Optimal bandwidth selection for retrieving Cu content in rock based on hyperspectral remote sensing Post-print

Authors: MA Xiumei, Kefa Zhou, WANG Jinlin, CUI Shichao, ZHOU Shuguang, Shanshan Wang, ZHANG Guanbin, Jinlin Wang

Date: 2022-01-30T00:00:00+00:00

Abstract

Hyperspectral remote sensing technology is widely used to detect element contents because of its multiple bands, high resolution, and abundant information. Although researchers have paid considerable attention to selecting the optimal bandwidth for the hyperspectral inversion of metal element contents in rocks, the influence of bandwidth on the inversion accuracy are ignored. In this study, we collected 258 rock samples in and near the Kalatage polymetallic ore concentration area in the southwestern part of Hami City, Xinjiang Uygur Autonomous Region, China and measured the ground spectra of these samples. The original spectra were resampled with different bandwidths. A Partial Least Squares Regression (PLSR) model was used to invert Cu contents of rock samples and then the influence of different bandwidths on Cu content inversion accuracy was explored. According to the results, the PLSR model obtains the highest Cu content inversion accuracy at a bandwidth of 35 nm, with the model determination coefficient (R2) of 0.5907. The PLSR inversion accuracy is relatively unaffected by the bandwidth within 5–80 nm, but the accuracy decreases significantly at 85 nm bandwidth (R2=0.5473), and the accuracy gradually decreased at bandwidths beyond 85 nm. Hence, bandwidth has a certain impact on the inversion accuracy of Cu content in rocks using the PLSR model. This study provides an indicator argument and theoretical basis for the future design of hyperspectral sensors for rock geochemistry.

Full Text

Preamble

Optimal Bandwidth Selection for Retrieving Cu Content in Rock Based on Hyperspectral Remote Sensing

MA Xiumei[1,2,3,4], ZHOU Kefa[1,2,3,4], WANG Jinlin[1,2,3,4]*, CUI Shichao[1,2,3,4], ZHOU Shuguang[1,2,3,4], WANG Shanshan[1,2,3,4], ZHANG Guanbin[5]

[1] State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China
[2] Xinjiang Key Laboratory of Mineral Resources and Digital Geology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China
[3] Xinjiang Research Centre for Mineral Resources, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China
[4] Xinjiang Academy of Science and Technology for Development Strategy, Urumqi 830011, China
[5] [Affiliation not provided in original text]

Abstract: Hyperspectral remote sensing technology is widely used to detect element contents due to its multiple bands, high resolution, and abundant information. Although researchers have paid considerable attention to selecting optimal bandwidths for the hyperspectral inversion of metal element contents in rocks, the influence of bandwidth on inversion accuracy has often been ignored. In this study, we collected 258 rock samples in and near the Kalatage polymetallic ore concentration area in the southwestern part of Hami City, Xinjiang Uygur Autonomous Region, China, and measured the ground spectra of these samples. The original spectra were resampled with different bandwidths. A Partial Least Squares Regression (PLSR) model was used to invert Cu contents of rock samples, and the influence of different bandwidths on Cu content inversion accuracy was explored. According to the results, the PLSR model achieves the highest Cu content inversion accuracy at a bandwidth of 35 nm, with a model determination coefficient (R^2) of 0.5907. The PLSR inversion accuracy is relatively unaffected by bandwidth within 5–80 nm, but the accuracy decreases significantly at 85 nm bandwidth (R^2=0.5473), and gradually declines at bandwidths beyond 85 nm. Hence, bandwidth has a certain impact on the inversion accuracy of Cu content in rocks using the PLSR model. This study provides an indicator argument and theoretical basis for the future design of hyperspectral sensors for rock geochemistry.

Keywords: hyperspectral remote sensing; Cu element; bandwidth; Partial Least Squares Regression; inversion accuracy; Kalatage polymetallic ore concentration area

Citation: MA Xiumei, ZHOU Kefa, WANG Jinlin, CUI Shichao, ZHOU Shuguang, WANG Shanshan, ZHANG Guanbin. 2022. Optimal bandwidth selection for retrieving Cu content in rock based on hyperspectral remote sensing. Journal of Arid Land, 14(1): 102–114. https://doi.org/10.1007/s40333-022-0050-8

1 Introduction

Mineral resources play significant roles in national economic development and guarantee national security (Xu and Wang, 2011). The scarcity of mineral resources is a serious problem faced by many countries around the world, especially China. With increasing difficulty and cost of prospecting, the development of shallow deposits has been abandoned, and efforts have gradually shifted to the exploration of concealed deposits (Cao et al., 2009).

