Front. Earth Sci. 2019, 13(3): 563–574
https://doi.org/10.1007/s11707-019-0753-0

RESEARCH ARTICLE

Estimation of copper concentration of rocks using hyperspectral technology

Shichao CUI[1,2,3,4], Kefa ZHOU (✉)[1,2,3,4], Rufu DING[5], Guo JIANG[1,2,3,4]

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, Urumqi 830011, China
3 Xinjiang Research Center for Mineral Resources, Chinese Academy of Sciences, Urumqi 830011, China
4 University of the Chinese Academy of Sciences, Beijing 100049, China
5 China Non-Ferrous Metals Resources Geological Survey, Beijing 100012, China

Abstract Rock geochemical information is important for mineral exploration and provides a theoretical basis for the rapid delineation of hidden minerals. Remote sensing technology provides the possibility of rapid and large-scale extraction of geochemical information from the earth's surface. This study analyzed the relationship between copper concentration and rock spectra by first collecting 222 rock samples, and then measuring the copper concentration of rock samples in the laboratory and reflectance spectra using an ASD FieldSpec3 portable spectrometer. It finally established quantitative relationships between the original spectra, first-order derivative spectra and second-order derivative spectra and copper concentration, respectively, using the partial least squares support vector machine method (PLS-SVM). The results show that 1) The estimation accuracy of using second-order derivatives spectra as input parameters to establish a model for estimating copper concentration is the highest, and the determined coefficient (R^2) between the predicted value and real value reaches 0.54. 2) When the copper concentration is less than 80 mg/kg, the inversion model of copper concentration established using PLS-SVM obtains a good result. The R^2 between the predicted copper concentration and the real copper concentration reached 0.70248. When the copper concentration is greater than 80 mg/kg, the inversion model of copper concentration established using partial least squares (PLS) obtains a good result. The R^2 between the predicted copper concentration and the real copper concentration reached 0.49. The R^2 between real copper concentration and copper predicted by the method of piecewise separate modeling reaches 0.816. Therefore, the method of segmental modeling has great potential to improve the accuracy of copper concentration inversion.

Keywords copper concentration, rock, geochemical information, PLS-SVM, remote sensing

Received August 13, 2018; accepted March 31, 2019

E-mail: zhoukf@ms.xjb.ac.cn

1 Introduction

In recent decades, remote sensing technology has been rapidly developing and maturing due to the launch of various new sensor satellites. The use of remote sensing technology has important theoretical significance and applications for mining prospecting information, narrowing the prospecting area, improving the prospecting efficiency and accelerating the progress of geological exploration due to advantages such as a wide range of vision, fast speed, and high efficiency. Therefore, an increasing number of scholars have tried using remote sensing technology for mineral exploration. There has been extensive research conducted on alteration information extraction and methods including the ratio method (Ranjbar et al., 2004; Salem et al., 2013; Son et al., 2014), principal component analysis (Qaid et al., 2009; Tangestani and Moore, 2011; Honarmand et al., 2012; Ibrahim et al., 2016), spectral angle classification (Wen et al., 2007; Shafaroudi et al., 2009; Azizi et al., 2010; Khaleghi et al., 2014) and extraction method based on machine learning (Xue et al., 2007; Yang et al., 2008; Yan et al., 2013; Zhou and Zhang, 2017), which were proposed to extract alteration information related to mineralization.

However, because of the limitations of spatial resolution, other objects easily interfere with remote sensing data, resulting in the uncertainty of extracting useful

prospecting information, and the process of remote sensing inversion itself has multiple solutions. Therefore, using mineral alteration information extracted from remote sensing data alone has not obtained satisfactory results in mineral exploration. In contrast, if remote sensing image data can be combined with geochemical data with a clear significance of ore indication, and with additional useful information, it can alleviate the uncertainty caused by single remote sensing data information and multiple solutions during the inversion process, using remote sensing image data can also enhance the reliability of mineral exploration. At present, some scholars have successfully estimated the metal concentration in plants by establishing the relationship between remote sensing and plant geochemical information. For instance, Hede, Zhang, Schuerge, and Sridhar have more accurately estimated the concentration of Cu and Zn in plants through constructing several vegetation indexes that can represent the degree of heavy metal stress (Schuerger et al., 2003; Sridhar et al., 2007; Hede et al., 2015; Zhang et al., 2017). Some scholars have used single variable regression, stepwise regression analysis, partial least square, support vector machine and random forest to establish the relationship between the reflectance spectra or the first-order derivative spectrum of soil and metal concentration of soil, and successfully estimate the Au, Cu, Pb, Zn, and As concentration of soil (Malley and Williams, 1997; Kooistra et al., 2001; Kemper and Sommer, 2002; Siebielec et al., 2004; Wu et al., 2007; Choe et al., 2009; Moros et al., 2009; Gannouni et al., 2012; Song et al., 2012; Shi et al., 2016). However, there are few academic studies on extracting geochemical information from rocks using remote sensing data. In fact, the geochemical anomalies of ore deposits are most fully preserved by the anomalies of rock and earth in various types of ore deposits, and the geochemical anomaly of rock is a component of the source of various types of secondary geochemical anomalies, and all kinds of secondary geochemical anomalies are derived from the abnormality of the primary antibody with its geochemical anomalies. Therefore, it is essential to rapidly extract rock geochemical information during mineral exploration. Related studies have shown that factors affecting spectral characteristics of rocks include the chemical composition of the rock, and the crystal structure, size, structure, and weathering degree of the rock surface (Yang et al., 2015). Among these, the influence of the rock's chemical composition on its spectral characteristics is the most important factor (Tang et al., 2006). This indicates that remote sensing technology can be used to rapidly extract geochemical information from rocks.

This study first collected the reflectance spectra and the concentration of copper of 222 rock samples in Ketebieteti, Fuyun County, Xinjiang, and then determined whether there is a quantitative relationship between the rock spectrum and copper concentration. If this relationship exists, the rock geochemical information can be rapidly and nondestructively extracted over a large area. This relationship is also significant for increasing prospecting efficiency and expanding prospecting space.

2 Methods and materials

2.1 Research area generalization and sample collection

The Ketebieteti mining area is located 60 km south of Fuyun County, in Xinjiang Uygur Autonomous Region, China. The geological structure of the study area is complex, magmatic activity is frequent, and mineral resources are rich, which makes it a favorable metallogenic zone. Many mineralized bodies have been found in the Ketebieteti rock mass. This rock mass is mineralized and altered strongly, and has good mineralization potential. On 10–15 September 2017, there were 222 rock samples collected along 5 sampling lines (as shown in Fig. 1), including 65 slate, 3 diorite-porphyrite, 7 mudstone, 11 granite, 12 granite diorite, 13 hornblende pyroxenite, 24 hornblende gabbro, 9 diorite, and 78 gabbro. The sampling interval of each sampling line is 20 m. Sampling lines 1, 2, 3, 4, and 5 contain 61, 24, 50, 44, and 43 sampling points, respectively. The length of sampling lines 1, 2, 3, 4, and 5 is 1204, 473, 992, 889, and 879 m, respectively.

