Feasibility of Estimating Heavy Metal Contaminations in Floodplain Soils Using Laboratory-Based Hyperspectral Data—A Case Study Along Le’an River, China

LIU Yaolin¹², LI Wei³, WU Guofeng¹², XU Xinguo¹
¹. School of Resource and Environmental Science, Wuhan University, 129 Luoyu Road, Wuhan 430079, China
². Key Laboratory of Geographic Information System, Ministry of Education, Wuhan University, 129 Luoyu Road, Wuhan 430079, China
³. Sichuan Zhongshui-chengkanyuan Surveying and Mapping Engineering Company, Chengdu 610072, China

Abstract It is necessary to estimate heavy metal concentrations within soils for understanding heavy metal contaminations and for keeping the sustainable developments of ecosystems. This study, with the floodplain along Le’an River and its two branches in Jiangxi Province of China as a case study, aimed to explore the feasibility of estimating concentrations of heavy metal lead (Pb), copper (Cu) and zinc (Zn) within soils using laboratory-based hyperspectral data. Thirty soil samples were collected, and their hyperspectral data, soil organic matters and Pb, Cu and Zn concentrations were measured in the laboratory. The potential relations among hyperspectral data, soil organic matter and Pb, Cu and Zn concentrations were explored and further used to estimate Pb, Cu and Zn concentrations from hyperspectral data with soil organic matter as a bridge. The results showed that the ratio of the first-order derivatives of spectral absorbance at wavelengths 624 and 564 nm could explain 52% of the variation of soil organic matter; the soil organic matter could explain 59%, 51% and 50% of the variation of Pb, Cu and Zn concentrations with estimated standard errors of 1.41, 48.27 and 45.15 mg•kg⁻¹; and the absolute estimation errors were 8%-56%, 12%-118% and 2%-22%, and 50%, 67% and 100% of them were less than 25% for Pb, Cu and Zn concentration estimations. We concluded that the laboratory-based hyperspectral data hold potentials in estimating concentrations of heavy metal Pb, Cu and Zn in soils. More sampling points or other potential linear and non-linear regression methods should be used for improving the stabilities and accuracies of the estimation models.

Keywords soil; heavy metal concentration; estimation; soil organic matter; hyperspectral data

CLC number P208

Introduction

Soil is an important resource for the survival of plants, animals and human races. With the social and economic developments, soil, however, is suffering from a series of problems, among which heavy metal contamination is especially outstanding for its...
potential harms to the health, survivals or activities of human or other living organisms.

Highly accumulated heavy metals within soils may disrupt the physical, chemical and biological balances of soils.[2] Heavy metals within soils could also be accumulated by plants and might be magnified within ecosystems because of their ability to be incorporated into food chains and their environmental persistence,[3] which might further impact the sustainable developments of ecosystems.[4] Thus, it is necessary to estimate heavy metal concentrations for understanding the heavy metal contaminations within soils and for keeping the sustainable developments of ecosystems.

Heavy metal concentrations within soils were traditionally estimated through soil sampling in field and sequent analyses in laboratory.[5,6,7] However, this method may be costly and time-consuming for intensive soil sampling in the field and analyses in laboratory. Moreover, such estimations can only provide limited information at certain locations and moments and could not describe the spatial and temporal dynamics of heavy metal concentrations over large areas.

Remote sensing techniques have many advantages compared to traditional method, such as lower cost, faster data acquisition and better spatial and temporal continuities; thus, they have been employed to estimate the heavy metal concentrations within soils. For example, Clevers et al.[6] used red-edge position of spectroscopic data to assess heavy metal contamination in river floodplains along the river Rhine in the Netherlands and concluded that spectral signatures with 1 to 3 nm spectral resolution might provide very fine spectral information for detecting metal-induced stress in river floodplains; Kemper and Sommer[8] estimated the heavy metal contamination in soils using reflectance spectroscopy in Spain and summarized that it was feasible to predict heavy metals in soils contaminated by mining residuals using the reflectance spectroscopy. Choe et al.[9] combined geochemistry, field spectroscopy and hyperspectral remote sensing to map heavy metal pollution in stream sediments in the Rodalquilar mining area, southeast Spain, and the results indicated the potential applicability of hyperspectral data in estimating heavy metal concentrations and the capacity of the image-derived spectral parameters to screen areas affected by heavy metals.

