Pol. J. Environ. Stud. Vol. 33, No. 4 (2024), 4997-5001
DOI: *10.15244/pjoes/176704* ONLINE PUBLICATION DATE: 2024-05-27

Short Communication

Hyperspectral Inversion of Chromium Content in Soils of the Tangshan Iron Tailings Area

Yuna Jia[1], Deqing Gan[1], Yang Bai[1]*, Jinhua Wang[1], Yuan Dong[2], Yujue Wang[1]

[1]School of mining engineering, North China University of Science and Technology, Tangshan, Hebei, China
[2]Bureau of Natural Resources and Planning of Tangshan, Tangshan, Hebei, China

Received: 8 February 2023
Accepted: 9 December 2023

Abstract

In this paper, a method of heavy metal chromium estimation was proposed, which can be used for soil heavy metal pollution evaluation in Anshan-type iron tailings. The content and spectral characteristics of chromium in 43 field samples were determined, and the optimal characteristic spectrum and model were obtained.

Our research shows that SPA selects 18 band points of the full spectrum, RMSEP is 6.2, RPD is 2.5, and the SPA method for heavy metal chromium had the best prediction performance. Moreover, the inversion model obtained via successive projections algorithm-multiple linear regression (SPA-MLR) was found to be the best inversion model. Collectively, the findings of the present study provide a basis for the estimation of soil heavy metal content in the Anshan tailings area.

Keywords: Tangshan, chromium content, successive projections algorithm

Introduction

Iron ore tailings are the main waste product from iron ore mining, and these waste materials are often stored in the open air. Many studies have demonstrated that this crude storage method leads to the spread of harmful substances in the tailings into the atmosphere, soil, surface water, groundwater, and other environments through external forces such as wind and rain, and produces erosive pollution in the surrounding environment [1-4]. Heavy metals are elements with an atomic density greater than 5.0 g/cm^3, such as Cr, Cu, Cd, Pb, Hg and Zn [5, 6]. Among them, chromium is one of the three heavy metal elements recognized by the international community as carcinogenic, and at the same time, chromium is very difficult to migrate from the soil, thus making chromium one of the priority pollutants for soil monitoring.

After years of stockpiling, heavy metals in tailings are spread by water and wind erosion, thus directly threatening the surrounding water and soil environment. Additionally, these heavy metals accumulate in animals, plants, and humans [7, 8], meaning that the heavy metal contents in tailings and surrounding soils pose a direct threat to human and environmental health [9].

Therefore, there is a pressing need to develop novel methods to quickly and accurately estimate heavy metal chromium contents in tailing ponds and surrounding soils. In turn, this would enable the establishment of more effective ecological protection and mine management strategies. Conventional methods for the quantification of soil heavy metal chromium content via field

*e-mail: 1246910071@qq.com
Tel.: +86-315-13703383116;
Fax: +86-315-8805053.

sampling and laboratory analysis are highly accurate. However, these methods are time-consuming and labor-intensive and are not well suited for the monitoring and assessment of soil heavy metal chromium pollution in large areas, which limits their wide applicability [10]. In contrast, the rapid development of hyperspectral technology provides a new way to conduct quantitative analysis of chromium based on the non-destructive, dynamic, macroscopic, and efficient acquisition of surface information [11]. Hyperspectral can obtain rich spectral information, which can enhance the spectral response characteristics. In addition, the continuous improvement of band selection algorithms and the application of different modeling methods are reported in many research results on soil heavy metal estimation. For example, in a study of arsenic concentration in the soil of an open-pit coal mine, a three-band spectral index was constructed by combining the spectral information and soil clay minerals, which led to the enhancement of arsenic estimation accuracy with the highest correlation coefficient and the lowest RMSE (r = 0.9732, RMSE = 0.0703) [12]. In another study, it was proposed that algorithms such as wavelet transform and random forest can effectively improve the inversion accuracy [13, 14], and all of these results show the applicability of hyperspectral technology in the evaluation of soil heavy metal pollution [15, 16]. However, the sources and distribution of heavy metals in tailing areas are very complex [17].

