water

MDPI

Article

An Indirect Inversion Scheme for Retrieving Toxic Metal Concentrations Using Ground-Based Spectral Data in a Reclamation Coal Mine, China

Yi Su [1,†], Bin Guo [1,*,†], Yongzhi Lei [2], Dingming Zhang [1], Xianan Guo [1], Liang Suo [1], Yonghua Zhao [3,4,*] and Yi Bian [1]

[1] College of Geomatics, Xi'an University of Science and Technology, Xi'an 710054, China
[2] China Power Construction Group Northwest Survey, Design and Research Institute Co., Ltd., Xi'an 710065, China
[3] School of Land Engineering, Chang'an University, Xi'an 710054, China
[4] Shaanxi Key Laboratory of Land Consolidation, Xi'an 710054, China
* Correspondence: yonghuaz@chd.edu.cn (Y.Z.); guobin12@xust.edu.cn (B.G.)
† These authors contributed equally to this work.

Abstract: A reclamation coal mine in Baishui County of Shaanxi Province, China, was selected as the study area to develop a fast survey method for estimating soil heavy metal concentrations using spectral data. A portable object spectrometer manufactured by Analytical Spectral Devices (ASD) was used to measure soil spectral reflectance, and an X-ray fluorescence device was utilized to obtain the content of heavy metals. The Savitzky-Golay filter, first derivative reflectance (FDR), second derivative reflectance (SDR), continuum removal (CR), and continuous wavelet transform (CWT) were used to transform the original reflectance (OR) spectra for enhancing the spectral characteristics, respectively. Furthermore, correlation analysis was introduced to determine the characteristic bands and the correlations of heavy metals. Partial least squares regression (PLSR), extremely learning machine (ELM), random forest (RF), and support vector machine (SVM) were implemented for quantitatively determining relations between heavy metal contents and spectral reflectance. The outcomes demonstrated that the spectral transformation methods could effectively capture the characteristic bands and increase the relations between heavy metal contents and spectral reflectance. The relation between Fe and Ni was close with a relatively high correlation coefficient (r = 0.741). RF combined with CWT at the decomposition scales of 9 demonstrated the best performance with the highest R_v^2 (0.71) and the lowest $RMSE_v$ (1019.1 mg/kg) for inferring Fe content. Ni content was inferred based on the close relationship between Fe and Ni. The result of RF was better than other methods with the highest R_v^2 (0.69) and the lowest $RMSE_v$ (1.94 mg/kg) for estimating Ni concentration. Therefore, the RF model was chosen for mapping Fe and Ni contents in the study area. The present study revealed that the indirect inversion methods using spectral data can be effectively used to predict heavy metal concentrations. The outcomes supply a new perspective for retrieving heavy metal content based on hyperspectral remotely sensed technology.

Keywords: hyperspectral; soil heavy metals; indirect inversion method; random forest

Citation: Su, Y.; Guo, B.; Lei, Y.; Zhang, D.; Guo, X.; Suo, L.; Zhao, Y.; Bian, Y. An Indirect Inversion Scheme for Retrieving Toxic Metal Concentrations Using Ground-Based Spectral Data in a Reclamation Coal Mine, China. *Water* **2022**, *14*, 2784. https://doi.org/10.3390/w14182784

Academic Editor: Domenico Cicchella

Received: 10 August 2022
Accepted: 5 September 2022
Published: 7 September 2022

1. Introduction

Soil is the basic natural resource for agricultural production and the largest carbon reservoir in terrestrial ecosystems [1,2]. The soil environment is a key factor for sustaining social and economic development and assuring healthy crop production for human beings [3]. In recent decades, heavy metal pollution in soils has aroused global concerns due to their poisonousness, persistence, non-degradation, and half-life period [4,5]. China, the largest developing country in the world, has been underlining the speedy urbanization and industrialization process [6–8]. A large number of mining resources have been

developed for meeting the demands of economic development [9]. However, mining and smelting activities are regarded as the main sources of heavy metal pollution in soil [10]. Moreover, plenty of abandoned mines without any preventative measures for heavy metals exhibited seriously negative effects on the soil environment [11]. In 2016, the Chinese government had to spend CNY 60.69 million (CNY 1 = USD 0.68) to investigate 6,150 km^2 of mining areas with the environmental deterioration generated by mine drainage [12]. Therefore, it is necessary to investigate heavy metals distribution in abandoned mines for land consolidation and reclamation in the future.

