Geoderma 418 (2022) 115845

Contents lists available at ScienceDirect

Geoderma

journal homepage: www.elsevier.com/locate/geoderma

Modified soil scattering coefficients for organic matter inversion based on Kubelka-Munk theory

Depin Ou [a], Kun Tan [b,c,*], Xue Wang [b,c], Zhifeng Wu [a], Jie Li [d], Jianwei Ding [e]

[a] School of Geography and Remote Sensing, Guangzhou University, Guangzhou, Guangdong 510006, China
[b] Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University, Shanghai 200241, China
[c] Key Laboratory of Spatial-temporal Big Data Analysis and Application of Natural Resources in Megacities, Ministry of Natural Resources, East China Normal University, Shanghai 200241, China
[d] South China Institute of Environmental Sciences, Ministry of Ecology and Environment of the People's Republic of China, Guangzhou, Guangdong 510655, China
[e] The Second Surveying and Mapping Institute of Hebei, Shijiazhuang 050037, China

ARTICLE INFO

Handling Editor: Kristin Piikki

Keywords:
Kubelka-Munk
Modified soil thickness model
Soil organic matter
Sensitive bands

ABSTRACT

The existing research into soil component inversion based on spectroscopic techniques has mainly focused on traditional statistical learning. However, the most prominent drawback of this approach is the difficulty in obtaining the soil components' sensitive bands with explanatory inversion mechanisms. Whether for soil organic matter inversion or soil heavy metal inversion, there is still a lack of inversion models based on the physical theory of remote sensing. Hence, in this paper, an inversion model based on thickness correction using Kubelka-Munk (K-M) theory is proposed. Firstly, in this study, a soil thickness observation experiment based on K-M theory was undertaken. The impact of the soil thickness and the material of the container on the spectra was explored by selecting different experimental samples with different background container materials. A modified K-M thickness model was then developed by combining indoor spectral data. This allows the corresponding scattering coefficients and absorption coefficients for soil samples with different organic matter contents to be calculated. The optimal organic matter inversion model can then be constructed by the scattering coefficients, with the sensitive band at 2.197 μm. The results obtained in this study demonstrate the feasibility and superiority of the proposed method and further explain the sensitive bands of soil organic matter in hyperspectral data, with a determination coefficient accuracy of up to 0.97. The experimental results also demonstrate that the recommended soil thickness for soil samples should be more than 7 mm. In addition, when selecting background container materials, materials with obvious reflectance peak and valley characteristics should be avoided.

1. Introduction

The application of hyperspectral techniques for the quantitative analysis of soil components is currently a hot topic (Wang et al., 2018). To date, most studies of hyperspectral applications for soil component inversion have focused on statistical methodology (Shi et al., 2014). Firstly, spectral denoising is performed using various pre-processing methods, such as the first-order derivative (Amigo et al., 2015) and continuum removal (Rezaei et al., 2008). The feature band combinations are then obtained by a band selection method, such as competitive adaptive reweighted sampling (Li et al., 2009). Finally, the soil component inversion model is obtained by partial least squares regression (PLSR) (Tan et al., 2021), support vector regression (SVR) (Tan et al., 2018), or a deep regression network (Ou et al., 2021). Although the data-driven statistical regression methods have been validated, but their generalization is poor. Besides, it is difficult to guarantee the accuracy and reliability of the models when applied in different places, data sources or even in different types of soil. Hence, one of the most the inadequate interpretation of the physical mechanism is their obvious drawback. Therefore, the current studies of both soil organic matter and soil heavy metal content inversion have encountered problems such as the unknown inversion mechanism and unclear sensitive bands (Angelopoulou et al., 2019). Therefore, it is of great significance to construct a mechanistic hyperspectral soil component inversion model based on a physical model, to explore the sensitive spectral characteristics of soil composition.

* Corresponding author at: Key Laboratory of Spatial-temporal Big Data Analysis and Application of Natural Resources in Megacities, Ministry of Natural Resources, East China Normal University, Shanghai 200241, China.
E-mail address: tankuncu@gmail.com (K. Tan).

https://doi.org/10.1016/j.geoderma.2022.115845
Received 11 October 2021; Received in revised form 13 February 2022; Accepted 23 March 2022
Available online 29 March 2022

The Hapke quadratic reflection model based on the bidirectional reflectance distribution function (BRDF) performs better than the analytical radiative transfer models for soil scattering properties (Hapke, 1981; Hapke, 2012). However, the application of this model requires precise multi-angle spectral data, which limits its application. Kubelka and Munk (1931) proposed the now famous Kubelka-Munk (K-M) theory for describing the optical behavior of radiation flux with matter. K-M theory is a two-flux approximation representation of general radiative transfer theory, with the advantage of physical interpretability and model simplicity when describing the scattering and absorption characteristics of matter(Yang and Kruse, 2004). As a result, K-M theory has been used in various scenarios, such as color prediction (de la Osa et al., 2020) and bandgap calculation for semiconductor materials (Murphy, 2007). To date, the application of K-M theory to hyperspectral inversion has mainly focused on soil moisture inversion, and although the application of K-M theory to other substances has been explored, the research work is limited. For example, Sadeghi et al. (2015) constructed a physically based soil moisture inversion model based on K-M theory to convert soil reflectance to a linear relationship model with soil moisture, and obtained stable and accurate results at 2.21 μm. Reich et al. (1953) transformed cotton fabric spectra into the K-M scattering theory space to estimate the content of soil adhering to cotton fabrics using a linear fitting relationship with log exponents. Barrón and Torrent (1986) constructed the color feature equation for a hematite and needle iron ore mixture based on K-M theory and used the soil color to estimate the iron oxide content in the soil. Gonçalves and Petter (2007) used the K-M model to find the absorption coefficient K and scattering coefficient S of kaolinite spectra by a linear decomposition method. The results of this study showed that the impurity content in kaolinite is highly correlated with coefficients S, and its determination coefficient can reach 0.9983 at 0.5 μm. The current K-M theory establishes the relationship between the ratio of the scattering and absorption features versus the diffuse reflectance spectrum, but it is difficult to decompose the scattering and absorption coefficients of the samples separately. A valid and straightforward assumption is that both the absorption coefficient and the scattering coefficient of a mixture consist of a linear superposition of the intrinsic absorption and scattering coefficients of the individual substances (Hu and Johnston, 2009; Yuan et al., 2019). This represents a potential breakthrough point for the construction of soil component physical inversion models based on K-M theory.

Soil reflectance spectra are mainly influenced by the physical properties, such as the soil surface roughness and particle size, as well as the soil moisture, organic matter, iron oxide, and calcium carbonate contents (Ben-Dor, 2002; Kokhanovsky, 2019). Soil organic matter and soil moisture have the most significant effect on soil spectra. With regard to soil moisture, numerous studies (Weidong et al., 2002; Zhang et al., 2020) have demonstrated that 1.4 μm, 1.9 μm, and 2.2 μm can be used as the feature bands, especially around 1.4 μm and 1.9 μm, which are heavily influenced by soil surface moisture. The reflectance spectra of soil decreases with the increase of the soil organic matter content. Especially when the soil organic matter content is below 2%, the reflectance decreases significantly with a slight increase of soil organic matter content (Al-Abbas et al., 1972). However, there is still no clear sensitive band for soil organic matter. Nevertheless, studies have shown that the feature bands are mainly concentrated in the intervals of 0.6–0.75 μm and 1.73–2.43 μm (Angelopoulou et al., 2019; Ben-Dor et al., 1997; Ting, 2006). The effect of iron oxide on soil reflectance spectra is mainly in the 0.52–0.62 μm spectral band (Camargo et al., 2015a; Stoner and Baumgardner, 1981). Heavy metals in soils are susceptible to adsorption by iron ion compounds, organic matter, and clay minerals. The spectral response bands of heavy metals in soil have not been clarified because the heavy metal content in most soil is too low (Shi et al., 2014; Tan et al., 2014; Wang et al., 2014).

Considering that laboratory soil spectra have been subject to air drying and sieving pretreatment, the effects of the particle size and moisture on the overall soil spectra can be considered to be negligible (Minasny et al., 2011). In small areas with homogeneous soils, the soil type is uniform and the dominant factor may be singular. Therefore, soil organic matter, soil texture, iron oxides and clay mineralogy or the other composition can be assumed to be the dominant factor influencing the differences in soil reflectance spectra for soil samples in the laboratory, respectively. Consequently, it can be assumed that the changes in soil organic matter components directly lead to the scattering and absorption coefficients in the K-M model. Therefore, constructing a method for obtaining the absorption and scattering coefficients of soil samples is the key to the successful application of K-M theory to soil component inversion (Vargas, 2002). A thickness equation based on K-M theory is an equation for the relationship among the scattering coefficient, material thickness and material spectrum, and the scattering coefficient of the material can be solved, theoretically, by obtaining multiple thickness observations (Kortüm, 2012; Kortüm et al., 1963). Therefore, in this study, we first designed a simple but effective observation experiment. This was then combined with the K-M thickness equation to obtain accurate scattering and absorption coefficients for the soil samples. Finally, the relationship model between the soil components and physical coefficients was established to achieve rapid detection of the soil components.

