World Scientific Research Journal (WSRJ) ISSN: 2472-3703
www.wsr-j.org

Study on Spectral Inversion Model of Soil Lead Content

Yuling Ba[1, 2, 3, 4, a], Jinbao Liu[1, 2, 3, 4, b, *]

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

[2]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 Land and Resources, Xi'an 710075, China

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

[a]790594557@qq.com, bliujinbao1212@163.com

Abstract: Chemical analysis showed that there was a serious phenomenon of Pb pollution in the study area, and Pb content approached the critical value. The results of the model show that: (1) the reflectance spectra of the reflectance spectra significantly increase the signal-to-noise ratio of the reflectance spectrum after being transformed by NOR, MSC and SNV. Combined with the differential transformation, it helps to improve the information of heavy metal elements in soils. Significantly improve model stability and predictive power. (2) The modeling accuracy of the optimal model of Pb element spectrum established by PLSR method is 0.6182, respectively. (3) The optimal estimation models established by different treatment methods of different elements have better stability and higher precision, and can quickly predict the Pb content of heavy metals in this area.
Keywords: Reflection spectra, soil heavy metals, partial least squares regression, visible – near-infrared.

---

1. INTRODUCTION

As the most precious natural resource for human survival, soil has a buffering and purifying effect on environmental pollutants (Soomro et al., 2014). In recent years, with the continuous development of industrialization and urbanization, the content of heavy metals in the soil has increased rapidly. Heavy metals are easily enriched and corrosive, seriously affecting crop yields and indirectly causing serious harm to the human body (Bachmann et al., 2014). At present, the determination of heavy metal content in soil is mainly based on spot-sampling in

the field, and the content of different heavy metal elements is obtained based on indoor chemical experiments. The traditional method for measuring heavy metal elements is costly and inefficient, and different elements require different chemical treatments, which cannot meet the monitoring of heavy metal pollution in large areas. Hyperspectral technology has been widely used in the prediction of heavy metal content due to its rich information, time saving, and no damage to sample structure.

In recent years, domestic and foreign scholars have conducted a large number of studies on the inversion of soil organic matter, nitrogen, phosphorus and potassium by hyperspectral (Zhang and Hu, 2013; Zheng, 2014), but there are few studies on the estimation of heavy metal content. Grzegorz et al. successfully predicted the content of Pb and Fe in the Spanish mining area by using reflectance spectroscopy (Grzegorz et al., 2014). Jiang Zhenlan et al. used the GWR hyperspectral model to predict the content of heavy metals Cd and Pb in Fuzhou soil (Jiang et al., 2017). Wang et al. successfully predicted the content of heavy metals such as Pb, Cd and Hg by partial least squares regression (Wang et al., 2007). In this study, the heavy metal elements in typical bauxite in Shaanxi Province were studied. The reflection spectra were preprocessed by standardization, multi-scattering correction and standard normal variable transformation. The Savitzky-Golay convolution smoothing method was used to smooth the spectral curve. Noise, and first-order differential, second-order differential, reciprocal logarithmic differential transformation, combined with partial least squares regression (PLSR) to establish a hyperspectral estimation model of Pb elements.

2. MATERIALS AND METHODS

2.1 2.1 Sample collection and determination of elemental content

The study area is located in Fufeng County, Yangling County and Wugong County of Shaanxi Province, and the soil type is bauxite. The soil sample was collected by the "S" point method, and the soil after removing the surface impurities was measured by a high-density reflection probe, and a total of 44 soil samples were used. The sample was naturally air-dried, decontaminated, and then mixed to obtain 200 g of soil sample and passed through a 100 mesh sieve for indoor heavy metal content determination. The statistical characteristics of heavy metal elements are shown in Table 1.

Table 1. Statistical result of heavy metal elements for soil samples

Element	samples	Max (mg kg-1)	Min (mg kg-1)	Mean (mg kg-1)	SD (mg kg-1)
Pb	44	25	19.3	21.0	1.1838

2.2 Determination of wild hyperspectral data

Soil reflectance spectroscopy was performed in the field using a high-density reflective probe equipped with an ASD Field Spec HR spectrometer. The spectrometer has a wavelength range of 350 to 2500 nm, a sampling bandwidth of 1.3 nm (350 to 1000 nm) and 2 nm (1000 to 2500 nm) with a resampling interval of 1 nm. High-density reflective probes can effectively avoid

the effects of stray light in the soil and eliminate the effects of weather. The 2cm field of view at the front end avoids stone particles and crop roots in the soil.

3. DATA ANALYSIS

3.1 Outliers culling

Soil samples will introduce different degrees of error in the process of collection, processing and analysis, which will affect the later data analysis and modeling accuracy. The Mahalanobis distance is based on a multivariate normal distribution, taking into account the covariance, mean and variance factors, and can comprehensively reflect the comprehensive index of soil samples. Therefore, this study used the Mahalanobis distance method to detect outliers in soil properties and spectral data.

