

# Estimate of Heavy Metals in Soil with Non-Soil Removed

Ji Jian, Yao Fang\*, Wei-Le Li, Qian-Yu Chen, Heng-Yu Tian, Shuang-Lin You

Key Laboratory of Geoscience Spatial Information Technology, Ministry of Land and Resources of the P.R. China, Chengdu University of Technology, Chengdu, China  
Email: \*fybyn221@163.com

**How to cite this paper:** Jian, J., Fang, Y., Li, W.-L., Chen, Q.-Y., Tian, H.-Y. and You, S.-L. (2017) Estimate of Heavy Metals in Soil with Non-Soil Removed. *Journal of Data Analysis and Information Processing*, 5, 140-155.

<https://doi.org/10.4236/jdaip.2017.54011>

**Received:** September 2, 2017

**Accepted:** November 6, 2017

**Published:** November 9, 2017

Copyright © 2017 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Quantifying and mapping heavy metals' concentrations in the soil are important in monitoring and managing heavy metal pollution in the mining areas. However, the cover on the soil acts as a balk when retrieving information from soil. In order to retrieve heavy metal pollution precisely and quickly from hyperspectral images, this study presents a new method to remove non-soil information based NDVI from hyper-spectral and multi-spectral images. The method assumed that the mixed objects in each pixel of remote sensing images are composed only of soil and vegetation-based non-soil end-generational endmembers, then, the soil information of each pixel can be compensated with the non-soil information removed based on its NDVI. Thus, the soil DN value can be corrected to retrieve soil information more precisely. The method has been used on the Hyperion image in June 8, 2002 and the Gaofen-2 (GF-2) image in February 14, 2016 to retrieve the heavy metals' contents in Bai-ma and De-sheng mining areas, Miyi County, Sichuan Province. From the non-soil information removed images, the  $R^2$  and RMSE of the models of estimating Cr, Ag, Cu and Ba in soil are 0.68, 0.724, 0.71, 0.695 and 75.96, 0.03, 52.88, 284.70 respectively. From the original images, the  $R^2$  and RMSE of the models of estimating Cr, Ag, Cu and Ba in soil are 0.67, 0.385, 0.425, 0.406 and 80.11, 0.18, 53.43, 396.49 respectively. The retrieval results show that the non-soil information removed images are superior to original images in soil heavy metals' contents retrieval. This indicates that this method is feasible, and it can be used in soil information retrieval.

## Keywords

Non-Soil Information Removal, Heavy Metal, NDVI, Spectral Unmixing

## 1. Introduction

The pollution of heavy metal in soil has aroused extensive attention in the min-

ing area. The main reason is that the soil contaminated by heavy metals make a great influence on the local citizens and a great threat to the ecological system [1] [2] [3]. However, the traditional method of estimating the heavy metals by collecting sampling points and analyzing in the laboratory is time-consuming and expensive. Therefore, new methods are needed for monitoring and managing heavy metal pollution in the mining areas. With the improvement in remote sensing technology, the new method of establishing inversion models with the heavy metals' concentrations and the feature spectra from hyperspectral and high spatial resolution images are popular to monitor heavy metal in soil [4] [5] [6] [7]. Many studies have shown that the Visible and Near-infrared spectra of soil obtained from remote sensing images were effectively used for retrieving real-time heavy metals' concentrations in large mining areas [3] [8] [9].

However, the spectral mix of the pixel is present on both high spatial resolution image and hyperspectral image. This will make a great influence on the results [10] [11]. The reasons are mainly: 1) the pixels usually contain a variety of ground information for the limited spatial resolution of the sensor; and 2) the pixels are of complexity in composition, organization and structure. Even though an image pixel size is smaller than the objects under study, the spectral features could still be mixed with those of adjacent objects when the pixels are on the edge of the objects [12] [13]. Therefore, the spectral unmixing methods are widely studied in the past decade. Broadly, the spectral unmixing methods can be categorized as linear mixture model (LMM) and nonlinear mixture model (NLMM) according to the mathematical formulation. However, NLMM is used not as widely as the LMM for its complexity and difficulty in obtaining parameters [10] [14]. The LMM is often used to map the urban impervious surface and the target materials with the high spatial resolution images, identify the exposed soil, and map the heavy metals' contents from remote sensing images [15] [16] [17] [18] [19]. And, most of the existing unmixing algorithms are based on the LMM, which assumes that the spectrum as a linear combination of pure spectra of all elements within a pixel and assumes that no significant multiple scattering exists among different elements [20], such as non-local sparse unmixing (NLSU), blind spectral unmixing method based on sparse component analysis (BSUSCA), Structured Sparse regularized Nonnegative Matrix Factorization (SS-NMF) and so on [21] [22] [23]. However, these studies are of low accuracy, the primary reason is that the non-soil information has not been removed effectively for that the vegetation and the soil have strong correlations when processing.

In this paper, a new method for removing non-soil information based on NDVI is proposed. The method tends to obtain a non-soil information removed image from original image by compensating the missed soil information. Then, the compensated soil information can be used to retrieve and map the heavy metals' concentrations more precisely. The highlights of this paper are as follows:

1) The non-soil information removal method supposes that there are only two endmembers: soil and non-soil in original images. This can avoid the influence of the correlation between the non-soils and lower the influence of non-soil information on the soil information extracting.

2) The non-soil information can be maximum remove by calculating the maximum value of the product of DN and the non-soil abundance.

3) The missed soil information caused by atmospheric and NDVI can be compensated to make sure the precision of the soil information be better.