Among most of these researches, the heavy metal concentrations were derived directly from the spectral measurements for the amounts of heavy metals that were large enough to be detected by spectroradiometer. In many cases, the low heavy metal concentrations, however, cannot influence the spectroscopic characterizations of soils and cannot be detected directly by spectroradiometer. Thus, it would be both meaningful and challenging to explore the possibility of assessing heavy metal concentrations in soils using field-measured hyperspectral data combined with indirect strategies, which would complement the methods for estimating heavy metal concentrations in soils using field-measured hyperspectral data, airborne or spaceborne hyperspectral remote sensing images.

Soil organic matter (SOM) is one of the most important parameters influencing the spectral reflectance intensity and spectral feature of soil.[10] Several studies explored the possibility of assessing soil organic matter using remote sensing techniques. For example, Kooistra et al.[11] reported that the hyperspectral data measured in the field and laboratory could explain 45% and 69% of the variation of soil organic matter. Melendez-Pastor et al.[12] applied hyperspectral data to assess soil properties (including soil organic matter) in semi-arid soils in a coastal zone of southeast Spain, and the results indicated the great potential of hyperspectral data in estimating soil organic matter.

Kooistra et al.[13] applied visible-near-infrared spectroscopy to assess soil contamination in river floodplains in the Netherlands and concluded that, in floodplain soils, metal concentrations depended on the exchange capacity of the soil, which largely depends on the clay and soil organic matter content; therefore, these soil characteristics were directly correlated to metal concentrations in the soil. Although metals at low concentration levels did not have spectral features within the visible-near-infrared region, the metal concentrations could be predicted by the correlations between metal concentrations and soil organic matter content that can be estimated using remote sensing techniques. Thus, the method could be also applicable in other regions, but few researches were found to explore this possibility.

This study, with the floodplain along the Le'an
River and its two branches in Jiangxi Province of China as a case study, aimed to explore the possibilities of estimating the concentrations of heavy metal lead (Pb), copper (Cu) and zinc (Zn) in soils through the relations among heavy metal concentrations, soil organic matter and laboratory-based hyperspectral data.

1 Materials and methods

1.1 Study area

Le’an River (Fig. 1) is 279 km long and flows into Poyang Lake, which is the largest freshwater lake in China presently. The drainage area of Le’an River is around 8989 km², and it covers several nonferrous metal mining regions, such as Dexing Copper Mine and Yishan Lead-zinc Mine. Dexing Copper Mine is the largest outcrop copper mine in Asia, and Dawu River crosses this mine area and drains into Le’an River. Yishan Lead-zinc Mine is located in the downstream of Jishui River, which is the largest branch of Le’an River.

The mining in Dexing and Yishan mining areas produces large amounts of acid mine drainages, which contain rich heavy metals such as Pb, Cu and Zn. The drainages with high heavy metal concentrations enter Le’an River and further Poyang Lake through Dawu or Jishui River, which could induce serious environmental and ecological impacts on the waters, sediments and plants in these regions.

1.2 Field sampling

The fieldwork was carried out on 5 to 7 September 2007. Thirty soil samples were collected at 21 sampling sites in the floodplain of the study area, nine of which are along Le’an River, three along Dawu River and nine along Jishui River (Fig. 1).

For each sampling, the surface soil at depths of 2 cm to 5 cm was collected from the four corners and centre of a 2 m×2 m square; then, 500 g well-mixed soil was put in a strictly cleaned container. All soil samples were stored in vacuum flasks filled with ice for keeping soil fresh. There was no laboratory available to the determination of soil organic matter in the field, so all soil samples were taken to laboratory on the third day of the fieldwork. The ice in the vacuum flasks was changed every day.