At present, many methods are employed to detect hidden deposits, including model prospecting prediction methods, physical detection methods, and chemical detection methods. Rock geochemistry, which combines petrology with geochemistry, is a chemical detection method that determines anomalous regions of ore-forming elements by detecting the contents of major elements and trace elements in rocks, thereby providing a basis for prospecting. However, traditional rock geochemistry has disadvantages of being time-consuming, labor-intensive, and costly, and is limited to point data measurements (Hecker et al., 2019). Nevertheless, with continuous development of high-resolution remote sensing technology, it has been widely applied to mineralization prediction (Ge, 2017). Hunt (1977) concluded that elements with characteristic absorption bands are limited in the visible-near-infrared (VNIR) spectral range, but include the most common mineralization indicator elements in metal mineral exploration, such as Fe, Cu, Mn, Cr, and Ni. This finding established the theoretical basis of quantitative inversion models for ore-forming elements based on reflectance spectroscopy. Due to its multiband nature, high resolution, and strong continuity, hyperspectral technology can quickly acquire spectral information of features without destroying samples, thus providing a powerful method for quantitative and semiquantitative prediction of ore-forming elements.

In recent years, considerable progress has been made in element content estimation models based on high spectral resolution data and element content data. Although current studies on estimation models for ore-forming elements remain in their beginning and exploratory stages, some scholars have leveraged successes from soil spectral reflectance data to estimate heavy metal element contents (Ben-Dor and Banin, 1994; Ben-Dor, 2000; Islam et al., 2003; Wu et al., 2007; Cheng et al., 2018). Hyperspectral data have many bands and high redundancy, and the Partial Least Squares Regression (PLSR) method can be applied to continuous spectra with many bands and serious autocorrelation, effectively improving quantitative inversion accuracy of metal elements (Wang et al., 2007). Based on the PLSR method, many researchers (e.g., Kemper and Sommer, 2002; Gong et al., 2010; Liu et al., 2010; Cong et al., 2013; Yang et al., 2017) have used characteristic bands to invert metal element content, focusing mainly on band position and spectral index while considering how to select optimal bands. However, these studies ignored the effect of bandwidth on inversion accuracy. Different types of hyperspectral data have bands with different centre positions, and the number of bands—especially their bandwidth—has significant influence on various hyperspectral applications (Wellburn, 1994; Zarco-Tejada

et al., 2001). Currently, hyperspectral data include terrestrial non-imaging hyperspectral data and imaging hyperspectral data acquired by ground-, aircraft-, and satellite-mounted sensors. Bandwidth of satellite platforms (such as Hyperion, Huan Jing (HJ), and other hyperspectral sensors) is 10 nm, whereas for airborne platforms (such as Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), Compact Airborne Spectrographic Imager (CASI), and Eagle II), bandwidth can range from 2 to 10 nm, and the Analytical Spectral Devices (ASD) spectrometer has bandwidth up to 1 nm on ground-mounted platforms. With continuous development of sensor technology, parameters of hyperspectral sensors have greatly improved, such as number of bands, band range, and bandwidth. However, changes in these parameters have yielded increases in data volume, storage burden, and computational complexity. Thus, it is imperative to choose appropriate bandwidth to accurately invert metal element content.

The effect of bandwidth on inversion accuracy of metal element content remains poorly understood. Huang et al. (2010) found that the optimal number of bands was 10 and the optimal bandwidth was 32 nm when predicting Cu content, and that narrowing bandwidth did not guarantee better results. In addition, the best inversion accuracy could be obtained by the PLSR method. However, previous studies focused only on selecting optimal bandwidth for suburban soils with relatively few samples, large spectral sampling intervals, and coarse results. For rock ore, Cong et al. (2013) established a Cu content inversion model with bandwidth of 5 nm and wavelength of 900 nm, but the effect was not significant.

