ISSN: 2688-8645

Research Progress in Remote Sensing Monitoring of Soil Heavy Metal Pollution

Tingyu Zhang[1, 2, 3, 4], Tianqing Chen[1, 2, 3, 4], Na Lei[1, 2, 3, 4], Yanan Li[1, 2, 3, 4], Yufei Xiong[1, 2, 3, 4], Panpan Zhang[1, 2, 3, 4], Jianglong Shen[1, 2, 3, 4], Xiaohu Fan[1, 2, 3, 4] and Ming Chang[1, 2, 3, 4]

[1]Shaanxi Provincial Land Engineering Construction Group Co., Ltd., Xi'an, 710075, China

[2]Institute of Land Engineering and Technology, Shaanxi Provincial Land Engineering Construction Group Co., Ltd., Xi'an, 710075, China

[3]Key Laboratory of Degraded and Unused Land Consolidation Engineering, the Ministry of Natural Resources, Xi'an, 710075, China

[4]Shaanxi Provincial Land Consolidation Engineering Technology Research Center, Xi'an, 710075, China

Abstract

In recent years, heavy metal pollution in farmland in my country has become more and more serious, causing great harm to crops and human health. Heavy metal pollutants move poorly in the soil, stay for a long time, and are difficult to be degraded by microorganisms. It can eventually affect human health through water, plants and other media, so quantitative monitoring of heavy metal pollution is very necessary and significant. The development of hyperspectral remote sensing technology provides a new opportunity for macroscopic and rapid acquisition of soil heavy metal element information. It has the characteristics of multiple bands, continuous and high resolution, and can dynamically monitor the distribution of heavy metals in farmland. At present, scholars at home and abroad have successfully predicted the content of various soil heavy metal elements based on the characteristics of soil reflectance spectrum and using a variety of statistical analysis methods. This paper focuses on the mechanism, influence factors, spectral characteristics and related prediction models of soil heavy metal hyperspectral remote sensing, and analyzes the existing problems of the current heavy metal hyperspectral remote sensing technology and prospects for it.

Keywords

Soil pollution; Heavy metals; Hyperspectral; Remote sensing.

1. Introduction

Soil is one of the main natural resources that mankind depends on for survival, and it is also an important part of the human ecological environment [1]. With the advancement of industrialization, urbanization and agricultural modernization, various chemicals, pesticides and fertilizers have been widely used, and pollutants represented by heavy metals have entered the farmland ecosystem, causing the quality of cultivated land to continue to decline. Heavy metals in soil pollution mainly refer to substances with significant biological toxicity such as mercury, cadmium, lead, chromium and metal-like arsenic, as well as general heavy metals with certain toxicity, such as zinc, copper, cobalt, nickel, tin, etc., which currently attract researchers the most attention The heavy metals are mercury, cadmium, lead and so on. At present, about 20% of my country's arable land is polluted by heavy metals, and the polluted area has reached more than 20 million hectares. Affected by heavy metal pollution, my country's annual grain

ISSN: 2688-8645

output is reduced by 10 million tons, and the direct economic loss caused by this is more than 20 billion yuan. The heavy metals Cr, Co, Ni, Cd, Zn, etc. have been proved to have carcinogenic effects. The "pain disease" caused by Japanese cadmium rice is a typical crop heavy metal pollution incident, which has extremely serious consequences for human life and health. In particular, heavy metals in farmland can spread along the food chain of crops-humans or crops-animals-humans and accumulate in the human body, causing great harm to human health. Therefore, the monitoring, repair and treatment of heavy metals in farmland are extremely urgent.

The qualitative and quantitative evaluation of soil heavy metal content and spatial distribution is a necessary prerequisite for soil heavy metal pollution remediation. Only after mastering the heavy metal distribution information can we carry out targeted remediation and treatment for different regions and different pollution levels. The traditional methods of soil heavy metal investigation include optical detection, electrochemical detection and biological detection. It has the advantages of high measurement accuracy and strong accuracy, but it is relatively time-consuming and labor-intensive, and it is difficult to obtain continuous pollutant content distribution information in a large area. Remote sensing technology has gradually been applied to the monitoring of soil heavy metal pollution due to its multi-temporal, multi-scale, all-weather and other characteristics. Hyperspectral remote sensing is based on the hyperspectral and high-altitude imaging technology of ground features. It has the characteristics of multiple bands, continuous and high resolution. It can record the reflectance of multiple narrow bands of ground features and perform more detailed soil heavy metal distribution mapping.