1.3 Spectral reflectance measurement

An ASD FieldSpec Pro FR portable spectroradi-
ometer with wavelength of 350 to 2500 nm was used to measure the spectral reflectance of soil samples, and its sampling interval and spectral resolution are 1.4 and 3 nm for the 350 to 1050 nm range and 2 and 10 nm for the 1050 to 2500 nm range, respectively (http://www.asdi.com). In the laboratory, for each soil sample, the spectral radiance over a standardized white spectralon panel was measured first; then, the spectral radiance over soils was measured 10 times and their mean value was calculated and recorded as the spectral radiance of the soil sample. Through dividing the mean radiance of leaves by the radiance over spectralon panel, the spectral reflectance of each soil sample was derived.

### 1.4 Soil organic matter and heavy metal concentration (HMC) laboratory analyses

The soil samples were first air-dried in the laboratory. The dried soil samples were then gently crushed in a porcelain mortar to break up large aggregates and sieved using a stainless steel sieve with 0.25 mm mesh. Finally, for each soil sample, the organic matter was analyzed using Walkley-Black method,[17] and the concentrations of Pb, Cu and Zn were determined with an TAS-990 atomic absorption spectrophotometer, respectively.

### 1.5 Sampling pre-processing

Thirty samples were randomly divided into two sub-datasets, one of which with 20 samples (sub-dataset 1) was used for developing models and another with 10 samples (sub-dataset 2) for validating the developed models. Five sampling points located at the upstreams of Jishui River and Dawu River were not used to analyze the relations between the heavy metal concentrates and soil organic matter because they were not affected by mining-induced heavy metals. A box-plot analysis was taken to explore the distribution of soil organic matter, Pb, Cu and Zn concentration values, respectively, and all the values larger than double variance were removed from the following analyses.

### 1.6 Model development and validation

The following method used by Krishnan et al.[18] was employed to explore the relation between the soil organic matter and the spectral reflectance using sub-dataset 1. For each soil sample, the spectral reflectance was transformed to spectral absorbance \((\log_{10}(1/\text{reflectance}))\), and the first-order derivative of absorbance was calculated. The simple linear model \((f_1)\) of soil organic matter against the ratio of first-order derivatives of spectral absorbance at wavelengths 624 and 564 nm \((\text{fod}(624)/\text{fod}(564))\) was used to predict soil organic matter as Mod 1:

\[
\text{SOM} = f_1(\text{fod}(624)/\text{fod}(564))
\]

Then, the linear regression model \((f_2)\) between the heavy metal concentrate (Pb, Cu and Zn) and soil organic matter was analyzed by using sub-dataset 1 with bootstrapping method, which is suitable for the regression with small sample size[19] as Mod 2:

\[
\text{HMC} = f_2(\text{SOM})
\]

Finally, the following combination of Mod 1 and Mod 2 was used to estimate the concentrations of heavy metal Pb, Cu and Zn from hyperspectral data.

We used sub-dataset 2 to evaluate the performance of Mod 3.

\[
\text{HMC} = f_2(f_1(\text{fod}(624)/\text{fod}(564)))
\]

The Mod 3 was first applied to sub-dataset 2 to estimate the concentrations of heavy metal Pb, Cu and Zn. Then, the performance of Mod 3 was assessed by comparing the estimated values with the measured ones through the coefficient of determination \((R^2)\) of the regression line between the estimated and the measured heavy metal concentrations and absolute estimation error.

### 2 Results

#### 2.1 Soil organic matter and heavy metal concentrations

Table 1 shows the variability of soil organic matter and Pb, Cu and Zn concentrations of 30 soil samples. The soil organic matter varies from 0.398% to 9.287%, and Pb and Cu concentrations in soils hold higher variation coefficients compared with Zn concentration, which means that the these two elements show larger variation in concentration than Zn.

