PAPER • OPEN ACCESS

Study on Hyperspectral Characteristics and Estimation Model of Soil Mercury Content

To cite this article: Jinbao Liu et al 2017 IOP Conf. Ser.: Mater. Sci. Eng. 274 012030

View the article online for updates and enhancements.

You may also like

- <u>Transformer fault diagnosis based on</u> <u>MPA-RF algorithm and LIF technology</u> Pengcheng Yan, JingBao Wang, Wenchang Wang et al.
- Determination of green and red spinach microgreen chlorophyll content using visible spectroscopy and wavelength selection E Nugroho, A M Zahra, R E Masithoh et al.
- <u>M-FISH evaluation of chromosome</u> <u>aberrations to examine for historical</u> <u>exposure to ionising radiation due to</u> <u>participation at British nuclear test sites</u> Kirsty Josephine Lawrence, Martin Scholze, Jose Seixo et al.





DISCOVER how sustainability intersects with electrochemistry & solid state science research



This content was downloaded from IP address 18.226.28.197 on 08/05/2024 at 14:10

IOP Publishing

Study on Hyperspectral Characteristics and Estimation Model of Soil Mercury Content

Jinbao Liu^{1, 3, 4, 5, *}, Zhenyu Dong², Zenghui Sun^{1, 3, 4, 5}, Hongchao Ma^{1, 3, 4, 5} and Lei Shi^{1, 3, 4, 5}

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

²College of Forestry, Shenyang Agricultural University, Shenyang, 110866, China ³Key Laboratory of Degraded and Unused Land Consolidation Engineering, the Ministry of Land and Resources, Xi'an 710075, China

⁴Shaanxi Provincial Land Engineering Construction Group Co., Ltd, Xi'an 710075, China

⁵Shaanxi Provincial Land Consolidation Engineering Technology Research Center, Xi'an 710075, China

*Corresponding author e-mail: liujinbao1212@163.com

Abstract. In this study, the mercury content of 44 soil samples in Guan Zhong area of Shaanxi Province was used as the data source, and the reflectance spectrum of soil was obtained by ASD Field Spec HR (350-2500 nm) Comparing the reflection characteristics of different contents and the effect of different pre-treatment methods on the establishment of soil heavy metal spectral inversion model. The first order differential, second order differential and reflectance logarithmic transformations were carried out after the pre-treatment of NOR, MSC and SNV, and the sensitive bands of reflectance and mercury content in different mathematical transformations were selected. A hyperspectral estimation model is established by regression method. The results of chemical analysis show that there is a serious Hg pollution in the study area. The results show that: (1) the reflectivity decreases with the increase of mercury content, and the sensitive regions of mercury are located at $392 \sim 455$ nm, 923 nm \sim 1040nm and 1806nm ~ 1969nm. (2) The combination of NOR, MSC and SNV transformations combined with differential transformations can improve the information of heavy metal elements in the soil, and the combination of high correlation band can improve the stability and prediction ability of the model. (3) The partial least squares regression model based on the logarithm of the original reflectance is better and the precision is higher, Rc2 = 0.9912, RMSEC = 0.665; Rv2 =0.9506, RMSEP = 1.93, which can achieve the mercury content in this region Quick forecast.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Published under licence by IOP Publishing Ltd 1

1. Introduction

In recent years, with the continuous development of industrialization and urbanization, soil heavy metal enrichment increased, seriously affecting crop yield (C.Y. HUANG, 2011; X.Z. LI, 2002). At present, the determination of heavy metal content in soil is mainly based on field sampling, based on indoor chemical experiments on the basis of different heavy metal content. Because of the high cost and low efficiency of traditional heavy metal element measurement methods, different elements need different chemical treatment, can't meet the large area of soil heavy metal pollution monitoring. Hyperspectral technology has been widely used in the prediction of heavy metal content because of its abundant information, time-saving and non-destructive sample structure, but less research on mercury (Hg) content prediction. Used the partial least squares regression model to estimate the Cr and Zn contents in the Rhine basin (KOOISTRA L, 2001). Gr, Cu, Pb, Ni, Zn in the soil of Tarnowskie Gory, Poland, in the near infrared and mid-infrared diffuse reflectance regions (GRZEGORZ S, 2004). Used the partial least squares method to establish the hyperspectral estimation model of soil mercury content in Zhundong coalfield, and the results show that the model based on the first order differential of reflectivity is the best (Y. SONG, et al, 2015). Other elements in Fuzhou soil Cu, Pb, Cr and other elements as the data source, the use of geographical weight regression model to achieve the region of soil heavy metal content inversion (Z.L. JIANG, et al. 2017). The results show that the Cr, Cd, Zn, Cu and Pb heavy metals in the soil of Xinzheng City, Henan Province are in good agreement (Q.X. ZHANG, et al. 2017). Established the regression model of heavy metals in the soil by using different forms of differential transformations combined with partial least squares regression and multiple stepwise linear regression method (K.TAN, et al, 2014).

