ISSN: 1813-4890

Hyperspectral Remote Sensing Monitoring Method for Soil Heavy Metal Pollution

Fangfang Liu

Shannxi Provincial Land Engineering Construction Group Land Survey Planning and Design Institute, Xi'an 710064, China

Abstract

Heavy metal pollution in soil is a serious threat to ecological environment and human health. Hyperspectral remote sensing technology has the ability to acquire ground object information quickly and macroscopically. It provides a scientific method for fast, efficiently and quantitatively acquiring the heavy metal contents in soil and pollution prevention. In this paper, the methods and principles of hyperspectral remote sensing monitoring of soil heavy metal pollution were summarized based on soil spectral analysis and vegetation spectral analysis. The process and modeling method of soil heavy metal content inversion were introduced. The prediction of heavy metal contents by soil spectral analysis is based on the laboratory soil spectrum, which has high prediction accuracy, but it is difficult to realize the large-scale monitoring of heavy metal pollution in soil. Vegetation spectral analysis is mostly based on field measuring spectral data, with relatively low prediction accuracy. However, it is easy to combine with hyperspectral images to predict regional soil heavy metal pollution. Therefore, it is a most popular research field in the future. Multiple stepwise linear regression and partial least squares regression are the most widely used modeling methods for inversion of soil heavy metal contents, but the inversion model often lacks universality and stability. With the development of hyperspectral sensors and acquisition platforms, modeling methods with better universality and stability will continue to emerge.

Keywords

Soil Heavy Metals; Hyperspectral Remote Sensing; Monitoring Method; Inversion Model.

1. Introduction

In recent years, China's industrial and agricultural production has developed rapidly, and the soil environmental pollution is becoming more and more serious, especially the heavy metal pollution with significant biological toxicity represented by copper (Cu), lead (PB), zinc (Zn), mercury (Hg), cadmium (Cd), chromium (Cr), arsenic (As), nickel (Ni). According to the 2014 national soil pollution survey bulletin, the total over standard rate of soil in China is 16.1%, of which inorganic pollutants account for 82.8% of the total over standard points. The quality of cultivated land in China is worrying due to heavy metal pollution. About 12 million tons of grain crops are lost due to soil heavy metal pollution every year. Soil heavy metal pollution will not only reduce the yield and quality of crops, but also affect human health through the food chain. Therefore, the monitoring of soil heavy metals is very necessary. How to quickly and efficiently detect the soil heavy metal pollution area and evaluate its pollution degree is not only the key to carry out soil pollution monitoring, but also an important premise for government departments to carry out soil environmental treatment.

Traditional soil heavy metal pollution detection methods usually carry out laboratory chemical analysis on soil samples collected in the field. This method has high measurement accuracy and strong accuracy. However, due to many detection links, long time-consuming and high cost, it is difficult to quickly obtain the content and distribution information of pollutants in a large area. Hyperspectral remote sensing has the ability to quickly and macroscopically obtain ground object information, and continues to show its unique advantages in the field of soil heavy metal pollution monitoring. This paper summarizes the principle, method and inversion model of soil heavy metal pollution monitoring using hyperspectral remote sensing technology, and analyzes the problems existing in soil heavy metal pollution monitoring and the development prospect in the future.

2. Overview of Hyperspectral Remote Sensing

Hyperspectral remote sensing, also known as imaging spectral remote sensing, is the cutting-edge technology of remote sensing. It combines spectral technology with imaging technology, and can obtain very narrow and continuous spectral information and spatial distribution information of the target. The spectral acquisition range covers the ultraviolet band to thermal infrared band of electromagnetic wave. Hyperspectral remote sensing has the characteristics of wide spectral range, high spectral resolution and "integration of maps", which greatly improves the ability of surface coverage detection and feature recognition, and makes it possible to identify feature categories semi quantitatively and quantitatively. In recent years, it has been gradually explored and applied to the field of soil heavy metal monitoring.

3. Hyperspectral Remote Sensing Monitoring Method of Soil Heavy Metal Pollution

Heavy metal elements in soil are trace elements, so it is difficult to capture heavy metal information in soil spectrum. Therefore, it is difficult to estimate the content of heavy metals by directly analyzing the absorption and reflection characteristics of spectral curve. According to the adsorption relationship between heavy metal elements and other substances and the stress characteristics of vegetation, the content of elements can be indirectly predicted and the degree of soil pollution can be determined. At present, there are two main methods to monitor soil heavy metal pollution by Hyperspectral Remote Sensing: (1) inverse estimation of element content directly according to the correlation between soil spectrum and heavy metal elements; (2) According to the spectral information of vegetation under heavy metal pollution stress, the pollution status of soil heavy metals is indirectly predicted.