2. Datasets and methodology

2.1. Derivation of the K-M thickness model

In this section, the K-M thickness model is introduced using Gustav Kortüm's (2012) derivation as an example. We consider a flat sample with both scattering and absorption characteristics, for which the thickness is d, as shown in Fig. 1. The bidirectional light flux in the x-direction is specified as I and J, respectively. And dx for a sample of infinite thickness.

Given a light absorption coefficient $K(\mathrm{cm}^{-1})$ and a light scattering coefficient $S(\mathrm{cm}^{-1})$:

$$\begin{aligned} -\frac{dI}{dx} &= -(K+S)I + SJ \\ \frac{dJ}{dx} &= -(K+S)J + SI \end{aligned} \tag{1}$$

By analytically solving this equation, and by assuming that $(S+K)/S \equiv a$ and $J/I \equiv r$, Equation (1) can be drastically simplified as:

$$\int \frac{dr}{r^2 - 2ar + 1} = S \int dx \tag{2}$$

When integrating Equation (2), two boundary quantities can be obtained. When $x = 0$, it means that the background reflectance $R_g = (J/I)_{x=0}$. When $x = d$, it means that the sample's reflectance with thickness d is $R = (J/I)_{x=d}$.

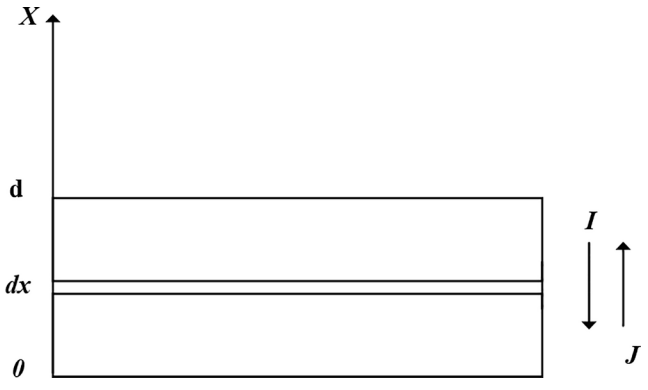


Fig. 1. Hypothesis of the Kubelka-Munk simultaneous equations.

After integrating Equation (2), the thickness relation can be obtained:

$$\ln\frac{\left(R-a-\sqrt{a^2-1}\right)\left(R_g-a+\sqrt{a^2-1}\right)}{\left(R_g-a-\sqrt{a^2-1}\right)\left(R-a+\sqrt{a^2-1}\right)}=2Sd\sqrt{a^2-1} \tag{3}$$

When $d=\infty$, the thickness of the flat sample is infinite. Therefore, its background reflectance can be handled as $R_g=0$. When the reflectance reaches the infinite reflectance value, R_∞ can be drastically simplified as:

$$R_\infty=\frac{1}{a+\sqrt{a^2-1}}=a-\sqrt{a^2-1}=1+\frac{K}{S}-\sqrt{\frac{K^2}{S^2}+2\frac{K}{S}} \tag{4}$$

By solving this equation for K/S, the so-called K-M function can be obtained:

$$\frac{K}{S}=\frac{(1-R_\infty)^2}{2R_\infty} \tag{5}$$

If we let the relationship between a and b have $\sqrt{a^2-1}\equiv b$, Equation (4) can be expressed as follows:

$$a=\frac{1}{2}\left(\frac{1}{R_\infty}+R_\infty\right) \tag{6}$$

$$\sqrt{a^2-1}\equiv b=\frac{1}{2}\left(\frac{1}{R_\infty}-R_\infty\right) \tag{7}$$

Combining Equations (3), (6), and (7) yields:

$$\ln\frac{(R-1/R_\infty)(R_g-R_\infty)}{(R_g-1/R_\infty)(R-R_\infty)}=Sd\left(\frac{1}{R_\infty}-R_\infty\right) \tag{8}$$

Equation (8) is the K-M thickness model. By solving this equation for R, it can be established that:

$$R=\frac{(1/R_\infty)(R_g-R_\infty)-R_\infty(R_g-1/R_\infty)\cdot\exp[Sd(1/R_\infty-R_\infty)]}{(R_g-R_\infty)-(R_g-1/R_\infty)\cdot\exp[Sd(1/R_\infty-R_\infty)]} \tag{9}$$

Equation (8) and Equation (9) show that the reflectance of a flat sample is related to the infinitely thick sample reflectance R_∞, the background reflectance R_g, and the "scattering power" Sd coefficient.

2.2. Soil organic matter inversion models

A simple and effective assumption is that the absorption and scattering coefficients of the mixed material are linearly superimposed by each individual material. Soil organic matter can be assumed to be the dominant factor influencing the differences in soil reflectance spectra for soil samples in a laboratory experiment. Therefore, the soil organic matter inversion model can be directly constructed as a linear model:

$$SOM=\mathrm{A}^*S+\mathrm{C} \tag{10}$$

where A and C are the linear fit parameters, and S represents the scattering coefficient. Therefore, obtaining accurate soil scattering coefficients S is the key to construct the inversion model. The value of k/s is easily obtained from the soil reflectance using the K-M model, but it is difficult to separate the scattering coefficient S and the absorption coefficient K. Hence, the following sections are the corresponding experimental methods that we designed.

2.3. Experimental design for the soil K-M thickness model

From Equation (8) and Equation (9), it can be seen that the sample's reflectance directly relates to the scattering power Sd. S is the scattering coefficient of the sample, which is determined by its material composition, particle size and homogeneity. S is constant for the same sample under the same conditions. Therefore, according to Equation (9), as the sample thickness d increases, the sample's reflectance R becomes larger. From Equation (8), for the same sample with determined R_∞ and R_g values, the scattering power Sd will have a linear property. When there is an ideal blackbody background material, $R_g=0$. However, in a realistic situation, it is not easy to find such a material, but the reflectance of R_g can be obtained in an experiment by directly measuring the spectrum of the container. The thickness of soil through which light can transmit is around 3–10 mm, so in a realistic situation, the reflectance obtained by direct measurement of the ground soil can be treated as R_∞. Furthermore, it is known from empirical knowledge that the reflectance of a thicker soil sample can be directly measured as R_∞.

According to Equation (8), and considering the characteristics of the soil sample, the calculation of the scattering coefficient of this soil sample can be achieved by directly measuring the reflectance spectrum of a soil sample with a large thickness as R_∞, measuring the reflectance spectrum of an empty container as R_g, and measuring at least one soil spectra at d thicknesses. Since the scattering coefficient is related to the sample's composition, it can be used to build the inversion model with the soil components, such as organic matter, which is the main component of the soil.

Therefore, we designed relevant experiments to obtain the scattering coefficients of soil samples as follows.

(1) The soil samples were pre-processed by air drying, impurity removal, grinding, and 100-mesh sieving.

(2) Cylindrical containers made of low reflectance materials were selected as the background containers.

(3) To ensure that the sample was flat and of consistent thickness in the container, a method of simulating the natural fall of soil was used. This was done by placing the sample on a 100-mesh sieve and simulating the soil particles' natural fall at a uniform speed.

(4) Since the soil sample's thickness is difficult to measure accurately, measuring the soil sample's weight was used instead. The specific operations were as follows. Firstly, the background container was placed on a high-precision balance for zeroing. The Petri dish container was filled with the soil sample, and its maximum weight M_t was measured. The height H of the container's inner wall was also measured using a ruler. Calculation of the soil thickness could then be achieved through Equation (11):

$$d=\frac{H}{M_t}{}^*M \tag{11}$$

where d is the calculated sample's thickness, M_t is the total weight of the Petri dish container that holds the soil sample, H is the total height of the container's inner wall, and M is the sample's weight with different thicknesses. Since the thickness d has a linear relationship with M, it is sufficient to use the weights of soil samples with different thicknesses as the parameters directly in the same sample.

(5) The spectrometer was secured and was warmed up for a sufficient period of time to prevent significant systematic errors. In the same position, the reflectance spectrum of the soil container was obtained, the spectra of the soil samples with at least five different thicknesses were obtained, and the reflectance spectra of the filled soil samples in the containers were obtained.

(6) The scattering coefficient of each sample was calculated using Equation (8).

2.4. Experiments with soil samples

(1) To study the effects of soil organic matter content on the soil samples' reflectance spectra, the comparison samples were subject to an organic matter removal operation. After the initial treatment, the soil samples were divided into two parts, one of which was calcined inside a muffle furnace at 600 °C for 8 h, thus achieving removal of soil organic

matter, while the other sample did not have additional treatment.