Although the influence of soil bandwidth on inversion effect has been roughly recognized when using spectral techniques to invert metal element content, the influence of bandwidth on the model was ignored for rocks. Additionally, large areas of original rocks are covered by wind-blown sediment due to environmental conditions, making large-scale geochemical exploration more difficult. Therefore, the Kalatage polymetallic ore concentration area, where similar work has not been carried out, was selected as the study area. This study establishes a PLSR model for rocks based on a small spectral sampling interval and compares and analyzes the effects of different bandwidths on inversion results of Cu content in rocks. Combined with hyperspectral remote sensing technology, this study provides a basis for simply and effectively predicting Cu content in rocks, establishes a foundation for inversion of regional remote sensing metal element contents across large areas, and provides a theoretical basis and demonstration for future design of hyperspectral sensors for rock geochemistry. These contributions will help promote geochemical exploration development in large arid areas such as the Gobi Desert.

2.1 Study Area

The study area is located in the Kalatage polymetallic ore concentration area in the southwestern part of Hami City, Xinjiang Uygur Autonomous Region,

China. The Kalatage polymetallic ore collection area was discovered in 1999, and is 180 km east of Hami City and 50 km north of Kuquan Station of the Lanxin Railway. The geographical coordinates of the Kalatage polymetallic ore concentration area are between 42°35'–42°42' N and 91°30'–92°05' E, covering an area of 700 km^2. Many geologists have carried out various studies on this area, including its metallogenic background, metallogenesis, and metallogenic regularity. The Early Paleozoic volcanogenic massive sulphide deposit of the Red Sea type was first discovered in northern Xinjiang in 2008, which had important geological significance (Mao et al., 2017). In 2011–2012, the Yudai porphyry Cu (Au) deposit was discovered by geological prospectors in the western section of the Kalatage polymetallic ore concentration area. The Yudai porphyry Cu (Au) collection area is located in the south of the Turpan-Hami (Tuha) Basin in the northwestern part of the Kalatage uplift, which is a Paleozoic tectonic uplift belt along the margin (Long et al., 2016; Yu et al., 2016). The northern and eastern parts of the mining area are Dananhu Formation (D1d) Clastic Sedimentary Strata [Figure 1: see original paper]. These clastic sedimentary rocks unconformably cover a set of volcanic clastic rocks, mainly consisting of conglomerate, tuff sandstone, and powdery sandstone. The southern mine is composed of brown and aubergine andesites, basaltic andesites, dacites, andesitic-dacite agglomerates, grey basalts, pyroclastic sedimentary rocks, volcanic breccia, and tuff (Mao et al., 2017).

2.2.1 Sample Collection

Since most of the study area is covered by wind-blown sand and the surface sand is poorly representative, this study collected rock samples in situ with little evidence of weathering and erosion in the vicinity of the Yudai porphyry Cu (Au) deposit as experimental samples in April 2018. Thirteen sampling lines were established; specifically, each sampling line was approximately 5 km long, the sampling line spacing was approximately 500 m, the point distance was approximately 200 m, and 3–5 pieces of rock debris with a diameter of 10 cm or less were collected at each point [Figure 2: see original paper]. A total of 338 rock samples were obtained. The collected samples were ground through a 200-mesh sieve (mesh diameter of 0.074 mm) to remove contaminated samples and small samples, and finally, 258 effective samples were obtained. The analyzed samples were directly used for outdoor spectral measurements, and the sample state was consistent, which is beneficial for comparative analysis.

2.2.2 Data Acquisition

The field test was performed using a portable field spectrometer (FieldSpec 4 Hi-Res, LICA United Technology Limited, Beijing, China) for spectral data acquisition. The FieldSpec 4 Hi-Res instrument is a general-purpose spectroradiometer

useful for many applications where measurements of reflectance, transmittance, radiance, or irradiance are required; it is a special type of spectrometer that can measure radiant energy (radiance and irradiance). The instrument, which is compact, portable, and precise, is specifically designed for remote sensing in a field environment to acquire VNIR and shortwave infrared (SWIR) spectra. We obtained the above information according to the manufacturer' s user manual (https://www.malvernpanalytical.com/en/).

