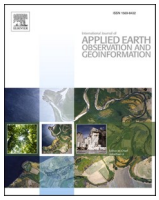
International Journal of Applied Earth Observation and Geoinformation 128 (2024) 103759

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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



A rapid soil Chromium pollution detection method based on hyperspectral remote sensing data

Lihan Chen [a], Kun Tan [b,c,d,*], Xue Wang [b,c,d], Yu Chen [a]

[a] Key Laboratory for Land Environment and Disaster Monitoring of NASG, China University of Mining and Technology, Xuzhou 221116, 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] School of Geographic Sciences, East China Normal University, Shanghai 200241, China

ARTICLE INFO

Keywords:
Hyperspectral remote sensing
Heavy metal pollution
Spectral threshold detection
Hilbert-Huang transform
Hapke model

ABSTRACT

Chromium (Cr) pollution in soil can cause serious harm to both human health and the environment. Cr (VI) is of particular concern due to its strong carcinogenic, teratogenic, and mutagenic effects on humans. It can also lead to soil hardening and crop yield reduction. Recently, hyperspectral approaches have become a popular way of analyzing heavy metal pollution in soil, and are both convenient and cost-effective, compared with chemical analysis. However, the complex composition of soil makes it challenging to directly analyze the influence of heavy metals on soil spectra and achieve efficient and rapid identification of heavy metal contamination in soil. In this research, a simulation experiment was designed for soil Cr content. Threshold detection-Hilbert-Huang transform (TD-HHT) was implemented to remove the complex environmental noise of soil spectra, to extract the weak heavy metal information. The results showed that the TD-HHT amplitude fluctuates at different Cr concentrations, and heavy metal pollution can be directly detected from the amplitude curve at around 600 mg/kg. To verify this threshold value, the correlation between the soil optical constant and the soil heavy metal concentration was explored through the bidirectional reflectance distribution function, based on the threshold validation-Hapke (TV-Hapke) model. At around 600 mg/kg, Cr can be directly detected using the soil radiative transfer model, which can be utilized to rapidly detect soil heavy metal pollution.

1. Introduction

The rise of industrialization and globalization has resulted in consumption patterns increasement, giving rise to substantial waste production and the release of diverse pollutants, which contributes to further damage on the environment and has hindered the progress in meeting the Sustainable Development Goals (SDGs) (UNEP, 2021). Soil constitutes a vital component of the overall environment, and the quality and safety of the soil environment is critical to maintain stable social and economic development and protect human health. Soil heavy metal pollution can have a serious impact on soil environmental quality. As a result of the extremely long degradation, without treatment, the soil heavy metal elements can lead to the deterioration of the soil environment and contamination of crops, as well as pose a threat to human health when these elements enter the food chain (Bolan et al., 2014). Heavy metal pollution is long-term, irreversible, non-degradable, and highly toxic. At present, the deterioration of the regional agricultural environment in China is very serious, and a large volume of Chromium (Cr) slag is unprotected, resulting in Cr slag dust being spread with the wind to the nearby soil and then into rivers, resulting in serious pollution of both soil and water bodies. Meanwhile, wastewater containing Cr is often arbitrarily discharged without treatment and, without effective control, Cr can even be found in city garbage and general waste (Zheng et al., 2017). Cr cannot be decomposed or degraded after entering the soil, and will always exist in the soil. Cr in soil exists in the form of Cr (III) and Cr (VI). The first form has a low concentration, poor activity, and represents a relatively minor hazard. Meanwhile, Cr (VI), which is a strong oxidant with high solubility and strong activity, and very toxic to animals and plants (Costa, 2003). In addition, its strong oxidation causes soil hardening and crop yield reduction.

Although traditional soil quality monitoring is highly accurate, it cannot meet the current demand for large-scale environmental

* Corresponding author at: Key Laboratory of Geographic Information Science (Ministry of Education), East China Normal University, Shanghai 200241, China.
E-mail address: tankuncu@gmail.com (K. Tan).

https://doi.org/10.1016/j.jag.2024.103759
Received 3 January 2024; Received in revised form 15 February 2024; Accepted 8 March 2024
Available online 13 March 2024

monitoring because of its inherent disadvantages, such as the time consumption and cost. The application of hyperspectral technology in estimating soil composition has become widespread, attributed to its rapid, non-destructive, and non-contact characteristics (Viscarra Rossel and Behrens, 2010; Viscarra Rossel et al., 2016; Tan et al., 2020a,b; Ou et al., 2021). However, since the low concentration of heavy metal content in soil, the response in the spectra is extremely weak. Obtaining optimal spectra that reflect the true properties of the sample is also a challenging task, as the result of the experimental conditions, including instrument noise, and it is challenging to capture the effective information of soil heavy metals directly. As a result, spectral preprocessing enhancement and spectral feature extraction are critical for soil heavy metal analysis with the hyperspectral technique.

Preprocessing methods which can remove the spectral noise and baseline drift and highlight the characteristic bands of the spectra have been extensively applied and have been proven to be generally effective. Typical preprocessing methods include the continuum removal, multiplicative scatter correction, standard normal variate correction, first-order derivative, and second-order derivative (Susi and Byler, 1983; Fearn et al., 2009; Asadzadeh and de Souza Filho, 2016; Zhang et al., 2019). When the mathematical transformation cannot capture the faint spectral differences, the spectrum can be converted to the frequency domain. In terms of frequency domain transformation, the predominant approach for detecting signal anomalies involves employing time–frequency analysis. This technique transforms signals from the temporal domain to the frequency domain, enables the identification of abnormal signal characteristics and changes in information within the frequency domain. In recent years, wavelet transform, harmonic analysis, and other methods have been applied in soil moisture and organic matter analysis and estimation (Jiang et al., 2017; Wang et al., 2018). The threshold denoising method based on wavelet transform can be utilized to denoise the first derivative spectra of soil, where the absorption area is then extracted from the denoised derivative spectra to characterize the change of organic matter content, so as to realize the detection of soil organic matter derivative spectra (Liu et al., 2011). Until now, limited research has been conducted on the extraction of information and spectral detection of signals related to soil heavy metal pollution. Fu et al. (Fu and Yang, 2018) converted the soil spectra into sparse spectra based on the second-order difference method and combined the sparse spectrum of soil and Gabor expansion theory to detect the weak differences of heavy metal stress spectra in different concentration soils in the frequency domain. Utilizing HHT in time–frequency analysis, Fu et al. (Fu et al., 2019) examined soil contaminated with varying lead concentrations. Information mining of the soil spectra was achieved by analyzing the frequency spectrum of Intrinsic Mode Function (IMF) component. It was found that there are differences in IMF2 of soil spectra under different concentrations of Pb pollution. Yang et al. (Yang et al., 2018) mined Cu pollution information in soil spectra on HHT and the conclusions showed that the spectra time–frequency results were non-identical when polluted by different Cu content. However, these studies only describe that there exited differences between the spectra time–frequency results under different pollution and were unable to determine a specific identifiable heavy metal pollution stress threshold, and these methods for signal decomposition possess limitations when elucidating the variation of reflectance in relation to the spectral mechanism.

In this study, in order to explore this problem, we attempted to analyze the interaction process and distribution of light radiation on the surface of a soil medium with different heavy metal contents through a radiative transfer model (RTM). The optical properties of each component of the soil determine the difference in the radiation transmission process of the incident light. The RTM provides an explicit connection between bidirectional reflectance and soil properties based on soil optical properties (Huete and Escadafal, 1991; Palacios-Orueta and Ustin, 1998). Sadeghi et al. (Sadeghi et al., 2015) developed a physically based soil water inversion model using the Kubelka-Munk theory, converting soil reflectance into a linear relationship model with soil water content. Yang et al. (Yang et al., 2011) integrated soil water content information into the SOILSPECT model, establishing the SWAP-Hapke model for inverting soil water content parameters under field conditions. Yao et al. (Yao et al., 2018) analyzed the impact of the different optical properties on the characteristics of soil duality through soil water content data from multiple angles, and performed quantitative inversion of the soil water content based on the Hapke model. Zhang et al. (Zhang et al., 2020) proposed the novel SMR-Hapke soil moisture retrieval approach, which was verified on soil moisture data, and reduced it to a linear form, which has great potential for estimating soil moisture content. The Hapke model has found extensive application in the estimation of soil moisture. However, to the best of our knowledge, the Hapke model has not been used to explore the optical properties of soil heavy metals.