## 2. Methodology

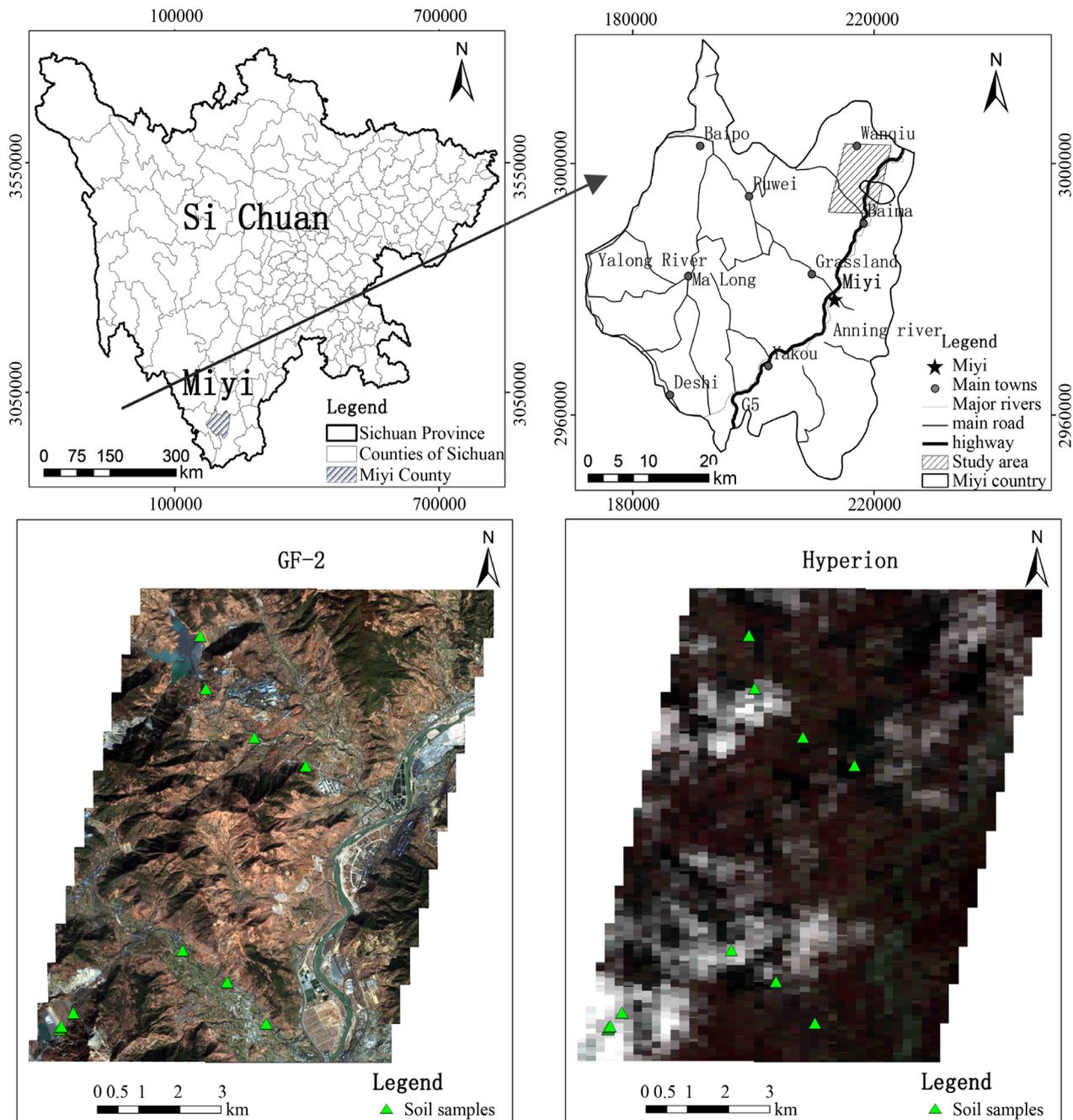
### 2.1. Study Area and Data

The mining areas of Bai-ma and De-sheng located in the northeast of Miyi County are selected as the study area (**Figure 1**). Miyi is located in the north of the Panzhihua city, the southwest of Sichuan Province. Bai-ma and De-sheng are the most important mining areas in Miyi Country. The study area is rich in different kinds of mineral resources, such as vanadium-titanium magnetite ore, coal, limestone, dolomite, refractory clay and so on. The waste water containing heavy metals such as Ag and Cu from the mining areas can pollute the local soil.

The spectra of 16 soil samples were collected at about 5 cm below the soil surface in June 2015; the locations of the sample points are shown in **Figure 1**. The contents of Cr, Ag, Ba and Cu of the 16 soil samples had been chemically analyzed by conventional digestion methods using Inductively Coupled Plasma Mass Spectrometer (ICP-MS) (**Table 1**). The ICP-MS is the most popular ion source in analytical chemistry for elemental mass spectrometry. In ICP-MS a mass spectrometer is coupled to an ICP torch by an interface including sampler and skimmer cones so that representative samples of the plasma can be transmitted through its orifices to the mass analyzer through its orifices to effectively eliminate atomic ions Interference, reduce the detection limit of susceptible elements [24].

The spectra of soil samples were obtained from a high spectral resolution ASD Fieldspec III spectroradiometer. Before taking any observation, the spectroradiometer was calibrated with white spectrum in order to minimize the effect of change in sun illumination. The spectral range of this instrument is from 350 nm to 2500 nm. In the range of 350 - 1000 nm, the spectral resolution is 3 nm with 1.4 nm of sampling interval, and in the range of 1000 - 2500 nm, the spatial resolution is 10 nm with 2 nm of sampling interval. Each sample was measured three times and the average value was calculated afterwards for the feature spectra selection.

Hyperion image in June 8, 2002 and GF-2 image in February 14, 2016 are collected as the remote sensing data source. The spectral range of Hyperion image is from 400 nm to 2500 nm, which is divided into 242 bands. The spectral range of GF-2 image is from 450 nm to 890 nm, which is divided into 4 bands. Soil sample's spectrum of images was obtained by using the position of sample's row



**Figure 1.** Study area and soil samples (Hyperion (R: band30, G: band21, B: band13) and GF-2 (R: band4, G: band3, B: band2)).

**Table 1.** Chemical analysis results of Cr, Ag, Cu and Ba (unit: mg/kg).

	Cr	Ag	Cu	Ba
Max	520.842	0.276	233.238	1757.989
Min	65.147	0.101	39.84	173.551
Mean	192.235	0.18	112.765	837.43
SD	135.36	0.0507	70.55	473.26

and column at the images.