3.1. Prediction of Heavy Metal Content by Soil Spectral Analysis

Although heavy metals are trace elements in soil and have no obvious characteristic spectrum, clay minerals, soil organic matter and other components in soil often have adsorption or occurrence relationship with them. The existence of these components leads to the change of soil spectral morphology and reflectance to a certain extent, which makes the soil spectral curve show specific reflection characteristics. The content of elements can be estimated by using the correlation between heavy metals and these components.

The realization method of estimating the content of heavy metals in soil by soil spectral analysis is to use the spectral data of soil samples measured in the laboratory or in the field, after pretreatment such as breakpoint repair and smoothing, and compare the transformed spectral indexes such as original spectrum and its mathematical transformation and continuum removal with the measured content of heavy metals in soil.

The best regression model between the content of heavy metals in soil and the characteristic band of spectral index was established, and the content of heavy metals in soil was quantitatively inversed by using the model.

It is found that there is a significant correlation between the contents of Cu, Pb, Zn, Co, Ni, Fe, CD, Cr and Mn in the polluted area of Guixi copper smelter in Jiangxi Province and the visible near infrared reflectance spectrum of the soil.

The reason is that organic matter, clay minerals, iron and manganese oxides and other soil components have adsorption on the above heavy metal elements. By collecting soil samples and analyzing soil spectral information, the inversion model of soil Cu, Pb and Zn elements in urban residential areas is established by using partial least square method. The inversion accuracy of Pb element is the highest (R^2 = 0.77, RMSE = 7.66), and the model has good prediction ability. Based on the soil hyperspectral data measured in the laboratory, the effective spectral variables 580, 810, 1410, 1910, 2160, 2260, 2270, 2350 and 2430nm related to iron oxides, organic matter and clay minerals are selected by stepwise regression method. The inversion model of Zn content is established by using random forest, support vector machine and partial least squares, The results show that the stochastic Senli model based on second-order differential transformation has the highest inversion accuracy (R^2 = 0.94, RMSE = 18.01), and is the best model for the inversion of zinc content in karst area.

The above research is based on laboratory soil hyperspectral data. The soil spectral data measured under laboratory conditions are least affected by external environmental conditions, and the spectral data processing conditions are relatively consistent. The research focuses on the selection of modeling methods. Therefore, the direct analysis of soil heavy metal content by laboratory soil spectrum is the most mature method in soil heavy metal hyperspectral inversion at present, However, due to the limitation of the number and scope of soil samples collected, it is difficult to realize the monitoring of soil heavy metal pollution in large areas based on laboratory soil hyperspectral.

3.2. Prediction of Heavy Metal Content by Vegetation Spectral Analysis

The research shows that heavy metals in soil can affect the physiological structure characteristics of plants, especially the synthesis of chlorophyll, to change the spectral characteristics of plants. When the vegetation is stressed by heavy metal pollution, the chlorophyll content in the vegetation leaves will decrease, making the "red edge" of the vegetation wave (referring to the sharp increase of the vegetation reflection spectrum characteristics from the red-light band to the near-infrared band (660 ~ 770nm) due to the absorption of chlorophyll in the plant) shift to the short-wave direction.