The soils in different regions are different due to different pollution sources, and the soil spectrum characteristics are different. The pre-processing method and the modeling method are different when the prediction model is established. The established inversion model is limited to the selected research area. Therefore, when predicting the heavy metal content of soil in the small area, based on the characteristics of the soil reflectance spectrum in the study area, a variety of pre-treatment and comparative analysis of the established inversion model can be used to establish the accurate heavy metal element content estimation model. In this study, the mercury elements in Fufeng County, Yangling County and Wugong County of Shaanxi Province were used as data sources. Normalization (NOR), Multiplicative Scatter Correction (MSC), Standard Normal (FD), second order differential (SD) and reciprocal logarithm (LOG) transformations were carried out by using the Savitzky - Golay convolution smoothing method to smooth and decrease the spectral curves. And the hyperspectral estimation model of Hg was established by using partial least squares regression (PLSR).

2. Materials and methods

2.1. sample collection and determination of elemental content

In this study, the main soil types were soil type in Fufeng County, Yangling County and Wugong County, Shaanxi Province. The soil samples were collected according to the "S" -shaped sampling method. The sampling depth was the thickness of the tillage layer, usually 0-30cm. A total of 44 soil samples were sampled. The samples were air-dried and passed through the 2-mm hole. A 200-g soil sample was mixed and passed through a 100-mesh sieve for indoor heavy metal content and another soil sample was used for soil reflectance spectroscopy. The content of mercury (Hg) was determined by inductively coupled plasma mass spectrometry (ICP-MS, Agilent 7700). The statistical characteristics are shown in Table 1.

Elements	Samples	Maximum (mg/kg)	Minimum (mg/kg)	Means (mg/kg)	SD (mg/kg)	C.V
Hg	44	23.9	6.5	12.6	5.2311	0.415163842

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

2.2. Hyperspectral data determination

The soil reflectance spectroscopy was performed in the field using a high density reflective probe equipped with an ASD Field Spec HR Terrestrial Spectrometer. The wavelength range of the spectrometer is $350 \sim 2500$ nm, the sampling bandwidth is 1.3 nm ($350 \sim 1000$ nm) and 2 nm ($1000 \sim 2500$ nm), and the sampling interval is 1 nm. High-density reflective probes can effectively avoid the effects of soil stray light and eliminate the effects of weather. 2cm front view area can avoid the soil in the stone tablets, crop roots, etc., see the parameters shown in Table 2.

Table 2. The technical spaces of high intensity contact prob

10"(25.4 cm)	
1.5 lbs	
12 ~ 18 VDC,6.5W	
Halogen bulb/1500h	
2901+/-10%K	
10mm	
	10"(25.4 cm) 1.5 lbs 12 ~ 18 VDC,6.5W Halogen bulb/1500h 2901+/-10%K 10mm

3. Data analysis

3.1. sample collection and determination of elemental content

Use the ViewSpePro software to remove the jump spectrum curve and calculate the average as the actual reflection spectrum of the soil sample. Soil samples in the collection, processing and analysis process will introduce different degrees of error, will affect the late data analysis and modeling accuracy. The Mahalanobis distance is based on the multiple normal distribution, taking into account the covariance, mean and variance of the three factors, can be a comprehensive response to soil samples of the comprehensive indicators (SHI. Z, 2014). (Normalization, NOR), Multiplication Scatter Correlation (MSC), Standard Normal Variate (SNV) were used to eliminate the scattering between soil samples using TQ Analyst. Caused by the impact.