This study analyzed the Sijiaying iron tailings reservoir in Luanzhou City, Hebei Province, and the spectral information of the soil was collected in the laboratory based on field sampling. Moreover, the amount of chromium in the soil samples was determined using inductively coupled plasma-mass spectrometry (ICP-MS). The inverse model of soil chromium content was established by using the partial least squares regression and multiple linear regression methods, and the accuracy of these approaches was verified and analyzed to confirm whether they could be applied for the rapid monitoring of soil heavy metal pollution in high-silica Anshan-type iron tailings.

Experimental

Sample Data

This study analyzed soil samples from the Sijiaying iron tailings pond and the surrounding area in Luanshi City, which is an Anshan-type sedimentary metamorphic iron ore deposit [18]. According to the geographical distribution characteristics of the tailing ponds in the study area, the sample points were taken approximately 10-20 cm below the surface layer. In a unit area, soil samples were collected according to the east, west, south, north, and central points, and then these 5 soil samples were mixed into 1 combined sample to obtain a total of 43 combined samples. The soil samples collected in the field were cleaned of roots, branches, leaves, stones, etc., and then put into glassware and dried in a dry and ventilated environment. After 24 h drying, each sample was weighed at 500 g, put into a triturator for grinding, and then filtered through a sieve with a diameter of 0.15 mm after grinding. Each soil sample is divided into two parts: one part is used to determine the chromium content, and the other part is used for spectral measurement.

The 4 mg sample was weighed by an analytical balance and placed in a polytetrafluoroethylene digestion tube. 3 mL nitric acid, 3 mL hydrochloric acid, and 0.5 mL hydrofluoric acid were dripped into the sample digestion tube and heated for digestion by an electric heating plate. After digestion, the digestion tube was naturally cooled to room temperature, and the digested samples were filtered and diluted with deionized water into a 10 mL volumetric bottle and mixed evenly. The heavy metal chromium content was determined by an inductively coupled plasma mass spectrometer (ICP-MS), in which each sample was measured twice in parallel, and the average value was selected as the final content.

Spectrum Collection and Measurement

Spectral profiles of soil samples were collected using an ASD hyperspectrometer at a spectral range of 350-2500 nm. Spectral data were obtained in a dark room by placing the soil samples in a 10 cm * 10 cm * 2 cm black box and placing them flat on a horizontal table with a black cloth background. Measurements were made with the probe touching the soil sample, and each sample was measured five times. The actual reflectance spectral curve for each sample was plotted using the average of the measured spectral curves. The inverse transformation of the spectral reflectance enhances the spectral differences in the visible region. Moreover, to eliminate baseline drift and background interference effects and improve the spectral resolution, the inverse transformed spectra were preprocessed using the second-order Savitzky-Golay (S-G) derivative method with a window of 5 [19].

Spectral Feature Preference

Based on the analysis of the response characteristics of soil reflectance to heavy metal content, the characteristic bands were selected using three methods: successive projections algorithm (SPA), competitive adaptive reweighted sampling (CARS), and correlation analysis. Sensitive bands are often determined by correlation analysis between soil heavy metal content and spectral reflectance, the higher the correlation, the more sensitive the band is. Accordingly, the sensitive band reflectance extracted from the correlation analysis and the soil heavy metal chromium content are subjected to stepwise and partial least squares regression analyses, and the inverse model for the heavy metal chromium

Table 1. Comparison of different band selection methods.

Modeling	Number of wavelengths	R_C	RMSEC	R_P	RMSEP	RPD
Full-partial least squares (FULL-PLS)	1075	0.9	4.5	0.2	14.9	1.0
Successive projections algorithm-multiple linear regression (SPA-MLR)	18	1.0	4.7	0.9	6.2	2.5
Competitive adaptive reweighted sampling-multiple linear regression (CARS-MLR)	8	0.3	13.5	0.6	15.8	1.0
Correlation analysis-partial least squares (CA-PLS)	70	0.8	5.8	0.8	7.6	2.0

content in the soil is established. The correlation is usually denoted by a letter and is used to measure the linear relationship between two variables [20].

$$r(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var[X]Var[Y]}} \quad (1)$$

where *Cov(X,Y)* is the covariance of *X* and *Y, Var[X]* is the variance of *X,* and *Var[Y]* is the variance of *Y.*

