources determined from the composition of mixed spectra.

To obtain the approximate $\boldsymbol{W}$ and $\boldsymbol{H}$, a cost function based on Euclidean distance was defined with (4) [18], [32]

$$\boldsymbol{E}(\boldsymbol{W},\boldsymbol{H}) = \|\boldsymbol{X} - \boldsymbol{W}\boldsymbol{H}\|^2 = \sum_{\boldsymbol{nm}} (\boldsymbol{X}_{\boldsymbol{nm}} - (\boldsymbol{W}\boldsymbol{H})_{\boldsymbol{nm}})^2 \tag{4}$$

where $\boldsymbol{E}(\boldsymbol{W},\boldsymbol{H})$ (size $n \times m$) was the residual or noise matrix. The process to acquire the approximate $\boldsymbol{W}$ and $\boldsymbol{H}$ was to minimize the $\boldsymbol{E}$ $(\boldsymbol{W},\boldsymbol{H})$. For NMF, there were various algorithms, and in this submission, we used an algorithm based on multiplicative update rules for $\boldsymbol{W}$ and $\boldsymbol{H}$. The Euclidean distance was no increasing under the updated rules and was calculated with (5)

$$\boldsymbol{H}_{\boldsymbol{am}} \leftarrow \boldsymbol{H}_{\boldsymbol{am}} \frac{(\boldsymbol{W}^{\boldsymbol{T}}\boldsymbol{X})_{\boldsymbol{am}}}{(\boldsymbol{W}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{H})_{\boldsymbol{am}}} \quad \boldsymbol{W}_{\boldsymbol{na}} \leftarrow \boldsymbol{W}_{\boldsymbol{na}} \frac{(\boldsymbol{X}\boldsymbol{H}^{\boldsymbol{T}})_{\boldsymbol{na}}}{(\boldsymbol{W}\boldsymbol{H}\boldsymbol{H})_{\boldsymbol{na}}} \tag{5}$$

where a was an integer variable $(a = 1, 2, \ldots, p)$. The Euclidean distance no longer changed under these updates if and only if $\boldsymbol{W}$ and $\boldsymbol{H}$ were at a stationary point of the distance [32]. Thus, the approximate $\boldsymbol{W}$ and $\boldsymbol{H}$ were obtained.

Spectra of 10 simulated vegetation coverage were calculated for each mixed (observed) spectrum, and then were used as input matrix $\boldsymbol{X}$ (size 1751×10) on which NMF was performed with the number of source spectra p equal to 2. After a certain times of updating (we chose 100 times), the output matrix $\boldsymbol{W}$ (size 1751×2) was obtained, which represented spectrum of soil and vegetation. Then the soil spectrum was discriminated through comparing the SA between each of the two extracted spectrum and the averaged spectrum of bare soil. The spectrum with a smaller SA value was identified as that of bare soil.

The NMF was conducted with MATLAB 7.9.0 (MathWorks Inc., Natick, MA, USA). The 147 spectra of Data set A after NMF processing were Data set B, which was also split into two subsets (Data set B_1 and Data set B_2), as in Section II-C, so the number of spectra in Data set B_1 and B_2 was the same with Data set A_1 and A_2.

F. Partial Least Squares Regression

PLSR with leave one-out cross validation was applied to establish a correlation between the extracted soil spectral values (independent variable) and SSCs (dependent variable). The PLSR model was implemented with Unscrambler X10.3 software (Computer-Aided Modeling, Trondheim, Norway).

To quantify the accuracy of the models, the coefficient of determination in the calibration (R_c^2), the coefficient of determination in the cross validation (R_{cv}^2), the root-mean-square error of the calibration ($RMSE_c$), the root-mean-square error of the cross validation ($RMSE_{cv}$), the ratio of the standard deviation to the root-mean-square error of the calibration (RPD_c), and the ratio of the standard deviation to the root-mean-square error of the cross validation (RPD_{cv}) were calculated. According to Viscarra Rossel *et al.*, six classes of models can be identified

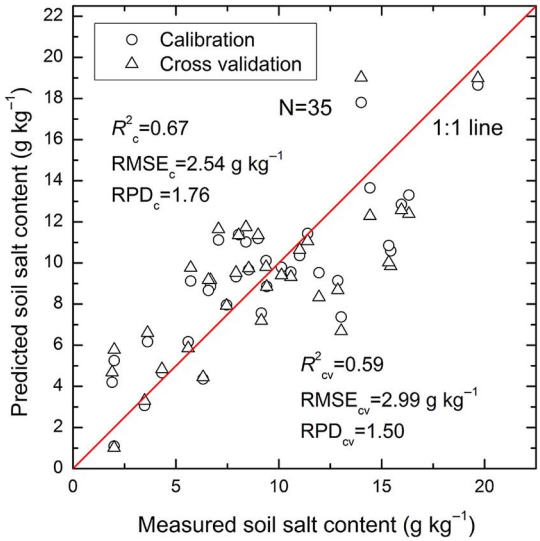

Fig. 3. Correlation between measured SSC and predicted SSC using PLSR calibrated with the bare soil spectra (vegetation cover <5% and SMC < 0.15 g g^{-1}).

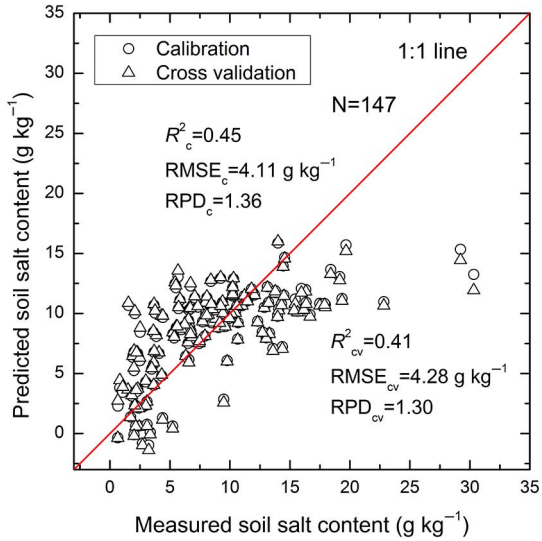

Fig. 4. Correlation between measured SSC and predicted SSC using PLSR calibrated with the original spectra (Data set A).

based on RPD values: RPD > 2.5 indicates excellent model, 2.0 < RPD < 2.5 indicates very good model, 1.8 < RPD < 2.0 indicates good model which predictions are possible, 1.4 < RPD < 1.8 indicates fair model which may be used for assessment, 1.0 < RPD < 1.4 indicates poor model which only high and low values are distinguishable, and RPD < 1.0 indicates very poor model which was not recommended to use [33].