This work is driven by the following motivations:

The current soil heavy metal Cr monitoring methods require field sampling, chemical analysis and model building to obtain the heavy metal content, and thus cannot be used to achieve efficient and rapid identification of soil heavy metal contamination in a study area. Therefore, in the proposed approach, threshold detection Hilbert-Huang transform (TD-HHT) and the threshold validation Hapke (TV-Hapke) model are utilized to obtain a threshold for the efficient identification of Cr contamination using only spectra to achieve rapid soil heavy metal pollution analysis.

The primary contributions of this study can be summarized as follows:

(1) We detect a Cr pollution threshold for soil spectra with TD-HHT using characteristic bands.
(2) We explore the optical properties of soil Cr and verify the Cr pollution threshold in soil spectra based on the TV-Hapke model.

Section II is divided into two parts. The first part describes the experimental design, and the second part introduces the proposed method. Section III details the experimental results. Finally, we draw our conclusions in Section IV. Fig. 1 illustrates the flowchart of the research.

2. Materials and methods

2.1. Experimental design

This experiment was devised with the artificial addition of Cr to soil. The metal compound $CrCl_3 \cdot 6H_2O$ was added based on the risk screening value and the risk intervention value stipulated in the soil environmental quality risk control standard for soil contamination of agricultural land (GB 15618–2018) (Regulation, 2018). If the soil Cr contamination of agricultural land exceeds the risk screening value (200 mg/kg), there are potential risks to the quality and safety of agricultural products, as well as the overall health of crop growth and soil ecology. Therefore, it is essential to implement measures to ensure the safe utilization of soil. If the soil Cr contamination in agricultural land exceeds the risk intervention value (1000 mg/kg), the edible agricultural products may fall short of meeting the quality and safety standards, and strict control measures should be taken, in principle. Therefore, in order to monitor Cr contamination, the added heavy metal values were set between the risk screening value and the risk intervention value. The concentrations of the artificially added Cr were 200, 300, 400, 500, 600, 700, 800, 900, and 1000 mg/kg.

2.2. Soil sample collection and treatment

We collected several samples in Xuzhou, sent them to the laboratory to measure the soil heavy metal concentration using inductively coupled plasma-mass spectrometry (ICP-MS), and selected an uncontaminated soil sample to conduct experiment.

Natural soil samples, belonging to cinnamon soil, were gathered at

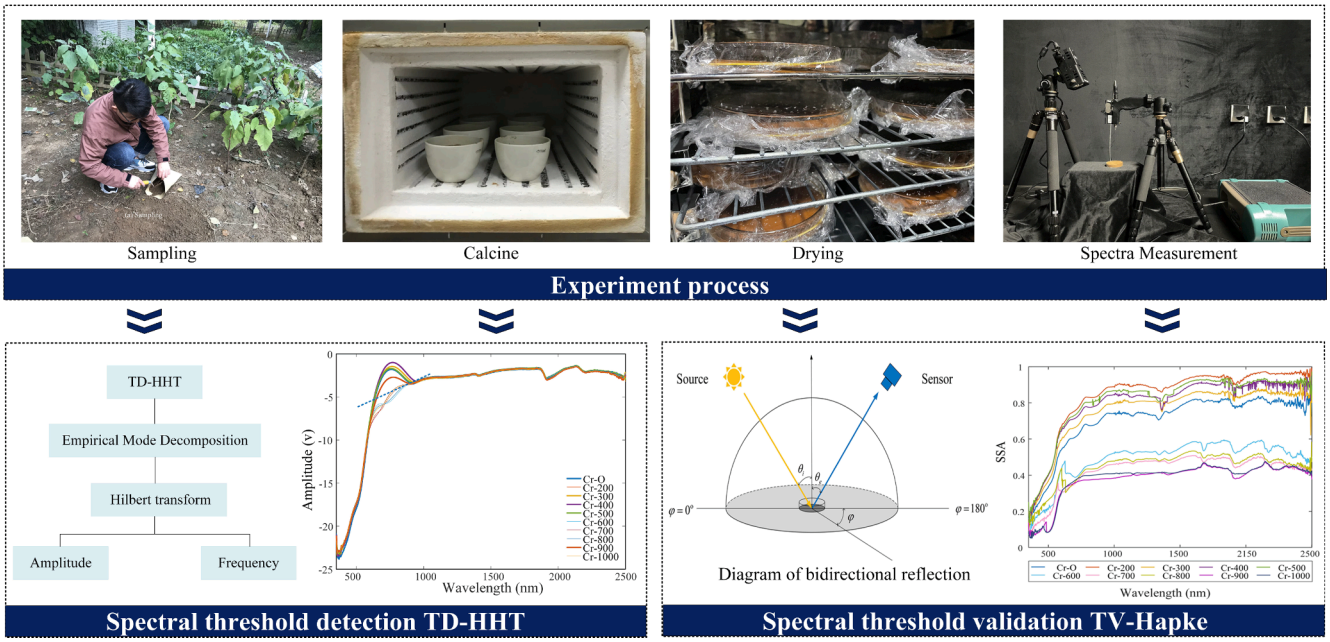


Fig. 1. Flowchart of the study.

the remote sensing experimental field, Xuzhou city, Jiangsu province. The climate of the experimental field is a warm temperate semi-humid monsoon climate. The experimental field has average annual temperature and precipitation of 14 °C and 847 mm, respectively. Following collection, the soil samples were carefully sealed and conveyed to the laboratory for subsequent analysis. In the laboratory, after removing any sundries in the soil samples, the soil samples underwent a process of drying, grinding, and sieving through a 100-mesh nylon sieve. As one of the important components of soil, soil organic matter significantly influences the characteristics of soil spectra. With the aim of exploring solely the effect of heavy metal on the soil spectra, the soil was calcinated in a muffle furnace at 600 °C for 8 h to remove organic matter. The value of the Cr concentration in the selected Cr-uncontaminated natural soil sample was found to be 76.2 mg/kg.

In the experiment, the controlled variable method was utilized to guarantee that the soil property was subject to only heavy metal stress. For each sample, 30 g of soil was placed in a conical flask and wetted with deionized water. The $CrCl_3{\cdot}6H_2O$ solution was then added to obtain the desired Cr concentration, additional water was added, and the sample was then placed in an oscillator for 24 h. After this, the material in the conical flask was placed in a forced air oven at 40 °C, covered with a hole-filled plastic wrap. When there was no more visible water, deionized water was added to the sample. This process was repeated twice and the soil was dried at 40 °C (Wu et al., 2007). After drying, grinding, and passing through a 100-mesh nylon sieve, the soil samples were placed into petri dishes, distributed evenly, and the soil surface was smoothed. The spectra of the artificially contaminated soil samples were then captured using an Analytical Spectral Devices (ASD) field spectrometer in a dark room. A 1000-lumen halogen lamp with a zenith angle of 45° was used as the light source. The detection fiber was placed vertically to the soil sample at a distance of 15 cm. The soil was placed in petri dish, which was set between light source and detection fiber. Each soil sample underwent five scans, and the resulting average spectrum was taken as the final spectrum.

2.3. Theories and methods

2.3.1. Threshold detection Hilbert-Huang transform (TD-HHT)

When the mathematical transformation cannot capture the faint spectral differences, the spectrum can be converted to the frequency domain. The Hilbert-Huang transform (HHT) has been investigated for soil spectral analysis because of its advantage of being able to handle 'nonlinear' and 'nonstationary' problems through Hilbert analysis. The TD-HHT process comprises two components: empirical mode decomposition (EMD) and Hilbert transformation.