Results and Discussion

This study uses multiple linear regression (MLR) or partial least squares regression (PLSR) to model the data obtained after processing using three characteristic band selection algorithms, after which their prediction accuracy was compared. The full spectrum has 1075 bands and the correlation analysis extraction 70 band points. The characteristic band selection of the spectral data is performed using the CARS algorithm

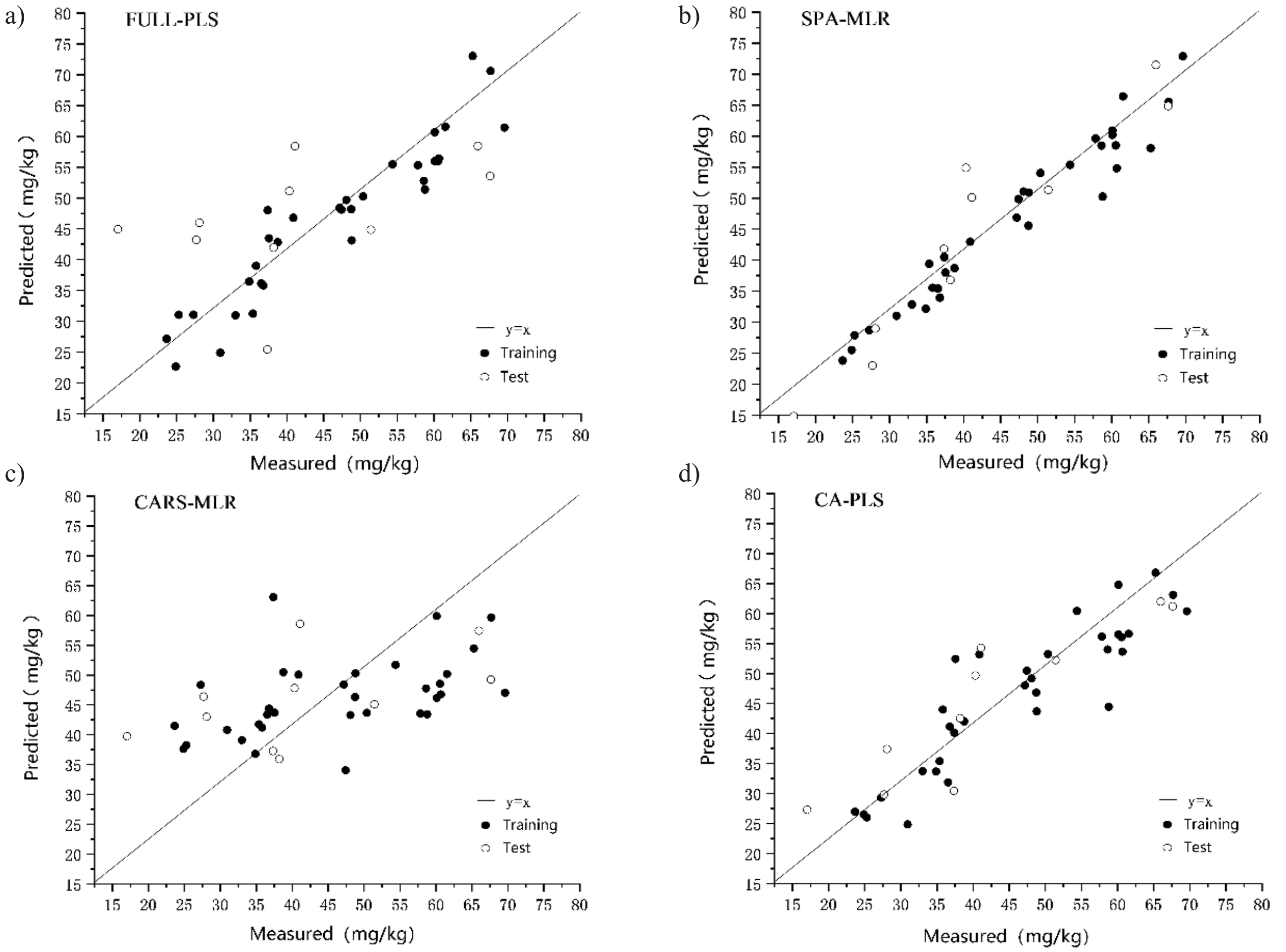

Fig. 1 Shows the effects of the inversion of different models. a) FULL-PLS inversion effect, b) SPA-MLR inversion effect, c) CARS-MLR inversion effect, d) CA-PLS inversion effect.