## 2.2. Data Pre-Processing

In general, most heavy metals would not be expected to produce spectral absorption in the visible and near-infrared regions. However, heavy metal ions can be adsorbed by inorganic and organic matter, therefore, the prediction of heavy metals' content can be performed indirectly by visible and near infrared [3] [25]. The field soil spectral measurements were pre-processed using the View SpecPro to enhance the spectral features in the mining area. The pre-processing of field spectra includes smoothing and averaging.

The digital remotely sensed image may contain noise or error that was introduced by the sensor system (e.g., electronic noise) or the environment (e.g., atmospheric scattering of light into the sensor's field of view) [26]. This can result in radiation and geometric distortion to the image [27]. Therefore, image pre-processing is required to correct the geometric and radiometric distortion in the original image of the study area. In this paper, five main pre-processing steps have been taken to correct Hyperion image including bad lines removal, uncalibrated and water vapor absorption bands removal, streaks removal, smile effect correction and atmospheric correction, two main pre-processing steps have been taken to correct the GF-2 image including radiation correction and atmospheric correction.

## 2.3. Non-Soil Information Removal Method Based on NDVI

NDVI is an effective index for detecting above-ground vegetation conditions, for that seasonal and inter-annual changes in vegetation growth and activity can be monitored, and the ratio reduces many forms of multiplicative noise (Sun illumination differences, cloud shadows, some atmospheric attenuation, some topographic variations) present in multiple bands of multiple-data image [26] [28]. NDVI is defined as:

$$\text{NDVI} = \frac{\text{NIR} - \text{R}}{\text{NIR} + \text{R}} \quad (1)$$

where NIR and R are the DN values of the near-infrared and red region respectively. In this paper, band53, band30 and band4, band3 are selected as the near-infrared and red bands to calculate the NDVI of Hyperion and GF-2 images respectively.

Generally, the greater the amount of healthy green vegetation, the greater the NDVI value. However, the NDVI has a negative value when the land covered with clouds, snow or water, and the vegetation and soil are the main elements of the pixels. Thus, it is assumed that the mixed objects in each pixel are composed of only soil and vegetation-based non-soil end-generational endmembers (e.g. vegetation, water, cloud and snow. The influence of the terrain and the soil texture are not taken into account.). Then, the absolute value of NDVI can be used to calculate the non-soil "pure" pixel radiation and soil abundance quickly and

efficiently, which can be seen as the non-soil abundance. And the non-soil DN value can be calculated. Then, the soil DN value can be calculated by removing non-soil information according to the principle of physics to reduce the influence caused by the strong correlation of vegetation and soil. Finally, the soil DN value will be corrected by using the probability principle to compensate the missed soil information during non-soil removal.

Thus, the total abundance of a pixel is considered to be 1, and the abundance calculation of a pixel is:

$$f_{\text{soil}} + f_{\text{non-soil}} = 1 \quad (2)$$

where  $f_{\text{non-soil}}$  and  $f_{\text{soil}}$  are the abundance of non-soil and soil respectively, and the absolute value of NDVI of the pixel is taken as the value of  $f_{\text{non-soil}}$ .

Next, the value of the non-soil “pure” pixel in a band ( $P_i$ ) can be calculated as Equation (3):

$$P_i = \max(D_i * f_{\text{non-soil}}) \quad (3)$$

where the  $D_i$  and  $f_{\text{non-soil}}$  are the DN value and the abundance value of non-soil of a pixel at  $i$  band,  $P_i$  is the maximum value of the product of  $f_{\text{non-soil}}$  (*i.e.* the value can be used to maximize the removal of non-soil after testing.), and  $D_i$  which is taken as the value of the non-soil “pure” pixel for  $i$  band.

Then, the missed soil information can be compensated with  $P_i$  as Equation (4):

$$D_{\text{soil}} = \sum_{i=1}^n \frac{D_i - f_{\text{non-soil}} * P_i}{f_{\text{soil}}} \quad (4)$$

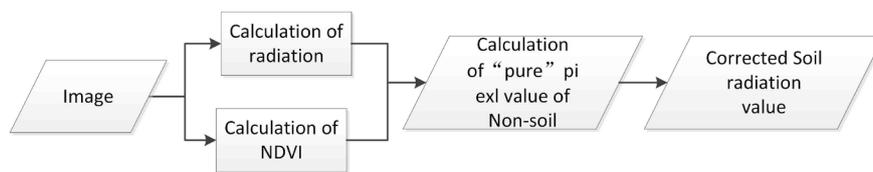
where  $n$  is the number of feature bands,  $D_i$  is the DN value of a pixel at  $i$  band,  $D_{\text{soil}}$  represents the total soil information of the pixel from all feature bands.

The flowchart of the method is shown as follows **Figure 2**.

## 2.4. Accuracy Assessment

Original and non-soil information removed images obtained by the above method are used to establish heavy metal retrieval regression models with the heavy metals contents in soil. Thus, the feasibility of the method of non-soil information removal can be verified by comparing the model accuracy from Original and non-soil information removed images.

Generally, the coefficient of determination ( $R^2$ ) is interpreted as the proportion of the variance in the dependent variable that is predictable from the independent variable; the Root mean square error (RMSE) is usually used to measure the deviation between the observed and predicted; and the model significance index is used to check if the linear statistical relationship exists between the



**Figure 2.** Non-soil information removal model based on NDVI.

response variable. Thus, according to relevant references, the  $R^2$ , RMSE and significant index are usually selected to verify the model accuracy and the feasibility of the method [3] [5]. In this paper, two methods are used to validate the performance of the non-soil information removal based on the NDVI:

1) Comparing the  $R^2$ , the RMSE and the significance index of models from original and non-soil information removed images, the greater the  $R^2$ , the higher correlation between the soil information and the heavy metal concentration. In the case of significant models, the larger the  $R^2$  or the smaller the RMSE, the higher the accuracy of the model. The RMSE is defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}} \tag{5}$$

where the  $P_i$  and  $O_i$  are the predicted and observed value of the heavy metals at the  $i$  band, and the  $n$  is the number of the soil samples.