2.2 Measurement of the rock spectra

The ASD FieldSpec3 portable spectrometer produced in 2010 was used to collect the spectra of rock samples. The selected spectral sampling interval was 1 nm, the wavelength range was 350 to 2500 nm, and there were 2151 bands in total. During the measurement process, we first spread all the small pieces of rock samples on the ground, and then used a 25-degree field-of-view probe to vertically observe the rock surface. The distance between the probe and the rock is about 30–35 cm. Finally, an in-circle average reflection spectrum with a radius of 0.075 m is obtained as the reflection spectrum of the sample. To reduce the noise that was caused by the external environment and uncorrected operation, the spectral average obtained by five measurements was used as the spectrum of the sample. To reduce the interference of external factors and obtain stable data, clear and wind-free weather should be selected and measured between 10:30–14:00 local time because the sun's height angle is more stable during this period.

2.3 Pretreatment of the rock spectra

Due to a low signal-to-noise ratio in a range of less than 400 nm and greater than 2400 nm, it is easily affected by external interference factors, resulting in a greater volatility of spectral data. Therefore, this paper deletes the data in

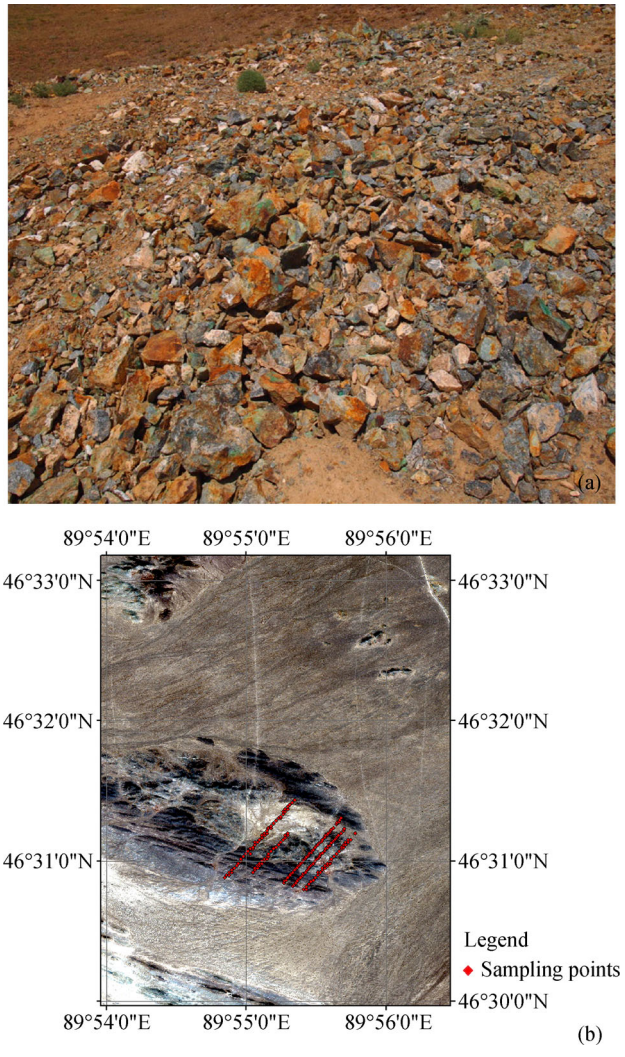


Fig. 1 Study area and sampling location; (b) Distribution of sampling points in Google Map; (a) field photographs of study area.

these two ranges. The data in the range of 1300–1400 nm and 1800–2000 nm are greatly influenced by atmospheric water vapor, resulting in the reflectivity appearing as abnormal, so the data of these two bands are also deleted. The number of bands remaining at the end is 1700 and the bandwidth is 1 nm. However, a narrow bandwidth will result in redundant data, and the rock spectrum will produce a "burr" phenomenon. Therefore, this paper smoothed the rock spectrum using the average value of the 10 adjacent bands, and the rock spectrum is finally left with 170 bands. The spectra of different types of rock with different concentration on copper are shown in Fig. 2.

2.4 Measurement of copper concentration in rock samples

After collecting the reflectance spectra of rocks, the samples were sent to the laboratory to measure their metal concentration. The copper concentration was measured at the Central Laboratory of North China Geological Exploration Bureau. The copper concentration of rock samples was measured with inductively coupled plasma mass spectrometry (ICP-MS). The histogram of copper concentration of rock samples is shown in Fig. 3, which shows that the range of copper concentration in rock samples is 14.06 to 123.69 mg/kg, the average copper concentration is 48.25 mg/kg and the standard derivation of copper concentration reaches 25.66 mg/kg.

2.5 Derivative spectrum

Figure 2 shows that the reflectance spectra of rocks are relatively flat, and the differences between the reflectance spectra of rocks with different metal concentrations are small. This makes it difficult to extract the characteristic bands that contribute greatly to the prediction of metal concentration, and may seriously affect the establishment of the estimation model of metal concentration. Derivative spectroscopy is a common means of change in hyperspectral analysis. This method can reduce or eliminate the influence of background noise, and also enhances the subtle change of the spectral curve on the slope, and improves the spectrum's multiple collinearity to a certain extent. The first and second derivatives can be used to determine the bending point of the plant spectral curve, and the wavelength position at the maximum and minimum reflectivity. Therefore, this study will analyze the relationships between the first-order derivatives, second-order derivatives and copper concentration of the rock samples, respectively. It will use the difference method to obtain the derivative spectrum of the rock samples because the reflection spectrum of rock is discrete data. The formulas are shown in Eqs. (1) and (2):