The phenomenon of "blue shift" of red edge is formed. "Red edge position" and vegetation index are important parameters to characterize the growth state of vegetation. The influence of heavy metal pollution on vegetation spectral characteristics and vegetation growth state parameters is the basis of predicting heavy metal pollution by vegetation spectral analysis. The realization method of predicting soil heavy metal content by vegetation spectral analysis is to measure the heavy metal content of surface soil and the chlorophyll content of vegetation leaves under the condition of heavy metal pollution (or form vegetation growth state parameters such as "red edge position" and vegetation index according to the spectral characteristics of vegetation), analyze the correlation between chlorophyll content or vegetation growth state parameters and soil heavy metal content, and establish an inversion model of soil heavy metal content, Predict the soil heavy metal pollution in the study area. One way to predict heavy metal pollution by vegetation spectral analysis is to establish a prediction model based on the hyperspectral reflection data of plant leaves collected by ground hyperspectral spectrometer and the measured chlorophyll content and heavy metal content of leaves. The canopy spectra of Winter Wheat in typical sewage irrigation areas were collected by field spec HH portable handheld ground object spectrometer, and the inversion models of Cr, Ni, Pb, Zn, Hg and CD elements were established to realize the inversion of heavy metal content and spatial distribution characteristics. Another way is to continuously obtain the spectral information of contaminated vegetation on the ground and establish an inversion model based on airborne and spaceborne hyperspectral spectrometer, to realize the real-time, large-area and in-situ monitoring of soil heavy metals. Hyperspectral images are obtained by the hyspexswir-384 imaging spectrometer carried by the UAV, and the ground spectral data are synchronously collected by the portable ground object spectrometer, to realize the rapid detection of CD elements in farmland; The modeling and inversion of as, Cr, Pb and other elements in farmland are carried out by using the GS hyperspectral data obtained by UAV, and the prediction accuracy of spectral modeling collected by ASD spectrometer is improved.

The results show that the UAV GS spectrum has the ability to monitor soil heavy metals, but the prediction accuracy is lower than ASD spectrum. Based on Hyperion hyperspectral image data, an estimation model of soil heavy metal Zn and Cd content in Yushu County Based on vegetation index was established. The ground measured spectral data has the characteristics of many bands and high precision. Therefore, the research on the content of soil heavy metals by using vegetation spectral analysis method mostly focuses on the analysis of the ground measured spectral data. However, due to the unique response characteristics of vegetation in hyperspectral images, and the advantages of fast acquisition speed and wide coverage of hyperspectral images, the research on the inversion of heavy metal content using airborne and spaceborne hyperspectral images has gradually increased in recent years. The vegetation spectral analysis method based on hyperspectral images has also become the focus of soil heavy metal content inversion research. Compared with the measured spectral data on the ground, the airborne and spaceborne hyperspectral data need more complex data preprocessing after acquisition. Such as noise removal (eliminating the influence of the sensor itself), radiation calibration, atmospheric correction (eliminating the influence of atmospheric, illumination and other factors), geometric correction (eliminating image geometric distortion) and other processing, in order to minimize the interference of atmospheric and other environmental factors on spectral information. The hyperspectral image data after data processing can be combined with the measured spectral data to predict regional soil heavy metals. However, due to the different observation scales of ground measured data and airborne (or spaceborne) hyperspectral data, there are differences between the two spectral parameters, resulting in the low accuracy of hyperspectral image prediction of soil heavy metals.

4. Establishment of Inversion Model

4.1. Spectral Processing

Because the soil background spectral signal is strong and the spectral noise is large, it is inevitable to be affected by the environment, instruments and human factors in the process of soil sample preparation and spectral data collection, so that the collected spectral data contains large noise. In addition, the soil background spectral signal is strong, which is easy to cover up the characterization of heavy metal elements in spectral information, and even the deviation or drift of the spectrum. Therefore, it is very necessary to process hyperspectral data before using soil spectral data to carry out correlation analysis of heavy metal content, to reduce the influence of background and noise.

(1) Spectral pretreatment

The function of spectral preprocessing is to reduce the noise caused by random factors in the process of spectral data acquisition. The main processing methods are as follows: breakpoint correction. The spectral data collected by the ground object spectrometer is usually composed of the spectra collected by several detection elements. Breakpoint correction is needed to eliminate the errors caused by different detection elements; Smoothing is used to weaken the influence of noise with high frequency and small value on the inversion model; Resampling can reduce the data redundancy between adjacent bands of hyperspectral data; Baseline correction can solve the problem of baseline offset or drift, standard normal variable correction and multivariate scattering correction, and eliminate the scattering effect caused by different particle sizes of soil samples in the spectrum.

(2) Spectral transformation

The function of spectral transformation is to weaken the influence of background noise and enhance the response of heavy metals in soil spectra. Common spectral transformation methods are: spectral differentiation, reciprocal logarithm, continuum elimination Except for others. Spectral differentiation technology can reduce the interference of environmental background on the spectrum, highlight the spectral information related to heavy metals in soil spectrum, and improve the accuracy of inversion model. First-order differentiation and second-order differentiation are the most used methods of spectral differentiation; The reciprocal logarithm can enhance the spectral difference and reduce the influence of random factors; Continuum removal can effectively highlight the absorption and reflection characteristics of the spectral curve and enhance the correlation with the content of heavy metals.