with 100 Monte Carlo (MC) sampling times and 8 characteristic bands. The feature bands were selected using the SPA method, at which time the number of feature bands was selected to be 18. The MLR model was constructed for the 18 feature band points extracted by SPA and the 8 feature band points extracted by CARS, and the RC correction, prediction correlation coefficient (RMSEC), RP correction, root mean square error of prediction (RMSEP), and RPD values of the model are shown in Table 1. Upon comparing the prediction performances of the three feature band selection algorithms, our findings indicated that SPA selects only 18 feature band points with an RMSEP of 6.2 and an RPD of 2.5, and therefore this approach was considered to have the best prediction performance. SPA not only reduces the complexity of modeling but also greatly improves the prediction accuracy of the model [21].

Fig. 1 shows the correlation plots between the measured and predicted values of soil chromium content for different models for the prediction sample sets. The oblique line in the figure is the trend line y = x, and the model fitting effect is mainly reflected by the distance between the scatter point and the trend line. The closer the distance, the higher the precision. As can be seen from Fig. 1b), only a small number of samples are far from the trend line, and most of the dispersion points are close to the trend line. Therefore, the SPA-MLR model of the chromium element has the best fitting effect.

Conclusions

This study measured the soil spectral information of the Tangshan iron tailing area by hyperspectral technology, and inverse modeling of soil chromium content in the tailing area was conducted using three characteristic waveband extraction methods combined with multiple linear regression and partial least squares regression. The following are the main conclusions of this study:

(1) The model established using the SPA-MLR method coupled with hyperspectroscopy was superior to the other models and could thus be used to accurately estimate the chromium content in soil samples collected from the Tangshan iron tailing area.

(2) The model constructed based on the SPA-MLR method exhibited an RC of 0.9, an RMSEC of 4.7, an RP of 0.9, an RMSEP of 6.2, and an RPD value of 2.5, which is better than the other models in terms of prediction accuracy.

(3) The number of samples directly determines the model accuracy. Our experiment only contained 43 samples, which cannot effectively express the relationship between the number of samples and the model estimation accuracy. Therefore, future studies will assess the degree of influence of the number of samples on model accuracy and root mean square error.

Acknowledgments

Hebei Provincial Education Department Youth Fund (QN2020149); Project of Tangshan Science and Technology Bureau (21130204C); The Cultivation Base of Shanxi Key Laboratory of Mining Area Ecological and Solid Wastes Utilization, Shanxi Institute of Technology (SXIT202304).

Conflict of Interest

The authors declare no conflict of interest.

References

1. BAI L., WANG J.Z., MA X.J., LU H.Y. Air Pollution Forecasts: An Overview. International Journal of Environmental Research and Public Health. **15** (4), 780, **2018**.
2. KŐRDEL W., RŐMBKE J. Requirements on physical, chemical and biological testing methods for estimating the quality of soils and soil substrates. Journal of Soils and Sediments, **1** (2), 98, **2001**.
3. CREAMER R.E., BAREL J.M., BONGIORNO G., ZWETSLOOT M.J. The life of soils: Integrating the who and how of multifunctionality. Soil Biology and Biochemistry, **166**, 108561, **2022**.
4. ZUBAIR M., ANWAR U., ASHFAQ M., ZAFAR M.N., FARID M., AHMAD F., AHMAD W., ALI S., RIZWAN M., ALSAHLI A.A., ALYEMENI M.N., WIJAYA L. Heavy Metals Occurrence, Seasonal Variation and Enrichment in Urban Soils Augmented with Industrial Waste. Polish Journal of Environmental Studies, **30** (5 Pt.2), 4873, **2021**.
5. WUANA R.A., OKIEIMEN F.E. Heavy Metals in Contaminated Soils: A Review of Sources, Chemistry, Risks and Best Available Strategies for Remediation. Isrn Ecology, **2011**, 1, **2011**.
6. BHAT U.N., KHAN A.B. Heavy Metals: An Ambiguous Category of Inorganic Contaminants, Nutrients and Toxins. Research Journal of Environmental Sciences, **5** (8), 682, **2011**.
7. REHMAN A.U., NAZIR S., IRSHAD R., TAHIR K., WAHAB Z. Toxicity of heavy metals in plants and animals and their uptake by magnetic iron oxide nanoparticles. Journal of Molecular Liquids, **321** (1), 114455, **2020**.
8. BORGULAT J., METRAK M., STASZEWSKI T., WILKOMIRSKI B., SUSKA-MALAWSKA M. Heavy Metals Accumulation in Soil and Plants of Polish Peat Bogs. Polish Journal of Environmental Studies, **27** (2), 1, **2018**.
9. REY J., MARTÍNEZ J., HIDALGO M.C., ROJAS D. Heavy metal pollution in the Quaternary Garza basin: A multidisciplinary study of the environmental risks posed by mining (Linares, southern Spain). CATENA, **110**, 234, **2013**.
10. XU D.M., FU R.B., LIU H.Q., GUO X.P. Current knowledge from heavy metal pollution in Chinese smelter contaminated soils, health risk implications and associated remediation progress in recent decades: A critical review. Journal of Cleaner Production, **286**, 124989, **2021**.