(2) To investigate the effect of different container materials on the soil thickness experiments, soil sample containers were selected made up of both transparent plastic Petri dishes made of polystyrene material and black Petri dishes treated with matte black spray paint, as shown in Fig. 2a.

In addition, a leveling protractor was used to level the samples uniformly, to simulate the soil's natural falling, to achieve a uniform distribution, as shown in Fig. 2b. The soil spectra were collected using a FieldSpec 3 spectrometer produced by ASD, USA, with a spectral range of 380–2500 nm. Five spectra were collected for each sample, and the collected spectra were then averaged to obtain the final sample spectra.

(3) According to the data sourced from National Soil Information Service Platform of China (https://www.soilinfo.cn) (1980–1990), organic matter is determined using the potassium dichromate volumetric method and divided into six classes (see more details in Table 1). To construct the soil organic matter inversion model, we selected nine soil samples collected in 2019 from the cultivated land of Yitong Autonomous Region, Jilin province, China According to the second national soil census classification criteria, the soil type in the study area belongs to dark brown loam with a slightly acidic to neutral response, pH 5.9–7.5. The texture is sandy loam with a through gravel content of about 30%. The effective cation exchange amount was 14.53me/100 g in the surface layer. Soil organic matter maintained at 3–5%. But in our research area, there is no natural soil organic matter content at Level 1, which is due to the degradation of soil organic matter in the area. The nine samples' organic matter content classification results are in Level 2 to Level 6, full details of which are listed in Table 1. The effects of the soil moisture and soil particle size on the soil spectra can be treated as consistent for all the soil samples after pre-processing. Fig. 3 shows the nine soil samples, where it can be seen that the soil with 9.6409 g/kg of organic matter exhibits a distinct orange-red color, indicating that the soil contains a high amount of iron. Petri dishes treated with matte black spray paint were then selected as the soil sample containers. The reflectance spectra of the soil samples were collected with at least five different thicknesses for all the soil samples using the ASD FieldSpec 3 spectrometer. Since variation in the lighting conditions can easily lead to variability of the soil samples' spectra, all the spectra needed to be measured together under the same lighting conditions.

Table 1
Soil organic matter classification and soil samples.

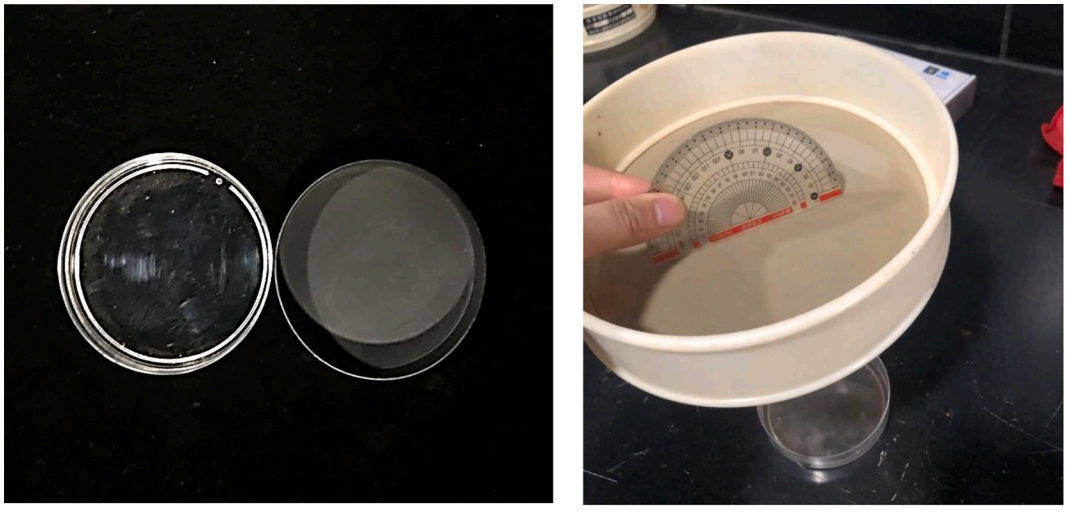
Level	Classification criteria	Soil samples (g/kg)
Level 1	SOM content >4.00%	–
Level 2	SOM content 3.01–4.00%	32.06, 35.59, 38.99
Level 3	SOM content 2.01–3.00%	26.04, 25.30
Level 4	SOM content 1.01–2.00%	19.44, 17.90
Level 5	SOM content 0.60–1.00%	9.64
Level 6	SOM content <0.60%	5.09

3. Results and analysis

3.1. Differences in the spectra of the soil samples before and after removal of organic matter

Figure 4 shows the soil spectra before and after removal of the organic matter. For Fig. 4a–c, the soil samples' containers were transparent plastic Petri dishes, and the soil spectral curves of the maximum soil mass were measured when the containers were filled. For Fig. 4d, the soil sample container was a black plastic Petri dish. As can be seen in both Fig. 3a and Fig. 4b, the reflectance of the soil increases as the thickness (mass) of the soil sample increases, but after reaching a certain level of soil thickness, the spectral curve either does not change much or undergoes small changes within the margin of error. When the soil thickness is at a thin level, the soil reflectance spectrum appears relatively low due to the light transmittance effect. The soil reflectance spectra before and after the removal of organic matter are compared in Fig. 4c. It should be noted that the valleys formed at 1.912 μm and 2.21 μm in the spectra of the soils containing organic matter are located at 1.906 μm and 2.184 μm, respectively, in the spectra of the soils with organic matter removed.

The Pearson correlation coefficient method was introduced to compare the linear relationship between the reflectance and sample thickness (mass) of the soil samples before and after the removal of organic matter, and the results are shown in Fig. 5. It can be seen that the soil sample reflectance before and after the removal of organic matter have a strong positive correlation with the soil thickness, while the spectra before 0.4 μm may show a negative correlation due to the systematic error of the spectroscopic instrument.



a b

Fig. 2. (a) Transparent Petri dish and Petri dish sprayed with matte black spray paint. (b) Simulating the natural fall of soil.

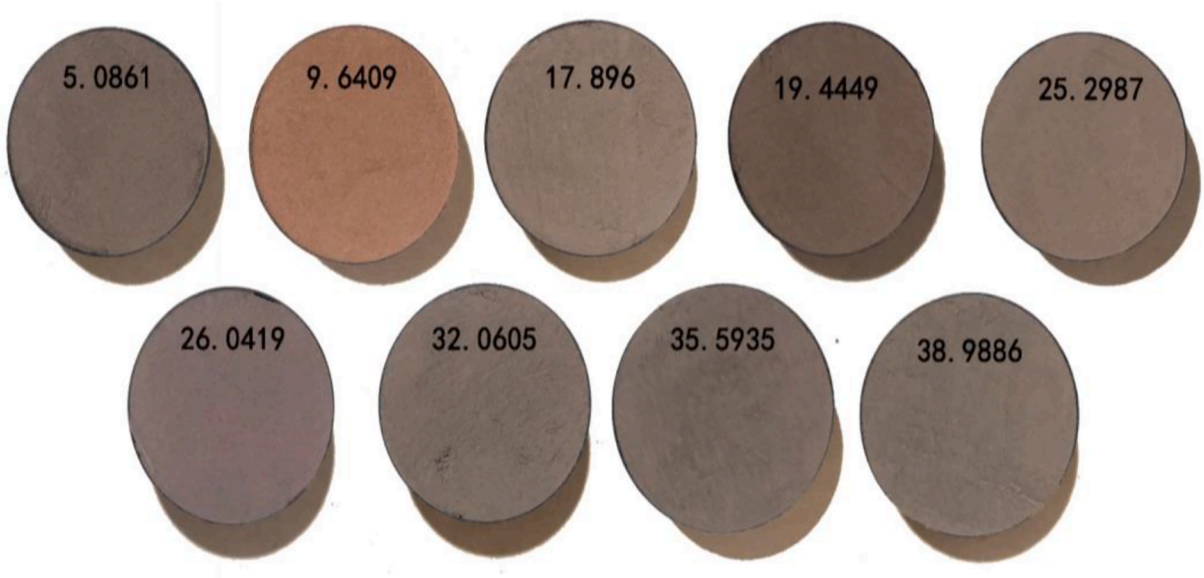


Fig. 3. The nine soil samples.