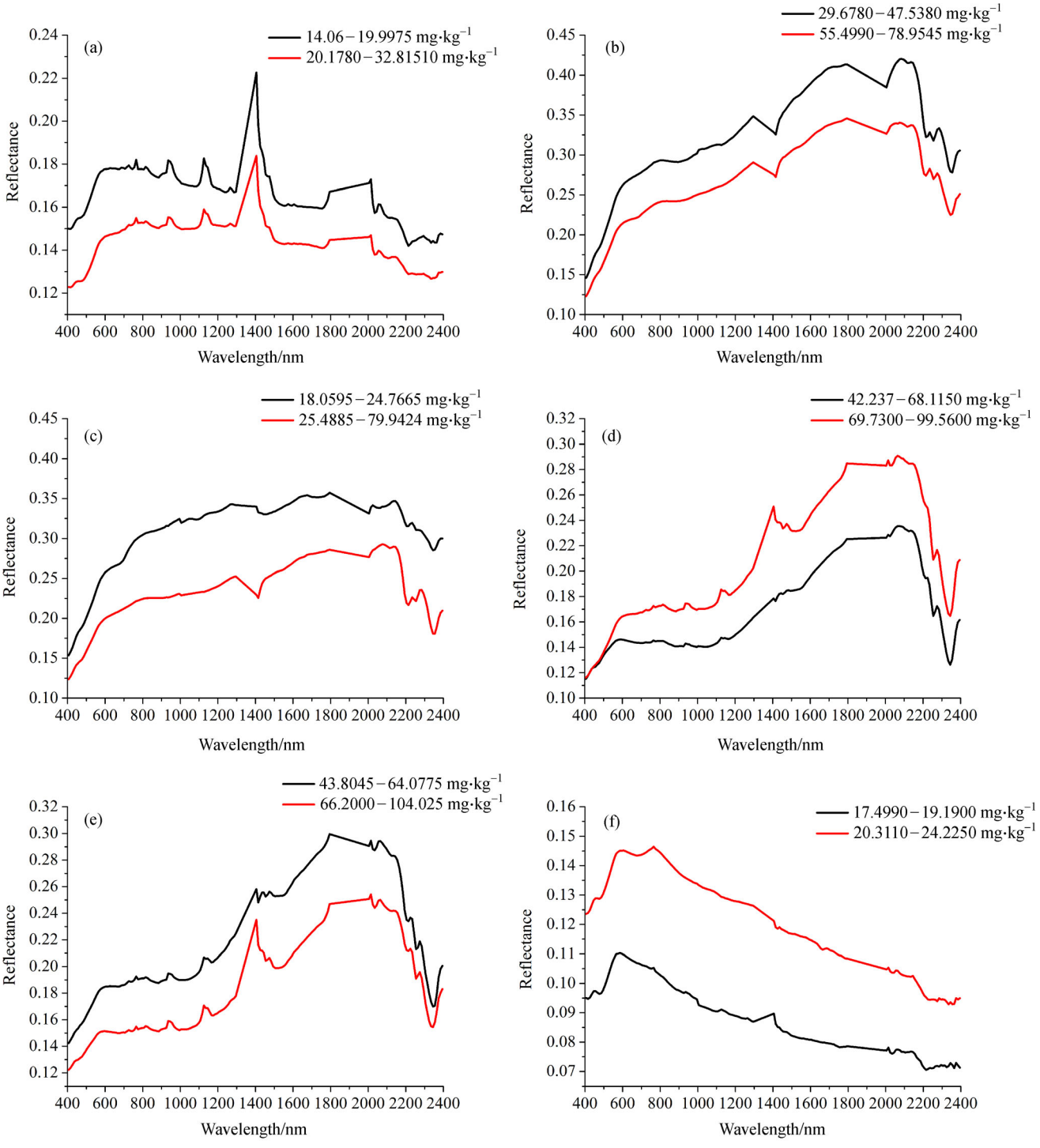
$$R'(\lambda_i) = \frac{R(\lambda_{i+1}) - R(\lambda_{i-1})}{\lambda_{i+1} - \lambda_{i-1}}, \tag{1}$$

$$R''(\lambda_i) = \frac{R'(\lambda_{i+1}) - R'(\lambda_{i-1})}{\lambda_{i+1} - \lambda_{i-1}}, \tag{2}$$

where $R(\lambda_{i+1})$ represents rock reflectance at wavelength λ_{i+1}; $R(\lambda_{i-1})$ represents rock reflectance at wavelength λ_{i-1}; $R'(\lambda_i)$ represents the first-order derivative value at wavelength λ_i; $R''(\lambda_i)$ represents second-order derivatives value at wavelength λ_i.

2.6 Partial least squares support vector machine

Partial least squares support vector machine (PLS-SVM) algorithm is a generalization of standard support vector machine (SVM) (Cortes and Vapnik, 1995; Wold et al., 2001). As with standard SVM, the starting point of PLS-SVM algorithm is to find the optimal hyperplane or


(a)
14.06–19.9975 mg·kg^{-1}
20.1780–32.81510 mg·kg^{-1}
(b)
29.6780–47.5380 mg·kg^{-1}
55.4990–78.9545 mg·kg^{-1}
(c)
18.0595–24.7665 mg·kg^{-1}
25.4885–79.9424 mg·kg^{-1}
(d)
42.237–68.1150 mg·kg^{-1}
69.7300–99.5600 mg·kg^{-1}
(e)
43.8045–64.0775 mg·kg^{-1}
66.2000–104.025 mg·kg^{-1}
(f)
17.4990–19.1900 mg·kg^{-1}
20.3110–24.2250 mg·kg^{-1}
Reflectance
Wavelength/nm

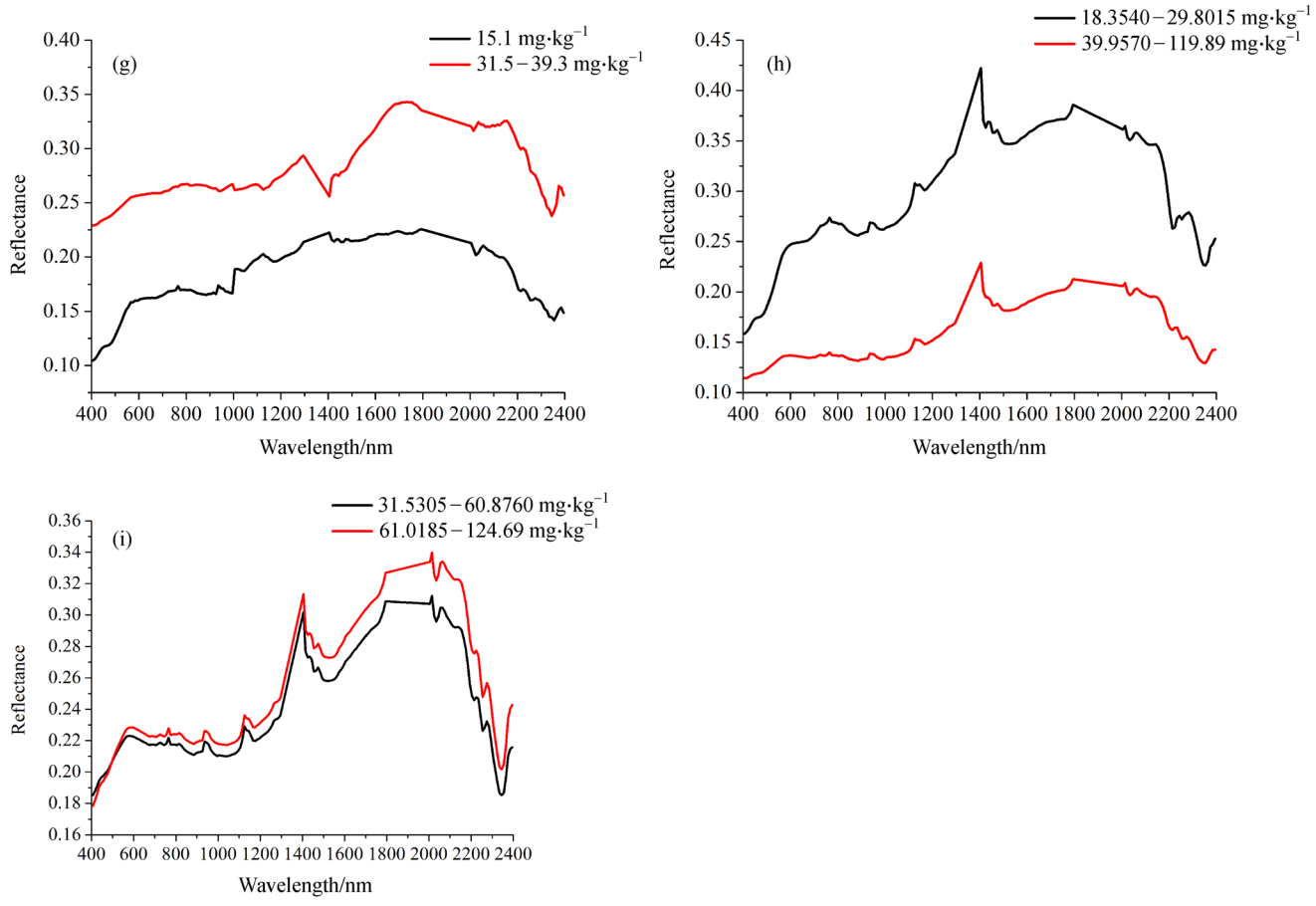


Fig. 2 Changes of the spectra of different types of rock with different concentration on copper. (a) Slate; (b) granite diorite; (c) granite; (d) hornblende pyroxenite; (e) hornblende gabbro; (f) mudstone; (g) diorite-porphyrite; (h) diorite; (i) gabbro.