3.2. sample collection and determination of elemental content

The first order differential, second order differential and reflectance reciprocal logarithmic transformation of the original reflection spectrum were used to smooth and remove the reflection spectrum by using Savitzky - Golay convolution smoothing (Savitzky, A., Golay, J. E. 1964). In addition to direct analysis of soil spectral reflectance, three transformations were made for finding the response regions of different heavy metal elements. First order differential and second order differential transformations can increase the correlation between reflectivity and heavy metal elements while eliminating or limiting the influence of partial linearity or near linear background.

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

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

3.3. Data modeling and verification

Partial least squares regression (PLSR) is a new method of multivariate statistical data analysis, which solves the problem that the number of samples is less than the number of variables, effectively reducing the number of variables High degree of linear correlation problems, with partial least squares regression method more effective. The model is model and verification set, and the prediction model is established by PLSR method. The model results are verified by the decision factor R2 and the root mean square error RMSE.

$$R^{2} = \frac{\sum_{i=1}^{n} (\hat{y}_{i} - \overline{y}_{i})^{2}}{\sum_{i=1}^{n} (y_{i} - \overline{y}_{i})^{2}}$$
(3)

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

Where is the mean value of the sample observations, m is the number of samples for the calibration set, and n is the number of samples to be verified.

4. Results and discussion

4.1. Overview of the study area

In this study area, soil types were soil and 44 soil samples were collected. Among them, the soil pH value between $7.47 \sim 8.38$, belonging to alkaline soil. The content of soil Hg is between 6.5 and 23.9 mg.kg-1, which is strongly enriched, and the content of all samples exceeds the secondary soil quality standard.

The spectral curves of five different mercury contents in 44 soil samples were randomly selected, and the trend of each spectral curve was similar. Compared with different mercury content of the spectral curve, the higher the mercury content, the lower the spectral reflectance. In the range of $350 \sim 543$ nm, the reflectivity increases rapidly with the increase of wavelength, and the curve is steep. In the range of $535 \sim 729$ nm, the curve growth trend is slightly slowed down. $729 \sim 1350$ nm and $1495 \sim 1797$ nm range, the reflectivity gradually increased, but the trend is relatively slow. In the 1400nm, 1900nm, 2200nm position appears three more significant absorption peak, is caused by the soil moisture.



Figure 1. The characteristics of spectral curves of different Hg content

4.2. Correlation analysis

In order to extract the correlation between mercury content and spectral reflectance under different mathematical transformations, the first order differential, second order differential and reflectance reciprocal logarithmic transformation are analyzed for the correlation with mercury content. The results are shown in Fig. 2 The Compared with the original reflection spectrum, the reflection spectrum of the differential transformation, the correlation significantly improved. From the

differential transformation method, the first order differential effect of the reflection spectrum is the best, followed by the second order differential.

The correlation between the original reflectance and the mercury content is the smallest, and the correlation coefficient is between $1350 \sim 1440$ and $1813 \sim 1969$ nm, but this interval is influenced by the water. In the visible region 450nm, 935nm, 1039,1141nm correlation coefficient peak, indicating that this interval is affected by mercury content, in the modeling of this band is conducive to the model construction. The correlation between the reflectance spectrum and the mercury content after differential transformation is obviously enhanced, and there is a strong correlation between the visible and near infrared regions. The correlation coefficients were significantly correlated in the range of 392nm ~ 407nm, $693 \sim 785$ nm, $923 \sim 929$ nm, $1806 \sim 1897$ nm, and the positive correlation coefficient reached the maximum at 2331nm, which was 0.81 The In the visible region, the positive correlation coefficient is larger than the negative correlation coefficient, and the correlation coefficient is positively correlated, indicating that the mercury content is positively correlated with the first order differential of the reflectance. After the second order differential transformation of the reflectivity, the correlation coefficient in the range of 760nm, 1363nm, $1846 \sim 1898$ nm, $2409 \sim 2468$ nm is above 0.4, which is significant correlation. Through the correlation coefficient curve, it can be found that the first order differential of the reflectivity, the second order differential transformation and the mercury content of the correlation coefficient between the existence of more 0 value, which shows that although the differential transform can enhance the signal, but also inhibit the spectrum itself Useful information. The reflectivity is negatively correlated with the mercury content after reciprocal logarithmic transformation. At the 433nm, 935nm, 1145nm, 1868nm there are four inflection points, and reached the maximum at 1868nm, -0.5. Based on the above results, it can be concluded that the sensitive regions of mercury are located at 392 ~ 455nm, 923nm ~ 1040nm and 1806nm ~ 1969nm.