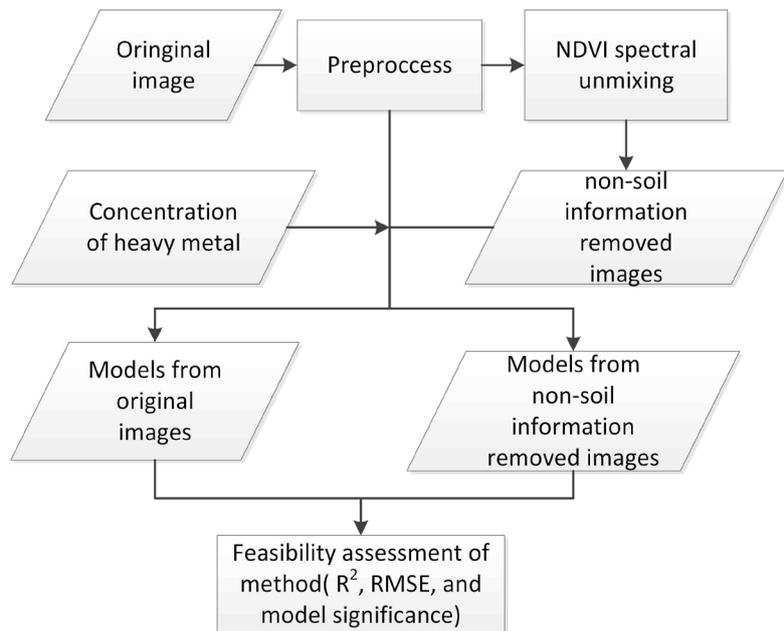
2) The distribution of the retrieved heavy metals' concentration from non-soil information removed images is compared with the field data. A more similar distribution trends to the distributions of field data will lead a higher accuracy of the models and the extracted soil information.

The total flowchart to access the accuracy is shown as **Figure 3**.

### 3. Results

#### 3.1. Non-Soil Information Removed Results

After a detailed examination on the crest and trough of the spectrum curves of the samples for retrieving heavy metals, band11, band13, band44, band120, band198, band202, band203, band205, band208, band209, band210, band217,



**Figure 3.** Verification of non-soil information removal model.

band218, band222 and band223 are selected as the feature bands for Hyperion, and band1, band2 and band4 of GF-2 are selected as feature bands. The non-soil “pure” pixel DN value ( $P_i$ ) of each feature band is calculated with Equations (1), (2) and (3), as listed in **Table 2**.

Then, with non-soil information removal method presented above, the soil information of Hyperion and GF-2 images can be calculated based on corresponding soil and non-soil abundances with Equation (4), as following Equations:

$$D_{\text{Hyper}_{\text{soil}}} = \frac{D_{11} - 2816 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_{13} - 2172 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_{44} - 4332 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_{120} - 1394 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_{198} - 1245 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_{202} - 1225 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_{203} - 1140 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_{205} - 1173 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_{208} - 1126 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_{209} - 1145 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_{210} - 1115 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_{217} - 941 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_{218} - 844 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_{222} - 2216 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_{223} - 992 * f_{\text{non-soil}}}{f_{\text{soil}}} \quad (6)$$

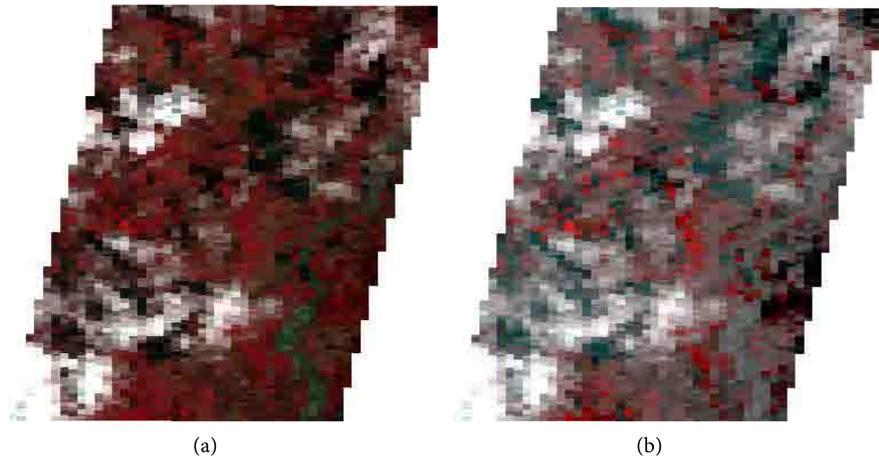
$$G_{\text{GF-2}_{\text{soil}}} = \frac{D_1 - 1423 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_2 - 1259 * f_{\text{non-soil}}}{f_{\text{soil}}} + \frac{D_4 - 4199 * f_{\text{non-soil}}}{f_{\text{soil}}} \quad (7)$$

where the  $D_{11}$ ,  $D_{13}$ ,  $D_{44}$  and so on are the DN value of a pixel at 11, 13, 44 band and so on, respectively.

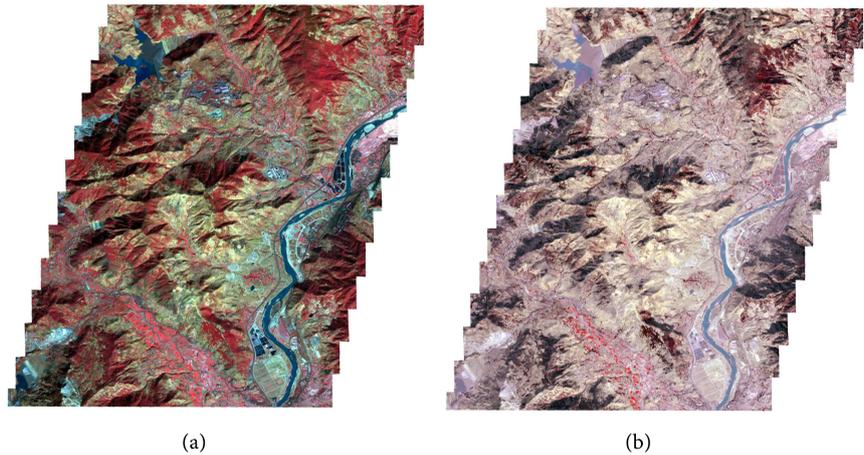
The pseudo color images from the results and the original data of Hyperion and GF-2 are shown in **Figure 4** and **Figure 5** respectively. From **Figure 4**, we can see that the non-soil information almost covered the entire study area (**Figure 4(a)**), while little non-soil information appeared in **Figure 4(b)**. And **Figure 5(a)** and **Figure 5(b)** has a similar pattern. That is to say, most non-soil information of the images had been removed after this process.