EMD can adaptively decompose the spectra into a finite number of Intrinsic Mode Functions (IMFs). Two essential conditions for generating IMF components include: 1) ensuring the number of extreme points and zero-crossing points for IMF needs to be the same or differ by at most one; and 2) the mean value of the envelope, as determined by the local maximum and minimum points at any point of IMF, should be zero, i.e., the waveform of the entire curve must be locally symmetric. Assuming that the original spectrum is $X(k)$, then the decomposition processes of EMD are as follows: 1) Extract all the local maximum and minimum points of the spectrum $X(k)$ to be analyzed, and then use the cubic spline interpolation method to fit the maximum and minimum points to form the maximum envelope $e_{max}(k)$ and the minimum envelope $e_{min}(k)$. The mean value m of the extremum envelope is then obtained. 2) Subtract m from $X(k)$ to obtain a new signal $y_1(k)$, where the signal $y_1(k)$ is in the form of IMF components. It is determined whether $y_1(k)$ is an IMF component according to the judgment criterion of the IMF component. If $y_1(k)$ does not meet the conditions, $y_1(k)$ is used as the original data, the above steps are repeated, and the filtering is continued until $y_1(k)$ meets the IMF conditions. At this time, $y_1(k)$, which is IMF_1, is the first IMF component of $X(k)$. 3) The remaining signal $r_1(k) = X(k) - IMF_1$ is used as the original signal. IMF_n and r_n follows a similar procedure. The termination criterion is met when r_n forms a monotonic curve. The original spectrum $X(k)$ can be represented as:

$$X(k) = \sum_{i=1}^{n} IMF_i(k) + r_i(k) \tag{1}$$

where n signifies the number of iterations, IMF_i stands for the ith IMF component during the iterative process, and r_i corresponds to the ith residual spectrum.

Assuming that the IMF component $IMF_i(k)$ is $s(k)$, then its conjugate spectrum $m(k)$ is calculated by Hilbert transformation performed on $s(k)$:

$$m(k) = \frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{s(\tau)}{k-\tau} \tag{2}$$

where k denotes the band number and τ is the integral variable.

$s(k)$ and $m(k)$ are utilized to construct the analytical spectrum $z(k)$:

$$z(k) = s(k) + jm(k) = a(k)e^{j\varphi(k)} \tag{3}$$

$$\alpha(k) = \sqrt{s(k)^2 + m(k)^2} \tag{4}$$

$$\varphi(k) = arctan\frac{m(k)}{s(k)} \tag{5}$$

where $\alpha(k)$ denotes the instantaneous amplitude or energy of the analytical spectrum, while $\varphi(k)$ signifies the instantaneous phase of the analytical spectrum.

The instantaneous frequency of $s(k)$ can be obtained from the instantaneous phase:

$$\omega(k) = \frac{d\varphi(k)}{dk}\# \tag{6}$$

If the amplitude is displayed on the plane of the frequency and band, the Hilbert spectrum $H(\omega, k)$ can be obtained.

2.3.2. Threshold validation Hapke (TV-Hapke) model

In addition to the signal decomposition method, in order to achieve Cr threshold detection, the RTM is introduced. The Hapke model, i.e., the bidirectional reflection model of the luminosity scattering properties on the surface of a dense medium, describes the interaction between light and the medium (Hapke, 2002), and can provide useful information related to soil surface properties and physicochemical properties.

In the TV-Hapke model, the radiant reflectance can be derived as:

$$r(u_i, u_e, g, \lambda) = \frac{\omega(\lambda)}{4}\frac{1}{u_i + u_e}(p(g)[1 + B(g)] + [H(u_i)H(u_e) - 1])\# \tag{7}$$

where $\omega(\lambda)$ is the soil Single Scattering Albedo (SSA).$\omega(\lambda)$ is defined as the probability that the photons would be scattered by the particle. u_i and u_e denote the cosine of the incidence angle and the emittance angle, respectively. The phase angle g represents the included angle between the incident light and the scattered light. $P(g)$, which is a phase function, reflects the scattering characteristics of the soil particles and describes the spatial distribution of scattered light. The second-order Legendre polynomial phase function equation is:

$$P(g) = 1 + bcosg + \frac{c[3cos^2(g) - 1]}{2} + b^{'}cosg^{'} + \frac{c^{'}[3cos^2(g^{'}) - 1]}{2}\# \tag{8}$$

$$cosg = cos\theta_i cos\theta_e + sin\theta_i sin\theta_e cos\varphi\# \tag{9}$$

$$cosg^{'} = cos\theta_i cos\theta_e - sin\theta_i sin\theta_e cos\varphi\# \tag{10}$$

where θ_i, θ_e signifie the solar zenith angle and the view zenith angle, respectively, and φ denotes the relative azimuth between the sun and the observation direction.

$B(g)$ is the backscatter function, which represents the effect of backscatter on the reflected light:

$$B(g) = \frac{B_0}{1 + (1/h)tan(g/2)} \tag{11}$$

where g is the scattering angle,h is the half-width parameter of the hot spot effect, and B_0 signifies the intensity size of the hot spot effect.

$H(x)$ is the multiple scattering function:

$$H(x) = \frac{1 + 2x}{1 + 2x\sqrt{1 - \omega}} \tag{12}$$

which x represents u_i and u_e.

3. Results

3.1. Spectral threshold detection

The $CrCl_3{\cdot}6H_2O$ solution was added to the soil samples in accordance with the experimental design. After the sample preparation, the spectra were measured the ASD spectrometer in a dark room. The soil spectra under different Cr concentrations are shown in Fig. 2. Cr-O represents the original soil spectra without additional contamination, and the contamination levels of the Cr-200, Cr-300, Cr-400, Cr-500, Cr-600, Cr-700, Cr-800, Cr-900, and Cr-1000 samples were set as 200, 300, 400, 500, 600, 700, 800, 900, and 1000 mg/kg, respectively. Between 510 nm and 800 nm, the spectral curves rise with a large slope. After 800 nm, the spectral curves rise only slowly. The spectral curves start to decrease from 2150 nm. There are obvious absorption peaks at 1400 nm, 1900 nm, and 2200 nm, with the 1400 nm and 1900 nm peaks corresponding to prominent water vapor absorption bands. The peaks at 1400 and 1900 nm can be attributed to the stretching vibration of the molecular H_2O and OH groups in the adsorbed H_2O (Bishop et al., 1994; Srasra et al., 1994), and the 2200 nm peak can be attributed to stretching vibration and bending vibration of AlOH, OH and organic matter (Post and Noble, 1993; Wight et al., 2016).

The spectra exhibit minor variations with varying heavy metal contents, making it challenging to extract subtle information related to heavy metals. Therefore, the spectra were mathematically transformed, primarily to enhance the existing spectra and detect the key information regarding heavy metal pollution. Fig. 3 shows the preprocessed spectra after first-order derivative (FD), second-order derivative (SD), logarithm of the reciprocal transformation (LR), and continuum removal (CR) processing.

In Fig. 3, it is clear that the FD reflectance, SD reflectance, and LR reflectance with different Cr contents are too similar to distinguish the different contents. The CR reflectance is distinctly different in the characteristic bands with different contents, but it can be challenging to distinguish between the varied concentrations of heavy metals directly, making it difficult to monitor soil heavy metal pollution rapidly and efficaciously. The presence of soil heavy metal pollution has the potential to induce localized alterations in spectra. Finding out the points that lead to mutations in the soil spectra is the key to solving the problem of mining the weak information of heavy metal pollution. In contrast to the spectral domain approach, the frequency domain method has the capability to decompose spectra in multiple dimensions, allowing for a comprehensive exploration of the detailed characteristics across various frequencies. According to signal theory, the soil light spectrum belongs to a 'nonlinear unstable' signal, for which Hilbert transformation cannot be directly applied. EMD can decompose the spectral signal into IMFs,

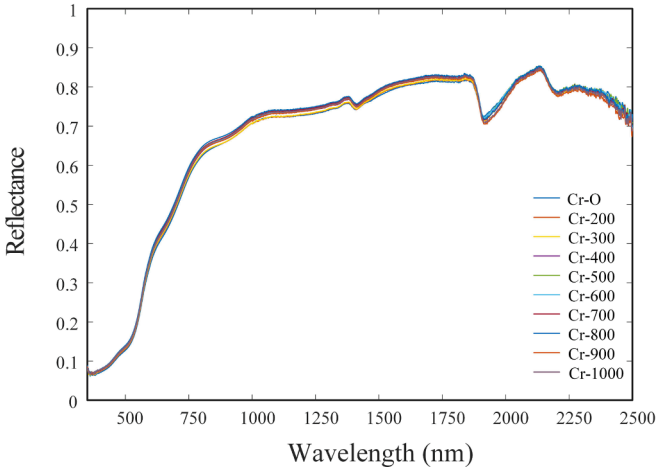


Fig. 2. Soil spectra with different levels of Cr contamination.