The spectral sampling interval was 1 nm in standard mode, the spectral acquisition time was 0.2 s, the spectral resolution was 3 or 8 nm, and the wavelength range was 350–2500 nm, for a total of 2151 bands. To use a natural light source for outdoor measurements while avoiding interference from other sources, we placed the samples on a black background cloth. We further flattened the surface of the black background cloth while loading the sample to reduce the influence of surface unevenness on the spectrum. Distance between the high-density dedicated probe and the rock powder sample used for ASD spectrometer measurement was controlled at approximately 2 cm. Before acquiring measurements, a whiteboard was used to calibrate the instrument, and then sample measurement was started. For every five samples, a whiteboard calibration was performed to ensure accuracy. A total of 258 rock powder samples were measured at a daily solar elevation angle close to vertical.

The measured spectrum corresponded to the reflectance value, and each sample was measured five times to obtain five spectral curves; finally, five spectral curves were averaged to obtain a spectral curve. Data extraction for drawing and text files was performed by ViewSpecPro version 5.6 software (LICA United Technology Limited, Beijing, China).

2.3.1 Raw Data Correction

The measured spectral reflectance curves were automatically corrected by ASD spectrometer software ViewSpecPro. The main steps included gain calibration of each spectral data channel, wavelength correction, and dark current removal from the original data radiation intensity (Liu, 2002).

2.3.2 Averaging Spectral Reflectance

The five spectral reflectance curves of each sample obtained by ViewSpecPro software were averaged, and the average value was taken as the spectrum of the sample. The measurement instability was evaluated by multiple measurements of the same target, making the measurements more natural and reliable while also enhancing the signal-to-noise ratio.

2.3.3 Removing the Water Vapour Absorption Band

Within wavelengths of 350–399, 1301–1399, 1790–1999, and 2401–2500 nm, water vapour has a significant absorption effect on the spectrum, leading to anomalies [Figure 3: see original paper]. We deleted some bands to eliminate the influence of water vapour absorption on element content inversion, leaving 1692 effective bands.

2.4 Spectral Resampling of Different Bandwidths

Bandwidth, also known as spectral resolution, refers to the recording width of the detector in the wavelength direction and is strictly defined as the wavelength width at which the instrument reaches 50% of the maximum spectral response (Dian et al., 2016). The finer the spectral resolution, the greater the number of bands; and the higher the resolution, the narrower the band. After deleting water vapour absorption bands, we studied the influence of bandwidth on inversion precision. Bandwidth was set as a variable while other factors remained unchanged. The original band was resampled to spectral intervals of 5, 10, 15, ⋯, 100, 150, and 200 nm.

Resampling was applied to downsample the high-resolution reflectivity to lower resolution. The original VNIR bandwidth was 3 nm, and the original SWIR bandwidth was 8 nm. The spectrometer output was resampled to a spectral band of 1 nm. On this basis, we resampled the original spectrum using a custom spectral response function. To resample the bandwidth to 5 nm, we resampled the original data using a 5 nm custom spectral response function in an Excel spreadsheet and ENVI 5.3 (Exelis Visual Information Solutions, Broomfield, Colorado, USA) [Figure 4: see original paper]. The number of bands was 339 after resampling. Finally, spectral responses of 10, 15, 20, 25, ⋯, 100, 150, and 200 nm were prepared, and the original spectral reflectance values were resampled using the corresponding spectral response functions. The numbers of bands after resampling were 169, 113, 85, 68, ⋯, 17, 12, and 9, respectively.

2.5 PLSR Model and Its Inversion

PLSR is a multivariate statistical analysis method first proposed by Svante Wold and Christer Albano (Wold et al., 1984). The PLSR method has been extensively used in different fields (Kawamura et al., 2010; Wang et al., 2010) because it can solve not only regression problems involving multiple dependent and independent variables but also problems of variable multicollinearity (Liu et al., 2019). The sum of squares of errors was used to minimize the best matching function for a set of data in the PLSR method. Here, the simplest method was used to find the true value that minimizes the sum of squares of errors. By

analogy, the optimal number of principal components was obtained using leave-one-out cross validation (LOOCV) method. The multivariate correction step in the PLSR method was employed to directly locate the target at prediction. Therefore, based on the principle and basis of minimum prediction error sum of squares (PRESS), we determined the number of extracted principal components. The principal components (t) were a linear combination of x, and thus a regression model of y and x could be established. Variables related to the original parameters could be transformed into a set of independent orthogonal variables after conducting PLSR. In many practical applications, the orthogonality of the variable system has many advantages because the information contained in each variable is complementary. Each variable contains information without crossover, which is highly convenient for statistical information (Wang et al., 2006).