III. Results and Analysis

A. Prediction of SSC Over Bare Soil

Thirty-five spectra under low vegetation coverage (lower than 5%) and low SMC (lower than 0.15 g g^{-1}) were selected to verify their ability in predicting SSC. The results were shown in Fig. 3. The predictions were acceptable, with $R^2_c = 0.67$ (N = 35), RMSE$_c$ = 2.54 g kg^{-1}, RPD$_c$ = 1.76, $R^2_{cv} = 0.59$ (N = 35), RMSE$_{cv}$ = 2.99 g kg^{-1} and RPD$_{cv}$ = 1.50.

The results indicated that SSC could be predicted over the bare soil samples selected for this study, which confirmed our previous work in the laboratory that SSC could be predicted from the spectral reflectance of air-dried soil samples [26].

B. Influence of Partial Vegetation Cover on the Estimation of SSC

The PLSR model was established using the original (mixed) spectra (Data set A), and the results showed that the accuracy of the SSC estimates was not satisfactory with vegetation coverage varying from 0% to 64.70%, and $R^2_c = 0.45$ (N = 147), RMSE$_c$ = 4.11 g kg^{-1}, RPD$_c$ = 1.36, $R^2_{cv} = 0.41$ (N = 147), RMSE$_{cv}$ = 4.28 g kg^{-1}, and RPD$_{cv}$ = 1.30 (Fig. 4). Thus, according to these parameters, the model developed from original spectra with different vegetation coverage did not perform well in estimating the SSC, identifying a common problem in practical application. Based on these results, removal of the vegetation effects before regression would be required to construct an accurate model.

C. NMF of the Mixed Spectra

In order to demonstrate the changes in spectral behaviors visually before and after NMF, three plots with vegetation coverage of 12.9%, 28.4%, and 41.9% were chosen as examples (Fig. 5). The reflectance spectra of bare soil and the three vegetated plots were shown in Fig. 5(a), while the extracted spectra after NMF corresponding to the original (mixed) spectra were shown in Fig. 5(b)–(d). The NMF could successfully separate each of the original spectra into two distinguishing spectra, one of which was very similar to soil spectrum, while the other was similar to vegetation spectrum. The results also demonstrated that the effectiveness of NMF decreased with the increasing vegetation coverage, as was shown in Fig. 5(d), the vegetation features, such as the red edge, remained more explicitly in both of the extracted spectra.

Fig. 6 showed all the SA values between the extracted spectra (vegetation and soil) and the averaged bare soil spectrum. It was clear that the SA values clustered into two groups, the higher one was identified as vegetation, and the lower one as soil, which confirmed the ability of NMF in separating mixed spectra. The SA values of the lower group were less than 10° (red line in Fig. 6) with vegetation coverage lower than approximately 30%. However, when vegetation coverage was more than 30%, the SA of extracted soil spectra became larger which meant that the extracted spectra were more different from the bare soil spectra. It also indicated that the ability of NMF in unmixing the mixed spectra weakened when it was applied to samples with higher vegetation coverage.

D. Effectiveness of the NMF-Extracted Soil Spectra for Estimation of SSC

To evaluate the effectiveness of the extracted soil spectra, PLSR models were established with Data set B to predict the SSC, and the leave one-out cross validation was performed on the data set to validate the model. The scatter plot of measured SSC against predicted SSC was shown in Fig. 7. Most of the

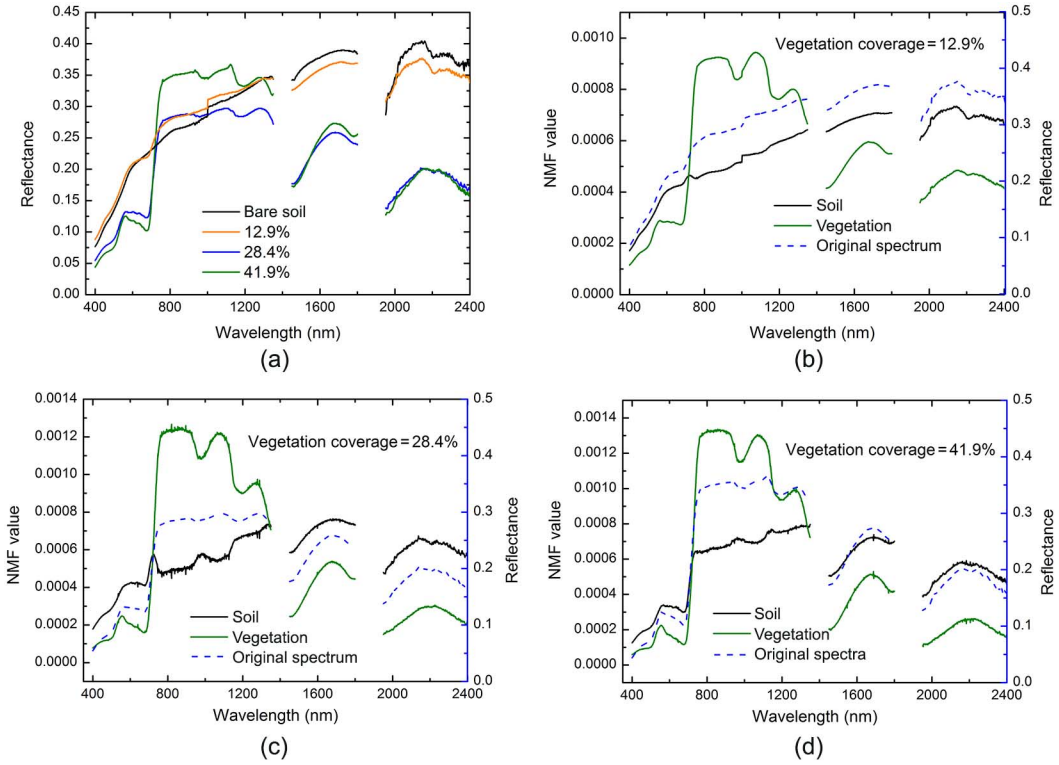

Fig. 5. (a) Original spectra with different vegetation coverage. (b–d) Separated spectra of the three original spectra using NMF.