Figure 2. The Correlation of Hg content based on different transformations

4.3. Model establishment and comparison

Since the number of samples in this study is only 44, in order to reduce the multiple correlation between the data, the partial least squares regression method is used to establish the estimation model of mercury content. Firstly, three kinds of pre-processing of NOR, MSC and SNV were carried out.

The partial least squares regression model was established by using the first order differential, second order differential and reflectance reciprocal logarithmic transformation. The Use the determination coefficient (R^2) and the root mean square error (RMSE) for testing. The larger the modeling decision coefficient, the smaller the root mean square error, indicating that the better the stability of the model. The larger the prediction coefficient, the smaller the root mean square error, indicating that the model predicts the stronger. At the same time, in order to avoid over- fitting, the model, the dimension of the model's independent variable is as small as possible.

The partial least squares regression results for different spectral indices R, FD, SD and log (1 / R)are shown in Table 3. By comparing the established regression model, it can be concluded that the model correlation coefficient based on the 16 transformations is 0.8 or higher in the modeling group, and the modeling effect is good. In the prediction group, except for S + C and SNV + LOG, the prediction correlation coefficient reached 0.8 or higher, and the inversion effect was better. In addition to MSC, the PLSR model based on the original spectral, NOR and SNV pre-treatment is better than the response based on the original data. Compared with different transformation methods modeling and prediction results can be drawn, modeling the effect of good model, the prediction effect is not necessarily the best. In the modeling group, the model established by combining the second order differential transform with SNV treatment is the best, $Rc^2 = 0.9932$ and RMSEC = 0.585, but the inversion result is $Rv^2 = 0.8876$ and RMSEP = 3.17. Compared with Rc2, Rv^2 , RMSEC and RMSEP, the best way to deal with the modeling effect is to convert the original reflection spectrum to Savitzky-Golay nine-point smoothing in all the optimal PLSR models: $Rc^2 = 0.9912$, $Rv^2 = 0.9506$, RMSEC = 0.665, RMSEP = 1.93. Based on this modeling method, the inversion model, the measured value and the predicted value are evenly distributed on both sides of the modeling set and the verification regression tract, and the regression lines of the two sets of data tend to be parallel (Fig. 3). The results show that the inversion model established by the logarithmic logarithmic transformation combined with the partial least squares regression of the reflectivity is smooth and the rapid detection of field mercury content can be achieved.

Elementa	Pre-treted	Calibration			Validation	
Elements		Rc^2	RMSEC	Pc	Rv2	RMSEP
	S+C	0.9366	1.76	7	0.6281	4.47
	C+FD	0.9641	1.34	4	0.9105	2.83
	C+SD	0.9654	1.47	4	0.9005	3.14
	C+LOG	0.9912	0.665	10	0.9506	1.93
	NOR+S	0.9883	0.766	10	0.9541	2.10
	NOR+FD	0.9699	1.22	5	0.8973	2.74
	NOR+SD	0.9870	0.807	7	0.9074	3.02
Цa	NOR+LOG	0.9826	0.935	9	0.9384	2.33
пg	MSC+S	0.9898	0.718	10	0.8389	3.17
	MSC+FD	0.8716	2.47	1	0.8748	4.03
	MSC+SD	0.9894	0.731	7	0.9041	2.82
	MSC+LOG	0.9235	1.93	6	0.8488	3.21
	SNV+S	0.9029	2.16	5	0.8441	3.27
	SNV+FD	0.9929	0.599	6	0.9317	2.20
	SNV+SD	0.9932	0.585	8	0.8876	3.17
	SNV+LOG	0.8068	2.97	4	0.6428	4.36