**Table 2.** The non-soil “pure” pixel DN value ( $P_i$ ) of each feature band.

Hyperion			GF-2		
feature bands	$P_i$	feature bands	$P_i$	feature bands	$P_i$
Band11	2816	Band208	1126	Band1	1423
Band13	2712	Band209	1145	Band2	1259
Band44	4332	Band210	1115	Band4	4199
Band120	1394	Band217	941	-	-
Band198	1245	Band218	844	-	-
Band202	1225	Band222	2216	-	-
Band203	1140	Band223	992	-	-
Band205	1173			-	-



**Figure 4.** Hyperion images ((a) original image, (b) non-soil information removed image. R: Band53, G: band30 and B: band21 are corresponding to the red, green, and blue channels respectively).



**Figure 5.** GF-2 images ((a) original image, (b) non-soil information removed image. R: Band4, G: band3 and B: band2 are corresponding to the red, green, and blue channels respectively).

### 3.2. Geochemical Analysis

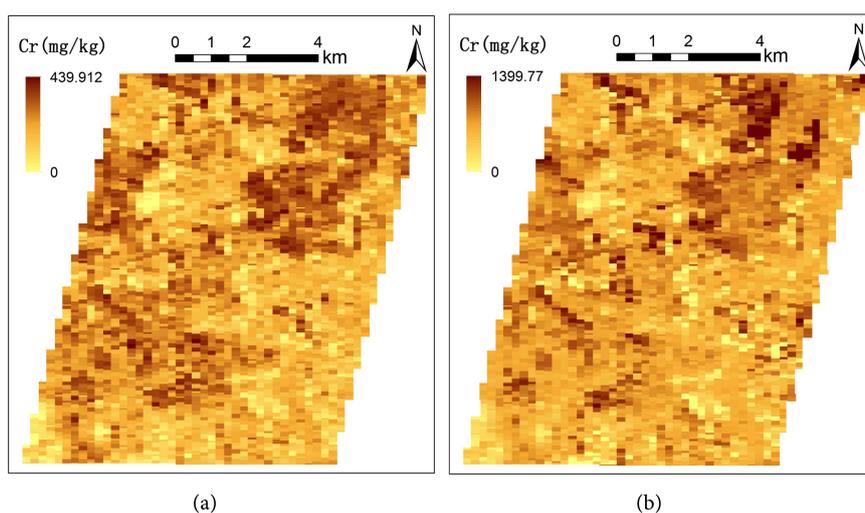
11 soil samples were selected to establish the inversion models randomly, and the remaining samples were used to verify these models. The heavy metals concentration (Cr, Ag, Cu and Ba) of the 11 soil samples are presented in **Table 1**. From **Table 1**, we can see that the standard deviation (SD) of Cr, Cu and Ba in the soil were relatively high (135.36, 70.55 and 473.26 respectively) except Ag, this indicates that their concentrations are of a little great difference in study area.

### 3.3. Heavy Metal Retrieval from Hyperion Images

The method of model building was described in another article. The parameters of inversion models of Cr and Ag from Hyperion are presented in **Table 3**. The  $D_{\text{Hyper-soil}}$  and RMSE are calculated with Equations (5) and (6). **Figure 6** and

**Table 3.** Regression models of Cr and Ag.

Hyperion	Original image		Non-soil information removed image	
Element	Cr	Ag	Cr	Ag
Modeling bands	b44, b120	b11	b44, b120	b11, b120
R <sup>2</sup>	0.667	0.385	0.68	0.724
RMSE	80.11	0.18	75.96	0.03
Significance	0.012	0.042	0.01	0.006
Equation	$Y = 505.774 - 0.134*b44 + 0.131*b120$	$Y = 0.239 - 8.482E - 6*b11$	$Y = 604.116 - 0.158*b44 + 0.15*b120$	$Y = 0.21 - 4.912E - 5*b11 + 7.83E - 5*b120$

**Figure 6.** Cr's concentrations in the study area ((a) Retrieved from original image, (b) Retrieved from non-soil information removed image).

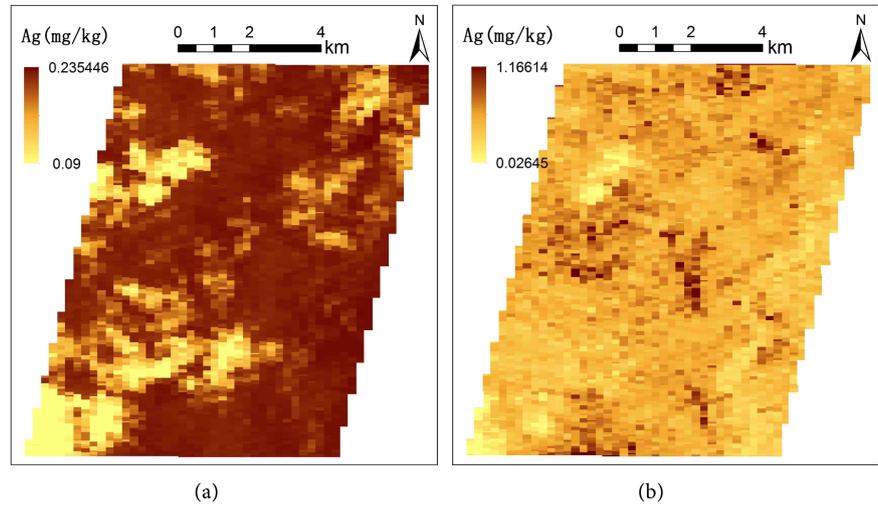
**Figure 7** are the retrieved concentration gradient maps of Cr and Ag. From **Table 3**, we can see that the significance of Cr and Ag are less than 0.05. This indicates that the models of Cr and Ag established from the original and non-soil information removed Hyperion image are significant.

Then, the other soil samples are used to verify the models by independent-samples T test. From **Table 4**, we can see that the P value are all less than 0.05. This indicates that the all models of Cr and Ag can be used to estimate the heavy metals concentration in the mining area.