11. PASQUINI C. Near infrared spectroscopy: A mature analytical technique with new perspectives-A review. Analytica Chimica Acta, **1026**, 8, **2018**.
12. FU P.J., YANG K.M., MENG F., ZHANG W., CUI Y., FENG F.S., YAO G.B. A new three-band spectral and metal element index for estimating soil arsenic content around the mining area. Process Safety and Environmental Protection, **157**, 27, **2022**.
13. DEPIN O., KUN T., WANG X., WU Z.F., LI J., DING J.W. Modified soil scattering coefficients for organic matter inversion based on Kubelka-Munk theory. Geoderma, **418**, 115845, **2022**.
14. GU X.H., WANG Y.C., SUN Q, YANG G.J., ZHANG C. Hyperspectral inversion of soil organic matter content in cultivated land based on wavelet transform. Computers and Electronics in Agriculture, **167**, 105053, **2019**.
15. ZHANG B., GUO B., ZOU B., WEI W., LEI Y.Z., LI T.Q. Retrieving soil heavy metals concentrations based on GaoFen-5 hyperspectral satellite image at an opencast coal mine, Inner Mongolia, China. Environmental Pollution, **300**, 118981, **2022**.
16. WANG Y.B., ZHANG X., SUN W.C., WANG J.N., DING S.T., LIU S.H. Effects of hyperspectral data with different spectral resolutions on the estimation of soil heavy metal content: From ground-based and airborne data to satellite-simulated data. Science of The Total Environment, **838**, 156129, **2022**.
17. GUO Q., CAO S., NAN L., ZHAN Z., WANG Z. Distribution Characteristics and Ecological Risk Assessment of Heavy Metals in Typical Farmland Soils from Baijiazui Village of Ningyuanbu Town, China. Polish Journal of Environmental Studies, **31** (4), 147056, **2022**.
18. SUN X.H., TANG H.S., LUAN Y., CHEN J.H., Geochronological constraints on the genesis of high-grade iron ore in the Gongchangling BIFs from the Anshan-Benxi area, North China Craton, Ore Geology Reviews, **122**, 103504, **2020**.
19. WATERLOT C., DOUAY F. Minimizing matrix effects and spectral interferences produced by Fe absorption lines in the determination of cadmium by electrothermal atomic absorption spectrometry: Application to the fractionation of cadmium in moderated contaminated soils. Measurement, **46** (8), 2348, **2013**.
20. RAHEEM A.M., OMAR N.Q. Investigation of distinctive physico-chemical soil correlations for Kirkuk city using spatial analysis technique incorporated with statistical modeling. International Journal of Geo-Engineering, **12** (1), 18, **2021**.
21. CEVOLI C., CECILIA L.D., FERRARI L., FABBRI A., MOLARI G. Evaluation of cut alfalfa moisture content and operative conditions by hyperspectral imaging combined with chemometric tools: In-field application. Biosystems Engineering, **222**, 132, **2022**.