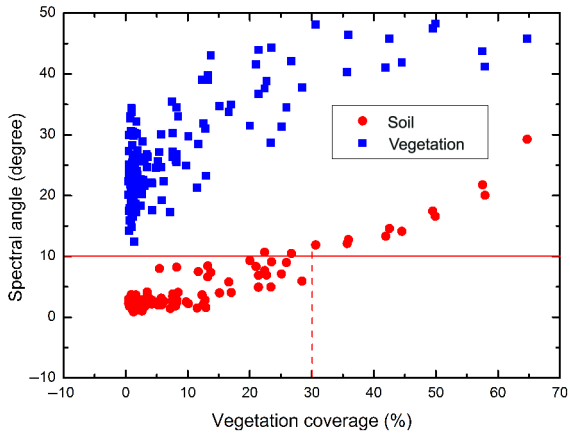

Fig. 6. SA values between extracted spectra and the averaged bare soil spectrum. The red line was used to indicate SA = 10°, and the red dashed line was used to indicate vegetation coverage = 30%.

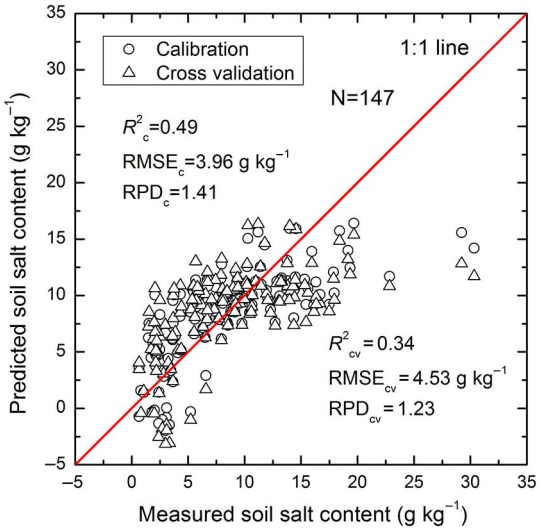

Fig. 7. Correlation between measured SSC and predicted SSC using PLSR calibrated with total NMF-extracted soil spectra (Data set B).

scatter points still deviated from the 1:1 line, and the prediction accuracy was not greatly improved, with $R^2_c = 0.49$ (N = 147), RMSE$_c$ = 3.96 g kg^{-1}, RPD$_c$ = 1.41, $R^2_{cv} = 0.34$ (N = 147), RMSE$_{cv}$ = 4.53 g kg^{-1}, and RPD$_{cv}$ = 1.23. These results were contrary to our initial expectation that the prediction accuracy would be greatly improved if the effects of vegetation coverage were removed from the mixed spectra.

We further divided the total data (Data set B) into two subsets (Data sets B_1 and B_2), according to SMC. The median SMC of 0.15 g g^{-1} was chosen as the threshold, as SMCs in our study were in the range of 0.09–0.20 g g^{-1}. Then, PLSR models were established using the two subsets, Data sets B_1 and B_2. The predicted versus measured plots for the PLSR calibration and the cross validation estimates for the different data sets were shown in Fig. 8. The scatter points of SSC were close to the 1:1 line, with $R^2_c = 0.91$ (N = 75), RMSE$_c$ = 1.51 g kg^{-1}, RPD$_c$ = 3.84, $R^2_{cv} = 0.63$ (N = 75), RMSE$_{cv}$ = 3.38 g kg^{-1}, and RPD$_{cv}$ = 1.72. The prediction accuracy was greatly improved compared with using the total data to establish the model. When SMCs were greater than 0.15 g g^{-1}, the accuracy of prediction was reduced rapidly, with $R^2_c = 0.54$ (N = 72), RMSE$_c$ = 3.59 g kg^{-1}, RPD$_c$ = 1.33, $R^2_{cv} = 0.37$ (N =

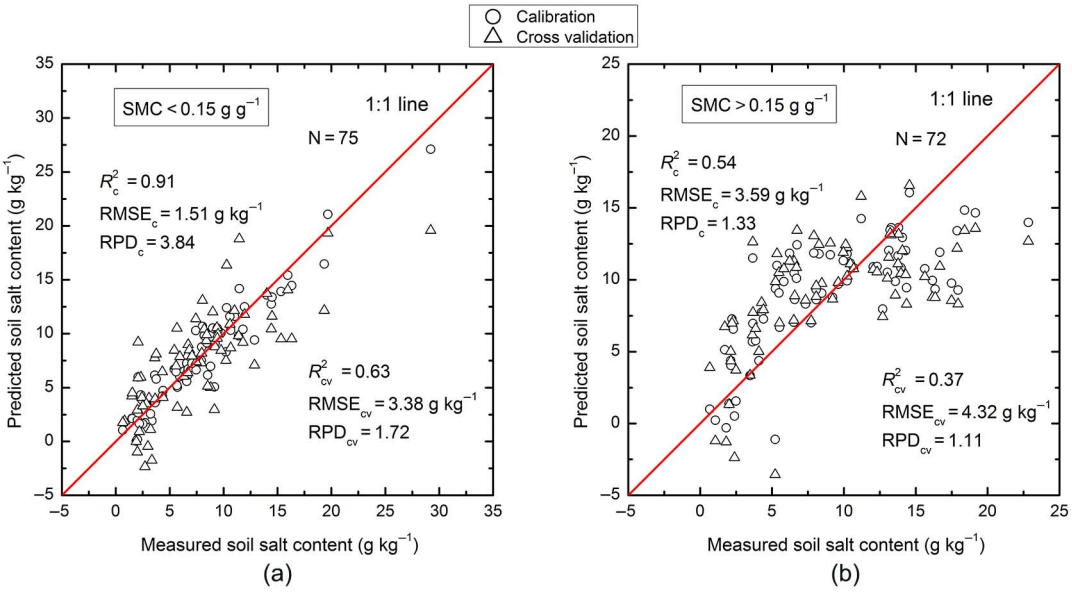

Fig. 8. Correlation between measured SSC and predicted SSC using PLSR calibrated with different data sets. (a) NMF-extracted soil spectra with SMC less than 0.15 g g^{-1} (Data set B_1). (b) NMF-extracted soil spectra with SMC greater than 0.15 g g^{-1} (Data set B_2).