Five sampling points located at the upstreams of Jishui River and Dawu River were not used to analyze the relations between the heavy metal concen-
trates and soil organic matter and the values larger than double variance were removed; thus, 19 samplings were used to develop the relation between soil organic matter and hyperspectral data, 14, 15 and 16 samplings were employed to regress soil organic matter to Pb, Cu and Zn concentrations and 6, 6 and 8 samplings were applied to validate the results of Pb, Cu and Zn concentration estimations, respectively.

Table 1  Statistics describing the variability of soil organic matter (%) and heavy metal (Pb, Cu and Zn) concentrations (mg/kg) of 30 soil samples

|       | SOM | Pb      | Cu     | Zn     |
|-------|-----|---------|--------|--------|
| Minimum | 0.398 | 2.480  | 1.670  | 73.76  |
| Average | 3.095 | 10.795 | 89.411 | 199.79 |
| StdDev  | 1.912 | 97.487 | 14.222 | 62.02  |
| CoeVar  | 0.618 | 1.090  | 1.317  | 0.310  |

Note: StdDev means standard derivation, and CoeVar is coefficient of variation.

2.2 Relation between soil organic matter and hyperspectral data

Using sub-dataset 1, we found that the ratio of the first-order derivatives of spectral absorbance at wavelengths 624 and 564 nm could explain 52% of the variation of soil organic matter with estimated standard error (s.e.) of 1.14% (Eq. (1)). F-test showed that the regression model was statistically significant at a significant level of 0.001.

\[ \text{SOM} = 11.5 \times \text{fod}(624) / \text{fod}(564) \ (R^2=0.53, \ s.e.=1.14, \ F=18.99, \ P<0.001, \ df=18) \]  (1)

2.3 Relation between heavy metal concentrations and soil organic matter

The positive correlations were found between Pb, Cu and Zn concentrations and soil organic matter from sub-dataset 1, and the regression models best fitting the relation between Pb, Cu and Zn concentrations and soil organic matter are shown in Eqs. (2), (3) and (4). The soil organic matter could explain 59%, 51% and 50% of the variation of Pb, Cu and Zn concentrations with estimated standard error (s.e.) of 1.41, 48.27 and 45.15 mg•kg⁻¹, respectively, and F-tests showed that the three regression models were statistically significant at a significant level of 0.05.

\[ \text{Pb} = 2.973 + 1.006 \times \text{SOM} \ (R^2=0.59, \ s.e.=1.41, \ F=17.05, \ P=0.001, \ df=13) \]  (2)
\[ \text{Cu} = -4.701 + 29.411 \times \text{SOM} \ (R^2=0.51, \ s.e.=48.27, \ F=13.57, \ P=0.003, \ df=14) \]  (3)
\[ \text{Zn} = 119.03 + 27.709 \times \text{SOM} \ (R^2=0.50, \ s.e.=45.15, \ F=14.12, \ P=0.002, \ df=15) \]  (4)

2.4 Estimation models of heavy metal concentrations

On the basis of the above-mentioned relations among heavy metal concentrations, soil organic matter and hyperspectral data, the Pb, Cu and Zn concentrations can be thus derived through the ratio of the first-order derivatives of spectral absorbance at wavelengths 624 and 564 nm as follows:

\[ \text{Pb} = 2.973 + 1.006 \times \text{fod}(624) / \text{fod}(564) \ (R^2=0.59, \ s.e.=1.41, \ F=17.05, \ P=0.001, \ df=13) \]  (5)
\[ \text{Cu} = -4.701 + 29.411 \times \text{fod}(624) / \text{fod}(564) \ (R^2=0.51, \ s.e.=48.27, \ F=13.57, \ P=0.003, \ df=14) \]  (6)
\[ \text{Zn} = 119.03 + 27.709 \times \text{fod}(624) / \text{fod}(564) \ (R^2=0.50, \ s.e.=45.15, \ F=14.12, \ P=0.002, \ df=15) \]  (7)

Eqs. (5), (6) and (7) were applied to sub-dataset 2 to estimate Pb, Cu and Zn concentrations from the measured hyperspectral data, respectively, and the

![Fig. 2 Scatter plots of measured vs. estimated Pb, Cu and Zn concentrations (mg•kg⁻¹)](image-url)
estimated concentrations of Pb, Cu and Zn were compared with the measured ones (Fig. 2). Significantly positive correlations between the estimated and the measured values were observed for Pb \( (r=0.81, p=0.05) \), Cu \( (r=0.89, p=0.02) \) and Zn \( (r=0.78, p=0.02) \) concentrations at a significant level of 0.05. The absolute estimation errors are 8%-56%, 12%-118% and 2%-22%, and 50%, 67% and 100% of them are less than 25% for Pb, Cu and Zn concentration estimations, respectively.