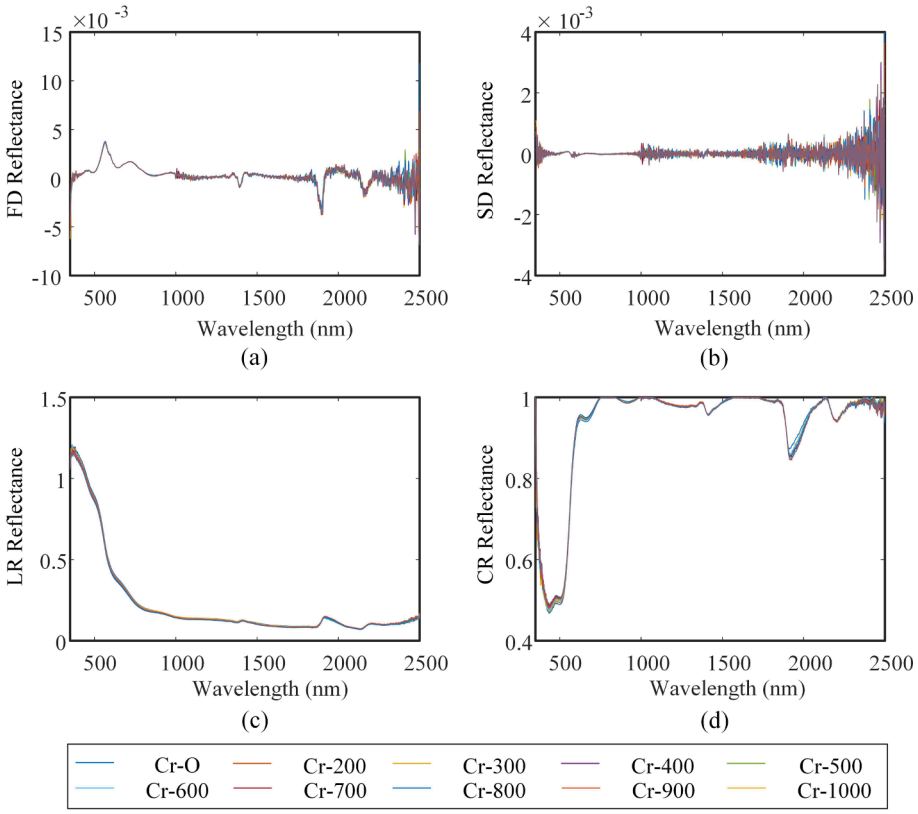


Fig. 3. Reflectance of the mathematically transformed results. (a) FD reflectance (b) SD reflectance (c) LR reflectance (d) CR reflectance.

providing an effective means to minimize the impact of noise in the original signal, and then Hilbert transformation and frequency spectrum analysis can be performed.

There are subtle differences in the soil spectra with the different Cr pollution levels, but the key point of soil pollution spectra cannot be fully expressed, nor can identifiable pollution thresholds be identified. In the experiment, EMD decomposition was performed on the 10 soil spectra samples, and 12 IMF components were obtained, but effective information could not be obtained from the IMF components. The Hilbert transformation was applied to calculate the frequency rate of each IMF component.

Fig. 4 shows the soil spectra frequency under different concentrations of Cr pollution. In Fig. 4, the frequency curves of IMF1 to IMF7 represent the high-frequency component, and the curves are messy. It is difficult to distinguish the curves for the different pollution levels and to extract useful information from these IMF components. The frequency curves from IMF8 onwards become smooth. The two curves of IMF8 and IMF9 continue to fluctuate within a small range. It can be seen from IMF8 to IMF11 that the trend of the curves is relatively similar. As a result, it is difficult to distinguish the differences between the different heavy metal concentrations.

To delve deeper into the distinctions among soil spectra corresponding to varying concentrations of heavy metals, and to explore a discernable threshold for soil heavy metal spectra, the amplitude curves were calculated and drawn. Fig. 5 shows the soil amplitude spectra under different concentrations of Cr pollution.

The amplitude curves of IMF1 to IMF7 represent the high-frequency component, and these curves fluctuate violently and irregularly. It is difficult to distinguish the curves of the different pollution levels and to extract useful information from these IMF components. The amplitude curves of IMF9 and IMF10 are relatively straight lines. The curves of IMF11 show an upward trend, but the slope gradually decreases, and there is little difference between the curves of the different heavy metal concentrations. These amplitude curves do not show enough differentiation to distinguish the different Cr concentrations. It is clear that the curves of IMF12 show different trends with the various concentrations. Consequently, IMF12 is taken for further investigation in Fig. 6.

Fig. 6(b) is the locally enlarged image of the IMF12 amplitude. It can be seen that the amplitude has different variation tendencies, with obvious absorption peaks from 600 nm to 940 nm. Based on the variation tendency, these amplitude curves can be categorized into two distinct groups. In the first group, the curves of Cr-O, Cr-300, Cr-400, Cr-500, and Cr-900 (the bold curves in Fig. 6) protrude obviously outward. In the other group, by contrast, these curves fluctuate gently, the four curves have Cr contents of greater than 600 mg/kg, and only one curve has a Cr content of less than 600 mg/kg. The correct rate is 80 %. Therefore, the results of the modeling experiment indicate that when the Cr content surpasses 600 mg/kg, Cr pollution can be detected from the amplitude curve of 600–940 nm with TD-HHT. Therefore, a key issue of heavy metal pollution determination is the identification of the effective spectral information.

Besides, another soil sample from Fujian Province was employed to evaluate the method validity. It was a Cr-contaminated soil sample which was determined using inductively coupled plasma optical emission spectrometer (ICP-OES) with Cr concentration of 14500 mg/kg. These samples were made by adding a certain amount of soil matrix. The design contents of heavy metal were the same as that of another soil samples, which were 200, 300, 400, 500, 600, 700, 800, 900, and 1000 mg/kg. The soil spectra under different Cr concentrations were shown in Fig. 7.

The reflectance steadily increases from the initial point up to 480 nm, followed by a brief decline before resuming its upward trend. Distinct absorption peaks are observed at 1400 nm, 1900 nm, and 2200 nm. These soil samples were conducted with TD-HHT. The amplitude curves of IMF12 were taken for further investigation in Fig. 8.

It can be seen that the amplitude has different variation tendencies between 610 nm and 830 nm. Based on the observed trends, these