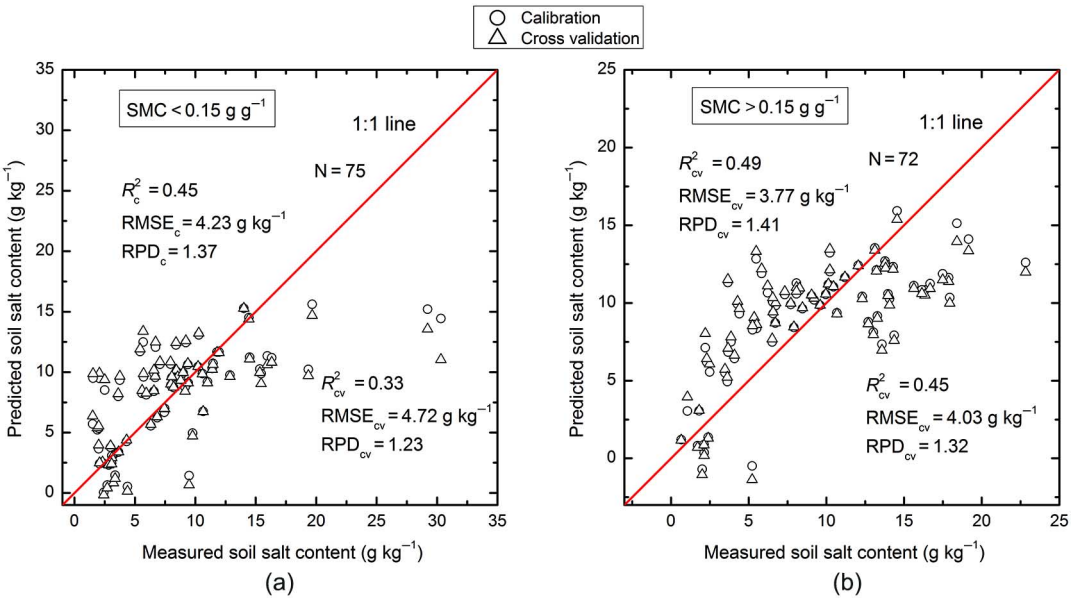

Fig. 9. Correlation between measured SSC and predicted SSC using PLSR calibrated with different data sets. (a) Original soil spectra with SMC less than 0.15 g g^{-1} (Data set A_1). (b) Original soil spectra with SMC greater than 0.15 g g^{-1} (Data set A_2).

72), $RMSE_{cv} = 4.32$ g kg^{-1}, and $RPD_{cv} = 1.11$ [Fig. 8(b)]. For NMF processed data, the accuracy of predicting SSC was higher when soil moisture was less than 0.15 g g^{-1}.

To compare the effects of SMC before and after NMF, we also divided the total original spectra data set (Data set A) into two subsets (Data sets A_1 and A_2) with the same rules that were applied to Data set B. Although the SMC was less than 0.15 g g^{-1}, the results were not satisfactory [$R_c^2 = 0.45$ (N = 75), $RMSE_c = 4.23$ g kg^{-1}, $RPD_c = 1.37$, $R_{cv}^2 = 0.33$ (N = 75), $RMSE_{cv} = 4.72$ g kg^{-1}, and $RPD_{cv} = 1.23$], and the scatter points of SSC seriously deviated from the 1:1 line [Fig. 9(a)]. The results were even worse when the SMC was more than 0.15 g g^{-1} [$R_c^2 = 0.49$ (N = 72), $RMSE_c = 3.77$ g kg^{-1}, $RPD_c = 1.41$, $R_{cv}^2 = 0.45$ (N = 72), $RMSE_{cv} = 4.03$ g kg^{-1}, and $RPD_{cv} = 1.32$], as shown in Fig. 9(b).

In order to demonstrate the effects of NMF, we compared the results derived from different analytical methods (including ICA and NMF) and no preprocess methods (original spectra). The comparisons were shown in Table II. For samples with all SMC, predictions of SSC with the NMF-, ICA-, and RSU-transferred spectra were nearly equal and worse than those with original spectra. For samples with low SMC, prediction of SSC with NMF-transferred spectra was better than other methods in both calibration and cross validation. For samples with high

TABLE II
STATISTICS FOR PREDICTIONS OF SSC USING A PLSR WITH SPECTRA AFTER DIFFERENT ANALYTICAL METHODS

Method	Data set	n	Calibration			Cross validation		
			R^2_c	$RMSE_c$	RPD_c	R^2_{cv}	$RMSE_{cv}$	RPD_{cv}
Original spectra	All samples	147	0.45	4.11	1.36	0.41	4.28	1.30
	Low SMC	75	0.45	4.23	1.37	0.33	4.72	1.23
	High SMC	72	0.49	3.77	1.41	0.45	4.03	1.32
NMF	All samples	147	0.49	3.96	1.41	0.34	4.53	1.23
	Low SMC	75	0.91	1.51	3.84	0.63	3.38	1.72
	High SMC	72	0.54	3.59	1.33	0.37	4.32	1.11
ICA	All samples	147	0.40	4.30	1.30	0.34	4.53	1.23
	Low SMC	75	0.79	2.66	2.18	0.28	4.94	1.17
	High SMC	72	0.50	2.78	1.42	0.31	4.51	1.19
RSU	All samples	147	0.40	4.31	1.29	0.34	4.55	1.22
	Low SMC	75	0.43	4.32	1.34	0.33	4.78	1.22
	High SMC	72	0.51	3.76	1.43	0.44	4.07	1.32

Note: All samples were samples with all levels of SMC; Low SMC was samples with SMC lower than 0.15 g g^{-1}; High SMC was samples with SMC higher than 0.15 g g^{-1}.

SMC, the prediction of SSC derived from original and RSU spectra were better than the results derived from NMF and ICA.

IV. DISCUSSION

A. Ability of NMF Compared With Other Methods

The results derived from the mixed spectra (Fig. 4) showed that the RPD values were less than 1.4, which indicated a poor model, especially the prediction for some samples with high salt content was worse. Because these samples usually had very low vegetation coverage whose spectra were very different from those with vegetated ones, after implementation with PLSR, the predictions were not good enough. The results further confirmed the conclusion of Bartholomeus *et al.* [6] and Ouerghemmi *et al.* [7] that vegetation cover seriously influenced the estimation of soil properties [6], [7]. Therefore, clearly, an effective method to separate the soil spectra from the mixed spectra to broaden the application of imaging spectrometry is required.

Only a few previous studies examined the separation of mixed spectral signals. The occurred RSU method was effective to obtain soil spectra, but with the prerequisite that the proportions of soil and vegetation should be known in advance [6]. While BSS can work without any prior knowledge of the mixed spectra, Ouerghemmi *et al.* [7] successfully implemented BSS using ICA technique to extract soil spectra from mixed spectra for clay prediction [7]. However, ICA demanded that the source spectral signals were independent from each other. Therefore, ICA may not apply to all cases. Furthermore, three parameters for ICA, all of which influence the results, should be optimized: the learning rate, the optimization function, and the number of iterations. The NMF applied in our study, however, has restrictions on neither the mixed nor the source spectral signals, and only requires the determination of the maximum number of iterations (100, in this study). Thus, the use of NMF may have wider applications, and the results of NMF may be more stable than those obtained with the ICA technique.