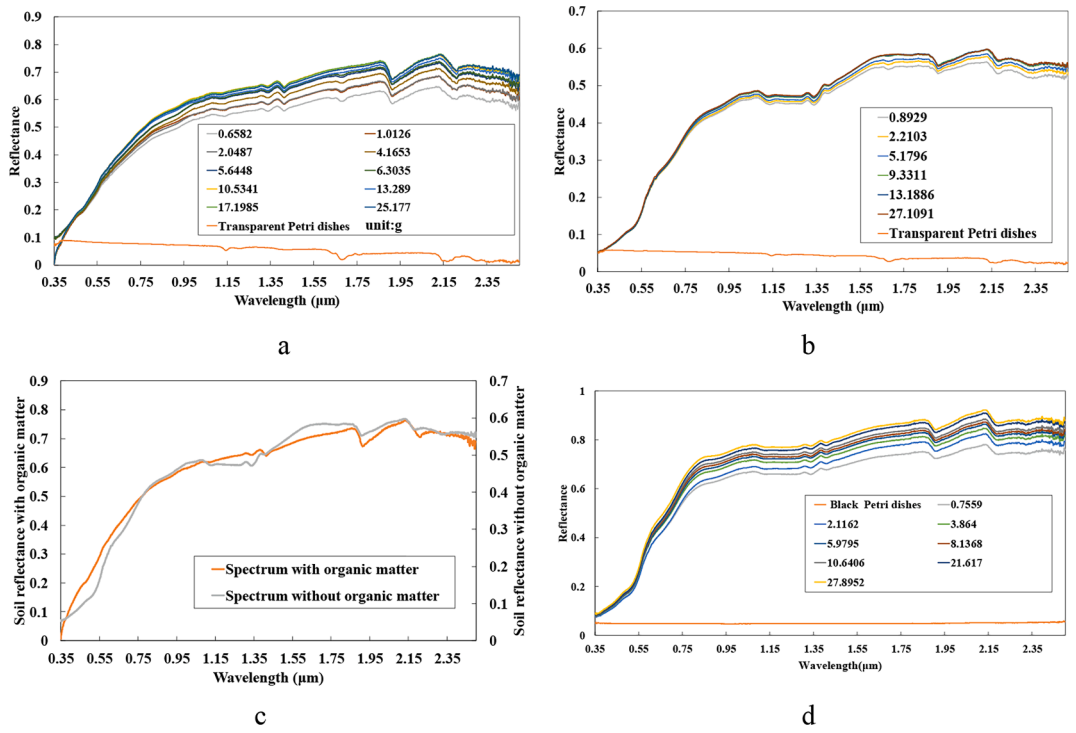


Fig. 4. Comparison of the soil spectra before and after the removal of the organic matter. (a) Soil spectra with organic matter in transparent plastic Petri dishes. (b) Soil spectra without organic matter in transparent plastic Petri dishes. (c) Two soil spectra before and after the removal of organic matter. (d) Soil spectra without organic matter in black plastic Petri dishes.

3.2. Modified scattering power Sd

3.2.1. Sd coefficients under different thickness

In this experiment, the Petri dishes treated with matte black spray paint were selected as the containers, and the soil spectra were collected from samples with different thicknesses (masses) after the soil was treated by organic matter removal, as shown in Fig. 4b. Equation (8) shows that, in order to obtain the *Sd* coefficient, it is necessary to know R_∞ and R_g. In an experiment, the measured spectrum of the black Petri dish can be used as R_g. The spectrum of the soil sample in the black Petri dish can be used as R_∞, and the *Sd* coefficient can be calculated using the rest of the measured spectra. Fig. 6 shows the calculated *Sd* coefficients at different thicknesses (masses) of soil. As the thickness (mass) of the soil sample increases, its *Sd* coefficient increases. From Equation (8), it can be seen that *Sd* should be, theoretically, linearly increasing, i.e., the scattering coefficient *S* is a constant under the same sample, and *Sd* varies linearly and multiplicatively with the thickness of the sample. Clearly, the calculated coefficients do not change multiplicatively with the change of thickness *d*.

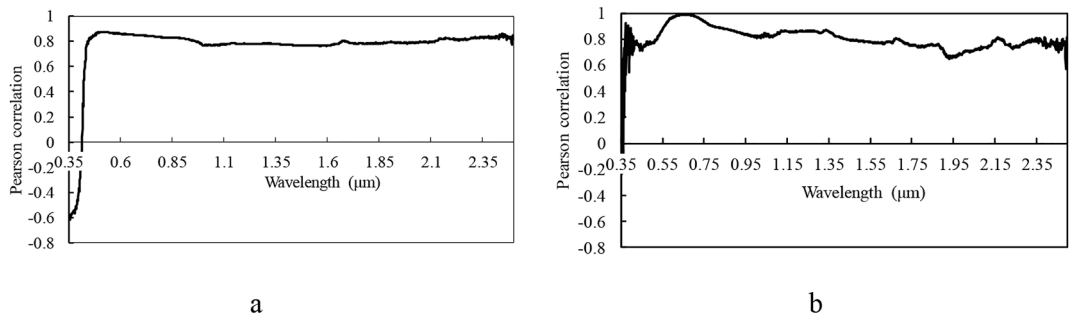


Fig. 5. Pearson correlation coefficients between the reflectance and soil sample thickness (mass). (a) Soil samples containing organic matter. (b) Soil samples after the removal of organic matter.

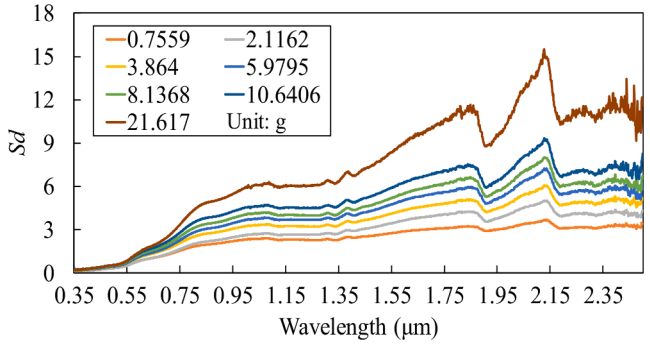


Fig. 6. *Sd* coefficients under different thickness.

3.2.2. Modified Sd

Figure 7 shows the scatter plot for the soil *Sd* coefficients versus soil thickness (mass) under the black Petri dish case. Fig. 8 shows the scatter plot of the soil *Sd* coefficients versus soil thickness (mass) under the transparent Petri dish case. By observing the scatter plot of the soil thickness (mass) versus *Sd* coefficient, it can be found that the *Sd* coefficient is linearly related to the soil thickness (mass) over the whole range of the wave spectrum, i.e.

$$Sd = \alpha m + \beta \tag{12}$$

where α and β are the coefficients and m is the thickness or mass of the soil sample.

From Equation (8), there should be no correction factor β under ideal conditions, just the *Sd* coefficient. However, there may be errors in the actual measurement processing, or error can be caused by the instrument's systematic error. Furthermore, error can also be caused by small variation of the distance of the instrument fiber from the sample. Although the samples had been passed through a 100-mesh sieve, some soil has smaller particle sizes than the 100-mesh aperture, which means that inconsistency in the particle size can also lead to this variation. However, this does not affect the fact that the *Sd* coefficients can be simulated with a higher accuracy using $\alpha m + \beta$ instead in these simple experiments. The coefficient α can be seen as the scattering coefficient S of the soil sample, while the coefficient β can be considered as the spectral error correction coefficient. That is, Equation (8) can be corrected to the following form:

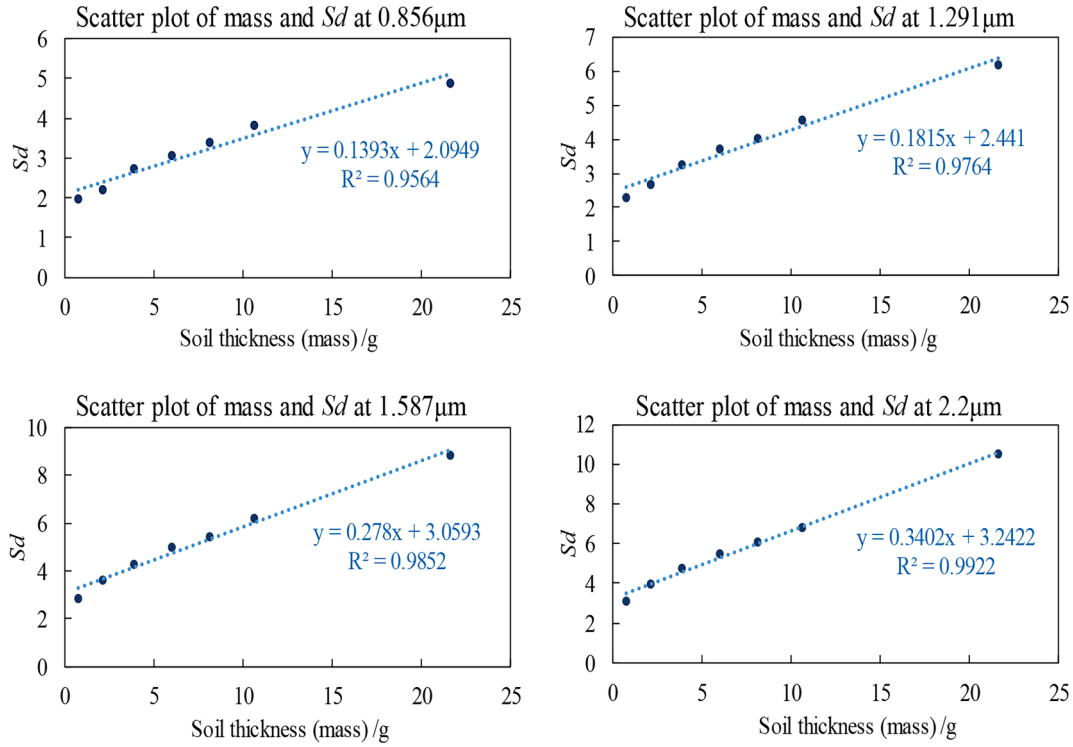


Fig. 7. Scatter plots of soil *Sd* and soil thickness (mass) in the black Petri dish.