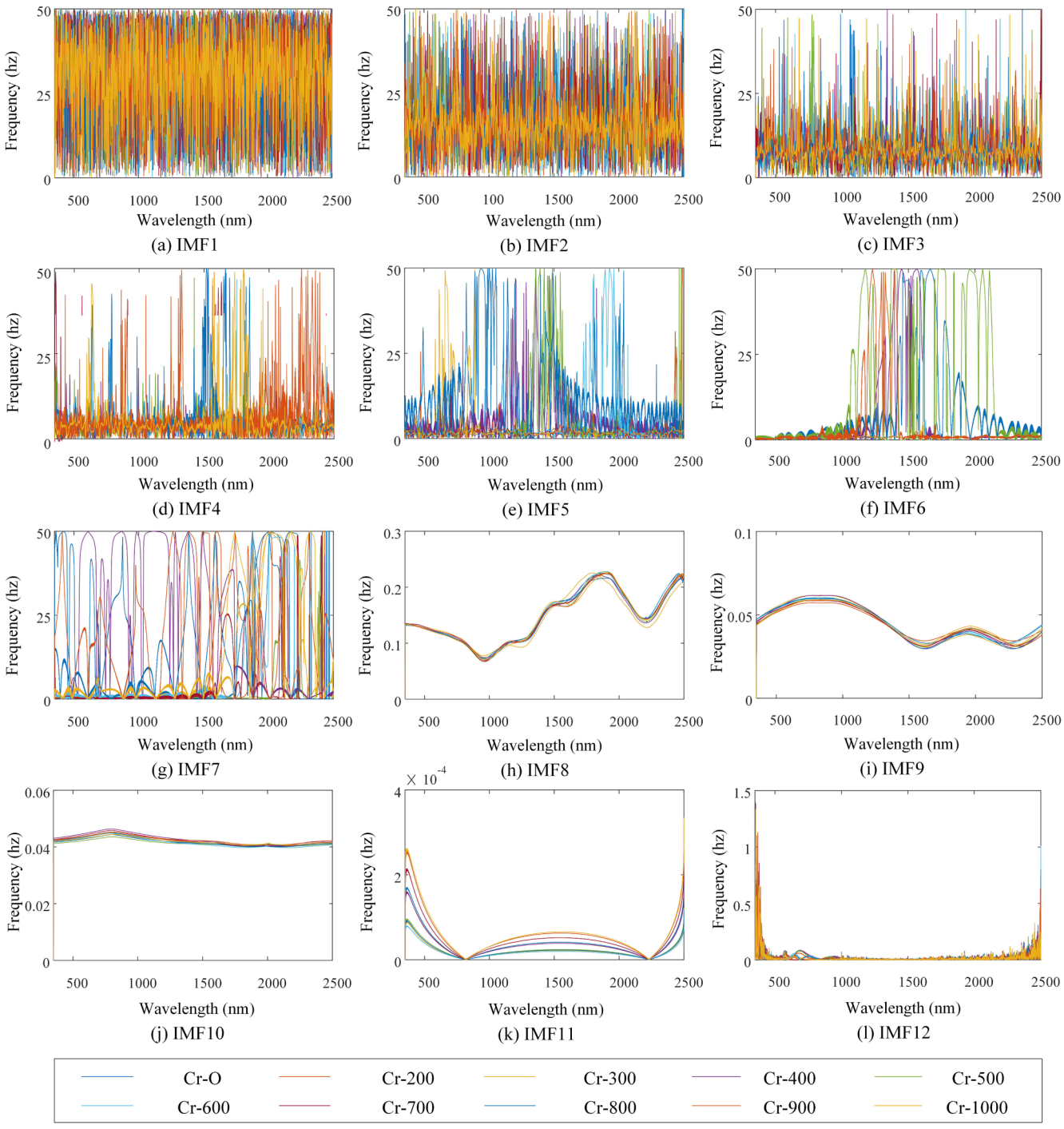


Fig. 4. The soil spectra frequency under different concentrations of Cr pollution. (a), (b) … (i) stand for IMF1, IMF2 … IMF12, respectively.

amplitude curves can be divided into two distinct groups. In the group of Cr content less than 600 mg/kg, the curves of Cr-400, Cr-500 (the bold curves in Fig. 8) protrude outward. For the group of Cr content greater than 600 mg /kg, by contrast, these curves fluctuate gently. Besides, there are two curves that cannot be classified (Cr-200 and Cr-300). Hence, the correct rate is about 77.8 %, which is similar with another soil sample.

In Fig. 6 and Fig. 8, the curves show differences between 600 nm and 940 nm, and the absorption characteristics of the soil spectra are related to the electronic transition of metal ions. Cr, being a transition metal with an unfilled d-shell, experiences energy level splitting of its d-orbitals within a crystal field. The movement of an electron from a lower level to a higher one leads to the absorption of electromagnetic energy, contributing to the distinctive absorption features of heavy metals in the spectra (Wu et al., 2007; Chen et al., 2022). In addition to the above reasons, in this range, the absorption characteristics are associated with goethite, hematite, Fe^{3+}, and ferric oxide, as described in previous studies (Hunt, 1977; Scheinost, 1998; Stenberg et al., 2010).

3.2. Spectral threshold validation

It is difficult to explain the change principle of reflectance with soil properties from the spectral mechanism level only by the signal decomposition method, and the essential characteristics of soil spectra

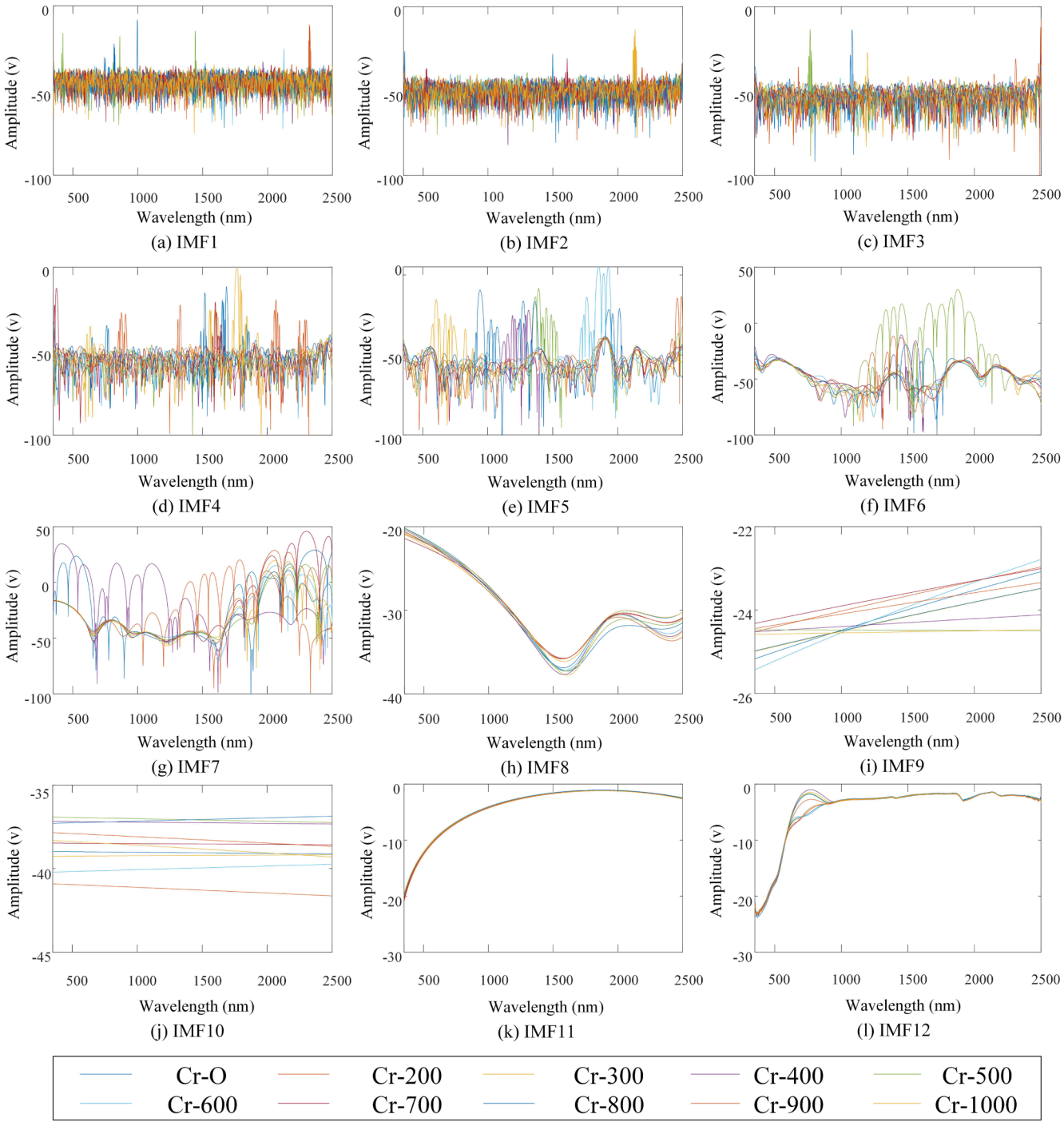


Fig. 5. The soil spectra amplitude under different concentrations of Cr pollution. (a), (b) … (i) stand for IMF1, IMF2 … IMF12, respectively.

cannot be deeply understood. The RTM explores the interaction process and distribution of light radiation on the surface of the soil medium, and can thus characterize the soil composition and soil surface structure. Furthermore, to validate the accuracy of the identification threshold and explore the optical properties of soil with the RTM, a laboratory-based bidirectional reflectance distribution function experiment was designed. In the design process, in order to meet the geometric conditions of the observation of directional reflectance, the experimental device was fixed on a circular base, and the semicircular zenith arc arm was used to fix the light source. The quarter-shaped zenith arm was utilized to hold the fiber optic probe. The two zenith arc arms were connected by a turntable, and both were engraved with angles, which could slide freely according to the experimental requirements. The relative azimuth angle, light source zenith angle, and view zenith angle were set as 0–360°, 0–90°, and 0–90°, respectively. The multi-angle reflectance was measured by setting different observation geometries. The experimental device is shown in Fig. 9.