We also compared the performance of NMF, ICA, and RSU to account for the differences between mixed and extracted soil spectra to enable the use of spectral unmixing methods to predict properties of soil. Without considering SMC effects, the results of all the three methods were approximately the same (Table II), but both ICA and RSU have some limitations in source spectra or mixed spectra, while NMF does not. Additionally, the NMF method largely improved the prediction accuracy of samples with low soil moisture.

As noted in Section III-C, when vegetation coverage was approximately more than 30%, the SA values between the extracted soil spectra and the averaged bare soil spectrum were higher than 10° (Fig. 6), and indicated that NMF could not completely separate soil spectra from mixed spectra (a conclusion reached from the SA values). Thus, the threshold of vegetation coverage from which the effectiveness of NMF weakened was approximately 30%, and we wanted to clarify that threshold of 30% for the vegetation coverage was an approximate value, just from the SA values in our study, not an absolute value for other cases or areas, because we judge it a little arbitrarily in our study area. When vegetation coverage was higher, even though vegetation effects were not removed completely by NMF, it did not mean that NMF did not work. Compared with the mixed spectra, the extracted soil spectra contained more useful information for the prediction of SSC with improved prediction accuracy, as confirmed in Fig. 8(a). The range of vegetation coverage in this experiment was from 0% to 64.7%, all spectra after NMF were used to predict SSCs, and the result was acceptable as shown in Fig. 8.

The effectiveness of NMF weakened when vegetation coverage was high, which was understandable because it cannot capture enough soil information from the mixed spectra. Thus, it was not unexpected that when the soil surface was fully covered by vegetation, the NMF did not work. As we know, currently, none of the methods of spectra separation can be applied to all levels of vegetation coverage, particularly to high vegetation coverage. Ouerghemmi *et al.* [7] simulated mixtures that varied from 15% to 30% and separated the spectra

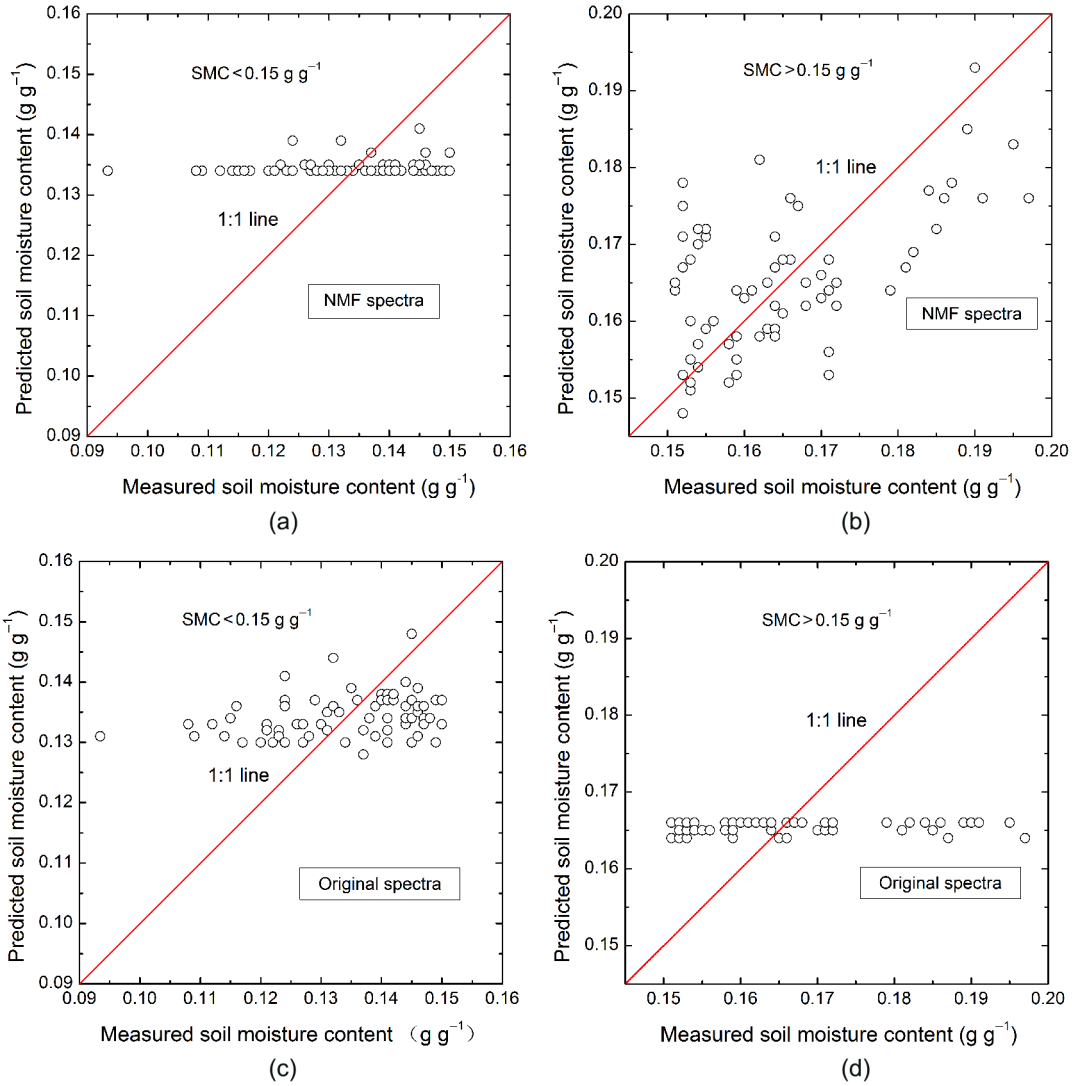

Fig. 10. Measured SMC versus predicted SMC using different data sets. (a) NMF-extracted soil spectra with SMC less than 0.15 g g^{-1} (Data set B_1). (b) NMF-extracted soil spectra with SMC greater than 0.15 g g^{-1} (Data set B_2). (c) Original soil spectra with SMC less than 0.15 g g^{-1} (Data set A_1). (d) Original soil spectra with SMC greater than 0.15 g g^{-1} (Data set A_2).

using the ICA method, but proposed that a minimum bare soil surface was required for the isolation of soil spectra [7]. In our case, NMF was highly effective until vegetation coverage reached 30%. When vegetation coverage exceeded 30%, the NMF method was unable to separate soil and vegetation completely [Fig. 5(d)], but the shape of extracted soil spectra was more similar to the bare soil compared with the original mixed spectra. Thus, though the vegetation effect was not totally removed, large parts were removed. Additionally, in most salt-affected areas, the vegetation cover typically is not high because salts seriously restrict the growth of plants [34].