4.2. Characteristic Band Selection

The hyperspectral spectrum used to study soil heavy metals includes the visible light to thermal infrared region, and the bands are narrow and many. It is necessary to select the characteristic bands related to the content of heavy metals from these bands as the variables of inversion modeling. At present, the commonly used method is to select the characteristic band by using the stepwise regression algorithm based on the correlation coefficient between soil heavy metal content and soil reflectance. Some studies have shown that the spectral curve after spectral transformation can extract the characteristic band of the spectrum more effectively than the original spectral curve. According to the correlation coefficient and significance level between heavy metal content and soil reflectance, the corresponding characteristic spectral band of heavy metal elements is selected, and the best band is selected by using the stepwise regression algorithm and the best fitness value of the model as the index. The spectral logarithm first-order differential inversion model of Zn, Cu, Ni and Cr elements is established; Based on the mathematical transformation of the original soil spectrum, the spectral band is preliminarily selected by analyzing the absorption characteristics of the soil spectral curve, and the characteristic spectral variables are further selected by using the correlation analysis between the content of heavy metals and the spectral band. Finally, the stepwise regression method is used to determine the best spectral variable for modeling. In addition to the above methods, there are genetic algorithm, non-information variable elimination, continuous projection algorithm, interval partial least squares, Kalman filter, window pls, competitive adaptive weighting algorithm and other methods to select the characteristic band.

Due to the different material components and adsorption mechanisms of heavy metals in the soil, there are also differences in the physical and chemical properties of the soil itself. In the hyperspectral Inversion Modeling, even the best inversion bands of the same heavy metal elements are quite different. At present, the characteristic bands selected by most scholars in the inversion of soil heavy metal content are concentrated in the range of near-infrared to short wave infrared. Some scholars also use thermal infrared spectroscopy to retrieve the content of heavy metals and achieve corresponding results.

4.3. Modeling Method

Soil spectrum is a comprehensive reflection of the spectral properties of various material components in soil. Due to the complexity of soil components, it is difficult to directly establish a mathematical and physical model between the content of some heavy metals and the spectral bands of soil characteristics. At present, the method of establishing heavy metal content estimation model based on soil hyperspectral is usually empirical statistical method, which mainly includes multiple stepwise linear regression, principal component regression, partial least squares regression, artificial neural network regression, support vector machine regression, random forest, multiple adaptive regression, adaptive neurofuzzy inference diagnosis index regression and many other methods, In a specific study area, the models established by these methods can meet the standard of quantitative calculation, among which multiple stepwise linear regression and partial least squares regression are the most widely used inversion methods at present.

Multiple stepwise linear regression is a method of multiple linear regression by using the stepwise method to select the best spectral variables participating in the establishment of the model. It is suitable for the modeling of variables affected by multiple factors. The regression effect can be optimized according to the correlation degree between various factors. Firstly, the spectral band significantly related to the content of heavy metals is found through correlation analysis as the spectral variable of modeling, and then the variables entering multiple linear regression are gradually selected according to the significance level of statistics. The regression model with the largest determination coefficient R2 and the smallest root mean square error RMSE is determined as the best model for heavy metal element inversion.

Based on the hyperspectral remote sensing research on the content of heavy metals in archaeological soils in different historical periods, an inversion model of soil heavy metal content based on spectral first-order differentiation is established by using multiple stepwise linear regression method. Based on the soil spectral data measured in the laboratory, a multiple stepwise linear regression model suitable for the estimation of soil mercury content in Zhundong coalfield is established through various spectral transformations. The results show that the first-order differential transformation of soil spectrum is the best index for establishing the inversion model of heavy metal content.

Partial least squares regression is a modeling method that combines the advantages of principal component regression and multiple linear regression. It is suitable for the case that the number of samples is small or even less than the number of variables and there is multiple correlation between variables. It has the advantages of strong stability, high prediction accuracy and convenient for quantitative interpretation. It is widely used in the inversion of soil heavy metal content. Based on the spectral diagnostic characteristics of soil organic matter, a multi spectral transformation index partial least squares regression model (M-PLSR) of Pb was established by using partial least squares regression method; Taking the coal mining area of Zhundong coalfield as the research area, based on various transformations of soil light, the inversion model of heavy metal content of Zn, Cu and Ni elements is constructed by using partial least square regression method.