The sample was placed in the central position of the device, and the distance between the optical fiber probe and the sample was kept at 10 cm, to avoid the influence of the shadow of the probe itself. The light source zenith angle was 30°, and the view zenith angles of the spectrometer probe were varied over 10°, 20°, 30°, 40°, and 50° for each orientation. The relative azimuth angles of the light source and view were 0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315°. Each sample was

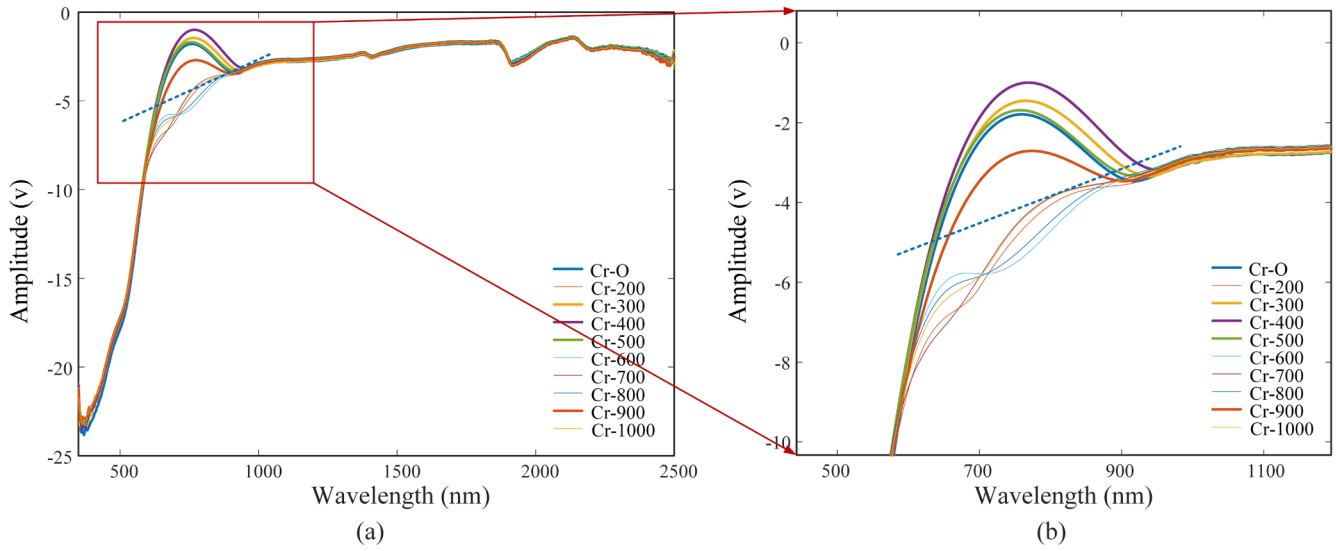


Fig. 6. (a) The IMF12 amplitude of soil spectra under different concentrations of Cr pollution. (b) Locally enlarged image of (a).

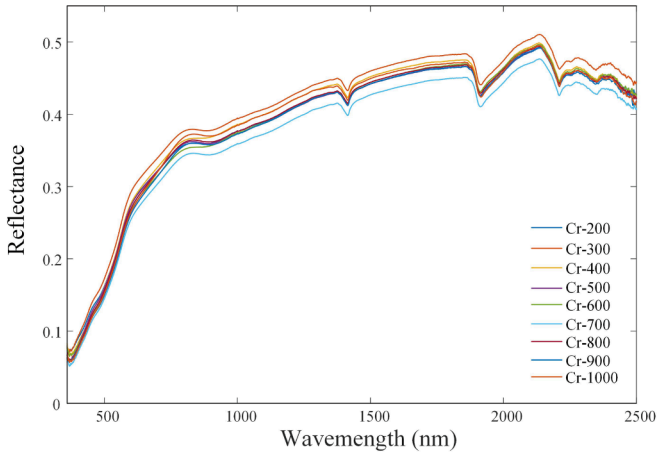


Fig. 7. Soil spectra with different levels of Cr contamination.

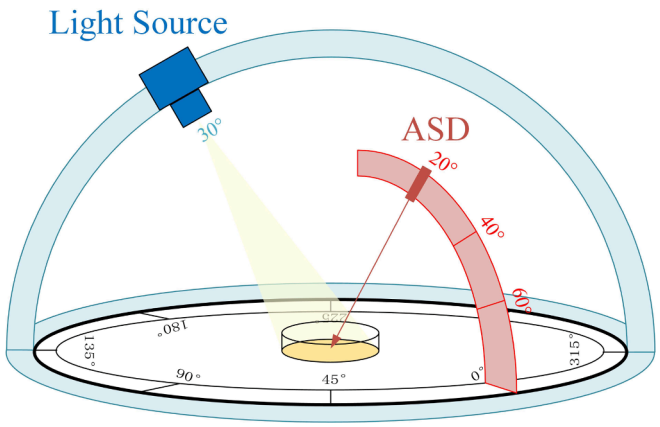


Fig. 9. Diagram of the bidirectional reflectance distribution function experimental device.

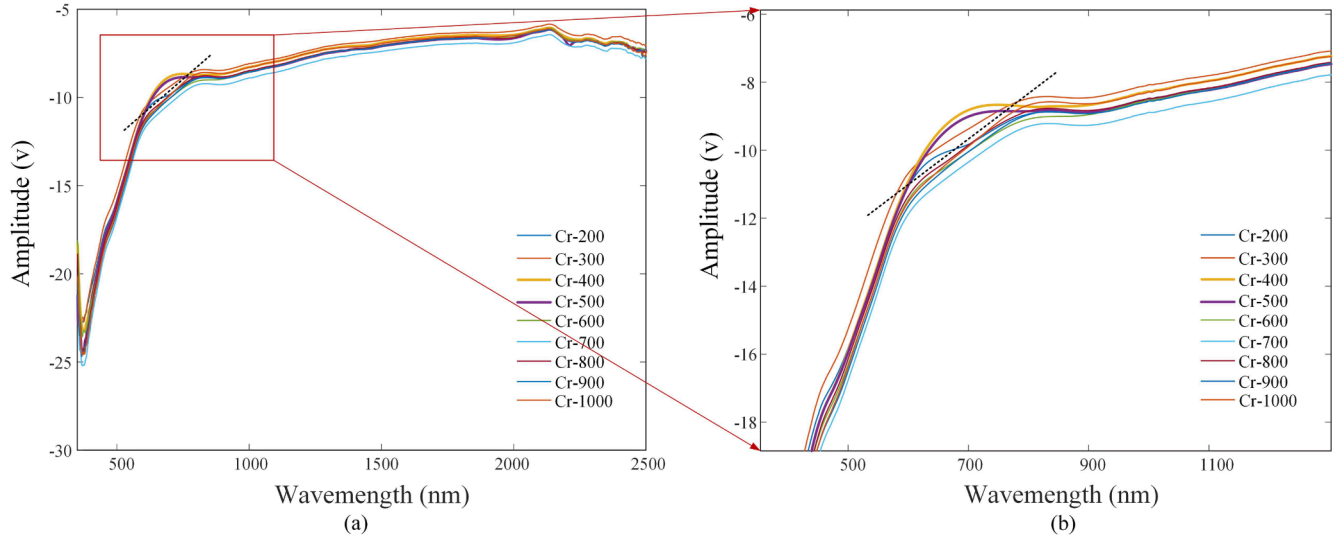


Fig. 8. (a) The IMF12 amplitude of soil spectra under different concentrations of Cr pollution. (b) Locally enlarged image of (a).

measured five times under each combination of angles. After removing the obvious abnormal spectra, the directional reflectance at the position was determined by calculating the average value. Fig. 10 shows the multi-angle soil spectra with Cr contamination.

From Eq. (7) to Eq. (12), there are six parameters to be settled:ω, h, b, c, b', and c'. The model parameters at each wavelength were calculated by fitting the multi-angle reflectance data. The model parameters were calculated using a nonlinear genetic algorithm, and the cost function δ^2 was defined to judge the performance of the optimization:

$$\delta^2 = \sum_{k=1}^{n} [R_k - R(i, e_k, g_k)]^2 \# \tag{13}$$

where R_k is the bidirectional reflectance measured at the angle combination i, e_k, g_k, $R(i, e_k, g_k)$ is the reflectance calculated by the model, and n is the total number.

The SSA (ω), which is a function of the optical constant, is mainly affected by the mineral composition, size, composition, structure, and other properties of soil particles, and is the only parameter related to wavelength in the model, reflecting the internal optical characteristics of soil particles. Therefore, the SSA was taken as the analysis object to study its distribution with different soil heavy metal contents. From the analysis in Fig. 6 of Section 3.1, it was found that most curves reach a peak at 760 nm. For effective heavy metal threshold identification, we considered the difference between the SSA of adjacent samples at 760 nm (Fig. 11).