optimal regression hyperplane in the high dimensional feature space based on structural risk minimization. Unlike standard SVM, the loss function in PLS-SVM is no longer related only to a small portion of the sample, but takes the two terms of the learning error of all samples to control experience risk. Therefore, PLS-SVM algorithm maintains standard SVM for many advantages, such as a small sample, nonlinear, high dimensional model and good generalization ability. At the same time, the convex two order programming problem of the inequality constraint condition of the standard SVM solution is transformed into a set of linear equations solving problem. The advantages of PLS-SVM are that it can solve large-scale problems, simplify the computation algorithm, increase the learning speed considerably increase learning speed and significantly reduce computation cost. Our study mainly attempts to establish the relationship between 170 bands of rock spectrum and copper concentration. However, the stability of the established model will be degraded due to the large number of spectral bands, data redundancy and high correlation between adjacent bands. The PLS method can simplify the structure of hyper-spectral data and effectively reduce its information redundancy. It can also solve the multi-correlation problem of spectral bands, which multiple stepwise regressions cannot solve. The hyper-spectral model established using this method could contain more spectral information and enhance the interpretation ability of heavy metals. At the same time, the spectrum of rocks is very complex, and there are many factors affecting the spectrum of rocks; including the influence of weathering, rock surface structure, rock surface color, and atmospheric environment on the spectral reflectance of rocks; which results in the relationship between the rock spectra and the copper concentration not being a simple linear relationship, but they may have a nonlinear relationship. Support Vector Machine (SVM) has unique advantages in solving small samples, nonlinearity and high-dimensional pattern recognition. To maximize the advantages of partial least squares and support vector machines, we introduced partial least squares support vector machines to construct the relationship between rock spectra and copper concentration.

2.7 Establishing a model and its accuracy evaluation

In this study, first the spectra of original, first-order derivative and second-order derivatives of the rock samples are used as independent variables, respectively,

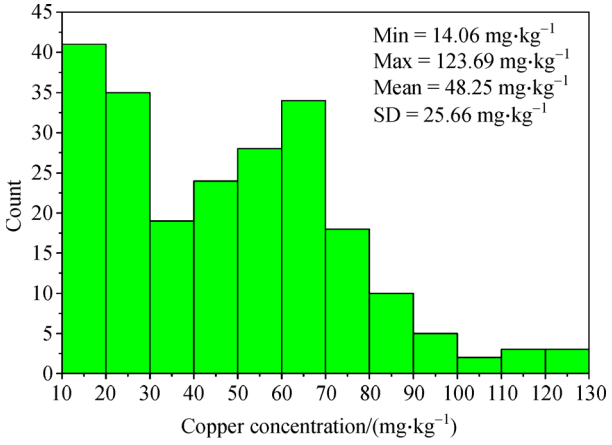


Fig. 3 Histogram of copper concentration of rock samples.

and the copper concentration of rock is used as the dependent variable input to PLS-SVM, and it is then assessed whether there is a nonlinear relationship between the rock spectrum and copper concentration. Using the determination coefficient (R^2) between the predicted value and actual value of copper concentration as the evaluation index of the inversion model's accuracy, a larger R^2 means there is a higher inversion accuracy.

However, when using machine learning to establish the model, fitting often occurs; that is to say, the precision of the training set is high, but the fitting accuracy of the data outside the training set is low, and the model generalization ability is poor and does not have generalization. To evaluate the reliability and stability of the model established, we used cross validation to evaluate the stability of the established model, using the determination coefficient (R^2) between the predicted value of copper concentration and the measured value as the evaluation index, where a higher R^2 value indicates a higher stability for the established model.

The basic idea of leave-out cross validation is to assume that the original data has N samples, and then each sample is used as a validation set alone, after which the rest of the N–1 samples are used as the training set, so the N model will be obtained. The final model evaluation precision of the prediction accuracy of the validation set of the N model is used. Compared with the previous cross validation method, leave-out cross validation is used for training models for almost all the samples in each round, as it is the closest to the distribution of the original samples so the evaluation results are more reliable.

3 Results and discussion

3.1 Correlation analysis of different spectral transformations and copper concentration of rock samples

The correlations of the copper concentration of rock samples with its original spectrum, first-order derivative and second-order derivatives spectrum were analyzed, respectively. The variation curve of the correlation coefficient with the wavelength is shown in Fig. 4, indicating that correlation between the original spectrum and the copper concentration is relatively low and smooth, and the correlation coefficient between all bands in the range of 400–2400 nm and copper concentration is not more than 0.45. In contrast, the volatility of changes of the correlation coefficients between rock derivative spectra and copper concentration of rock with the wavelength obviously increased. The correlation coefficients between several bands and copper concentration have been greatly improved, even up to 0.6. These bands include the first-order derivatives of 2243, 2293, and 2303 nm, as well as the second-order derivatives of 2223, 2233, and 2283 nm. These bands can be used as the characteristic bands for retrieving copper concentration in rocks, due to the strong correlation between these bands and copper. These findings are consistent with previous studies (Liu et al., 2010). They also found that the bands with strong correlation with copper content are mainly concentrated at 2275 to 2295 nm. In contrast, the correlation between some bands and copper concentration is reduced. However, in general, derivative spectroscopy is more helpful in extracting characteristic bands that can characterize copper in rocks. The correlation analysis between copper concentration and reflectance spectrum and derivative spectra lays the theoretical foundation for the establishment of the copper concentration estimation model based on hyperspectral technology. At the same time, Figure 4 also shows that the correlation coefficient between rock spectra and copper concentration varies greatly under different spectral transformations. Therefore, it is essential to choose the appropriate spectral transformation before we use the rock spectrum. A good spectral transformation can

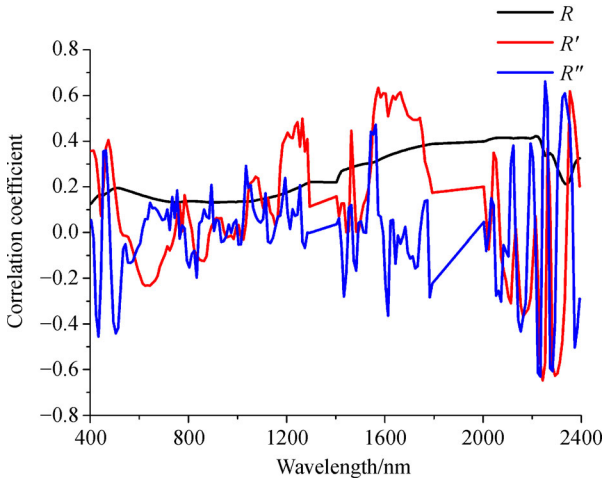


Fig. 4 Correlations between three types of transformed spectra and the copper concentration of rock samples.

more effectively remove the noise information in the spectrum, and can maintain the information that is closely related to copper concentration.