2. Inversion Mechanism of Soil Heavy Metal Content

Due to the low content of heavy metals in the soil, it is difficult to directly study the characteristic spectra of heavy metals. The main components such as organic matter, clay minerals and iron-manganese oxides in the soil can adsorb different types of heavy metals, thereby affecting the characteristic spectra of these heavy metals. By using the adsorption or occurrence relationship between heavy metal elements and soil organic matter, clay minerals, iron-manganese oxides, and carbonate minerals, the content of heavy metal elements in the soil can be indirectly inverted. The accuracy of the inversion depends to a certain extent on the content of heavy metal elements and The correlation between these components [2].

For example, Moros et al. found the correlation between soil heavy metal elements and organic matter in their research. Based on soil visible-near-infrared and mid-infrared reflectance spectra. The partial least squares regression model was established to realize the quantitative monitoring of the pollution level of As, Cd, Co and other elements in the river beach soil. Xie Xianli et al. studied the visible-near-infrared reflectance spectra of the polluted area of copper smelters. As the main pollutant in the study area, Cu is active in many spectral regions of the soil reflectance spectrum. The correlation with the reflectance spectrum is mainly affected by organic matter. Impact. Pb, Zn, Co, Ni are mainly affected by clay minerals and iron-manganese oxides, and the correlation between Cr and reflectance spectra is also affected by organic matter and clay minerals. The study by Xia et al. found that the heavy metal Cd content in the sediments of the Baguazhou shoal of the Yangtze River is related to the visible-near-infrared spectrum characteristics of organic matter, clay minerals, and iron oxides. Xu Mingxing et al. analyzed the spectral information of soil in different historical periods and found that the spectral diagnostic band of As content exists in the characteristic absorption region of iron oxides, carbonate minerals and organic matter; Cd is mainly characterized by the influence of iron oxide components. Soil reflectance spectra; Cu content inversion mainly involves the sensitive band of soil clay minerals; Ni content inversion is related to the characteristic band of organic matter; Pb content needs to also consider the characteristic bands of iron oxide and clay minerals; Cr

ISSN: 2688-8645

inversion The evolution is more complicated and needs to take into account the influence of many factors, such as soil organic matter and moisture.

3. Spectroscopic Instrument and Spectral Data Preprocessing

At present, FieldSpec, PIMA and SPECTRO are the most widely used spectrometers in the world. The most commonly used domestically are FieldSpec portable spectroradiometric spectrometers, which mainly measure soil reflectance and emissivity; ISI921VF-128 surface spectrometer is used in vegetation spectroscopy research. Commonly used instruments; In addition, with the advancement of science and technology, airborne and spaceborne hyperspectral instruments have gradually developed and matured. Aeronautical spectroscopy instruments such as AVIRIS, Hymap and TG-1 play an important role in the acquisition of hyperspectral data.