Fig. 11 illustrates that the SSA difference peaks at sample five, indicating the largest difference between Cr-500 and Cr-600. When the Cr content in soil exceeds 600 mg/kg, Cr pollution can be identified from the SSA curve. Fig. 12 displays the SSA curves for different levels of Cr pollution.

It can be seen from Fig. 12 that the curves can be divided into two groups, each with a consistent trend. The high SSA group indicates low heavy metal contents, while the low SSA group indicates heavy metal contents exceeding 600 mg/kg. Consistent with the findings for TD-HHT, Cr pollution can be detected from 600 mg/kg. $CrCl_3$ crystals are dark green in color, with a low reflectance in the visible region (400–710 nm). In the short-NIR region (710–900 nm), the reflectance curves rise rapidly, and then decrease sharply after 1200 nm. These spectral reflectance features of Cr are determined by its chemical properties and coordination structures, and are related to the off-core electron arrangement. The electronic structure formula of Cr^{3+} is $1s^2 2s^2 2p^6 3s^2 3p^6 3d^3$, and its 3d orbital is not filled with electrons (d is the orbital type, and the d-orbital can hold up to 10 electrons). When white light irradiates $CrCl_3$, the electrons in the 3d orbital acquire light energy, and under the action of the crystal field, the 3d orbital will split

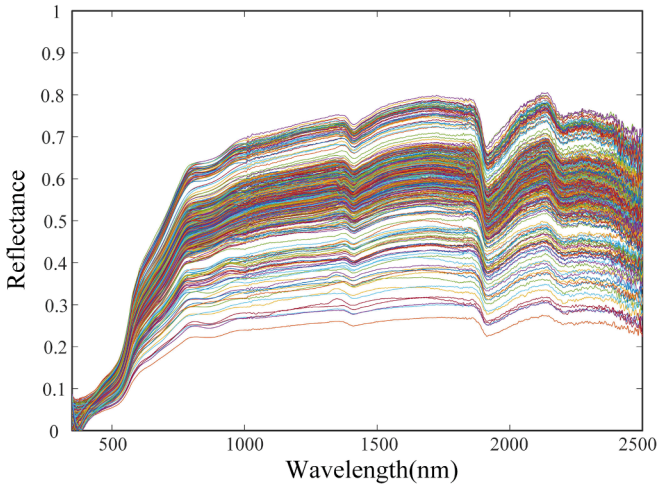


Fig. 10. Multi-angle soil spectra with Cr contamination.

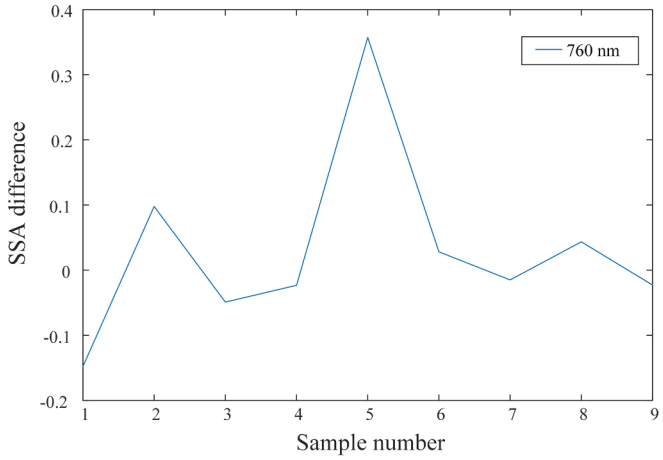


Fig. 11. SSA difference of adjacent samples under Cr pollution (sample one is the SSA difference of Cr-O and Cr-200, sample two is the SSA difference of Cr-200 and Cr-300, etc.).

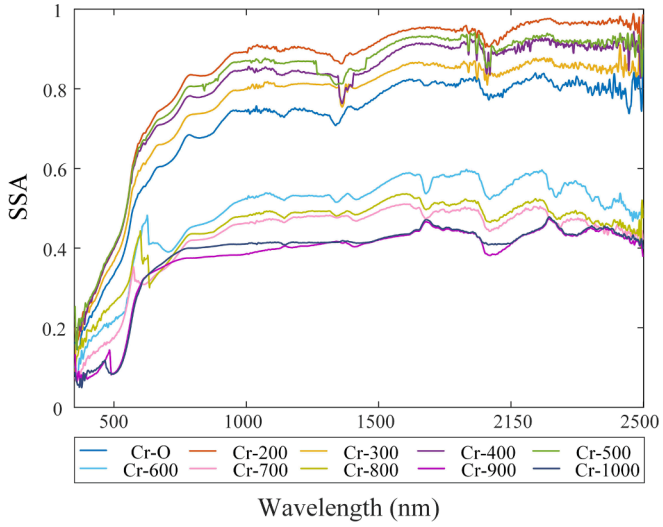


Fig. 12. The SSA under different concentrations of Cr pollution.

the energy levels. The electron jump (d-d jump) from the low-energy d-orbital to the high-energy d-orbital is generated. The energy difference is generally 1.99×10^{-19} to 5.96×10^{-19}J, and the wavelength is between 300 nm and 1000 nm, so $CrCl_3$ selectively absorbs the light in the visible region (Liu et al., 1994; Cheng et al., 2018).

The current soil heavy metal Cr monitoring methods require field sampling, laboratory testing, spectral measurement, and model building to obtain the heavy metal distribution tend, and thus cannot be used to achieve rapid and efficacious identification of soil heavy metal contamination. This method enables rapid analysis of soil heavy metal contamination in the study area using only spectroscopy and without chemical analysis, and identifies areas that may exceed the heavy metal thresholds, so that areas exceeding the heavy metal thresholds can be intensively sampled or monitored in subsequent studies.

4. Conclusion

In this study, we explored and verified the Cr contamination detection threshold for soil spectra. First of all, a simulation experiment with different soil Cr contents was designed based on the soil environmental quality risk control standard, to obtain gradient spectra. TD-HHT transformation was then introduced to conduct time–frequency

analysis. The Hilbert transformation curve of IMFs obtained by EMD with the TD-HHT algorithm can remove the complex environmental noise of soil spectra and effectively detect the differences between soil contaminated with different heavy metal concentrations. When the Cr content in soil reaches about 600 mg/kg, the heavy metal pollution can be identified between 600 and 900 nm from the amplitude curve. In addition to the signal decomposition method, the TV-Hapke model was utilized to verify the conclusion. The correlation between the soil optical constant and soil heavy metal was explored, and it was found that the optical constant SSA with different Cr contents also behaves differently at this value. This research centered on examining the theoretical foundation of the model and validating its viability through experiment conducted in optimal indoor conditions.

In our future research, we will attempt to verify the application of the proposed method on other heavy metals, further on other samples. At present, this method is aimed at heavy metal information mining and pollution threshold identification in laboratory soil spectra. We will overcome current application limitations which result from the complex imaging conditions and mixed pixels of remote sensing images, and so on, and apply the proposed method in airborne and satellite hyperspectral images. This method will provide a novel way for the direct detection of soil Cr pollution from hyperspectral remote sensing technology, and effectively expand the application of hyperspectral remote sensing technology in soil heavy metal pollution monitoring. It improves the efficiency of localization and early warning monitoring of soil heavy metal pollution, provides technical support for the emergency investigation of soil heavy metal pollution, and realizes the rapid classification of soil heavy metal pollution levels, which provides important application for soil pollution prevention and control.

CRediT authorship contribution statement

Lihan Chen: Writing – original draft, Methodology, Data curation. **Kun Tan:** Writing – review & editing, Resources, Conceptualization. **Xue Wang:** Writing – review & editing, Methodology. **Yu Chen:** Supervision, Conceptualization.

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.

Data availability

Data will be made available on request.

Acknowledgements

We express our gratitude to Prof. Jihong Dong in China University of Mining and Technology and Prof. Erkai He in East China Normal University for generously providing the original soil samples. This research is jointly supported by the Shanghai Municipal Science and Technology Major Project (No. 22511102800), National Natural Science Foundation of China (No. 42171335),and National Civil Aerospace Project of China (No. D040102).