3.2 Comparison of inversion accuracy of copper concentration under different spectral transformations

R^2 represents the determined coefficient between the predicted value and real value. It is used to evaluate the performance of the proposed inversion model. Table 1 compares R^2 of three different types of spectra that are used as input parameters to establish the inversion model, which shows that the R^2 of the inversion models for copper concentration in rocks, whether they are original spectra or first-order or second-order derivatives as input parameters, are all above 0.5. The original spectrum that has undergone the first and second derivations can slightly increase copper concentration inversion accuracy. This is probably due to the elimination of some noise from the original spectrum using the derivative method. The model established using the second-order derivatives as the input parameter has the highest inversion precision, and the determined coefficient between the predicted copper concentration and the measured copper concentration reaches 0.543. This indicates that using the rock spectrum and partial least squares support vector machine to estimate copper concentration in rocks is feasible. At the same time, the model is established based on 9 different types of rocks. This shows that the model established by this method is robust and can be applied to different types of rocks, not just a single type of rock. However, there are many types of rocks in nature, far more than these 9 types. It is also necessary to further discuss whether the model established in this study also applies to other types of rock. At the same time, Figure 5 also shows that when the copper concentration of the rock is less than 80 mg/kg, the inversion model constructed by the rock spectrum and PLS-SVM often overestimates the copper concentration in the rock. When the concentration of the rock is more than 80 mg/kg, the constructed inversion model often underestimates the copper concentration.

The transformation of spectral data are very useful for data mining (Sawut et al., 2014; Abdel-Rahman et al., 2014; Rady et al., 2014). Spectral data transformation is important for eliminating background noise, abnormal values, reducing the interference of internal and external environmental factors, and amplifying the characteristics of spectral absorption (Zhu et al., 2017). Appropriate spectral transformation methods can reduce or even eliminate irrelevant information of spectral data and highlight subtle spectral characteristics, which can help improve the prediction accuracy and robustness of the inversion model (Seasholtz and Kowalski, 1993). However, in this study, we only use three forms of spectral transformation as the input parameters to establish the model. In fact, there are many other spectral transformations, such as logarithmic spectrum, reciprocal spectrum and so on. Each of these has different advantages for the mining of spectral information. For instance, the logarithmic spectrum can transform multiplicative noise into additive noise (Zhou et al., 2009; Wang et al., 2012). It also can reduce the effects of changes in illumination conditions (Yoder and Pettigrewcrosby, 1995; Serrano et al., 2002). Differential transformation of spectral data after logarithmic transformation can reduce the effects of additive random noise (Gong and Yu, 2001). Scattering correction of spectra can reduce physical phenomena and baseline drift caused by scattering. Therefore, in future research, we can try to establish the inversion model by using more spectral transformations, and find the optimal spectral transformation by comparison. We can even try making a variety of spectral transformations on a spectral curve to form a mixed spectrum.

In this study, we only use PLS-SVM to establish the inversion model, but there have recently been some new machine learning methods such as random forest, deep learning and other methods. Therefore, in future studies, we can analyze the advantages and disadvantages of different machine learning methods for the inversion of the metal concentration of rock, and find the best model for this inversion.

3.3 Inversion model construction when the rock copper concentration is less than 80 mg/kg

Figure 6(a) shows that the model established by PLS-SVM cannot obtain a good result when the copper concentration of rock is more than 80 mg/kg, as the copper concentration predicted in this interval is seriously deviated from the 1:1 line. Therefore, we will try dividing the copper concentration of rock into two intervals larger than 80 mg/kg and less than 80 mg/kg, and then construct the inversion model in each interval. Figure 6(b) shows that when the copper concentration of rock is less than 80 mg/kg, the inversion model of copper concentration established using PLS-SVM can obtain a good result. The coefficient of determination between the predicted copper concentration and the actual copper concentration reached 0.70248.

Table 1 Comparison of inversion accuracy of copper concentration using three different types spectral transformations

Spectral transformation	Training set (R^2)	Leave-one-out cross validation (R^2)
Original spectra (R)	0.698	0.525
First-order derivative spectra (R')	0.761	0.530
Second-order derivatives spectra (R'')	0.791	0.543

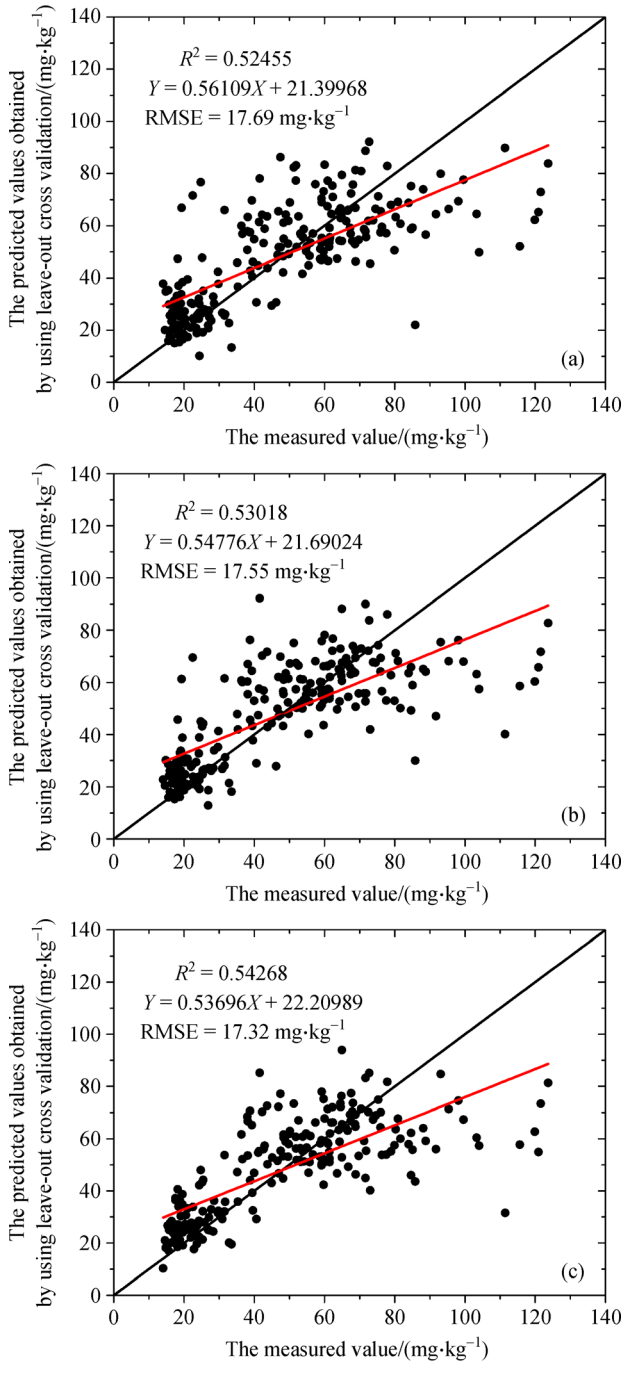


Fig. 5 Scatter plots between the predicted value of model established using different spectral transformations and the true value. (a) Original spectra; (b) first-order derivative spectra; (c) second-order derivatives spectra.