In the process of heavy metal contaminated soil sampling, it is inevitable to be affected by environmental random factors and the noise of the spectrometer itself, so it is necessary to preprocess the collected spectral data. In view of the possible deviations in the joints of different detection elements, the spectral data can be corrected for breakpoints; for the spectral data without such problems, the convolution method can be used to directly smooth and denoise the spectral data. To extract the heavy metals in the soil spectrum, the common spectral processing methods are as follows. 1) Resampling processing. In order to eliminate the data redundancy between adjacent bands in the process of spectral data collection, 10nm interval re-sampling is generally used to perform arithmetic average calculation. The resulting spectral curve is smoother and can better maintain the original curve spectral characteristics. 2) Baseline correction. The main purpose of baseline correction is to eliminate the influence of background or drift on the spectral data. The first-order differential processing that has a greater impact on the correction result is often used. Due to the fitting phenomenon, a very good regression is obtained, and the predictive behavior index is very high. In this case, the model built is not very good, and the prediction of the sample will produce a large deviation. 3) Variable standardization. Variable standardization is suitable for deducting the linear translation caused by the particle size and additional scattering of the sample in the sample spectrum, so that the feature with the smallest error probability of the classification result is selected as the most effective feature. For different spectrometers, the spectrum measurement range and spectrum data will have certain differences. The standard orthogonal transformation method and binomial interpolation method are commonly used for variable standardization to reduce the difference between the spectra within the class, increase the distance between the classes, and help improve the discrimination. The accuracy of the analysis. 4) Differential processing. Differential method is also called derivation method, which can eliminate part of the baseline and other background interference, distinguish overlapping peaks, thereby improving spectral sensitivity. Studies have shown that the low-order differential processing of the spectrum is less sensitive to noise and is more effective in practical applications. It is generally believed that the first-order differential processing technology can be used to remove part of the linear or nearly linear background and noise spectrum's influence on the target spectrum. 5) The reciprocal logarithmic processing method of the spectrum also has an ideal processing effect. The logarithmic change of soil reflectance not only enhances the spectral difference in the visible light region (the original spectrum in the visible light region is generally low), but also tends to reduce the influence of random factors caused by changes in lighting conditions, terrain, etc., and improve the accuracy of the model.

ISSN: 2688-8645

4. Spectral Image Factor

As a complex mixture, soil is composed of various substances with different physical and chemical properties, and these substances have varying degrees of influence on the soil spectrum. Soil-forming minerals (especially iron oxide), water content, organic matter and soil texture are the main influencing factors of soil spectral characteristics. Different types of soils have different spectral characteristics. For soils of the same type or similar composition, their spectral characteristics are related to soil particle size, water content and soil organic matter content. Therefore, the research on the content of heavy metals in soil needs to minimize or ignore the effects of soil particle size, water content and organic matter. Laboratory methods are commonly used to grind, filter, and dry soil samples, while the detection of heavy metal content in soil in the field selects areas with a single soil background value and close to vegetation and man-made influences for research, that is, it is assumed that the size of soil particles in the study area, The water content is similar, and the organic matter composition and content are basically the same. The main factor affecting the difference in soil spectrum is heavy metals [3].

5. Spectral Characteristics of Soil Polluted by Different Heavy Metals

In the hyperspectral retrieval of soil heavy metals, the main characteristic spectral bands used are visible light, near infrared, mid-infrared and thermal infrared. Due to the differences in the composition of heavy metal adsorbates in the soil, their own properties and their adsorption mechanism for heavy metals, the optimal inversion bands for soil heavy metal hyperspectral modeling are different. Siebielec et al. found that near-infrared spectroscopy is not as effective as mid-infrared spectroscopy in predicting Cd and Cu in mining areas. For soils contaminated by different types of heavy metals, the main characteristic spectrum of such heavy metal pollution should be selected to make it distinguishable. Li Jubao et al. studied the indoor spectral reflectance and the relationship between the Fe, Zn, and Se content of 51 soil surface samples from farmland on both sides of the Fuyang River, and found that the best spectral interval for predicting Fe is 16nm, and that of Zn and Se is 8nm. Although Zn and Se have no spectral characteristics, they are significantly related to Fe, which has spectral characteristics, so that they can be predicted by reflectance spectra [4]. Gan Fuping found that in general, Mn^{3+}, Fe^{2+}, and Fe^{3+} have characteristic bands in the visible light region or near infrared due to electronic transitions. The characteristic bands of manganese are mainly 0.45 μm and 0.55 μm, and those near 0.36 to 0.41 μm. Three absorption features, or a relatively steep semi-absorption feature in the 0.4-0.6μm region. Fe^{2+} produces a common strong and broad band around 1.0～1.1μm; Fe^{3+} produces a strong absorption band between 0.6～0.9μm, but the characteristic spectral shapes of different types of iron compounds are quite different.

Due to the universality of pollution and the particularity of the spectrum, there are many studies on the heavy metals Fe and Cu at home and abroad. This article will focus on the summary. For other serious and common heavy metals, such as Cr, Pb, Cd, Zn, Hg, etc., this article also summarizes the existing conclusions.