Table 3. The result of Partial least squares regression based on different pre-treatment methods



Figure 3. Comparison between measured and predicted of soil Hg content

5. Conclusion

In this study, the original hyperspectral estimation model of Hg was established by performing the first order differential, second order differential and reflectance reciprocal logarithm transformation, and the partial least squares regression method. By comparing the effects of different pre-treatment methods on the establishment of soil heavy metal spectral inversion model, the following conclusions are obtained:

(1) The Hg content in this study area was seriously exceeded, which posed a serious threat to the normal growth of animals and plants in the area.

(2) The reflection spectra are processed by NOR, MSC and SNV respectively. The first order differential, second order differential and reflectance reciprocal logarithmic transformation are used to reduce the influence of external factors such as surface particle size and other factors. At the same time, the differential transformation can help to improve the correlation between the heavy metal elements and the reflection spectrum in the soil, and the use of the higher correlation band can significantly improve the stability and prediction ability of the model. After differential transformation, it can be concluded that the sensitive regions of mercury are located at $392 \sim 455$ nm, 923nm ~ 1040 nm and 1806nm ~ 1969 nm.

(3) According to the modeling and prediction results can be drawn, modeling good effect of the model, the forecast is not necessarily the best. The partial least squares regression model established by the reciprocal logarithmic transformation of Savitzky-Golay nine-point smoothing is the highest, and the prediction effect is better than 0.95, $Rc^2 = 0.9912$, $Rv^2 = 0.9506$, RMSEC = 0.665, RMSEP = 0.193, can achieve rapid prediction of mercury content in the region.

Acknowledgments

This work was financially supported by The Key of Science and Technology Innovation Team Projects of Shaanxi Province, China (2016KCT-23). We greatfully acknowledge the Institute of Land Engineering and Technology, Shaanxi Provincial Land Engineering Construction Group, Xi'an, China.

References

- [1] C.Y. HUANG, J.M. XU. Soil Science. China Agriculture Press, 2011.
- [2] X.Z. LI. The Damage of Heavy Metals Lead and Cadmium to Soil and Crops with Reference to the Control Measures. Economic Forest Researches, 2002, 18 (4):12 - 13.
- [3] KOOISTRA L, WEHRENS R, LEUVEN R S E W, et al. Possibilities of visible-near-infrared spectroscopy for the assessment of soil contamination in river floodplains [J]. Analytica Chimica Acta, 2001, 446: 97–105.
- [4] GRZEGORZ S, GREGORY W M, TOMASZ I S, et al. Near and mid-infrared diffuse reflectance spectroscopy for measuring soil metal content [J]. Journal of Environment Quality, 2004, 33: 2056 – 2069.
- [5] Y. SONG, C.B. LI, Y.J. HOU, et al. PLSR Based Hyperspectral Remote Sensing Retrieval of

Soil Hg Content. Geography and Geo-Information Science, 2015,31 (3): 44 – 53.

- [6] Z.L. JIANG, Y.S YANG, J.M SHA. Application of GWR modeling hyperspectral prediction of soil heavy metals. ACTA GEOGRAPHICA SINICA, 2017,72 (3):533 – 544.
- [7] Q.X. ZHANG, H.B. ZHANG, H.J. ZHANG, et al. Hybrid Inversion Model of Heavy Metals with Hyperspectral Reflectance in Cultivated Soils of Main Grain Producing Areas. Transaction of the Chinese Society for Agricultural Machinery, 2017, 48 (3):148 – 155.
- [8] K.TAN, Y.Y. YE, P.J. Du, et al. Estimation of Heavy Metal Concentrations in Reclaimed Mining Soils Using Reflectance Spectroscopy. Spectroscopy and Spectral Analysis. 2014, 34 (12): 3317 – 3312.
- [9] SHI, Z, et al. 2014. Principle and Method of Hyperspectral Remote Sensing for Soil. Beijing: Science Press.
- [10] Savitzky, A., & Golay, J. E. (1964). Smoothing and differentiation of data by simplified least square procedure. Analytical Chemistry, 36, 1627 1639.