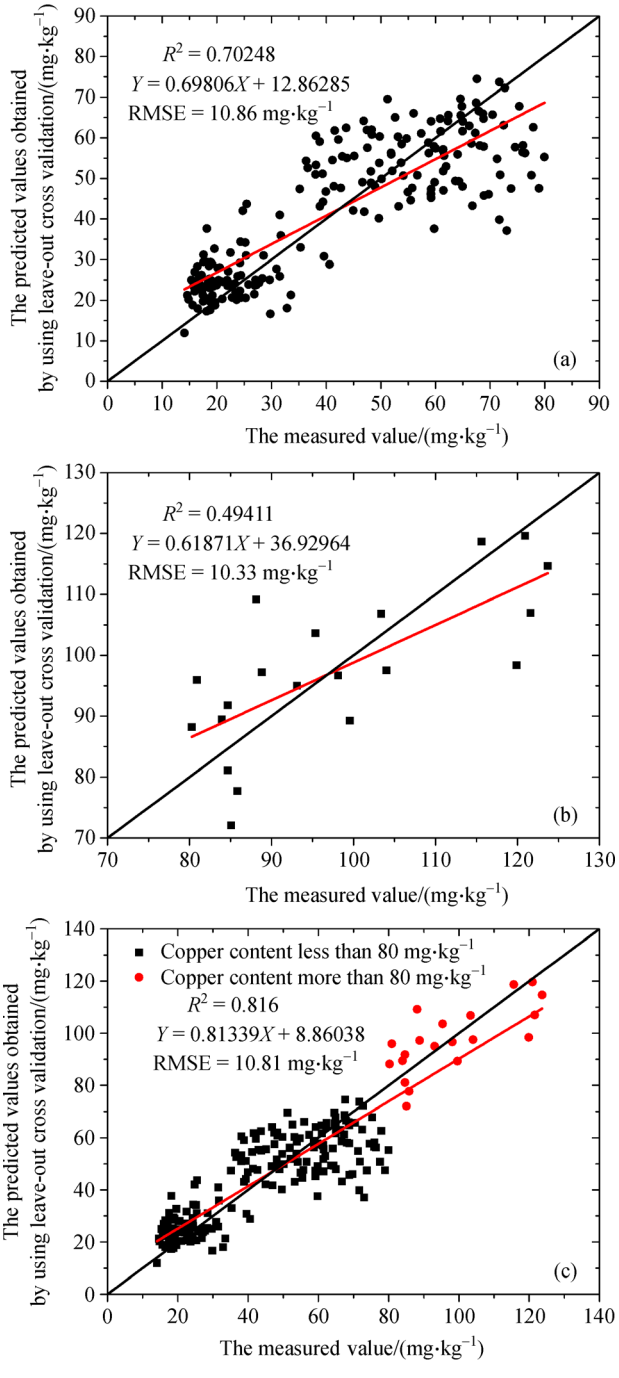


Fig. 6 Comparison of inversion accuracy under different levels of copper concentration. (a) Copper concentration of rock samples is less than 80 mg/kg; (b) copper concentration of rock samples is more than 80 mg/kg; (c) the scatter plot between real copper concentration and copper predicted by the method of piecewise separate modeling.

When the copper concentration in the rock is more than 80 mg/kg, we found that using the PLS method to establish a model can obtain a better result, and the predicted copper concentration and the measured copper concentration are shown in Fig. 6(b). From this, we can see that the determinant coefficient between the predicted copper concentration and the actual copper concentration also reaches 0.49441. Figure 6(c) is a scatter plot between real copper concentration and copper predicted with the method of piecewise separate modeling. From this, we can see that the determination coefficient reaches 0.816,

which greatly improves the inversion precision of the copper concentration of rock. Therefore, we believe that when the copper concentration of rock is less than 80 mg/kg, using the second-order derivatives and PLS-SVM to establish an inversion model can achieve a better result, and when the copper concentration of rock is more than 80 mg/kg, the result is better when the second-order derivatives and PLS method are used to build the model. The method of segmental modeling can significantly improve the accuracy of copper concentration inversion.

3.4 Uncertainty analysis of the model for estimating copper concentration

In this study, we used the original and derivative spectra of rocks to construct statistical models for estimating copper concentration. Concretely speaking, the 170 bands of spectra of rock samples are recorded as X1, X2, X3,…, X170 respectively and copper concentration is recorded as Y and input into the PLS-SVM method to establish the model for estimating the copper concentration. The biggest disadvantage of this model is that its accuracy depends heavily on the accuracy of independent variables. That is to say, small changes in rock spectra can easily affect the accuracy of copper content estimation. There are many factors affecting the spectrum of rocks, including weathering, surface structure, rock surface color and atmospheric environment, which results in the unusually complex nature of rock spectra. These factors can be regarded as interfering information for estimating the copper concentration of rock. The influence of these interference factors is why there is a low accuracy of model building. Therefore, how to exclude the influence of these interference factors from rock spectra and extract spectral characteristics caused by copper elements to the greatest extent is very important for improving the estimation accuracy of the model and is the focus of future research.

In this paper, all of the bands in ranges of 400 to 2400 nm were input into PLS-SVM to establish the model for estimating the copper concentration in rock. However, the whole band contains a large amount of spectral information, which contains more invalid information. Using these variables to build the model may reduce the accuracy and reliability of the model. Therefore, using appropriate spectral preprocessing and band selection methods can simplify and extract characteristic spectral information and improve the prediction ability of quantitative models. At present, the commonly used feature band selection methods include stepwise regression analysis, continuous projection algorithm, non-information variable elimination method, genetic algorithm, ant colony algorithm and so on. Therefore, in future research, we can combine these methods with PLS-SVM to judge whether they can improve the accuracy and stability of the model.

In this experiment, we used an ASD spectrometer to measure rock spectrum in outdoor environments. Owing to the influence of water vapor content, the reflectance anomalies occur in the range of 1300–1400 nm and 1800–2000 nm, so these two parts of the spectrum are not considered in the modeling process. The contribution of these two parts to the estimation of copper content in rocks needs further verification. Therefore, in the future, we can measure spectrum indoors and then try to model it.