References

Asadzadeh, S., de Souza Filho, C.R., 2016. A review on spectral processing methods for geological remote sensing. Int. J. Appl. Earth Obs. Geoinf. 47, 69–90.
Bishop, J.L., Pieters, C.M., Edwards, J.O., 1994. Infrared spectroscopic analyses on the nature of water in montmorillonite. Clay Miner. 42, 707–716.
Bolan, N., Kunhikrishnan, A., Thangarajan, R., Kumpiene, J., Park, J., Makino, T., Kirkham, M.B., Scheckel, K., 2014. Remediation of heavy metal (loid)s contaminated soils - to mobilize or to immobilize. J. Hazard. Mater. 266, 141–166.
Chen, L., Lai, J., Tan, K., Wang, X., Chen, Y., Ding, J., 2022. Development of a soil heavy metal estimation method based on a spectral index: combining fractional-order derivative pretreatment and the absorption mechanism. Sci. Total Environ. 813, 1–12.
Cheng, H., Wan, Y., Chen, Y., Wan, Q., Shi, T., Shen, R., Guo, K., Hu, J., 2018. Study on the characteristics and mechanism of visible and near infrared reflectance spectra of soil heavy metals. Spectrosc. Spectr. Anal. 38 (3), 771–778.
Costa, M., 2003. Potential hazards of hexavalent chromate in our drinking water. Toxicol. Appl. Pharmacol. 188 (1), 1–5.
Fearn, T., Riccioli, C., Garrido-Varo, A., Guerrero-Ginel, J.E., 2009. On the geometry of SNV and MSC. Chemom. Intel. Lab. Syst. 96 (1), 22–26.
Fu, P.J., Yang, K.M., 2018. A spectroscopic second-order differential gabor expansion method for copper, Lead pollution detection in soil. Spectrosc. Spectr. Anal. 38 (10), 3245–3253.
Fu, P., Yang, K., Cheng, L., Wang, M., 2019. HHT identification and BC-PLSR prediction model of soil Lead pollution Spectrum. Spectrosc. Spectr. Anal. 39 (5), 1543–1550.
Hapke, B.J.I., 2002. Bidirectional reflectance spectroscopy: 5. The Coherent Backscatter Opposition Effect and Anisotropic Scattering. 157 (2), 523–534.
Huete, A.R., Escadafal, R., 1991. Assessment of biophysical soil properties through spectral decomposition techniques. Remote Sens. Environ. 35 (2–3), 149–159.
Hunt, G.R., 1977. SPECTRAL signatures of particulate minerals in the visible and near infrared. Geophysics 42, 501–513.
Jiang, X., Ye, Q., Lin, Y., Li, X., 2017. Inverting study on soil water content based on harmonic analysis and hyperspectral remote sensing. Acta Opt. Sin. 37 (10), 300–310.
Liu, W., Chang, Q., Guo, M., Xing, D., Yuan, Y., 2011. Extraction of first derivative Spectrum features of soil organic matter via wavelet de-noising. Spectrosc. Spectr. Anal. 31 (1), 100–104.
Liu, H., Li, B., Sun, J., 1994. Study on eleetronic structure and d-d excited energies of Cr3 + complexes. Acta Phys. Chim. Sin. 10 (11), 978–985.
Ou, D., Tan, K., Lai, J., Jia, X., Li, J., 2021. Semi-supervised DNN regression on airborne hyperspectral imagery for improved spatial soil properties prediction. Geoderma 385 (X), 114875.
Palacios-Orueta, A., Ustin, S.L., 1998. Remote sensing of soil properties in the Santa Monica Mountains I. Spectral Analysis. Remote Sensing of Environment. 65 (2), 170–183.
Post, J.L., Noble, P.N., 1993. The near-infrared combination band frequencies of dioctahedral smectites, micas, and illites. Clays Clay Minerals. 41 (6), 639–644.
Regulation, S.A.M., 2018. Soil environmental quality risk control standard for soil contamination of agricultural land. Ministry of Ecological Environment of the people's Republic of China, Beijing, pp. 1–7.
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.
Scheinost, A.C., 1998. Use and limitations of second-derivative diffuse reflectance spectroscopy in the visible to near-infrared range to identify and quantify fe oxide minerals in soils. Clay Clay Miner. 46 (5), 528–536.
Srasra, E., Bergaya, F., Fripiat, J.J., 1994. Infrared spectroscopy study of tetrahedral and octahedral substitutions in an interstratified illite-smectite clay. Clays Clay Minerals. 42 (3), 237–241.
Stenberg, B., Rossel, R.A.V., Mouazen, A.M., Wetterlind, J.J.A..IA., 2010. Visible and near infrared spectroscopy in soil science. Adv. Agron. 107, 163–215.
Susi, H., Byler, D.M., 1983. Protein structure by fourier transform infrared spectroscopy: second derivative spectra. Biochem. Biophys. Res. Commun. 115 (1), 391–397.
Tan, K., Ma, W., Chen, L., Wang, H., Du, Q., Du, P., Yan, B., Liu, R., Li, H., 2020a. 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., Chen, L., Du, Q., Du, P., Pan, C., 2020b. Estimation of the spatial distribution of heavy metal in agricultural soils using airborne hyperspectral imaging and random forest. J. Hazard. Mater. 382, 120987.
Unep, 2021. Measuring Progress of environment and the SDGs. United Nations Environment Programme.
Viscarra Rossel, R.A., Behrens, T., 2010. Using data mining to model and interpret soil diffuse reflectance spectra. Geoderma 158 (1–2), 46–54.
Viscarra Rossel, R.A., Behrens, T., Ben-Dor, E., Brown, D.J., Dematte, J.A.M., Shepherd, K.D., Shi, Z., Stenberg, B., Stevens, A., Adamchuk, V., Aichi, H., Barthes, B.G., Bartholomeus, H.M., Bayer, A.D., Bernoux, M., Bottcher, K., Brodsky, L., Du, C.W., Chappell, A., Fouad, Y., Genot, V., Gomez, C., Grunwald, S., Gubler, A., Guerrero, C., Hedley, C.B., Knadel, M., Morras, H.J.M., Nocita, M., Ramirez-Lopez, L., Roudier, P., Campos, E.M.R., Sanborn, P., Sellitto, V.M., Sudduth, K.A., Rawlins, B.G., Walter, C., Winowiecki, L.A., Hong, S.Y., Ji, W., 2016. A global spectral library to characterize the world's soil. Earth Sci. Rev. 155, 198–230.
Wang, Y., Zhang, L., Wang, H., Gu, X., Zhuang, L., Duan, L., Li, J., Li, J., 2018. Quantitative inversion of soil organic matter content based on continuous wavelet transform. Spectrosc. Spectr. Anal. 38 (11), 3521–3527.
Wight, J.P., Ashworth, A.J., Allen, F.L., 2016. Organic substrate, clay type, texture, and water influence on NIR carbon measurements. Geoderma 261, 36–43.
Wu, Y., Chen, J., Ji, J.-F., Gong, P., Liao, Q.-L., Tian, Q., Ma, H., 2007. A mechanism study of reflectance spectroscopy for investigating heavy metals in soils. Soil Sci. Soc. Am. J. 71, 918–926.
Yang, K.M., Wang, G.P., Fu, P.J., Zhang, W., Wang, X.F., 2018. A model on extracting the pollution information of heavy metal copper ion based on the soil spectra analyzed by HHT in time-frequency. Spectrosc. Spectr. Anal. 38 (2), 564–569.
Yang, G.J., Zhao, C.J., Huang, W.J., Wang, J.H., 2011. Extension of the hapke bidirectional reflectance model to retrieve soil water content. Hydrol. Earth Syst. Sci. 15 (7), 2317–2326.

Yao, Y. M., Liu, Y., Gao, M. F. and Chen, Z. X. (2018). Hyperspectral Inversion of Soil Moisture Content Based on SOILSPECT Model. 7th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), George Mason Univ, Ctr Spatial Informat Sci & Syst, Hangzhou, PEOPLES R CHINA.

Zhang, S., Shen, Q., Nie, C., Huang, Y., Wang, J., Hu, Q., Ding, X., Zhou, Y., Chen, Y., 2019. Hyperspectral inversion of heavy metal content in reclaimed soil from a mining wasteland based on different spectral transformation and modeling methods. Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy. 211, 393–400.

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. (Basel) 12 (14), 1–21.

Zheng, S., Qiu, J., Zheng, C., Zhan, G., Bao, Y., Zhang, W., Liu, W., He, W., 2017. Review on remediation technologies on chromium contaminated soil. Henan Sci. Technol. 9, 157–158.