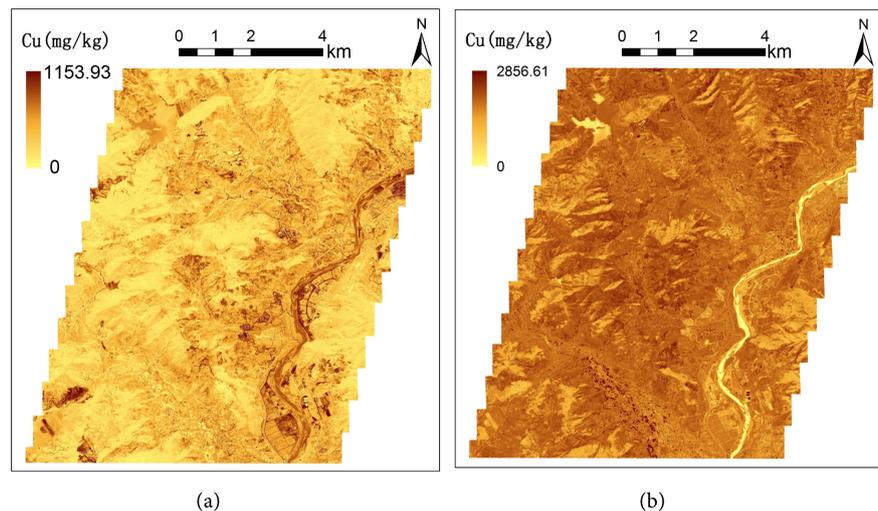
### 3.4. Heavy Metal Retrieval from GF-2 Images

The parameters of inversion retrieval models of Cu and Ba from GF-2 are presented in **Table 5**. The  $D_{GF-2soil}$  are calculate with Equation (7). **Figure 8** and **Figure 9** are the gradient maps of Cu and Ba from the calculated results.

Then, the other soil samples are used to verify the models by independent-samples T test. From **Table 6**, we can see that the P value are all less than



**Figure 7.** Ag’s concentrations in the study area ((a) Retrieved from original image, (b) Retrieved from non-soil information removed image).



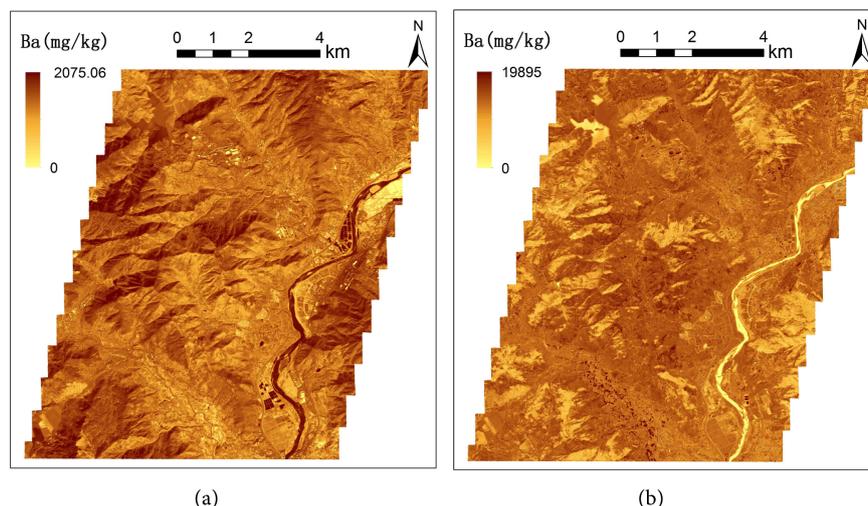
**Figure 8.** Cu’s concentrations in the study area ((a) Retrieved from original image, (b) Retrieved from non-soil information removed image).

0.05. This indicates that the all models of Cu and Ba can be used to estimate the heavy metals concentration in the mining area.

### 3.5. Validation of Non-Soil Information Removal Results

From **Figure 4** and **Figure 5**, we can see that the original images have more vegetation-based non-soil information than the non-soil information removed images. This indicates that most of the vegetation-based non-soil information has been removed by the calculation of non-soil “pure” pixel and abundance.

From **Table 3** and **Table 5**, we can see that the  $R^2$  (Cr: 0.68, Ag: 0.724, Cu: 0.71 and Ba: 0.695) of inversion models (Cr, Ag, Cu and Ba) from non-soil information removed images are higher than that from original images (Cr: 0.667, Ag: 0.385, Cu: 0.425). The RMSE (Cr: 75.96, Ag: 0.03, Cu: 52.88 and Ba: 284.70)



**Figure 9.** Ba’s concentrations in the study area ((a) Retrieved from original image, (b) Retrieved from non-soil information removed image).

**Table 4.** The reference of the T test of the Cr and Ag.

Hyperion	Original image		Non-soil information removed image	
Element	Cr	Ag	Cr	Ag
P value	0.039	0.011	0.039	0.007

**Table 5.** Regression models of Cu and Ba.

GF-2	Original image		Non-soil information removed image	
Element	Cu	Ba	Cu	Ba
Modeling bands	Band1, 2, 4		Band1, 2, 4	
R <sup>2</sup>	0.425	0.406	0.71	0.695
RESE	53.43	396.49	52.88	284.70
Significance	0.249	0.274	0.027	0.032
Equation	$Y = 10.18 - 0.34*b1 + 0.49*b2 - 0.05*b4$	$Y = 2212.2 - 2.92*b1 + 1.741*b2 - 0.54*b4$	$Y = 142.9 + 0.46*b1 - 1*b2 + 0.26*b4$	$Y = 2634 + 3.6*b1 - 7*b2 + 1.8*b4$

**Table 6.** The reference of the T test of the Cu and Ba.

GF-2	Original image		Non-soil information removed image	
Element	Cu	Ba	Cu	Ba
P value	0.014	0.0003	0.015	0.0004

from non-soil information removed images are lower than that from original images (Cr: 80.11, Ag: 0.18, Cu: 53.43 and Ba: 396.49). The model significance indexes of Ag, Cu and Ba from original images are greater than 0.04 and the models are not significant. While, the inversion models for Cr, Ag, Cu and Ba from non-soil information removed images are less than 0.032 and the models are significant. From **Figures 6-9**, we can see that the retrieved heavy metal

concentrate (Cr, Ag, Cu and Ba) from non-soil information removed images are closer to the field measurements than that from original images. These indicate that the soil information is more accurate by the methods of non-soil removing based on NDVI and missing soil compensation, and it has a stronger correlation with the heavy metals of the soil in the mining area. Thus, the method of non-soil information removal based on NDVI for obtaining soil is feasible.