The majority of our samples have a rather low vegetation cover, because the germination rates in the plots with high SSC were low and the vegetation coverage measured in the early period of growth was also very low, both the two reasons resulted in the majority of vegetation coverage was low. Even though most vegetation coverage was low, our samples still contained vegetation coverage ranged from 0% to 64.7%, and we thought the range of vegetation coverage was enough to prove the effectiveness of our method. Comparing the PLSR results derived from the spectra before and after NMF (covered the same vegetation coverage), the results could prove the effectiveness of our methods. Additionally, for saline area, vegetation coverage was usually not very high for grain crops.

For the original (mixed) spectra, the major factor to influence the accuracy of SSC prediction was vegetation cover because the soil surface was partially covered by vegetation and thus the dominating factor to influence the spectral reflectance. However, for extracted soil spectra by NMF, ICA, and RSU, as vegetation effects were alleviated and the spectra were viewed as bare soil spectra, soil moisture played an important role in determining the spectral reflectance and the accuracy of the prediction. When the SMC was low, the effects of moisture were not strong enough to be detected; however, with an increase in the SMC, the effects were more significant and eventually

decreased the accuracy of predictions of SSC. That was the reason why when SMC was high, all the spectral unmixing methods cannot improve the prediction accuracy.

B. Effects of SMC on NMF-Extracted Spectra

Theoretically, the extracted soil spectra should more accurately predict SSC than the original mixed spectra, according to previous studies [3], [35], [36]. However, soil moisture played an important role in decreasing the accuracy of soil salt estimation based on soil spectra.

In the recent decades, many researchers have examined the estimations of soil properties with various SMC conditions [25], [29], [30], [37], [38], and all demonstrated that the prediction accuracy dropped when the SMC was high, which was in accordance with our results. Nearly all those studies were conducted in the laboratory with artificial moisture contents obtained by the rewetting of soils, and the maximum SMC was set at approximately 0.20 and 0.25 g g^{-1}. In our study, the SMCs used were from the field, and the maximum SMC was also approximately 0.20 g g^{-1}. The results from this study demonstrated clearly that for partially vegetated surfaces, when SMC was below 0.15 g g^{-1}, the SSC can be predicted reasonably well *in situ*, after NMF preprocessing.

Further analyses were performed to support the above conclusion. We calibrated the SMC and spectra (original and NMF-extracted soil spectra) using PLSR. As shown in Fig. 10(a) and (c), for both the original and NMF spectra, when the SMCs were less than 0.15 g g^{-1}, the SMC was not predicted accurately, thus indicating that there was poor correlation between SMC and spectral values. When the SMCs were greater than 0.15 g g^{-1}, the SMC could not be predicted from original spectra [Fig. 10(d)] but were accurately predicted from NMF-extracted soil spectra, as shown in Fig. 10(b). These results provided further evidence for the conclusion that the original spectra were not affected by SMC over the entire range from 0.09 to 0.20 g g^{-1}, but that the NMF spectra were seriously affected by SMC if SMCs were higher than 0.15 g g^{-1}.

In our study area, the field moisture capacity was approximately 0.20 g g^{-1} [39], and thus 0.15 g g^{-1} was near the field moisture capacity. Generally, the SMC did not exceed 0.15 g g^{-1} in the field unless special circumstances occurred, such as rainy weather or irrigation. In our study area, the proposed method was appropriate. For higher SMC, we did not attempt to address the problem in this study, but further research to eliminate the effects of moisture using some other effective algorithms such as the EPO will be explored [40]. Many previous studies have demonstrated that SMC effects on the estimation of soil properties could be removed [37], [41], [42], so we foresee that NMF has great potential to improve the prediction accuracy of SSC over partially vegetated areas.

The studies conducted by Bartholomeus *et al.* [6] and Ouerghemmi *et al.* [7] used soil samples and spectra collected on one date to avoid the effects of variable SMC, whereas in our study, soil samples and spectra were collected on three different dates, which was more realistic and reflected actual conditions [6], [7]. Therefore, the inevitable variable SMCs should be considered.

C. Critical Points in NMF Application and Perspective

According to previous studies [17], [19], [43], the number of mixed spectra m must be equal to or greater than the number of source spectra p because the ultimate purpose of NMF was the compression of the original data set.

This study was performed on a field experiment, so vegetation type (barley) and soil texture (silt loam) were not variable, although both might influence results. In the future, more vegetation types should be tested to determine if NMF can be broadly applied. According to our previous study [26], the soils in the 50×20 km region were all silt loams without a difference in texture, and we plan to verify the effectiveness of NMF in areas with more heterogeneous soils in further studies.

V. Conclusion

The NMF effectively alleviated the effects of vegetation on the spectra and estimation of SSC, and broadened the use of spectroscopy which was limited to bare soil previously. Additionally, soil moisture was an important factor to be considered for NMF-extracted soil spectra. The use of NMF would facilitate the mapping of partially vegetated areas and would lead to the improvement of digital soil mapping, particularly where vegetation conditions vary with time or among different areas. In the future, we will test this method for removing the effects of partial vegetation cover on the estimation of SSC using remote sensing images.

References

[1] G. I. Metternicht and J. A. Zinck, "Remote sensing of soil salinity: Potentials and constraints," *Remote Sens. Environ.*, vol. 85, no. 1, pp. 1–20, Nov. 2003.

[2] A. Abbas, S. Khan, N. Hussain, M. A. Hanjra, and S. Akbar, "Characterizing soil salinity in irrigated agriculture using a remote sensing approach," *Phys. Chem. Earth. A B C*, vol. 55–57, pp. 43–52, Dec. 2013.

[3] J. Farifteh, F. Van der Meer, C. Atzberger, and E. J. M. Carranza, "Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods (PLSR and ANN)," *Remote Sens. Environ.*, vol. 110, no. 1, pp. 59–78, Oct. 2007.

[4] N. Fernández-Buces, C. Siebe, S. Cram, and J. L. Palacio, "Mapping soil salinity using a combined spectral response index for bare soil and vegetation: A case study in the former lake Texcoco, Mexico," *J. Arid Environ.*, vol. 65, no. 4, pp. 644–667, Jun. 2006.