3 Discussion

Many studies have been taken to explore the relation between spectrum and soil organic matter. For example, Krishnan et al.\(^{[18]}\) reported that the visible spectrum was more suitable than near-infrared spectrum for estimating soil organic matter. Lu et al.\(^{[20]}\) indicated that the soil organic matter content had a significant correlation with spectral reflectance in 545 to 830 nm at 0.05 significant level and also found that the first-order derivative of spectral reflectance had stronger correlation with soil organic matter than the original spectral reflectance.

In this study, we found that the ratio of the first-order derivatives of spectral absorbance at wavelengths 624 and 564 nm could explain 52% of the variation of soil organic matter. Such result confirms the former research results; also, it is comparable with the one obtained by Kooistra et al.\(^{[11]}\) who reported that the hyperspectral data measured in the field and laboratory could explain 45% and 69% of the variation of soil organic matter, respectively.

Many researchers have found potential relations between heavy metal contamination and soil organic matter. For instances, dissolved soil organic matter is one of important factors affecting heavy metal uptake by sludge particulates,\(^{[21]}\) and the level of heavy metal concentration depends on the exchange capacity of the soil, which is largely related to clay and organic matter content.\(^{[22]}\) Therefore, it is reasonable to observe the significantly positive relations of soil organic matter against Pb, Cu and Zn concentrations in this study.

While validating the models, we found that all estimated concentrations of Pb, Cu and Zn were significantly and positively correlated to the measured ones at a significant level of 0.05, while 50%, 67% and 100% of the absolute estimation errors were less than 25% for Pb, Cu and Zn concentration estimations, which indicates the feasibility and also reveals the great potentials of laboratory-based hyperspectral data in estimating concentrations of heavy metal Pb, Cu and Zn in soils.

Although we found the potentials of laboratory-based hyperspectral data in estimating concentrations of heavy metal Pb, Cu and Zn in soils in this study, the used method should be validated for being extended to other studies areas or heavy metal types. Only 30 samplings combined the process described by Krishnan et al.\(^{[18]}\) were used for model development and validation in this study, which might have impact on the stabilities of the models, and more sampling points or other potential linear and non-linear regression methods should be employed in the future to further improve the models.

4 Conclusion

In this study, we explored the feasibility of estimating heavy metal concentrations in soils using laboratory-based hyperspectral data along Le'an River, China. The principal results obtained can be summarized as follows:

(1) The ratio of the first-order derivatives of spectral absorbance at wavelengths 624 and 564 nm could explain 52% of the variation of soil organic matter.

(2) The Pb, Cu and Zn concentrations are significantly and positively related to the soil organic matter, which could explain 59%, 51% and 50% of the variation of Pb, Cu and Zn concentrations with estimated standard error of 1.41, 48.27 and 45.15 mg·kg\(^{-1}\).

(3) The absolute estimation errors are 8%-56%, 12%-118% and 2%-22%, and 50%, 67% and 100% of them are less than 25% for Pb, Cu and Zn concentration estimations using laboratory-based hyperspectral data.

We concluded that the laboratory-based hyperspectral data hold potentials in estimating concentrations of heavy metal Pb, Cu and Zn in soils with low heavy metal concentrations. More sampling points and other potential linear or non-linear regression methods should be employed for improving the stabilities and
accuracies of the estimation models.