#### 4. Discussions

Many spectral unmixing methods have been applied to estimate heavy metals contents from hyperspectral and high spatial resolution images. However, these methods require complex endmembers selection and abundance calculation, which are time-consuming and do not consider the effect of vegetation on soil. This paper presented a fast and convenient method based on NDVI to remove the vegetation-based non-soil information from the soil surface of the Hyperion and GF-2 images in the mining area.

Comparing the model parameters and the retrieved heavy metals concentration from the non-soil information and original images, from the non-soil information removed images, the relationship between the soil information and the heavy metals concentration are stronger and the deviation between the predicted and the observed value of the heavy metals are less than that from original images. Thus, compared with other unmixing methods, in this method, the multiple linear relationship between non-soil components can be reduced based on the NDVI, and the non-soil information can be maximum removed by the  $P_i$ . Then the missed soil information can be compensation by the probability theory. This step makes the obtained soil more accuracy and the relationship between the soil information and the heavy metal concentration stronger.

However, the retrieved heavy metals in soil with non-soil removed method based on NDVI may be influenced by the additive noise effects such as atmospheric path radiance, the scaling problems with saturated signals often encountered in high-biomass conditions and the canopy background variations. Thus, the method has some shortness need to be improved, such as the impact of cloud and snow, the selection of the optimal NDVI value and non-soil “pure” DN value, etc. Then, the non-soil information cannot be removed completely. We will try to address these shortcomings and make the non-soil information removed method can be used widely.

#### 5. Conclusion

A non-soil information removed method based on NDVI from hyper-spectral and multi-spectral images was developed for retrieving heavy metal method in mining soil. The proposed approach assumes that the mixed objects in each pixel of remote sensing images are composed only of soil and vegetation-based non-soil end-generational endmembers, then, the soil information of each pixel

can be compensated with the non-soil information removed based on its NDVI. It makes the non-soil information can be maximum removed and the mission soil information can be compensation. Thus, a new non-soil information removed method was developed to retrieve heavy metal concentration in mining area more precisely.

## Acknowledgements

This work was supported by Natural Science Foundation of China under Grant 41201440; Science and Technology Projects of Land Resources and Housing Authority of Chongqing under Grant CIGMR1103; and Natural Science Key Project of Sichuan Provincial Education Department under Grant 15ZA0072. The authors would like to thank WANG Jun-wei from Space Star Technology Co., LTD and YANG Wen-bing from China Nonferrous Metals in Changsha Survey and Design Institute Co., LTD for the help in the collection of spectra data. The authors would also like to thank SONG Lian from Nanjing University and GAO Xiao-jie from National Geographic Information Bureau of Surveying and Mapping for the help in the paper's literal correction.

## References

- [1] Zhang, G., Guo, X., Zhao, Z., He, Q., Wang, S., Zhu, Y., *et al.* (2016) Effects of Biochars on the Availability of Heavy Metals to Ryegrass in an Alkaline Contaminated Soil. *Environmental Pollution*, **218**, 513-522. <https://doi.org/10.1016/j.envpol.2016.07.031>
- [2] Wu, S., Shi, Y., Zhou, S., Wang, C. and Chen, H. (2016) Modeling and Mapping of Critical Loads for Heavy Metals in Kunshan Soil. *Science of The Total Environment* **569-570**, 191-200. <https://doi.org/10.1016/j.scitotenv.2016.06.072>
- [3] Song, L. Jian, J., Tan, D.-J., Xie, H.-B., Luo, Z.-F. and Gao, B. (2015) Estimate of heavy metals in soil and streams using combined geochemistry and field spectroscopy in Wan-sheng mining area, Chongqing, China. *International Journal of Applied Earth Observation and Geoinformation*, **34**, 1-9. <https://doi.org/10.1016/j.jag.2014.06.013>
- [4] Choe, E., van der Meer, F., van Ruitenbeek, F., van der Werff, H., de Smeth, B. and Kim, K.-W. (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. *Remote Sensing of Environment*, **112**, 3222-3233. <https://doi.org/10.1016/j.rse.2008.03.017>
- [5] Liu, M., Liu, X., Ding, W. and Wu, L. (2011) Monitoring Stress Levels on Rice with Heavy Metal Pollution from Hyperspectral Reflectance Data Using Wavelet-Fractal Analysis. *International Journal of Applied Earth Observation and Geoinformation*, **13**, 246-255. <https://doi.org/10.1016/j.jag.2010.12.006>
- [6] Chen, H., Yuan, X., Li, T., Hu, S., Ji, J. and Wang, C. (2016) Characteristics of Heavy Metal Transfer and Their Influencing Factors in Different Soil-Crop Systems of the Industrialization Region, China. *Ecotoxicology and Environmental Safety*, **126**, 193-201. <https://doi.org/10.1016/j.ecoenv.2015.12.042>
- [7] Zhao, Y., Pu, R., Bell, S.S., Meyer, C., Baggett, L.P. and Geng, X. (2013) Hyperion Image Optimization in Coastal Waters. *IEEE Transactions on Geoscience and Remote Sensing*, **51**, 1025-1036. <https://doi.org/10.1109/TGRS.2012.2205262>