3.5 The difficulties and prospects of models applied to aerospace platforms

The purpose of establishing the model is to apply it to the sensors carried on an aeronautic or space platform, to estimate the metal concentration in the rock quickly and in a large area, and to solve the disadvantages of time-intensive and laborious work using the conventional method. However, the remote sensing method has its own limitations such as the mixed pixel problem. In fact, the spectrum in the remote sensing image is often mixed, in which a large amount of soil and plant information is combined, causing a great level of uncertainty in the rock spectrum. We have built a statistical model, and its precision has a significant impact on the input spectrum. Even a small deviation in the spectrum often leads to a large deviation in inversion accuracy. Therefore, when using remote sensing technology to extract large amounts of metal content in rocks, it is very important to effectively reduce the influence of mixed pixels. Although some of the current methods can somewhat reduce the impact of mixed pixels, such as linear spectral mixed model (Quarmby et al., 1992; Haertel and Shimabukuro, 2005; Freitas et al., 2008; Fassoniandrade et al., 2017), and nonlinear spectral mixed model (Borel and Gerstl, 1994; Ray and Murray, 1996; Altmann et al., 2012; Heylen and Scheunders, 2016). The simplest and most effective way is using images with high spatial resolution. Therefore, we use the power triangle wing and Hyspex hyperspectral sensor to build the ultra low altitude detection platform (Fig. 7), which can achieve the observation of the “submeter level” and effectively solve the mixed pixel problem. Our next goal was to solve the effective use of the model established in this paper in the ultra low altitude detection platform to achieve the rapid extraction of metal content in the rock.

Our team plans to use the Ultra-Low Altitude Exploration Platform (Fig. 7) to acquire hyper-spectral images of the research area next year. When the flying altitude of the dynamic delta wing reaches 200 m, the spatial resolution of the image can reach 0.15 m. The reason for choosing this height is to ensure that the size of image pixels is consistent with the detection range of each sample in this experiment, that is to say, to ensure that the spatial scale of the ground experiment and aerial experiment is consistent. It then judges whether the rules summarized in this paper are applicable to hyper-spectral sensors mounted on aircraft platforms. Finally, a set of rock metal content estimation

Fig. 7 Ultra low altitude detection platform.

methods based on ultra-low altitude detection platforms are formed.

4 Conclusions

This study first measured the copper concentration and reflectance spectra of 222 rock samples of 9 types of rocks, and then tried to establish the relationships between the original, first-order derivative, and second-order derivatives spectra and copper concentration of rock, respectively, using the PLS-SVM method. The results show that:

1) Compared to the original and first-order derivative spectra, the correlation between the second-order derivatives spectra and copper concentration is significantly increased;

2) It is feasible to use rock spectra to retrieve copper concentration of rocks, which provides a possibility for rapid and large scale extraction of copper elements. The next step is how to effectively solve the problem of spatial and spectral scales, and how the model built by the ground test field can be effectively applied to the hyperspectral sensors on the aeronautics and space platforms;

3) The method of segmental modeling has great potential to improve the accuracy of copper concentration inversion.

Acknowledgements This research is funded by Xinjiang Uygur Autonomous Region Key Laboratory Open Subject (No. 2018D04025), National Natural Science Foundation of China (Grant Nos. U1503291 and 41402296), Key Laboratory fund of Xinjiang Uygur Autonomous Region (No. 2016D03006), The "Belt and Road" team of the Chinese Academy of Sciences (2017-XBZG-BR-002), Key R&D Program of Xinjiang Uygur Autonomous Region (No. 2017B03017-2).

References

Abdel-Rahman E M, Mutanga O, Odindi J, Adam E, Odindo A, Ismail R (2014). A comparison of partial least squares (PLS) and sparse PLS regressions for predicting yield of swiss chard grown under different irrigation water sources using hyperspectral data. Comput Electron Agric, 106: 11–19

Altmann Y, Halimi A, Dobigeon N, Tourneret J Y (2012). Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery. IEEE Trans Image Process, 21(6): 3017–3025

Azizi H, Tarverdi M A, Akbarpour A (2010). Extraction of hydrothermal alterations from ASTER SWIR data from east Zanjan, Northern Iran. Adv Space Res, 46(1): 99–109

Borel C C, Gerstl S A W (1994). Nonlinear spectral mixing models for vegetative and soil surfaces. Remote Sens Environ, 47(3): 403–416

Choe E, Kim K W, Bang S, Yoon I H, Lee K Y (2009). Qualitative analysis and mapping of heavy metals in an abandoned Au–Ag mine area using NIR spectroscopy. Environ Geol, 58(3): 477–482

Cortes C, Vapnik V N (1995). Support vector networks. Mach Learn, 20(3): 273–297

Fassoniandrade A C, Zanotta D C, Guasselli L A, de Andrade A M (2017). Linear spectral mixing model for estimating optically active components in estuarine waters of Patos Lagoon, Brazil. Int J Remote Sens, 38(17): 4767–4781

Freitas R M D, Haertel V, Shimabukuro Y E (2008). Linear spectral mixture model in moderate spatial resolution image data. Bol Ciênc Geod, 14(1): 55–71

Gannouni S, Rebai N, Abdeljaoued S (2012). A spectroscopic approach to assess heavy metals contents of the mine waste of Jalta and Bougrine in the north of Tunisia. J Geogr Inf Syst, 4(3): 242–253

Gong P, Yu B (2001). Conifer species recognition: effects of data transformation. Int J Remote Sens, 22(17): 3471–3481

Haertel V F, Shimabukuro Y E (2005). Spectral linear mixing model in low spatial resolution image data. IEEE Trans Geosci Remote Sens, 43(11): 2555–2562

Hede A N H, Kashiwaya K, Koike K, Sakurai S (2015). A new vegetation index for detecting vegetation anomalies due to mineral deposits with application to a tropical forest area. Remote Sens Environ, 171: 83–97

Honarmand M, Ranjbar H, Shahabpour J (2012). Application of principal component analysis and spectral angle mapper in the mapping of hydrothermal alteration in the Jebal–Barez area, Southeastern Iran. Resour Geol, 62(2): 119–139

Heylen R, Scheunders P (2016). A multilinear mixing model for nonlinear spectral unmixing. IEEE Trans Geosci Remote Sens, 54(1): 240–251

Ibrahim W S, Watanabe K, Yonezu K (2016). Structural and lithotectonic controls on Neoproterozoic base metal sulfide and gold mineralization in north Hamisana shear zone, south eastern desert, Egypt: the integrated field, structural, Landsat 7 ETM + and ASTER data approach. Ore Geol Rev, 79: 62–77

Kemper T, Sommer S (2002). Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy. Environ Sci Technol, 36(12): 2742–2747

Khaleghi M, Ranjbar H, Shahabpour J, Honarmand M (2014). Spectral angle mapping, spectral information divergence, and principal component analysis of the ASTER SWIR data for exploration of porphyry copper mineralization in the Sarduiyeh area, Kerman Province, Iran. Appl Geomat, 6(1): 49–58

Kooistra L, Wehrens R, Leuven R S E W, Buydens L M C (2001). Possibilities of visible–near-infrared spectroscopy for the assessment of soil contamination in river floodplains. Anal Chim Acta, 446(1–2): 97–105

Liu M, Lin Q Z, Wang Q J, Li H (2010). Study on the geochemical anomaly of copper element based on reflectance spectra. Spectrosc Spect Anal, 30(5): 1320–1323

Malley D F, Williams P C (1997). Use of near-infrared reflectance spectroscopy in prediction of heavy metals in freshwater sediment by their association with organic matter. Environ Sci Technol, 31(12): 3461–3467