5.1. Spectral Characteristics of Soil Polluted by Heavy Metal Iron

Iron exists in the soil mainly in the form of iron oxide. Iron oxide is an important soil component that affects the spectral reflectance characteristics of the soil, and the increase in its content will reduce the reflectivity. Generally speaking, there is a negative correlation between the iron oxide content of the soil and the reflectivity. When the iron oxide content increases, the absorption of visible light and near-infrared is enhanced, while the absorption enhancement in the wavelength range of 0.5-0.7μm is not large, and the correlation is not obvious. The sensitive band of heavy metal iron is the same as that of organic matter. It is mainly located in visible light and near infrared. It is difficult to quantitatively distinguish the contribution of organic matter and iron oxide to spectral reflectance. Therefore, it is very difficult to accurately estimate soil iron oxide content through remote sensing technology. In 1964, Obulhov and Orlov pointed out that there are absorption bands of ferric oxide at 0.7 μm and 0.9 μm of the soil spectral curve. In 1976, Montgomery's research showed that the presence of free iron oxide has a great influence on the entire visible and near-infrared spectrum. Research by He Ting et al. showed that: ①The iron oxide content in the soil is negatively correlated with the reflectivity, and the increase in the iron oxide content will cause the soil reflectivity to decrease; ②After the first-order differential transformation, the square root of the reflectivity has a better effect on iron oxide at 1674nm. Sensitivity; ③The logarithm of the average reflectivity of the red band and the blue band can be used to form the soil iron oxide index, which has a good correlation with soil iron oxide.

5.2. Spectral Characteristics of Copper Contaminated Soil

In 2005, Wu Yunzhao studied the spectral characteristics of heavy metal elements, and found that the heavy metal element Cu in the soil exceeds 4000mg/kg concentration, it will respond in the soil spectrum, otherwise the spectral characteristics will be affected by the characteristics of other components cover. In 2008, Yang Lu et al. conducted a hyperspectral experiment on heavy metal contaminated soil with a copper content of 1,000 to 5000 mg/kg, and concluded that it is not feasible to directly monitor heavy metal copper contaminated soil by using hyperspectral remote sensing. In 2008, Dr. Ren Hongyan conducted spectral detection of heavy metal pollution in the farmland soil-rice system of Baoshan mining area and concluded that 540nm, 670nm and 2100-2190nm are the key bands for predicting heavy metal Cu models. In 2010, Huang Changping conducted research on the band selection and the best spectral resolution for remote sensing inversion of soil heavy metal Cu content. The best number of bands for predicting the heavy metal Cu content was 10, and the best spectral sampling interval was 32nm. In 2011, Wang Wei et al. based on the hyperspectral inversion of soil heavy metal copper, and found that the soil heavy metal Cu has a significant correlation with Mg and Fe (n-34, P-0.05), but has a poor correlation with soil organic matter.

5.3. Spectral Characteristics of Soil Contaminated by Other Heavy Metals

In 1997, Malley and Williams used reflectance spectroscopy to quickly predict the content of heavy metals in lake sediments. Based on the positive correlation between Cd and Zn content and soil organic matter, Kooistra et al. used reflectance spectroscopy to predict soil Cd and Zn pollution in the Rhine River Basin. Using the correlation between heavy metal elements and Fe, Kemper used reflectance spectroscopy to successfully predict the As, Fe, Hg, and Pb content in the soil of the mining area. Through the analysis of heavy metal elements and soil spectral characteristics, the heavy metal elements Ni, Cr and Cu have diagnostic spectral characteristics caused by crystal field effects in the VNIR band, while heavy metal elements Zn, Hg, Cd and Pb have no spectral characteristics. Among them, Ni, Cr, Cu have the highest correlation with reflectivity, and Hg and Cd have the lowest correlation with reflectivity. Gong Shaoqi et al. studied the occurrence of soil heavy metals Cr, Cu, Ni and soil clay minerals, iron and manganese oxides and carbonates through the measurement of soil reflectance spectra and simultaneous soil chemical analysis, and concluded that: These three heavy metals have a good correlation with the spectral variables at wavelengths of 429nm, 470nm, 490nm, 1430nm, 2398nm, and 2455nm.