3.2 Spectral differential transformation

First-order differential, second-order differential, and reciprocal logarithmic transformation of the original reflection spectrum are performed. In addition to directly analyzing the spectral reflectance of the soil, three transformations are performed to find the response regions of different heavy metal elements. First-order and second-order differential transformations increase the correlation between reflectivity and heavy metal elements while eliminating or limiting the effects of partially linear or near-linear backgrounds. First-order differential ρ'(λi) and second-order differential ρ" (λi) transformation are performed on the original spectral reflectance of the soil, where λi is the wavelength band of i nm, Δλ=λi+1-λi=10 nm, i =400,410, 2450 nm. The calculation formula is:

$$\rho^{'}(\lambda_i) = \left[\rho(\lambda_{i+1}) - \rho(\lambda_{i-1})\ \right] / \Delta\lambda \tag{1}$$

$$\rho^{''}(\lambda_i) = \left[\rho^{'}(\lambda_{i+1}) - \rho^{'}(\lambda_{i-1})\ \right] / \Delta\lambda \tag{2}$$

$$R^2 = \frac{\sum_{i=1}^{n}\left(\hat{y}_i - \overline{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \overline{y}_i\right)^2} \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{4}$$

4. RESULTS AND DISCUSSION

4.1 4Correlation analysis

In order to further analyze the correlation between heavy metal elements and spectral reflectance, the first-order differential, second-order differential, and reflectance reciprocal logarithmic transformation of the reflectance spectra were correlated with the heavy metal element content. The results are shown in Fig. 1. Compared with the original reflection spectrum, the correlation is significantly improved after differential transformation of the reflection spectrum. From the point of view of the differential transformation method, the

first-order differential effect on the reflection spectrum is best, and the second-order differential is second. Among them, the correlation coefficient between Pb and the first-order differential of the reflection spectrum is -0.69, which is significantly correlated. The soil reflection spectrum can significantly highlight the absorption characteristics after differential transformation, and enhance the correlation between heavy metal elements and reflectivity.

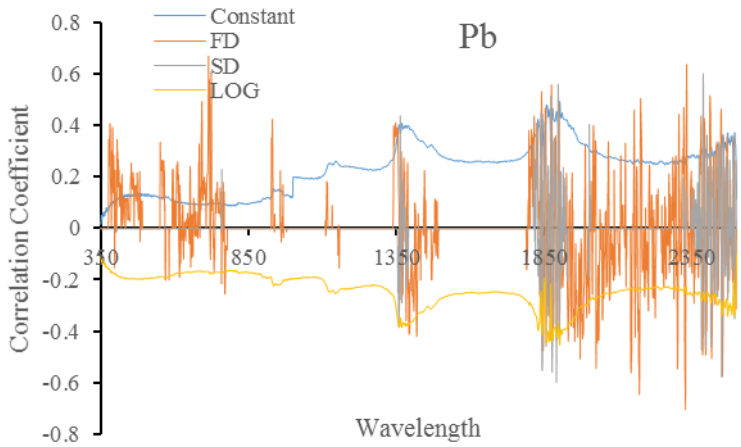

Fig 1. Correlation coefficient between Pb and spectra and their different transformations

4.2 Model establishment and comparison

The original reflection spectrum was used as the control group. After the spectral data was preprocessed by NOR, MSC and SNV, the first-order differential, second-order differential and reflectance reciprocal logarithmic transformation were performed. Savitzky-Golay nine-point smoothing was performed, and PLSR was used to establish the corresponding the heavy metal estimation model is tested using the decision coefficient and the root mean square error. The larger the modeling decision coefficient, the smaller the root mean square error, indicating that the stability of the estimated model is better. The larger the prediction coefficient is, the smaller the root mean square error is, indicating that the model prediction ability is stronger. At the same time, in order to avoid over-fitting of the model, the dimension of the independent variable of the model is as small as possible. Table 3 lists all modeling results for heavy metal elements, and Figure 3 shows the measured and predicted scatter plots for the optimal modeling effect of heavy metals.

5. CONCLUSION

In this study, the optimal hyperspectral estimation model of Pb elements was established by partial least squares regression method. By comparing the effects of different pretreatment methods on the establishment of soil heavy metal spectral inversion models, the following conclusions were obtained:

After the reflection spectrum is processed by NOR, MSC and SNV, the first- order differential, second-order differential and reflectance reciprocal logarithmic transformation are respectively carried out, which effectively reduces the influence of external factors such as soil particle size and surface scattering on the spectrum. At the same time, the differential transformation helps to improve the correlation between the heavy metal elements in the soil

and the reflection spectrum, and the combination of the higher correlation bands can significantly improve the stability and prediction ability of the model. From the accuracy of the estimated model, the correlation coefficients of modeling and prediction exceeded 0.6, Rc2= 0.6182, Rv2= 0.9359, RMSEC= 0.875, RMSEP= 1.00. Although the modeling coefficient of cadmium is not high, the prediction coefficient is above 0.9, and the predicted root mean square error is small, which can realize the rapid monitoring of soil Pb content in the study area.

6. REFERENCES

[1] Soomro A, Siyal A A, Mirjat M S, et al. Seasonal variations of trace elements and heavy metal concentrations in Phuleli Canal water (Sindh), Pakistan.[J]. Sarhad Journal of Agriculture, 2014, 30:73-82.

[2] Bachmann R T, Johnson a C, Edyvean R G J. Biotechnology in the petroleum industry: An overview [J]. International Biodeterioration & Biodegradation, 2014, 86 (Part C (0)): 225-237.

[3] Zhang H, You-Ping H U. Acupuncture Clinical Application Overview in Chronic Obstructive Pulmonary Disease [J]. Journal of Jiangxi University of Traditional Chinese Medicine, 2013.

[4] Zheng W. Research Summary of Domestic and Foreign Scholars about 1949-1978 Years of Cross-Strait Relations [J]. Taiwan Research Journal, 2014, 64(3):573-582.

[5] GRZEGORZ S, GREGORY W M, TOMASZ I S, et al. Near and mid-infrared diffuse reflectance spectroscopy for measuring soil metal content [J]. Journal of Environment Quality, 2004, 33: 2056-2069.

[6] Jiang Zhenlan, YANG Yusheng, SHA Jinxing. Application of GWR model in hyperspectral prediction of soil heavy metals [J]. Acta Geographica Sinica, 2017, 72(3): 533-544.

[7] Wang L, Lin Q Z, Jia D, et al. Study on the Prediction of Soil Heavy Metal Elements Content Based on Reflectance Spectra[J]. Journal of Remote Sensing, 2007.