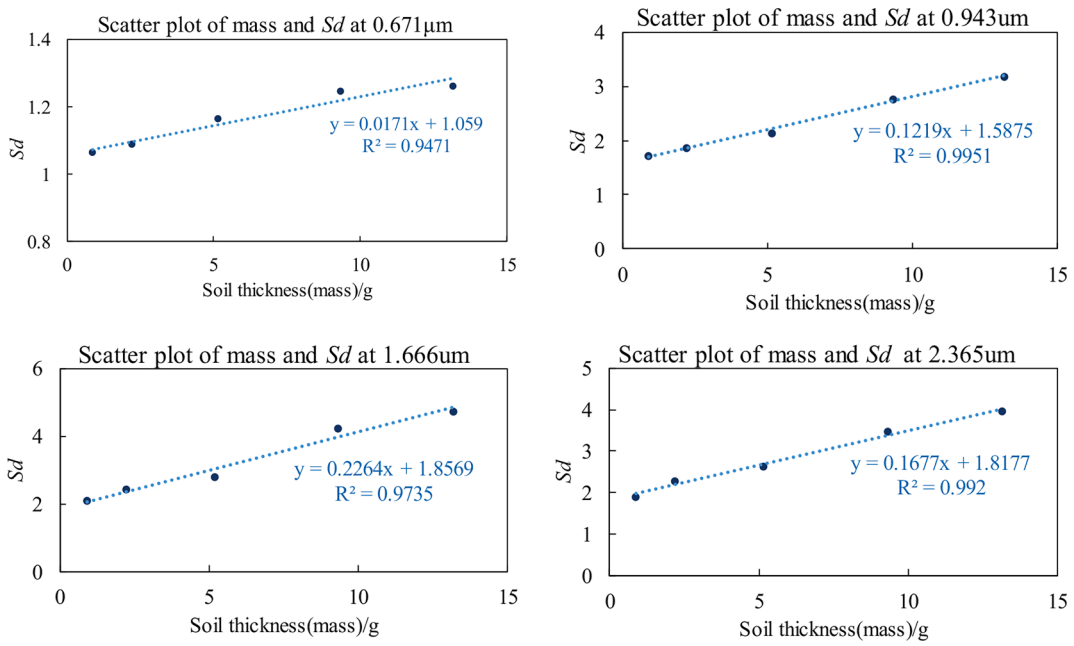


Fig. 8. Scatter plots of soil *Sd* and soil thickness (mass) in the transparent Petri dish.

$$\ln\frac{\left(R-\frac{1}{R_\infty}\right)\left(R_g-R_\infty\right)}{\left(R_g-\frac{1}{R_\infty}\right)\left(R-R_\infty\right)}=(\alpha d+\beta)\left(\frac{1}{R_\infty}-R_\infty\right) \tag{13}$$

Each sample has a scattering coefficient α with a correction factor β, which is independent of the wavelength and is not related to the background material. Equation (13) shows that the solution to α and β can be obtained by measuring the spectra of at least two soil samples with known thicknesses, the spectrum of an infinite thickness soil sample, and the background container spectrum.

3.2.3. *The performance of α and β*

Figure 9 shows the coefficients α and β, as calculated by the decomposed *Sd* coefficients according to Equation (13). From Fig. 9a, it can be seen that the coefficients α and β are consistent with the original spectrum, and there is also similarity in the peak-to-trough performance. From Fig. 9b, the correlation between the coefficient *Sd* and the coefficient after linear fitting shows that all the spectra after 0.7 μm can reach a correlation of 0.9 and above, while the correlations before 0.4 μm show an unstable state due to the instrument error problem. In general, the coefficients α and β can accurately describe the variation of the soil samples with thickness. By comparing the coefficients α with the coefficients *Sd* in Fig. 6, it can be found that the waveforms are consistent.

Figure 10 shows the scatter plots of the coefficients α, β and the infinitely thick original spectral reflectance, respectively. The coefficient α can be treated as the scattering coefficient *S*. As the coefficient α increases, the soil spectral reflectance first increases and then plateaus to a steady state. A simple fit by the logarithmic relation between the original spectral reflectance and the coefficient α can reach an absolute coefficient of 0.993, while the highest absolute coefficient of 0.9978 can be achieved by using a quadratic fit of the coefficient β. In summary, the scattering coefficient calculated by the actual measured spectra has some difference with the theoretical value, but can be effectively corrected by the use of Equation (13).

3.3. *Relationship between soil organic matter and the scattering coefficient*

The scattering coefficient α and correction factor β for each soil sample were obtained via Equation (13). Fig. 11 shows the scattering coefficients α for all the soil samples. The scattering coefficients α for soil samples with low organic matter content are relatively low, especially in the short-wave infrared interval. Furthermore, the visible spectrum shows linearity, except for the sample with an organic matter content of 9.6409 g/kg, which is anomalous. For the wavelength range before 0.4 μm or after 2.45 μm, the increase in noise due to the instrument system error is reflected in the scattering coefficient α. The scattering coefficient

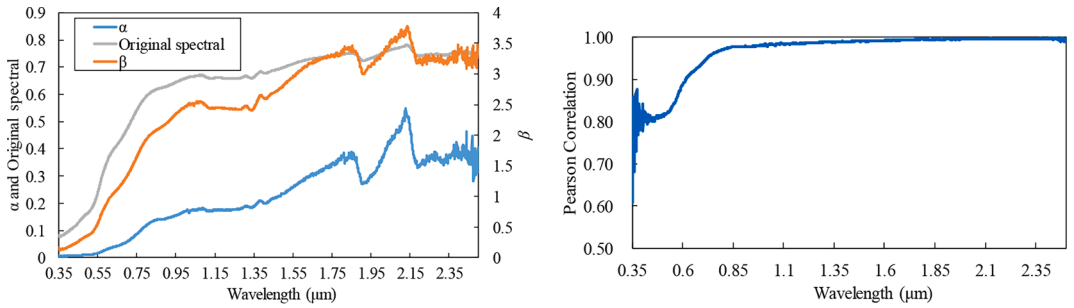


Fig. 9. (a) Comparison chart of α, β and the original spectrum. (b) The correlation between coefficient *Sd* and the coefficient after linear fitting.

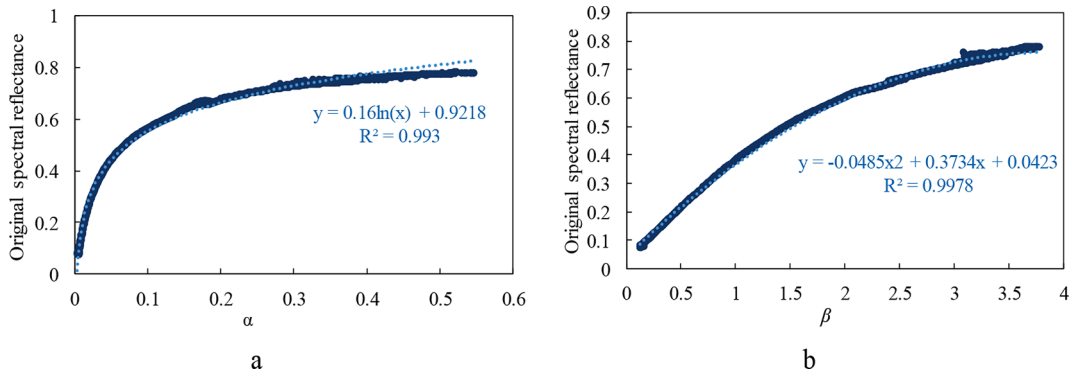


Fig. 10. Scatter plots of coefficients α, β and the infinite thickness original spectral reflectance.