6. Hyperspectral Estimation Model of Soil Heavy Metals

At present, there are two main types of hyperspectral models for estimating soil heavy metals: empirical statistical models and physical models. The empirical statistical model is an indirect

ISSN: 2688-8645

method, mainly a regression method. The physical model is a direct method, such as the radiative transfer model (RTM) and the geometric optical model (GOM). The content of heavy metal elements in the soil is very low, there is no obvious absorption characteristic in each band of the soil reflection spectrum, and the composition of the soil is complex, and the influence of each component on the reflection spectrum is nonlinear mixing, which makes the reflection radiation process of the soil complicated. It is difficult to invert with physical models. Statistical methods are usually used to analyze the correlation between soil heavy metal content and reflectance spectrum characteristics, and indirectly estimate the soil heavy metal element content. Regression methods mainly include univariate regression (UR), multiple linear regression (MLR), principal component regression (PCR) and partial least square regression (PLSR).

Univariate statistical methods mainly use correlation analysis methods to explore whether there is a significant correlation between soil heavy metal content and spectral reflectance. Choose the most significant wave band to build a model to predict the content of heavy metals in the soil. According to different waveband selection methods, it can be divided into single-band analysis method and analysis method after effective waveband conversion. For example, Li Shumin and others used the method of spectral analysis to explore the correlation between the content of heavy metals in agricultural soils in Beijing area and the reflectance of visible-near infrared spectroscopy. The original reflectance spectra of soil samples and their first- and second-order differential spectra were compared with the content of heavy metals in the soil. Single-band analysis determined the characteristic spectra of eight soil heavy metals including Cr, Ni, Cu, and established a regression model for estimating the content of soil heavy metals. Ren Hongyan researched and analyzed the original reflectance spectrum of the farmland soil in the mining area and the spectral information after continuum removal, and determined the band with the greatest correlation between the soil spectral reflectance and the content of heavy metal elements, and obtained the best inversion of the content of heavy metal elements such as Cu and Cd. Fit the model.

Multivariate statistical analysis is a common method for predicting spectral characteristic substances in spectroscopy. Because of its comprehensive use of more bands, the accuracy of statistical prediction is improved. Multiple stepwise regression analysis, principal component regression (PCR) analysis and partial least square regression (PLSR) analysis are currently commonly used statistical methods to analyze the relationship between soil composition and reflectance spectra.

Principal component regression analysis is a multiple regression analysis method. It is widely used in chemistry and spectroscopy analysis. It uses all spectral information and compresses it, assigns highly correlated wavelength points to an independent variable, extracts a few independent variables to establish a regression equation, and prevents overfitting through internal testing. The root mean square error of prediction (RMSEP) is used to evaluate the predictive ability of the model. Some scholars have achieved good prediction results by using the principal component regression analysis method. Wu et al. established an inversion model of indoor soil spectrum and Hg content by PCR method, the correlation coefficient between the two was 0.69, and the root mean square error RMSE was 0.15.

Partial least squares regression method as an effective spectral analysis method has been widely used in spectral data processing. This method provides a many-to-many linear regression modeling method, when the number of two sets of variables is large. And when there are multiple correlations and the amount of observation data is small, compared with the traditional multiple linear regression analysis method, the PLSR method solves the multiple collinearity problem faced by the multiple linear regression method, and can summarize and extract spectral information. Thereby, the content of heavy metal elements can be inverted quantitatively more accurately. Moreover, compared with the principal component regression analysis, the PLSR method not only summarizes the spectral information well, but also requires the newly generated components to have the strongest explanatory properties for the dependent variable (heavy metals). In a sense, the PLSR model combines two methods of multiple linear regression and principal component analysis. In addition, compared with the artificial neural network method (ANN), the factor loading of PLSR can vividly reveal the relationship between the independent variable and the dependent variable, thereby helping to understand the mechanism of using reflectance spectroscopy to retrieve heavy metals without obvious spectral characteristics. Based on the above reasons, PLSR is widely used in the current research to invert the content of soil heavy metal elements, and good inversion results have been achieved. Domestic Ren and others apply the PLSR method. Analyzing the reflectance spectrum of the salt marsh soil of the Yangtze River Estuary, quantitatively inverted the content of heavy metals As and Cu in the soil, and obtained extremely significant correlation results. Zheng Guanghui and others used the PLSR method to establish a model between the reflectance spectrum and the soil As content, and verified the modeling accuracy through cross-validation and estimation. It proves the feasibility of using reflectance spectrum to retrieve the As content in soil. Foreign Kooistra et al. found that using the reflectance spectra of river beach soils can better invert the pollution levels of soil heavy metals Zn and Cd, and pointed out that the PLSR model established by using soil visible-near infrared reflectance spectroscopy is an effective way to quantitatively analyze the content of river beach soils and heavy metals.