- [8] Liu, Y., Wang, H., Li, X. and Li, J. (2015) Heavy Metal Contamination of Agricultural Soils in Taiyuan, China. *Pedosphere*, **25**, 901-909. [https://doi.org/10.1016/S1002-0160\(15\)30070-9](https://doi.org/10.1016/S1002-0160(15)30070-9)
- [9] Liu, M., Liu, X., Zhang, B. and Ding, C. (2016) Regional Heavy Metal Pollution in Crops by Integrating Physiological Function Variability with Spatio-Temporal Stability Using Multi-Temporal Thermal Remote Sensing. *International Journal of Applied Earth Observation and Geoinformation*, **51**, 91-102. <https://doi.org/10.1016/j.jag.2016.05.003>
- [10] Yang, J. and He, Y. (2017) Automated Mapping Of Impervious Surfaces In Urban And Suburban Areas: Linear Spectral Unmixing of High Spatial Resolution Imagery. *International Journal of Applied Earth Observation and Geoinformation*, **54**, 53-64. <https://doi.org/10.1016/j.jag.2016.09.006>
- [11] Du, B., Wang, S., Wang, N., Zhang, L., Tao, D. and Zhang, L. (2016) Hyperspectral Signal Unmixing Based on Constrained Non-Negative Matrix Factorization Approach. *Neurocomputing*, **204**, 153-161. <https://doi.org/10.1016/j.neucom.2015.10.132>
- [12] Huang, X., Lu, Q. and Zhang, L. (2014) A Multi-Index Learning Approach for Classification of High-Resolution Remotely Sensed Images over Urban Areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, **90**, 36-48. <https://doi.org/10.1016/j.isprsjprs.2014.01.008>
- [13] Yang, J., He, Y. and Oguchi, T. (2014) An Endmember Optimization Approach for Linear Spectral Unmixing of Fine-Scale Urban Imagery. *International Journal of Applied Earth Observation and Geoinformation*, **27**, 137-146. <https://doi.org/10.1016/j.jag.2013.09.013>
- [14] Shi, C. and Wang, L. (2014) Incorporating Spatial Information in Spectral Unmixing: A Review. *Remote Sensing of Environment*, **149**, 70-87. <https://doi.org/10.1016/j.rse.2014.03.034>
- [15] Yang, J. and Li, P. (2015) Impervious Surface Extraction in Urban Areas from High Spatial Resolution Imagery Using Linear Spectral Unmixing. *Remote Sensing Applications: Society and Environment*, **1**, 61-71. <https://doi.org/10.1016/j.rsase.2015.06.005>
- [16] Deng, C. and Wu, C. (2013) Examining the Impacts of Urban Biophysical Compositions on Surface Urban Heat Island: A Spectral Unmixing and Thermal Mixing Approach. *Remote Sensing of Environment*, **131**, 262-274. <https://doi.org/10.1016/j.rse.2012.12.020>
- [17] Magendran, T. and Sanjeevi, S. (2014) Hyperion Image Analysis and Linear Spectral Unmixing to Evaluate the Grades of Iron Ores in Parts of Noamundi, Eastern India. *International Journal of Applied Earth Observation and Geoinformation*, **26**, 413-426. <https://doi.org/10.1016/j.jag.2013.09.004>
- [18] Kopačková, V. (2014) Using Multiple Spectral Feature Analysis for Quantitative pH Mapping in a Mining Environment. *International Journal of Applied Earth Observation and Geoinformation*, **28**, 28-42. <https://doi.org/10.1016/j.jag.2013.10.008>
- [19] Luis Clenio Jario, M., Adunias dos Santos, T. and Lênio Soares, G. (2015) Potential of Multispectral and Hyperspectral Data to Detect Saline-Exposed Soils in Brazil. *GIScience & Remote Sensing*, No. 4, 416-436.
- [20] Dutta, D., Rahman, A. and Kundu, A. (2015) Growth of Dehradun City: An Application of Linear Spectral Unmixing (LSU) Technique Using Multi-Temporal Landsat Satellite Data Sets. *Remote Sensing Applications: Society and Environment*, **1**,

98-111. <https://doi.org/10.1016/j.rsase.2015.07.001>

- [21] Zhong, Y., Wang, X., Zhao, L., Feng, R., Zhang, L. and Xu, Y. (2016) Blind Spectral Unmixing Based on Sparse Component Analysis for Hyperspectral Remote Sensing Imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, **119**, 49-63. <https://doi.org/10.1016/j.isprsjprs.2016.04.008>
- [22] Feng, R., Zhong, Y. and Zhang, L. (2014) Adaptive Non-Local Euclidean Medians Sparse Unmixing for Hyperspectral Imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, **97**, 9-24. <https://doi.org/10.1016/j.isprsjprs.2014.07.009>
- [23] Zhu, F., Wang, Y., Xiang, S., Fan, B. and Pan, C. (2014) Structured Sparse Method for Hyperspectral Unmixing. *ISPRS Journal of Photogrammetry and Remote Sensing*, **88**, 101-118. <https://doi.org/10.1016/j.isprsjprs.2013.11.014>
- [24] Aghaei, M., Lindner, H. and Bogaerts, A. (2012) Optimization of Operating Parameters for Inductively Coupled Plasma Mass Spectrometry: A Computational Study. *Spectrochimica Acta Part B: Atomic Spectroscopy*, **76**, 56-64. <https://doi.org/10.1016/j.sab.2012.06.006>
- [25] Ning, Y., Li, J., Cai, W. and Shao, X. (2012) Simultaneous Determination of Heavy Metal Ions in Water Using Near-Infrared Spectroscopy with Preconcentration by Nano-Hydroxyapatite. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, **96**, 289-294. <https://doi.org/10.1016/j.saa.2012.05.034>
- [26] John, R.J. (2007) *Introductory Digital Image Processing: A Remote Sensing Perspective*. 3rd Edition, Pearson Education, Columbus.
- [27] Liu, M., Liu, X., Li, J. and Li, T. (2012) Estimating Regional Heavy Metal Concentrations in Rice by Scaling up a Field-Scale Heavy Metal Assessment Model. *International Journal of Applied Earth Observation and Geoinformation*, **19**, 12-23. <https://doi.org/10.1016/j.jag.2012.04.014>
- [28] Li, Y., Cao, Z., Long, H., Liu, Y. and Li, W. (2017) Dynamic Analysis of Ecological Environment Combined with Land Cover and NDVI Changes and Implications for Sustainable Urban-Rural Development: The Case of Mu Us Sandy Land, China. *Journal of Cleaner Production*, **142**, 697-715. <https://doi.org/10.1016/j.jclepro.2016.09.011>