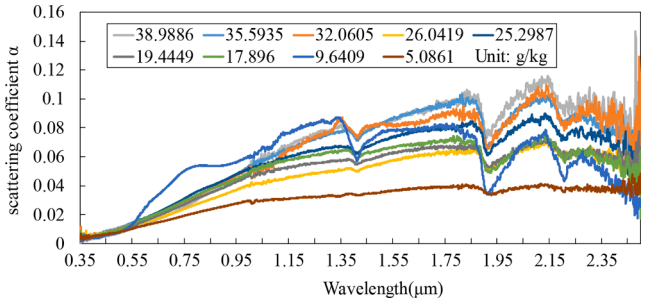


Fig. 11. Scattering coefficient α of all the soil samples.

α of the sample with 9.6409 g/kg of soil organic matter is anomalous in the visible wavelength range relative to the other samples. The picture of the soil sample in Fig. 3 shows that this sample contains a high amount of iron. Most of the iron oxide absorption peaks in the soil are mainly concentrated around 0.35–1.1 μm, leading to a significant scattering coefficient in the visible spectral range (Camargo et al., 2015b).

The soil spectrum of the filled Petri dish container was selected as the infinite thickness spectral reflectance, and the scattering coefficient α was substituted into Equation (5) as S to obtain the absorption coefficient K, as shown in Fig. 12. The absorption of light by the soil samples is mainly concentrated in the visible range. The absorption value decreases from the visible to the short-wave infrared spectral range. As the organic matter content of the soil increases, its absorption coefficient also tends to increase. Pronounced absorption peaks can be observed around 1.4 μm, 1.9 μm, and 2.2 μm, while the noise phenomenon is more noticeable in the bands before 0.4 μm and after 2.45 μm, due to instrumentation errors. In the visible spectral range, the scattering coefficient α and absorption coefficient K have opposite curve trends. Related studies have suggested that this phenomenon may be caused by the composition of soil organic matter, i.e., mainly humic acid, which has a strong absorption effect in the visible spectrum(Ben-Dor, 2002).

The Pearson correlation coefficients between scattering coefficient α,

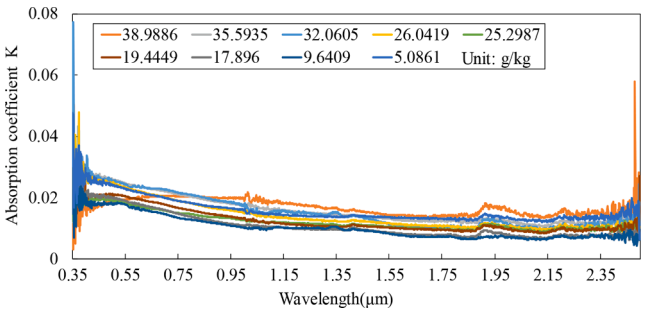


Fig. 12. Absorption coefficient K of all the 9 soil samples.

absorption coefficient K, coefficient β, the original spectrum, and the soil organic matter content are shown in Fig. 13. The scattering coefficient α correlation curve shows that the scattering coefficient α in the short-wave infrared spectral range has a strong correlation with the soil organic matter content, especially in the peaks around 1.4 μm, 1.9 μm, and 2.2 μm. Moreover, at 0.56–0.85 μm, a peak with a lower correlation is observed. The correlation reaches a maximum value of 0.9850 at 2.197 μm. At the same time, there is also a peak at 1.902 μm, with 0.9823. As for 1.412 μm, the peak is relatively low, at 0.8348. The 1.4 μm and 1.9 μm peaks are mainly the characteristic absorption bands of water vapor, which are important for soil moisture monitoring. The soil samples in the laboratory were treated by air drying, so that the effect of moisture was negligible.

The Pearson correlation coefficient curves between the absorption coefficient K and organic matter show relatively good correlation within the visible spectrum. In the range of 0.71–1.3 μm, the Pearson correlation coefficient is higher than 0.6 and can reach about 0.7. In contrast to the scattering coefficient, the absorption coefficient has a higher correlation in the visible spectral range, but performs poorly in the short-wave infrared spectral range. The highest value of 0.8160 is reached at 0.351 μm, which is due to instrument error. Hence, this can be recognized as an anomaly.

The absolute Pearson correlation coefficient value of the correction coefficient β and the original spectrum are low, and the correlation curves of the two are similar. The Pearson correlation coefficients of both curves are mostly lower than 0.7, indicating that it is difficult to invert the soil organic matter content directly using the original spectrum and the correction coefficient β.

Figure 14 shows scatter plots for scattering coefficient α at 2.197 μm, scattering coefficient α at 1.902 μm, scattering coefficient α at 1.412 μm, and absorption coefficient K at 0.76 μm with soil organic matter content. The scattering coefficients α at 2.197 μm, 1.902 μm, and 1.412 μm are at the peak of the correlation with soil organic matter content, while absorption coefficient K at 0.76 μm represents the best result with high correlation between absorption coefficient K and soil organic matter content. The scattering coefficient α shows better results for the linear fit at both 2.197 μm and 1.902 μm, where it has the largest determination coefficient of 0.9702 at 2.197 μm, whereas it performs poorly at 1.412 μm, with a determination coefficient of 0.6969. In contrast to the absorption coefficient K, which has strong correlation in the visible wavelength band, the determination coefficient at 0.76 μm reaches only 0.4024, as shown by the scatter plot.

4. Discussion

4.1. Effect of the soil background container on the spectrum

When there is an ideal blackbody background, for which $R_g = 0$, the effect of the background container on the soil spectrum disappears. However, in reality, it is not easy to find such an ideal material.

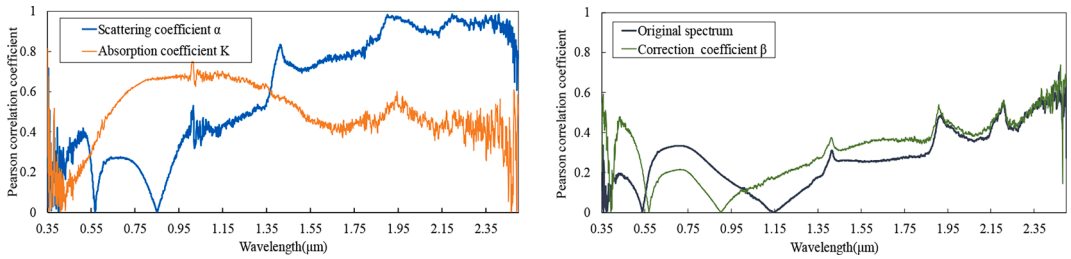


Fig. 13. Correlation diagrams of scattering coefficient α, absorption coefficient *K*, coefficient *β*, and the original spectrum with soil organic matter content.

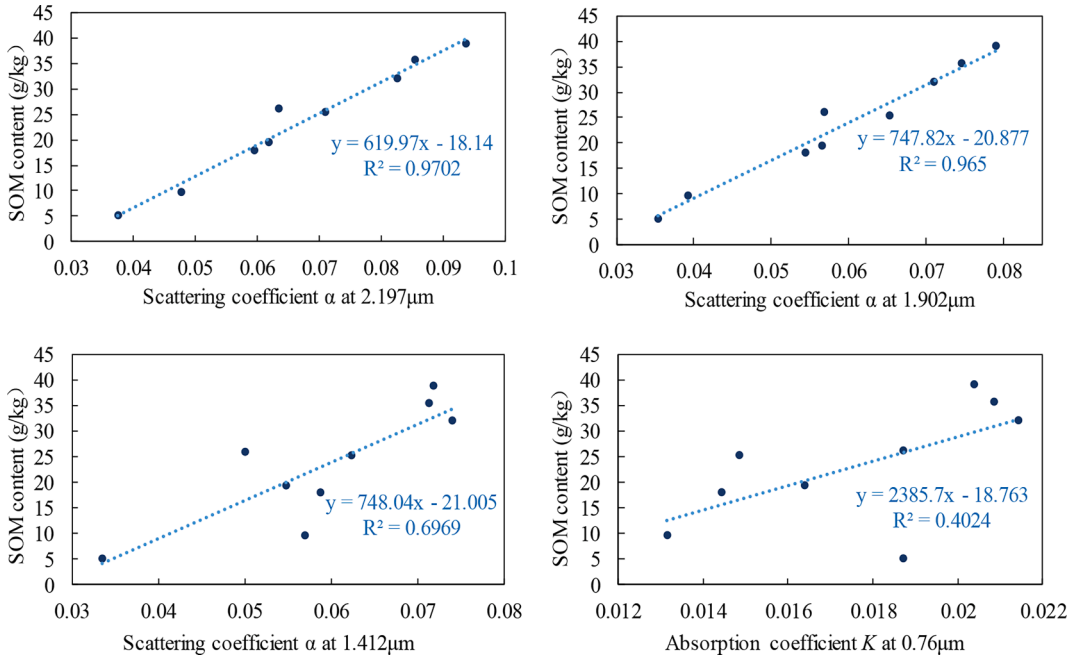


Fig. 14. Scatter plots for scattering coefficient *α* at 2.197 μm, scattering coefficient *α* at 1.902 μm, scattering coefficient *α* at 1.412 μm, and absorption coefficient *K* at 0.76 μm with soil organic matter content.