[5] I. O. Odeh and A. Onus, "Spatial analysis of soil salinity and soil structural stability in a semiarid region of New South Wales, Australia," *Environ. Manage.*, vol. 42, no. 2, pp. 265–278, Aug. 2008.

[6] H. Bartholomeus, L. Kooistra, A. Stevens, M. Van Leeuwen, B. Van Wesemael, and E. Ben-Dor, "Soil organic carbon mapping of partially vegetated agricultural fields with imaging spectroscopy," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 13, no. 1, pp. 81–88, Feb. 2011.

[7] W. Ouerghemmi, C. Gomez, S. Naceur, and P. Lagacherie, "Applying blind source separation on hyperspectral data for clay content estimation over partially vegetated surfaces," *Geoderma*, vol. 163, no. 3–4, pp. 227–237, Jul. 2011.

[8] K. Wester, B. Lundén, and G. Bax, "Analytically processed Landsat TM images for visual geological interpretation in the northern Scandinavian Caledonides," *ISPRS J. Photogramm.*, vol. 45, no. 4–5, pp. 442–460, Oct. 1990.

[9] C. L. Wiegand, J. D. Rhoades, D. E. Escobar, and J. H. Everitt, "Photographic and video-graphic observations for determining and mapping the response of cotton to soil-salinity," *Remote Sens. Environ.*, vol. 49, no. 4, pp. 212–223, Sep. 1994.

[10] D. R. Tilley, M. Ahmed, J. H. Son, and H. Badrinarayanan, "Hyperspectral reflectance response of freshwater macrophytes to salinity in a brackish subtropical marsh," *J. Environ. Qual.*, vol. 36, no. 3, pp. 780–789, May/Jun. 2007.
[11] A. Thorhaug, A. D. Richardson, and G. P. Berlyn, "Spectral reflectance of Thalassia testudinum (Hydrocharitaceae) seagrass: Low salinity effects," *Amer. J. Bot.*, vol. 93, no. 1, pp. 110–117, Jan. 2006.
[12] T. T. Zhang, S. L. Zeng, Y. Gao, Z. T. Ouyang, B Li, C. M. Fang, and B. Zhao, "Using hyperspectral vegetation indices as a proxy to monitor soil salinity," *Ecol. Indic.*, vol. 11, no. 6, pp. 1552–1562, Nov. 2011.
[13] M. Zhang, S. L. Ustin, E. Rejmankova, and E. W. Sanderson, "Monitoring pacific coast salt marshes using remote sensing," *Ecol. Appl.*, vol. 7, no. 3, pp. 1039–1053, Aug. 1997.
[14] A. E. K. Douaoui, H. Nicolas, and C. Walter, "Detecting salinity hazards within a semiarid context by means of combining soil and remote-sensing data," *Geoderma*, vol. 134, no. 1–2, pp. 217–230, Sep. 2006.
[15] G. Ghosh, S. Kumar, and S. K. Saha, "Hyperspectral Satellite Data in Mapping Salt-Affected Soils Using Linear Spectral Unmixing Analysis," *J. Indian Soc. Remote.*, vol. 40, no. 1, pp. 129–136, Mar. 2012.
[16] D. Guillamet, J. Vitrià, and B. Schiele, "Introducing a weighted non-negative matrix factorization for image classification," *Pattern Recognit. Lett.*, vol. 24, no. 14, pp. 2447–2454, Oct. 2003.
[17] H. J. Oh, K. M. Lee, and S. U. Lee, "Occlusion invariant face recognition using selective local non-negative matrix factorization basis images," *Image Vis. Comput.*, vol. 26, no. 11, pp. 1515–1523, Nov. 2008.
[18] X. C. Xiong, X. Fang, Z. Ouyang, Y. Jiang, Z. J. Huang, and Y. K. Zhang, "Feature extraction approach for mass spectrometry imaging data using non-negative matrix factorization," *Chin. J. Anal. Chem.*, vol. 40, no. 5, pp. 663–669, May 2012.
[19] V. P. Pauca, J. Piper, and R. J. Plemmons, "Nonnegative matrix factorization for spectral data analysis," *Linear Algebra Appl.*, vol. 416, no. 1, pp. 29–47, Jul. 2006.
[20] W. D. Liu, F. Baret, X. F. Gu, Q. X. Tong, L. F. Zheng, and B. Zhang, "Relating soil surface moisture to reflectance," *Remote Sens. Environ.*, vol. 81, no. 2–3, pp. 238–246, Aug. 2002.
[21] D. B. Lobell and G. P. Asner, "Moisture effects on soil reflectance," *Soil Sci. Soc. Amer. J.*, vol. 66, no. 3, pp. 722–727, May/Jun. 2002.
[22] C. L. S. Morgan, T. H. Waiser, D. J. Brown, and C. T. Hallmark, "Simulated in situ characterization of soil organic and inorganic carbon with visible near-infrared diffuse reflectance spectroscopy," *Geoderma*, vol. 151, no. 3–4, pp. 249–256, Jul. 2009.
[23] C. K. Wang, X. Z. Pan, M. Wang, Y. Liu, Y. L. Li, X. L. Xie, R. Zhou, and R. J. Shi, "Prediction of soil organic matter content under moist conditions using VIS-NIR diffuse reflectance spectroscopy," *Soil Sci.*, vol. 178, no. 4, pp. 189–193, Apr. 2013.
[24] S. Kanzari, M. Hachicha, R. Bouhlila, and J. Battle-Sales, "Characterization and modeling of water movement and salts transfer in a semi-arid region of Tunisia (Bou Hajla, Kairouan)—Salinization risk of soils and aquifers," *Comput. Electron. Agric.*, vol. 86, pp. 34–42, Aug. 2012.
[25] Y. Liu, X. Z. Pan, C. K. Wang, Y. L. Li, R. J. Shi, R. Zhou, and X. L. Xie, "Predicting soil salinity based on spectral symmetry under wet soil condition," *Spectrosc. Spect. Anal.*, vol. 33, no. 10, pp. 2771–2776, Oct. 2013.
[26] Y. Liu, "Spectral response of saline soil to the change of soil moisture and salt content during the simulated evaporation," Institute of Soil Science, Chinese Academy of Sciences, Graduate University of Chinese Academy of Sciences, Nanjing, Jiangsu Province, China, 2012.
[27] Y. Wang, D. Wang, G. Zhang, and J. Wang, "Estimating nitrogen status of rice using the image segmentation of G-R thresholding method," *Field Crops Res.*, vol. 149, pp. 33–39, Aug. 2013.
[28] R. J. Barnes, M. S. Dhanoa, and S. J. Lister, "Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra," *Appl. Spectrosc.*, vol. 43, no. 5, pp. 772–777, 1989.
[29] M. Nocita, A. Stevens, G. Toth, P. Panagos, B. van Wesemael, and L. Montanarella, "Prediction of soil organic carbon for different levels of soil moisture using Vis-NIR spectroscopy," *Geoderma*, vol. 199, pp. 37–42, May 2013.
[30] E. A. Rienzi, B. Mijatovic, T. G. Mueller, C. J. Matocha, F. J. Sikora, and A. Castrignanò, "Prediction of soil organic carbon under varying moisture levels using reflectance spectroscopy," *Soil Sci. Soc. Amer. J.*, vol. 78, no. 3, p. 958, May 2014.
[31] D. D. Lee and H. S. Seung, "Learning the parts of objects by nonnegative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999.
[32] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Inf. Process. Syst.*, 2000, pp. 556–562.
[33] R. A. Viscarra Rossel, R. N. McGlynn, and A. B. McBratney, "Determining the composition of mineral-organic mixes using UV-vis-NIR diffuse reflectance spectroscopy," *Geoderma*, vol. 137, no. 1–2, pp. 70–82, Dec. 2006.
[34] H. E. Hayward and L. Bernstein, "Plant-growth relationships on salt-affected soils," *Bot. Rev.*, vol. 24, no. 8–10, pp. 584–635, 1958.
[35] N. H. Broge, A. G. Thomsen, and M. H. Greve, "Prediction of topsoil organic matter and clay content from measurements of spectral reflectance and electrical conductivity," *Acta Agric. Scand. B Soil Plant*, vol. 54, no. 4, pp. 232–240, 2004.
[36] Y. Liu, X. Z. Pan, C. K. Wang, Y. L. Li, R. Zhou, X. L. Xie, and M. Wang, "Prediction of coastal saline soil salinity based on VIS-NIR reflectance spectroscopy," *Acta Pedol. Sin.*, vol. 49, no. 4, pp. 824–829, Jul. 2012.
[37] Y. Ge, C. L. S. Morgan, and J. P. Ackerson, "VisNIR spectra of dried ground soils predict properties of soils scanned moist and intact," *Geoderma*, vol. 221–222, pp. 61–69, Jun. 2014.
[38] Q. Wang, P. Li, and X. Chen, "Modeling salinity effects on soil reflectance under various moisture conditions and its inverse application: A laboratory experiment," *Geoderma*, vol. 170, pp. 103–111, Jan. 2012.
[39] R. J. Yao, J. S. Yang, D. H. Wu, F. R. Li, P. Gao, and X. P. Wang, "Evaluation of pedotransfer functions for estimating saturated hydraulic conductivity in coastal salt-affected mud farmland," *J. Soils Sediments*, vol. 15, no. 4, pp. 1–15, Jan. 2015.
[40] J. M. Roger, F. Chauchard, and V. Bellon-Maurel, "EPO-PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits," *Chemometr. Intell. Lab. Syst.*, vol. 66, no. 2, pp. 191–204, Jun. 2003.
[41] B. Minasny, A. B. Mcbratney, V. Bellon-Maurel, J.-M. Roger, A. Gobrecht, L. Ferrand, and S. Joalland, "Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon," *Geoderma*, vol. 167–168, pp. 118–124, Nov. 2011.
[42] C. K. Wang, X. Z. Pan, Y. Liu, Y. L. Li, R. J. Shi, R. Zhou, and X. L. Xie, "Alleviating moisture effects on remote sensing estimation of crop residue cover," *Agron. J.*, vol. 105, no. 4, p. 967, Jul./Aug. 2013.
[43] X. Wang, H. Ling, and X. Xu, "Parts-based face super-resolution via non-negative matrix factorization," *Comput. Elect. Eng.*, vol. 40, no. 8, pp. 130–141, Nov. 2014.