Moros J, Vallejuelo S F O, Gredilla A, Diego A, Madariaga J M, Garrigues S, Guardia M. (2009). Use of reflectance infrared spectroscopy for monitoring the metal content of the estuarine sediments of the Nerbioi-Ibaizabal River (Metropolitan Bilbao, Bay of Biscay, Basque Country). Environ Sci Technol, 43(24): 9314–9320

Qaid A M, Basavarajappa H T, Ranjbar H (2009). Application of principal component analysis to ASTER and ETM + data for mapping the alteration zones in north east of Hajjah, Yemen. Southeast Asian J Surg, 9(2): 15–21

Quarmby N A, Townshend J R G, Settle J J, White K H, Milnes M, Hindle T L, Silleos N (1992). Linear mixture modelling applied to AVHRR data for crop area estimation. Int J Remote Sens, 13(3): 415–425

Ray T W, Murray B C (1996). Nonlinear spectral mixing in desert vegetation. Remote Sens Environ, 55(1): 59–64

Rady A M, Guyer D E, Kirk W, Donis-González I R (2014). The potential use of visible/near infrared spectroscopy and hyperspectral imaging to predict processing-related constituents of potatoes. J Food Eng, 135(2): 11–25

Ranjbar H, Honarmand M, Moezifar Z (2004). Application of the Crosta technique for porphyry copper alteration mapping, using ETM data in the southern part of the Iranian volcanic sedimentary belt. J Asian Earth Sci, 24(2): 237–243

Salem S M, Arafa S A, Ramadan T M, El Gammal E S A(2013). Exploration of copper deposits in Wadi El Regeita area, Southern Sinai, Egypt, with contribution of remote sensing and geophysical data. Arab J Geosci, 6(2): 321–335

Sawut M, Ghulam A, Tiyip T, Zhang Y J, Ding J L, Zhang F, Maimaitiyiming M (2014). Estimating soil sand content using thermal infrared spectra in arid lands. Int J Appl Earth Obs Geoinf, 33(12): 203–210

Schuerger A C, Capelle G A, Di BenedettoJ A, Mao C, Thai C N, Evans M D, et al. (2003). Comparison of two hyperspectral imaging and two laser-induced fluorescence instruments for the detection of zinc stress and chlorophyll concentration in bahia grass (*Paspalum notatum*, Flugge.). Remote Sens Environ, 84(4): 572–588

Seasholtz M B, Kowalski B (1993). The parsimony principle applied to multivariate calibration. Anal Chim Acta, 277(2): 165–177

Serrano L, Peñuelas J, Ustin S L (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: decomposing biochemical from structural signals. Remote Sens Environ, 81(2–3): 355–364

Malekzadeh Shafaroudi A, Karimpour M H, Stern C R, Mazaheri S A (2009). Hydrothermal alteration mapping in SW Birjand, Iran, using the advanced spaceborne thermal emission and reflection radiometer (ASTER) image processing. J Appl Sci (Faisalabad), 9(5): 829–842

Shi T, Liu H, Chen Y, Wang J, Wu G (2016). Estimation of arsenic in agricultural soils using hyperspectral vegetation indices of rice. J Hazard Mater, 308: 243–252

Siebielec G, McCarty G W, Stuczynski T I, Reeves J B 3rd (2004). Near- and mid-infrared diffuse reflectance spectroscopy for measuring soil metal content. J Environ Qual, 33(6): 2056–2069

Song Y, Li F, Yang Z, Ayoko G A, Frost R L, Ji J (2012). Diffuse reflectance spectroscopy for monitoring potentially toxic elements in the agricultural soils of Changjiang River Delta, China. Appl Clay Sci, 64(4): 75–83

Son Y S, Kang M K, Yoon W J (2014). Lithological and mineralogical survey of the Oyu Tolgoi region, Southeastern Gobi, Mongolia using ASTER reflectance and emissivity data. Int J Appl Earth Obs Geoinf, 26(1): 205–216

Sridhar B B M, Han F X, Diehl S V, Monts D L, Su Y (2007). Spectral reflectance and leaf internal structure changes of barley plants due to phytoextraction of zinc and cadmium. Int J Remote Sens, 28(5): 1041–1054

Tang L L, Zhang X F, Liu Y H (2006). Research progress in the foundation and technology of remote sensing lithology information extraction. Mining Research and Development, 26(3): 68–73 (in Chinese)

Tangestani M H, Moore F (2001). Comparison of three principal component analysis techniques to porphyry copper alteration mapping: a case study, Meiduk Area, Kerman, Iran. Can J Rem Sens, 27(2): 176–182

Wang L, Bai Y L, Lu Y, Wang H (2012). Effect on retrieval precision for corn N content by spectrum data transformation. Remote Sensing Technology and Application, 26(2): 220–225 (in Chinese)

Wen X, Hu G, Yang X (2007). A Simplified Method for Extracting Mineral Information from Hyperspectral Remote Sensing Image Using SAM Algorithm. In: Zhao P D, Agterberg F, Cheng Q M, eds. The 12th Conference of the international Association for mathematical Geology. China University of Geosciences, 526–529

Wold S, Sjöström M, Eriksson L (2001). PLS-regression: a basic tool of chemometrics. Chemom Intell Lab Syst, 58(2): 109–130

Wu Y, Chen J, Ji J, Gong P, Liao Q, Tian Q, Ma H (2007). A mechanism study of reflectance spectroscopy for investigating heavy metals in soils. Soil Sci Soc Am J, 71(3): 918–926

Xue Y, Dai T G, Zou L, Xia H D, Liu J L (2007). Extracting alteration information based on SVM—taking Jidi area as an example. Geological Survey Re, 30(4): 315–320 (in Chinese)

Yan J, Zhou K, Wang J, Wang S, Wang W, Dong L I (2013). Extraction of hyper-spectral remote sensing alteration information based on SAM and SVM. Computer Engineering and Applications, 49(19): 141–146 (in Chinese)

Yang T, Xue Y, Dai T G (2008). Extracting alteration information by SVM based on spectrum and texture—taking Jidi area as the example. Earth & Environment, 36(1): 81–86 (in Chinese)

Yang C B, Zhang C X, Liu F, Jiang Q G (2015). Study on the relationship between the depth of spectral absorption and the content of the mineral composition of biotite. Spectrosc Spect Anal, 35(9): 2583–2587

Yoder B J, Pettigrewcrosby R E (1995). Predicting nitrogen and

chlorophyll content and concentrations from reflectance spectra (400–2500 nm) at leaf and canopy scales. Remote Sens Environ, 53 (3): 199–211

Zhang C, Ren H, Qin Q, Ersoy O K (2017). A new narrow band vegetation index for characterizing the degree of vegetation stress due to copper: the copper stress vegetation index (CSVI). Remote Sens Lett, 8(6): 576–585

Zhou K, Zhang N (2017). Extraction of alteration mineral information from moderate remote sensing images using MPS method. Photonirvachak (Dehra Dun), 46(2): 1–8

Zhou G Z, Wang C Z, Yang F J, Li Y M (2009). Field collected plant spectrum denoising by logarithm transform and wavelet transform. J Infrared Millim W, 28(4): 316–320

Zhu Y, Shen G R, Xiang Q Q, Wu Y (2017). Spectral characteristics of soil salinity based on different pre-processing methods. Chinese Journal of Soil Science, 48(3): 560–568 (in Chinese)