Therefore when the soil sample is thin, the light will be transmitted onto the background container, and the absorption and scattering of the background material or other effects will be represented in the soil sample, resulting in spectral anomaly for this soil sample. In this study, thickness experiments were conducted on soil samples after organic matter removal, using background containers made up of different materials, i.e., a matte black painted Petri dish, a transparent Petri dish, and a Petri dish painted with matte black spray ink, to investigate the effect of different background containers on the spectra of the soil samples.

Figure 15 shows the spectral curves for the Petri dishes of different materials. The matte black painted Petri dish is more stable throughout the spectral range and has stronger absorption properties than the Petri dish made of black velvet fabric material. It can also be seen that the reflectivity of the transparent Petri dish is relatively high. On the one hand, the surface of the transparent Petri dish is smoother and flatter, which leads to specular reflection, but, on the other hand, the absorption properties of this material are inferior, in comparison. In addition, the transparent Petri dish has more pronounced absorption characteristics around 1.1 μm, 1.68 μm, and 2.17 μm.

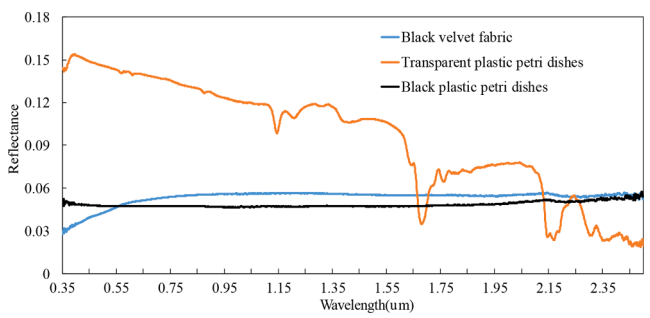


Fig. 15. Spectra comparison with different Petri dishes.

When the background container of the soil sample is a transparent Petri dish, a clear anomalous absorption valley in the band around 1.68 μm can be observed in Fig. 4a and b. In Fig. 4a, it can be seen that the spectra of the soil samples containing organic matter show significant absorption valley characteristics at masses below 13.289 g (which converts to a thickness of 6.33 mm), but this phenomenon is not detected in the spectra of the soil samples when the thickness is more than 6.33 mm. While it can be seen from Fig. 4b that the soil samples without organic matter also have noticeable absorption valley characteristics in the spectra at masses below 13.1886 g (which converts to a thickness of 5.84 mm), this phenomenon is not observed for soil samples above this thickness. In addition, the gradient of the absorption valley becomes steeper and more pronounced as the soil thickness becomes thinner. From the transparent Petri dish spectral curve in Figure 17, it is clear that the transparent Petri dish spectrum also has a distinguishable absorption valley feature at 1.68 μm. To further illustrate this phenomenon, the soil spectra for the transparent and matte black painted Petri dishes were compared, as shown in Fig. 16. The red lines mark the

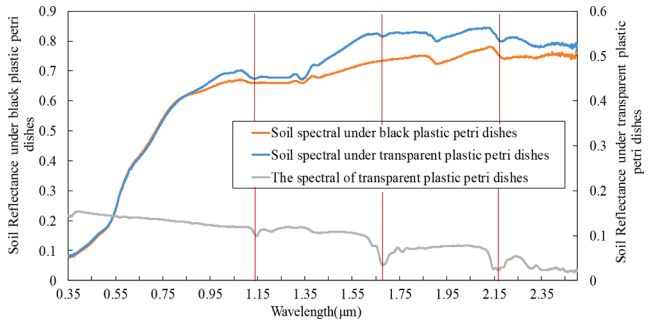


Fig. 16. Comparison of the thin-layer soil spectra for the transparent Petri dish and matte black painted Petri dish.

distinctive absorption peaks for the transparent Petri dish at 1.1 μm, 1.68 μm, and 2.17 μm. A relatively noticeable absorption feature due to the absorption of the background material reflected in the soil sample is also present at 1.1 μm, while it was weaker around 2.17 μm. However, at 2.25 μm, the reflection feature of the Petri dish also has some influence on the soil, and at 1.68 μm, the absorption feature of the Petri dish has the most significant influence on the soil spectrum. This indicates that light is highly transmissive in thin soil samples, and that the absorption features of the background material are directly reflected in the final soil sample spectrum. Studies have suggested that soil samples with a thickness of 2–5 mm can generally be treated as infinitely thick, but it depends on the size of the soil density. For samples with tiny particles, a thickness of up to 10 mm may be required (Kortüm, 2012). For the soil samples in this study area, when measuring the spectra, the spectra of the soil in the natural fallen state after grinding with a thickness greater than 7 mm can be treated as infinitely thick soil spectra that are no longer influenced by the absorption characteristics of the background container material.

4.2. Comparison and shortcomings

To verify the effectiveness of the modified K-M thickness model for soil organic matter inversion, it was compared with the "deviation of arch" (DOA)-based regression model (which is referred to as "DOAR" here), the normalized difference vegetation index (NDVI), and the original spectra at 2.197 μm for organic matter content fitting. The DOAR index is defined as shown in Equation (14). In the study by Zheng et al. (2016), it was shown that the DOAR index can be significantly correlated with soil organic matter by fitting a linear function, for which the determination coefficient R^2 reached 0.55. In order to observe the scattering feature at broad bands, we re-sample the spectrum to Landsat 8 spectral resolution (2.11–2.29 μm average for SWIR2 band).

$$DOAR = R_{0.6\mu m} - 0.5 * R_{0.65\mu m} + R_{0.55\mu m} \quad (14)$$

The comparison results are listed in Table 2. The scattering coefficients obtained in this study using the modified K-M thickness model for linear inversion with organic matter result in a significantly improved accuracy. The accuracy of the determination coefficient obtained by the conventional methods is only about 0.3, while the accuracy of the proposed method reaches 0.97. Among the different methods, the determination coefficient of the DOAR index is the worst, but the mean absolute error (MAE) is lower. The accuracy at SWIR2 (Landsat 8) can reach 0.8895, but the accuracy is much lower than that at 2.197 μm. Therefore, if this inversion method is performed directly on the satellite SWIR band, the accuracy of soil organic matter may be reduced. And the influence of various environmental factors also needs to be taken into consideration. Therefore, the modified K-M soil thickness model can be used to effectively and accurately calculate the scattering coefficient and successfully perform inversion of soil organic matter content. In addition, since the organic matter content of the soil samples has a large range, the proposed model is robust. Specifically, the band around 2.197 μm can obtain optimal results for the inversion of soil organic matter content.

Table 2
Comparison of the inversion results of the different methods.

Method	Inversion expression	R^2	MAE	RMSE
DOAR	SOM = −17.07*ln (*DOAR*) − 25.672	0.2327	6.7005	9.4518
Reflectance (*R*) at 2.197 μm	SOM = 120.25**R* − 45.115	0.3000	7.5089	9.0278
Scattering α at 2.197 μm	SOM = 619.97*α−18.14	**0.9702**	**1.2794**	**1.8625**
Scattering α at SWIR2	SOM = 551.08*α−16.852	0.8895	2.5373	3.5864

When compared to the Hapke model and other methods used in soil composition inversion, which require accurate multi-angle data, the K-M thickness model only requires the user to measure the spectra of soil samples with different thicknesses. Furthermore, the method presented in this paper represents a more straightforward method of using soil mass instead of thickness, with the characteristics of high accuracy and good generalizability.

Although the proposed method is both simple and effective, it is challenging to apply to hyperspectral imaging data in practice. On the one hand, it is difficult to collect soil sample reflectance data with a known thickness because the soil samples under imaging conditions are equivalent to infinite thickness. On the other hand, soil moisture cannot be neglected in actual field conditions, and soil moisture content has a more significant effect on soil reflectance than soil organic matter. In addition, the variability due to inconsistent lighting conditions in the field can limit the use of this method. Further exploratory studies should be conducted to clarify all these possible influencing factors. Despite the limited application for imaging data, this method represents a novel idea for soil composition inversion.

5. Conclusion

The K-M model is widely used in industrial applications, especially in the color prediction of mixed pigments or paint mixing composition prediction, but is less commonly used in soil composition inversion. In this study, the K-M thickness model was used to study soil scattering properties and provide essential support for future soil composition inversion.