3.1 PLSR Model Accuracy of Different Bandwidths

To establish the PLSR model and obtain a reliable and stable model, we adopted the LOOCV method to prevent overfitting caused by an excessive number of principal components. In the LOOCV method, one sample was left as a verification sample while the rest were used as training samples; this step was continually cycled until all samples were verified. The PLSR model was built with the training samples and predicted by the verification sample, and the PRESS was recorded until all samples were predicted (only once). The LOOCV method was used to detect the PRESS of each principal component, and the number of principal components with minimum PRESS was selected. Values of PRESS corresponding to the number of different principal components detected at a bandwidth of 35 nm are shown in Figure 5. The PLSR method directly locates the target of multivariate correction on prediction, and thus the principle of determining the number of principal components in the PLSR model is to minimize prediction error. The prediction error value is minimum when eight principal components are extracted; therefore, eight principal components were determined [Figure 5: see original paper].

To investigate the optimal bandwidth required to predict Cu content in rocks, this study applied the PLSR model for the above sample sets at all resampled resolutions. The regression results of the PLSR model are shown in Table 1. It is clear that the accuracy of the PLSR model reaches the highest when the bandwidth is 35 nm. Figure 6 shows the accuracy of the PLSR model on Cu content prediction with increasing bandwidth, displaying high volatility of the curve and an overall decreasing trend. The spectral resolution within the resampled bandwidth range is 35 nm, corresponding to the highest R^2 value. As bandwidth gradually increases, the accuracy of the PLSR model changes. Specifically, the accuracy changes slightly before 35 nm but irregularly after 35 nm and drops sharply after 85 nm. This also means that metallic elements are resolvable at high spectral resolution, whereas multispectral accuracy is low.

The specific PLSR model analysis results for Cu content prediction under different bandwidths included the extracted optimal number of principal components and the PRESS, in addition to the inversion model determination coefficient R^2. The smaller the PRESS value, the smaller the fitting error (Table 1). The number of principal components extracted from the PLSR model is determined by the fitting error; when the number is eight, the resolution is the highest. The number of principal components higher or lower than eight will decrease the prediction accuracy of the PLSR model.

When Cu content within the 258 rock samples in the mining area was inspected by the VNIR ground-measured spectrum, the influence of bandwidth on spectral reflectance data was in a small range (Table 1; Fig. 6). Specifically, at 85 nm bandwidth, R^2 value decreased considerably; at 100 nm bandwidth, the accuracy increased again, but the overall trend continued to decline; and at 150 and 200 nm bandwidth, R^2 values were close and continued to decline. Thus, for Cu content detection in rocks, it is not true that the smaller the spectral bandwidth, the better the prediction result of the PLSR method. Likewise, it is also not true that the larger the spectral bandwidth, the better the prediction result. During 5–80 nm bandwidth, the inversion prediction accuracy exhibited small amplitude fluctuation, but the best effect was found at 35 nm within the resampled bandwidth range [Figure 6: see original paper].

3.2 Cu Content Prediction by the PLSR Model

We applied the PLSR model based on different bandwidths obtained from resampled original data. Specifically, we established the PLSR model for sample data with a bandwidth of 35 nm and determined the optimal number of principal components as eight according to LOOCV method accuracy. In the PLSR calculation process, the extracted independent variable tn represents as much variation in x as possible and is associated with y as much as possible to interpret y information. Variable t1 had the strongest interpretation ability. In other words, the multivariate correction step in the PLSR method was used to directly locate the target on prediction. Based on the principle and basis of PRESS, the principal components were extracted.