In the above studies, although the inversion methods used are basically the same, the best regression models obtained from each study are different due to the differences in the types of heavy metals in soil and soil texture. Therefore, although the above research methods have achieved good inversion results in specific areas, the adaptability of the inversion model is very poor.

5. Problems and Prospects

5.1. Existing Problems

(1) At present, the spectral measurements of soil samples are mostly laboratory measurements, which are carried out in an ideal state. It is difficult to apply the empirical results to aerospace remote sensing images to retrieve the spatial distribution of soil heavy metals. It is necessary to further strengthen the research on field spectral measurement, study and eliminate the influence of environmental factors such as atmosphere, light and vegetation, and realize the application of the combination of ground soil spectral measurement and remote sensing image.

(2) The modeling methods used to retrieve the content of heavy metals in soil are usually empirical statistical method and vegetation index method. The inversion model is often difficult to be widely used because of its lack of universality and stability. It is necessary to strengthen the research on the model and theoretical method of remote sensing hyperspectral inversion of soil heavy metal content, to enhance the universality, stability and prediction accuracy of soil heavy metal inversion model.

(3) Soil spectral characteristics are extremely complex, and there are many factors affecting soil spectral characteristics. Different soils have different adsorption states for different heavy metals, and the degree of adsorption of heavy metals by the same soil in different environments is also different. It is necessary to further study the attachment mechanism of heavy metals in soil and the effect of heavy metal content on soil spectral changes.

(4) At present, the research on soil heavy metals mainly focuses on the inversion of heavy metal content in a specific region and a specific period. Few scholars use long-time series hyperspectral remote sensing data to study the evolution characteristics and migration law of soil heavy metals in a certain region.

5.2. Outlook

(1) The data acquisition platform is more abundant. With the progress of technology, the acquisition methods of hyperspectral remote sensing data are more diversified, forming a "Star Air Ground" three-dimensional acquisition situation of spaceborne hyperspectral, Airborne Hyperspectral and ground hyperspectral. In particular, UAV hyperspectral has been widely used, which greatly improves the data acquisition efficiency and is more conducive to the monitoring of soil heavy metal pollution.

(2) Greatly improve sensor performance

With the rapid development of hyperspectral sensors and the continuous improvement of the accuracy of hyperspectral image data, the methods of soil heavy metal pollution monitoring have developed from laboratory hyperspectral analysis to hyperspectral image analysis. The research on hyperspectral and soil physical and chemical properties on a large spatial scale will continue to emerge, which will help to realize a wide range of soil quality monitoring.

(3) More diversified modeling methods

Modeling methods determine the accuracy of soil heavy metal information inversion. At present, there have been experiments to optimize and improve the existing models or apply new methods to the inversion of soil heavy metal content. With the development of hyperspectral data, new modeling methods will continue to appear, making the inversion model have stronger stability, higher inversion accuracy and better adaptability.

Acknowledgments

Project: Shaanxi Provincial Natural Science Basic Research Program (2022JQ-457).

Shaanxi Provincial Land Engineering Construction Group Internal Scientific Research Project (DJNY2021-10) (DJNY2022-16).

References

[1] Bonham-carter G F, 1988. Numerical procedures and computer program for fitting an inverted gaussian model to vegetation reflectance data [J]. Computers & Geosciences, 14(3):339-356.

[2] Gholizadeh A, Borůvka L, Vašát R, et al., 2015. Estimation of potentially toxic elements Contamination in anthropogenic soils on a brown coal mining dumpsite by reflectance spectroscopy: A case study [J]. Plos One, 10(2): e117457. https://doi.org/10.1371/journal. pone. 0117457.

ISSN: 1813-4890

[3] Kemper T, Sommer S, 2002. Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy [J]. Environmental Science & Technology, 36(12): 2742-2747.

[4] Teng Yanguo, Ni Shijun, Wang Jinsheng, et al., 2010. A geochemical survey of trace elements in agricultural and non- agricultural topsoil in Dexing area, China [J]. Journal of Geochemical Exploration, 104(3):118-127.

[5] WU Yunzhao, CHEN Jun, JI Junfeng, et al., 2007. A mechanism study of reflectance spectroscopy for investigating heavy metals in soils[J]. Soil Science Society of America Journal, 71(3):918-926.