Spectral analysis of soil samples with different thicknesses revealed that the smaller the thickness, the lower the reflectance. The absorption and scattering characteristics of the background container material were also found to affect the reflectance for thinner soil samples. However, it was found that the container material has no influence on the soil spectra when the soil sample is thicker than a certain level. Therefore, in order to ensure the spectral reliability of the samples, when measuring different soil samples, it is necessary to ensure that the thickness of the samples is sufficient and constant. For example, the recommended soil thickness for this approach is more than 7 mm. In addition, when selecting background container materials, the user should avoid choosing materials with obvious reflectance peak and valley characteristics.

Meanwhile, in the proposed approach, the scattering coefficient is calculated by the thickness equation based on K-M theory. We found that the theoretical model requires correction in the actual spectral calculation. The modified K-M thickness model was established to simulate the soil spectral information for different thicknesses. The modified K-M thickness model was used for the first time to calculate the scattering coefficient α for soil organic matter content inversion. The soil organic matter inversion model established by the scattering coefficient α at 2.197 μm obtained the highest accuracy in this study, with a determination coefficient of 0.97. The proposed model is characterized

by high accuracy and robustness, and could provide essential technical support for future studies of soil organic matter inversion.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was supported in part by the Natural Science Foundation of China (No. 41871337, No. 42171335), National Key Research and Development Program of China (Grant No. 2019YFC1804900) and in part by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

References

Al-Abbas, A.H., Swain, P.H., Baumgardner, M.F., 1972. Relating organic matter and clay content to the multispectral radiance of soils. Soil Sci. 114 (6), 477–485.
Amigo, J.M., Babamoradi, H., Elcoroaristizabal, S., 2015. Hyperspectral image analysis. A tutorial. Analyt. Chim. Acta 896, 34–51.
Angelopoulou, T., Tziolas, N., Balafoutis, A., Zalidis, G., Bochtis, D., 2019. Remote sensing techniques for soil organic carbon estimation: a review. Remote Sens. 11 (6), 676.
Barrón, V., Torrent, J., 1986. Use of the Kubelka-Munk theory to study the influence of iron oxides on soil colour. J. Soil Sci. 37 (4), 499–510.
Ben-Dor, E., 2002. Quantitative remote sensing of soil properties. Adv. Agron. 75 (02), 173–243.
Ben-Dor, E., Inbar, Y., Chen, Y., 1997. The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a controlled decomposition process. Remote Sens. Environ. 61 (1), 1–15.
Camargo, L.A., Júnior, J.M., Barrón, V., Alleoni, L.R.F., Barbosa, R.S., Pereira, G.T., 2015a. Mapping of clay, iron oxide and adsorbed phosphate in Oxisols using diffuse reflectance spectroscopy. Geoderma 251, 124–132.
Camargo, L.A., Júnior, J.M., Barrón, V., Alleoni, L.R.F., Barbosa, R.S., Pereira, G.T., 2015b. Mapping of clay, iron oxide and adsorbed phosphate in Oxisols using diffuse reflectance spectroscopy. Geoderma 251–252, 124–132.
de la Osa, R.A., Iparragirre, I., Ortiz, D., Saiz, J., 2020. The extended Kubelka-Munk theory and its application to spectroscopy. ChemTexts 6 (1), 1–14.
Gonçalves, Í.G., Petter, C.O., 2007. Teoria de Kubelka-Munk aplicada na indústria de minerais industriais: predição do teor de contaminantes em caolim. Rem: Revista Escola de Minas 60 (3), 491–496.
Hapke, B., 1981. Bidirectional reflectance spectroscopy: 1. Theory. J. Geophys. Res.: Solid Earth 86 (4), 3039–3054.
Hapke, B., 2012. Theory of Reflectance and Emittance Spectroscopy. Cambridge University Press, Cambridge.
Hu, X., Johnston, W.M., 2009. Concentration additivity of coefficients for maxillofacial elastomer pigmented to skin colors. Dent. Mater. 25 (11), 1468–1473.
Kokhanovsky, A., 2019. Springer Series in Light Scattering. Springer, New York.
Kortüm, G., 2012. Reflectance Spectroscopy: Principles, Methods, Applications. Springer Science & Business Media, New York.
Kortüm, G., Braun, W., Herzog, G., 1963. Principles and techniques of diffuse-reflectance spectroscopy. Angew. Chem., Int. Ed. Engl. 2 (7), 333–341.
Kubelka, P., Munk, F., 1931. Ein Beitrag zur Optik der Farbanstriche. Zeitschrift fürTechnische Physik 12, 593–601.
Li, H., Liang, Y., Xu, Q., Cao, D., 2009. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration. Anal. Chim. Acta 648 (1), 77–84.
Minasny, B., McBratney, A.B., Bellon-Maurel, V., Roger, J.-M., Gobrecht, A., Ferrand, L., Joalland, S., 2011. Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon. Geoderma 167, 118–124.
Murphy, A.B., 2007. Band-gap determination from diffuse reflectance measurements of semiconductor films, and application to photoelectrochemical water-splitting. Sol. Energy Mater. Sol. Cells 91 (14), 1326–1337.
Ou, D., Tan, K., Lai, J., Jia, X., Wang, X., Chen, Y., Li, J., 2021. Semi-supervised DNN regression on airborne hyperspectral imagery for improved spatial soil properties prediction. Geoderma 385, 114875.
Reich, I., Snell, F.D., Osipow, L., 1953. Reflectance as a measure of the soil content of cotton fabric. Ind. Eng. Chem. 45 (1), 137–141.
Rezaei, Y., Mobasheri, M.R., Zoej, M.J.V., 2008. Unsupervised information extraction using absorption line in Hyperion images. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37, 383-388.
Sadeghi, M., Jones, S.B., Philpot, W.D., 2015. A linear physically-based model for remote sensing of soil moisture using short wave infrared bands. Remote Sens. Environ. 164, 66–76.
Shi, T., Chen, Y., Liu, Y., Wu, G., 2014. Visible and near-infrared reflectance spectroscopy-An alternative for monitoring soil contamination by heavy metals. J. Hazard. Mater. 265, 166–176.
Stoner, E.R., Baumgardner, M., 1981. Characteristic variations in reflectance of surface soils 1. Soil Sci. Soc. Am. J. 45 (6), 1161–1165.
Tan, K., Ma, W., Chen, L., Wang, H., Du, Q., Du, P., Yan, B., Liu, R., Li, H., 2021. Estimating the distribution trend of soil heavy metals in mining area from HyMap airborne hyperspectral imagery based on ensemble learning. J. Hazard. Mater. 401, 123288.
Tan, K., Wang, H., Zhang, Q., Jia, X., 2018. An improved estimation model for soil heavy metal (loid) concentration retrieval in mining areas using reflectance spectroscopy. J. Soils Sediments 18 (5), 2008–2022.
Tan, K., Ye, Y.-Y., Du, P.-J., Zhang, Q.-Q., 2014. Estimation of heavy metal concentrations in reclaimed mining soils using reflectance spectroscopy. Spectrosc. Spectral Analysis 34 (12), 3317–3322.
Ting, H.E., 2006. Spectral Features of Soil Organic Matter. Geomatics & Information Science of Wuhan University 12(1), 33-40.
Vargas, W.E., 2002. Inversion methods from Kubelka-Munk analysis. J. Opt. A: Pure Appl. Opt. 4 (4), 452.
Wang, F., Gao, J., Zha, Y., 2018. Hyperspectral sensing of heavy metals in soil and vegetation: feasibility and challenges. ISPRS J. Photogramm. Remote Sens. 136, 73–84.
Wang, J., Cui, L., Gao, W., Shi, T., Chen, Y., Gao, Y., 2014. Prediction of low heavy metal concentrations in agricultural soils using visible and near-infrared reflectance spectroscopy. Geoderma 216, 1–9.
Weidong, L., Baret, F., Xingfa, G., Qingxi, T., Lanfen, Z., Bing, Z., 2002. Relating soil surface moisture to reflectance. Remote Sens. Environ. 81 (2–3), 238–246.
Yang, L., Kruse, B., 2004. Revised Kubelka-Munk theory. I. Theory and application. JOSA A 21 (10), 1933–1941.
Yuan, J., Wang, X., Yan, C.-X., Wang, S.-R., Ju, X.-P., Li, Y., 2019. Soil moisture retrieval model for remote sensing using reflected hyperspectral information. Remote Sens. 11 (3), 366.
Zhang, Y., Tan, K., Wang, X., Chen, Y., 2020. Retrieval of soil moisture content based on a modified Hapke photometric model: a novel method applied to laboratory hyperspectral and Sentinel-2 MSI data. Remote Sens. 12 (14), 2239.
Zheng, G., Dongryeol, R., Caixia, J., Changqiao, H., 2016. Estimation of organic matter content in coastal soil using reflectance spectroscopy. Pedosphere 26 (1), 130–136.