Based on the optimal number of principal components determined above, we predicted Cu content in each experimental sample using the established PLSR prediction model. There are two ways to predict Cu content: one is the combination of PLSR model and LOOCV method, the other is direct fitting prediction using the PLSR model, hereinafter referred to as LOOCV prediction and partial fitting prediction [Figure 7: see original paper]. These predicted contents and Cu contents measured by chemical analysis were then fitted [Figure 8: see original paper]. Prediction accuracy is expressed by the determination coefficient R^2 of the PLSR model. The larger the determination coefficient R^2, the better the prediction effect. Cu element content predicted by partial fitting clustered

around measured values without large outliers. The linear equation and fitting accuracy are shown in Figure 8, and the bias of this bandwidth (35 nm) to fitted prediction values is depicted in Figure 9. Unlike Figure 8, Figure 9 shows fitting results using LOOCV. It can be seen from the figure that accuracy of results fitted after cross-validation is decreased but more convincing. Both figures show similar point distribution.

The distribution trends of measured and predicted Cu contents are generally consistent [Figure 10: see original paper]. However, we need to use spatial interpolation to compare measured and predicted Cu contents. Deterministic local fitting methods (such as inverse distance weighting) are needed to avoid excessive interpolation differences in order to more intuitively visualize spatial distribution differences of Cu content. In this study, inverse distance weighting was applied to interpolate Cu contents retrieved from the PLSR model established at 35 nm bandwidth.

These results indicate that when inverting Cu content in rocks using the PLSR method, result accuracy is affected by bandwidth [Figure 6: see original paper]. During 5–80 nm bandwidth, inversion precision changes are similar, and inversion accuracy reaches the highest at 35 nm bandwidth. Inverse distance weighting interpolation can be performed based on inversion results at 35 nm; although accuracy is not high enough, results can still indicate Cu content distribution.

---

4 Discussion

Cu is a common metallogenic element with certain significance as an indicator for ore prospecting. Hyperspectral remote sensing technology can help invert element content over large areas. Due to high resolution of hyperspectral data, different numbers of bands, wavelengths, and bandwidths correspond to different types of hyperspectral data. Therefore, selecting the optimal bandwidth with the best resolution is very important. In this study, the accuracy of inverting Cu content in rocks with different bandwidths was analyzed.

Results demonstrate that bandwidth influences inversion results. At bandwidth lower than 85 nm, inversion accuracy fluctuates little. Beyond 85 nm bandwidth, however, accuracy drops dramatically. This indicates that inversion results are reliable as long as bandwidth is less than 85 nm when inverting Cu content. At 35 nm bandwidth, accuracy is highest and remains almost the same at smaller bandwidths. However, accuracy changes considerably between 35 and 85 nm [Figure 6: see original paper]. This implies that compared with uncertainty in precision variation between 35 and 85 nm, it is best to invert Cu content at bandwidth below 35 nm. The spectral characteristics of metal elements are relatively broad, and lower spectral resolution can increase signal-to-noise ratio, thereby improving prediction accuracy. This indicates that spectral resolution should be within 80 nm when designing hyperspectral sensors for rock geochem-

istry in the future, if there is no necessity to require excessively high resolution. Instead, in fine instrument design, we can refer to design with spectral resolution near 35 nm, but whether the most appropriate spectral resolution is 35 nm remains to be discussed.

According to R^2 value of the linear fitting equation, R^2 reaches its largest value at 35 nm bandwidth within the resampled bandwidth range. This result is similar to Huang et al. (2010), who inverted heavy metal Cu content in soil. Results from other researchers also show that bandwidth has certain effect on chlorophyll in vegetation (Gitelson et al., 2006; Zhao et al., 2007; Peng et al., 2011; Dian et al., 2016). These findings demonstrate that bandwidth influences ground object inversion, and inversion does not improve as spectral resolution increases. Different types of research objects and inversion elements produce different results, and thus, as selection of optimal bandwidth for metal element content in rocks is still in its preliminary exploration stage, more comprehensive research on various metal elements is needed in future.

In this study, the influence of bandwidth on inversion accuracy of Cu content was discussed. Since reflectance data were obtained by ground-based hyperspectral instruments, the water vapor absorption factor was considered, but other factors involving airborne and satellite hyperspectral data were not considered.