7. Problems and Prospects

There are many factors that affect the reflection spectrum acquisition in the field or the application of high-altitude remote sensing. Such as surface conditions (thickness, soil moisture, vegetation coverage, etc.), atmospheric absorption and light conditions, the current research on farmland heavy metal hyperspectral remote sensing is mainly based on field sampling and laboratory analysis. This method has poor timeliness and small monitoring range. It is greatly affected by the location and number of sampling points; while there are few spectroscopic studies based on high-altitude/satellite remote sensing images. Among them, single-time image analysis is often used, and continuous time-phase images are not used to analyze the distribution of heavy metals in large-scale farmland soils.

In recent years, the research on soil/vegetation hyperspectral inversion of heavy metal content has made great progress, and related inversion models have been widely used. However, most of the current hyperspectral inversion models are empirical models, such as the relationship between heavy metal adsorbents and heavy metals. Issues such as the quantitative relationship between soil heavy metal adsorption in different types of farmland soil, and the extent to which various factors affect the spectral characteristics of soil/vegetation have not yet been resolved, and physical modeling is very difficult; and at present, most of the heavy metals in the soil can only be retrieved by hyperspectral Determining the total amount of some heavy metals cannot predict the proportion of the forms of each heavy metal in the soil. The modeling form is single, the characteristic spectrum database data is insufficient and the accuracy is not high, the application range is narrow, and the degree of noise reduction is not enough to cause problems such as low accuracy of the quantitative inversion of the target. In general, using hyperspectral technology to estimate the content of heavy metal elements in the soil, the overall accuracy of the model fitting can reach 75% to 80%, with an average relative error of 30%. 40%, the verification accuracy is 60% to 70%.

With the continuous innovation of remote sensing technology, there are more and more methods of remote sensing inversion of soil heavy metal content. Establishing a model that is more suitable for the existing remote sensing technology to retrieve the soil heavy metal

ISSN: 2688-8645

content and improve the simulation accuracy of the model will be the main goal of future research.

Acknowledgments

This work was supported by the Fundamental Research Funds for the Central Universities (300102351502); Inner scientific research project of Shaanxi Land Engineering Construction Group (number SXDJ2021-10, SXDJ2021-30, SXDJ2020-22).

References

[1] Zhijuan Li, Zongwei Ma, Tsering Janvan der Kuijp. A review of soil heavy metal pollution from mines in China: Pollution and health risk assessment. Science of The Total Environment. Vol. 3 (2014) No. 3, p. 845-853.

[2] Ioan Suciu, Constantin Cosma, Mihai Todică. Analysis of Soil Heavy Metal Pollution and Pattern in Central Transylvania. International Journal of Molecular Sciences. Vol. 1 (2008) No. 7, p. 37-39.

[3] Xudong Jiao, Yanguo Teng, Yanhong Zhan. Soil Heavy Metal Pollution and Risk Assessment in Shenyang Industrial District, Northeast China. Plos One. Vol. 21 (2005) No. 5, p. 127-136.

[4] Facchinelli, Sacchi, Mallen. Multivariate statistical and GIS-based approach to identify heavy metal sources in soils. Environmental Pollution. Vol. 11 (2019) No. 9, p. 313-324.