Ya Liu was born in Heze, Shandong Province, China, 1986. She received the B.Sc. degree in geography information systems from Shandong Jianzhu University, Jinan, China, in 2009, the M.S. and Ph.D. degrees in agricultural resources and environment from the University of Chinese Academy of Sciences, Beijing, China, 2012, and 2015, respectively.

Currently, she is a Postdoctoral Researcher with the Institute of Soil Science, Chinese Academy of Sciences, Nanjing, China. Her research interests include the hyperspectral remote sensing of soils and proximal soil sensing.

Xian-Zhang Pan was born in Anhui, China, in 1965. He received the B.Sc. degree in soil agrochemistry from Huazhong Agricultural University, Wuhan, China, in 1987, the M.S. degree in soil science from the University of Chinese Academy of Sciences, Beijing, China, in 2000, and the Ph.D. degree in agricultural resources and environment from the University of Chinese Academy of Sciences, Beijing, China, in 2005.

Currently, he is a Professor with the Institute of Soil Science, Chinese Academy of Sciences. His research interests include remote sensing of soils especially in digital mapping of soil properties, land use changes, and soil quality assessment.

Dr. Pan is a member of the Soil Science Society of China. He was the recipient of the First class of S&T Progress Award of CAS, and the Special class of S&T Progress Award of National Bureau of Statistics of China.

Rong-Jie Shi received the B.Sc. degree in geography from Shandong Normal University, Jinan, China, and the M.S. degree in agricultural resources and environment from the University of Chinese Academy of Sciences, Beijing, China, in 2012 and 2015, respectively.

His research interests include monitoring heavy metal pollution in soil and vegetation using hyperspectral remote sensing.

Yan-Li Li received the B.Sc. degree in geography information system from Shandong Jianzhu University, Jinan, China, the M.S. degree in agricultural resources and environment from the University of Chinese Academy of Sciences, Beijing, China, in 2010 and 2013, respectively. She is currently pursuing the Ph.D. degree in agriculture at the University of Chinese Academy of Sciences.

Her research interest includes estimation of biomass by remote sensing.

Chang-Kun Wang received the B.Sc. degree in geography information system from Shandong Agricultural University, Tai'an, China, the Ph.D. degree in agricultural resources and environment from the University of Chinese Academy of Sciences, Beijing, China, in 2008 and 2013, respectively.

Currently, he is a Researcher with the Institute of Soil Science, Chinese Academy of Sciences. His research interests include estimation of crop residual cover by remote sensing.

Zhi-Ting Li received the B.Sc. degree in geography information system from Shandong Jianzhu University, Jinan, China, in 2013. She is currently pursuing the M.S. degree in geography information systems at the University of Chinese Academy of Sciences.

Her research interests include the estimation of crop residual cover by hyperspectral data.