---

5 Conclusions

In this study, we measured ground spectra of rock powder samples collected near the Yudai porphyry Cu (Au) deposit in the Kalatage polymetallic ore concentration area and investigated Cu content inversion accuracy at different bandwidths. Results show that accuracy of metal element content inversion near 80 nm bandwidth is small, but inversion results are superior at 35 nm bandwidth. Inversion model precision shows sharp decline at 85 nm bandwidth. Application of metal element content inversion in soil by reflection spectroscopy has received considerable attention, and related techniques tend to be mature. However, due to heterogeneity of chemical compositions in rocks and numerous rock types, complexity of ground spectra leads to some difficulties in inversion of element contents using reflection spectra. This study provides a basis for future design of hyperspectral sensors for rock geochemistry and provides ideas for further research on rock reflection spectra.

The subject of this study is ground spectra measurement of rock powders in a mining area, and we explore only the influence of bandwidth on the PLSR method. Future studies should consider other driving factors, such as characteristic band and band position. Additionally, aerial hyperspectral remote sensing is the focus and difficulty of future research.

Acknowledgements

This work was supported by the Science and Technology Major Project of Xinjiang Uygur Autonomous Region (2021A03001-3), the Key Area Deployment Project of the Chinese Academy of Sciences (ZDRW-ZS-2020-4-30), and the National Natural Science Foundation of China (U1803117).

References

Ben-Dor E, Banin A. 1994. Visible and near-infrared (0.4–1.1 m) analysis of arid and semi arid soils. Remote Sensing of Environment, 48(3): 261–274.

Ben-Dor E. 2002. Quantitative remote sensing of soil properties. Advances in Agronomy, 75: 173–243.

Cao X Z, Zhang W S, Sun H S. 2009. Progress in the study of deep exploration in China. Geological Science and Technology Information, 28(2): 104–109. (in Chinese)

Cheng H, Wan Y, Chen Y Y, et al. 2018. Study on characteristics and mechanisms of visible near infrared reflectance spectra of soil heavy metals. Spectroscopy and Spectral Analysis, 38(3): 771–778. (in Chinese)

Cong L J, Cen K, Yu X Z, et al. 2013. Relationship between spectral characteristics and geochemical composition of Zhulazhaga gold deposit. Journal of Central South University (Science and Technology), 44(1): 266–274. (in Chinese)

Dian Y Y, Le Y, Fang S H, et al. 2016. Influence of spectral bandwidth and position on chlorophyll content retrieval at leaf and canopy levels. Journal of the Indian Society of Remote Sensing, 44: 583–593.

Ge J M. 2017. Study on prospecting criteria and prediction methods of concealed ore-deposit. Western Resources, 4: 27–28. (in Chinese)

Gitelson A A, Keydan G P, Merzlyak M N. 2006. Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. Geophysical Research Letters, 33(11): L11402, doi: 10.1029/2006GL026457.

Gong S Q, Wang X, Shen R P, et al. 2010. Study on heavy metal content in the coastal saline soil by hyperspectral remote sensing. Remote Sensing Technology and Application, 25(2): 169–177. (in Chinese)

Hecker C, van Ruitenbeek F J A, van der Werff H M A, et al. 2019. Spectral absorption feature analysis for finding ore: a tutorial on using the method in geological remote sensing. IEEE Geoscience and Remote Sensing Magazine, 7(2): 51–71.

Huang C P, Liu B, Zhang X, et al. 2010. Study on band selection and optimal spectral resolution for prediction of Cu contamination in soils. Remote Sensing Technology and Application, 25(3): 353-341, 357. (in Chinese)

Hunt G R. 1977. Spectral signatures of particulate minerals in the visible and near infrared. Geophysics, 42(3): 501-513.

Islam K, Singh B, McBratney A. 2003. Simultaneous estimation of several soil properties by ultra-violet, visible and near-infrared reflectance spectroscopy. Australian Journal of Soil Research, 41: 1101-1114.

Kawamura K, Watanabe N, Sakanoue S, et al. 2010. Testing genetic algorithm as a tool to select relevant wavebands from field hyperspectral data for estimating pasture mass and quality in a mixed sown pasture using partial least squares regression. Grassland Science, 56(4): 205-216.