References

[1] Wild A (1993) Soils and the environment: an introduction [M]. New York: Cambridge University Press

[2] Alloway B J (1994) Heavy metals in soils [M]. Berlin: Springer

[3] Kishe M A, Machiwa J F (2003) Distribution of heavy metals in sediments of Mwanza Gulf of Lake Victoria, Tanzania[J]. Environment International, 28(7): 619-625

[4] Ozmen H, Kulahci F, Cukurovali A, et al. (2004) Concentrations of heavy metal and radioactivity in surface water and sediment of Hazar Lake (Elazig, Turkey)[J]. Chemosphere, 55(3): 401-408

[5] UNESCO (1996) Ecological effects of heavy-metal pollution in the Daxing copper mine region (Jian Xi province, China)[R]. The United nations Educational, Scientific and Cultural Organization.

[6] Clevers, Kooistra, Salas (2004) Study of heavy metal contamination in river floodplains using the red-edge position in spectroscopic data[J]. International Journal of Remote Sensing, 25(19): 3883-3895

[7] Dauvalter V, Rognerud S (2001) Heavy metal pollution in sediments of the Pasvik River drainage[J]. Chemosphere, 42(1): 9-18

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

[9] Choe E, van der Meer F, van Ruitenbeek F, et al. (2008) Mapping of heavy metal pollution in stream sediments using combined geochemistry, field spectroscopy, and hyperspectral remote sensing: a case study of the Rodalquilar mining area, SE Spain[J]. Remote Sensing of Environment, 112(7): 3222-3233

[10] Dematté J A M, Campos R C, Alves M C, et al. (2004) Visible-NIR reflectance: A new approach on soil evaluation[J]. Geoderma, 121(1-2): 95-112

[11] Kooistra L, Wanders J, Epema G F, et al. (2003) The potential of field spectroscopy for the assessment of sediment properties in river floodplains[J]. Analytica Chimica Acta, 484(2): 189-200

[12] Melendez-Pastor I, Navarro-Pedreño J, Gómez I, et al. (2008) Identifying optimal spectral bands to assess soil properties with VNIR radiometry in semi-arid soils[J]. Geoderma, 147(3-4): 126-132

[13] Kooistra L, Wehrens R, Leuven R S, et al. (2001) Possibilities of visible-near-infrared spectroscopy for the assessment of soil contamination in river floodplains[J]. Analytica Chimica Acta, 446(1-2): 97-105

[14] Liu W X, Coveney R M, Chen J L (2003) Environmental quality assessment on a river system polluted by mining activities[J]. Applied Geochemistry, 18(5): 749-764

[15] Zeng F, Xiao H, Zhou W (2007) Spatial and temporal variations and their source analysis of Copper, Lead and Zinc in river-waters and sediments of the Le An River (in Chinese)[J]. Research of Environmental Sciences, 20(6): 14-20

[16] He M, Wang Z, Tang H (1998) The chemical, toxicological and ecological studies in assessing the heavy metal pollution in Le An River, China[J]. Water Research, 32(2): 510-518

[17] Walkley A L, Black A (1947) A critical examination of a rapid method for determination of organic carbon in soils-effect of variations in digestion conditions and of inorganic soil constituents[J]. Soil Science, 63: 251-163

[18] Krishnan P, Alexander J D, Butler B J, et al. (1980) Reflectance technique for predicting soil organic matter[J]. Soil Science, 44(6): 1282-1285

[19] Efron B (1979) Bootstrap methods: another look at the Jackknife[J]. Annals of Statistics, 7(1): 1-26

[20] Lu Y, Bai Y, Yang L, et al. (2007) Prediction and validation of soil organic matter content based on hyperspectrum (in Chinese) [J]. Scientia Agricultura Sinica, 40(9): 1989-1995

[21] Wang J, Huang C P, Allen H E, et al. (1999) Effects of dissolved organic matter and pH on heavy metal uptake by sludge particulates exemplified by Copper (II) and Nickel (II): three-variable model [J]. Water Environment Research, 71: 139-147

[22] Middelkoop H (1997) Embanked floodplains in the Netherlands: geomorphological evolution over various time scales [D]. Universiteit Utrecht