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

Liu M, Lin Q Z, Wang Q J, et al. 2010. Study on the geochemical anomaly of copper element based on reflectance spectra. Spectroscopy and Spectral Analysis, 30(5): 1320-1323. (in Chinese)

Liu W D. 2002. Study on extraction of soil information and data mining by hyperspectral remote sensing. PhD Dissertation. Beijing: Institute of Remote Sensing Applications Chinese Academy of Sciences. (in Chinese)

Liu Y H, Wang Q Q, Gao X W, et al. 2019. Total phenolic content prediction in Flos Lonicerae using hyperspectral imaging combined with wavelengths selection methods. Journal of Food Process Engineering, 42(6): e13224, doi: 10.1111/jfpe.13224.

Long L L, Wang J B, Wang Y W, et al. 2016. Discussion on the age of ore-host volcanic strata in the Kalatage ore concentration area, eastern Tianshan: Evidence of SHRIMP zircon U-Pb dating. Mineral Exploration, 7(1): 31-37. (in Chinese)

Mao Q G, Wang J B, Fang T H, et al. 2017. Discovery of the Middle Devonian yudai porphyric Cu (Au) deposit in the Kalatage area of eastern Tianshan Mountain, Xinjiang and its geological prospecting significance. Geology and Exploration, 53(1): 1-11. (in Chinese)

Peng Y, Gitelson A A, Keydan G, et al. 2011. Remote estimation of gross primary production in maize and support for a new paradigm based on total crop chlorophyll content. Remote Sensing of Environment, 115(4): 978-989.

Wang H W, Wu Z B, Meng J. 2006. Partial Least-Squares Regression—Liner and Nonliner Methods. Beijing: National Defense Industry Press, 97-98. (in Chinese)

Wang L, Lin Q Z, Jia D, et al. 2007. Study on the prediction of soil heavy metal elements content based on reflectance spectra. Journal of Remote Sensing, 11(6): 906–913. (in Chinese)

Wang W, Peng Y K, Ma W, et al. 2010. Prediction of chlorophyll content of winter wheat using leaf level hyperspectral data. Transactions of the Chinese Society of Agricultural Machinery, 41(5): 172–177. (in Chinese)

Wellburn A R. 1994. The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. Journal of Plant Physiology, 144(3): 307–313.

Wold S, Albano C, Dunn III W J, et al. 1984. Multivariate data analysis in chemistry. In: Kowalski B R. Chemometrics. NATO ASI Series (Series C: Mathematical and Physical Sciences). Dordrecht: Springer, 17–95.

Wu Y Z, Chen J, Ji J F, et al. 2007. A mechanism study of reflectance spectroscopy for investigating heavy metals in soils. Soil Science Society of America Journal, 71(3): 918–926.

Xu Y J, Wang L. 2011. The importance of mineral resources. Science & Technology Communication, (19): 62–83. (in Chinese)

Yang C B, Liu N, Zhou Z C, et al. 2017. Research on the relationship between main rock metal element content, physical parameters and spectral features in Tahe Area. Spectroscopy and Spectral Analysis, 37(8): 2569–2574. (in Chinese)

Yu M J, Wang J B, Mao Q G, et al. 2016. Pyroelectric characteristics of pyrite from the No. M1 orebody in Meiling copper-gold deposit and its geological significance in the Kalatage area, eastern Tianshan. Mineral Exploration, 7(1): 149–156. (in Chinese)

Zarco-Tejada P J, Miller J R, Noland T L, et al. 2001. Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing, 39(7): 1491–1507.

Zhao D H, Huang L M, Li J L, et al. 2007. A comparative analysis of broadband and narrowband derived vegetation indices in predicting LAI and CCD of a cotton canopy. ISPRS Journal of Photogrammetry and Remote Sensing, 62(1): 25–33.

Zhou S G, Liao S B, Zhou K F, et al. 2018. Application of portable X-ray fluorescence spectrometer in the analysis of rock samples. Rock and Mineral Analysis, 37(1): 56–63. (in Chinese)

Note: Figure translations are in progress. See original paper for figures.