Traditional methods with high precision for obtaining heavy metals distribution are in-situ sampling combined with laboratory chemical analysis, whereas traditional approaches have obvious deficiencies, including low efficiency, high cost, time-consuming, and small scope of application. Meanwhile, the chemical analysis methods may generate secondary pollution in the environment due to the materials used in the lab process. In comparison, spectrum technology, based on visible and near-infrared reflectance (VNIR) spectroscopy, supplies a new perspective for rapidly and efficiently detecting heavy metal concentrations. At present, VNIR technology has been widely used in retrieving soil heavy metal contents. For example, many studies demonstrated the feasibility of using spectrum technology to detect soil lead concentration and predicted the spatial distribution of lead [13]. Chen et al. examined the reliability of using soil spectral reflectance to estimate the concentration of Cd, Pb, As, Cr, Cu, and Zn in suburban soils [14]. Han et al. inferred As concentration based on hyperspectral remote sensing data [15]. Tan et al. established the model between the selected spectral features and the soil heavy metal data based on an ensemble learning method [16]. Most published studies constructed the inversion model directly. However, soil heavy metals are trace elements that may hardly be detected in spectral response compared to water, organic matter, and clay mineral. Various challenges were met for inversing heavy metal contents using VNIR due to the low correlation between trace metal and spectra and the complication of soil's physical and chemical properties [4,17]. Therefore, it is really hard to capture the spectral response signals through direct inversion methods. Fortunately, previous research proved that heavy metals are always related to organic matter, clay minerals, hyperspectral indices, and iron oxides [18]. For example, some studies confirmed that the relations between organic matter and heavy metals were close [18,19]. In addition, other research proved that the relations between heavy metals and Fe are close [20]. To our knowledge, Fe is one of the dominant elements in soils [21]. Hence, the close relationship between heavy metals and related materials in the soils supplies a chance to detect heavy metal contents indirectly. For example, Zhang et el. took an indirect approach to predict soil heavy metal contents using hyperspectral indices [22]. Wang et al. took the indirect inversion model based on iron oxide content to estimate As, Zn, and Cd quantitatively [23]. Shen et al. proved the feasibility of indirect inversion based on the Fe element to infer Cu concentration [17]. Therefore, this paper attempts to use the Fe element as a proxy to infer soil heavy metal elements based on VNIR indirectly.

One of the key processes of predicting heavy metal concentrations in this study is to capture the spectral response signals of heavy metals in the soil. However, it is hard to detect the spectral response signals of heavy metals in the soil owing to spectral noises, mixed overlapping peaks, and other uncertainties [24–26]. Therefore, the suitable spectral preprocessing methods that can be used for removing undesirable effects and enhancing interesting features are valuable for inversing heavy metals in soils. A variety of methods, including first (FD) and second derivatives (SD), continuum removal (CR), square root (SR), logarithm, reciprocal of the logarithm, and reciprocal have been used for spectral preprocessing by previous studies. For example, some researchers used FD and SD to extract spectral information and assess their relation to soil heavy metals, such as As, Cd, Cu, Pb, and Zn [24,26–28]. Similarly, the related researchers reviewed spectral derivative techniques and reported that FD and SD are frequently applied [4]. Although the spectral derivative transformations can remove the baseline effect and enhance minor absorption features, the FD and SD may ignore some detailed spectral signals [29,30]. Meanwhile,

some other transformation methods were adopted for soil spectral data, including min-max normalization (MMN), multiplicative scatter correction (MSC), spectral optimized indices, Gaussian filtering (GF), standard normal variate (SNV), and Fractional-order derivative (FOD) [31–33]. However, the efficiency of previous methods concerning spectral preprocessing was relatively weak and not very significant because the heavy metals were trace elements in the soil. The spectral preprocessing approaches need to be further strengthened. Therefore, it is necessary to develop an efficient spectral transformation method to enhance the spectral response of heavy metals in the soil. Published studies confirmed that the capabilities of wavelet transform (WT) are powerful and reliable in extracting spectral features of heavy metals in the soil [12]. WT has been widely used in signal and image analysis, denoising, compression, and decomposition. In addition, continuous wavelet transforms (CWT) demonstrated outstanding performance in enhancing the spectral response of heavy metals in the soil [34,35]. Moreover, CWT combined with inversion models can effectively improve the coefficient value of determination for calibrating heavy metal contents. According to the published studies, the capability of using spectral data based on machine learning methods in estimating heavy metals concentrations has been confirmed by many previous studies [36–39], but the feasibility and reliability of using a combination method of CWT and machine learning based on spectral data to indirectly estimate heavy metal concentrations at a reclamation coal mine are still uncertain.

The objectives of this study are to (1) examine the difference in characteristic bands selection in different experimental conditions based on Fe element, including in-situ, lab-based processed (the soil samples were processed in the lab before surveying reflectance spectra of soil), and lab-based unprocessed (the soil samples were unprocessed before measuring soil spectral reflectance); (2) find the optimal combination between transformation method and machine learning by the accuracy of inversion; (3) explore the feasibility of inversion indirectly.

2. Materials and Methods

A flowchart of this study is summarized as follows (Figure 1): heavy metal concentrations were surveyed in the study area, including Cr, Zn, Ni, Cu, Mn, and Fe. Meanwhile, soil spectra were collected in field samples, lab-processed samples, and lab unprocessed samples from a mining area. Spectral preprocessing was implemented based on the removal of wavelength in noisy regions and Savitzky-Golay smoothing methods. FDR, SDR, CR, and CWT transformation were applied to the smoothed spectra. Correlation analysis was used to determine the correlation between Fe and the other metals and extract the characteristic bands. Additionally, PLSR, ELM, RF, and SVM models were used to calibrate for estimating Fe concentration, inferring Ni concentration indirectly based on the optimal model.

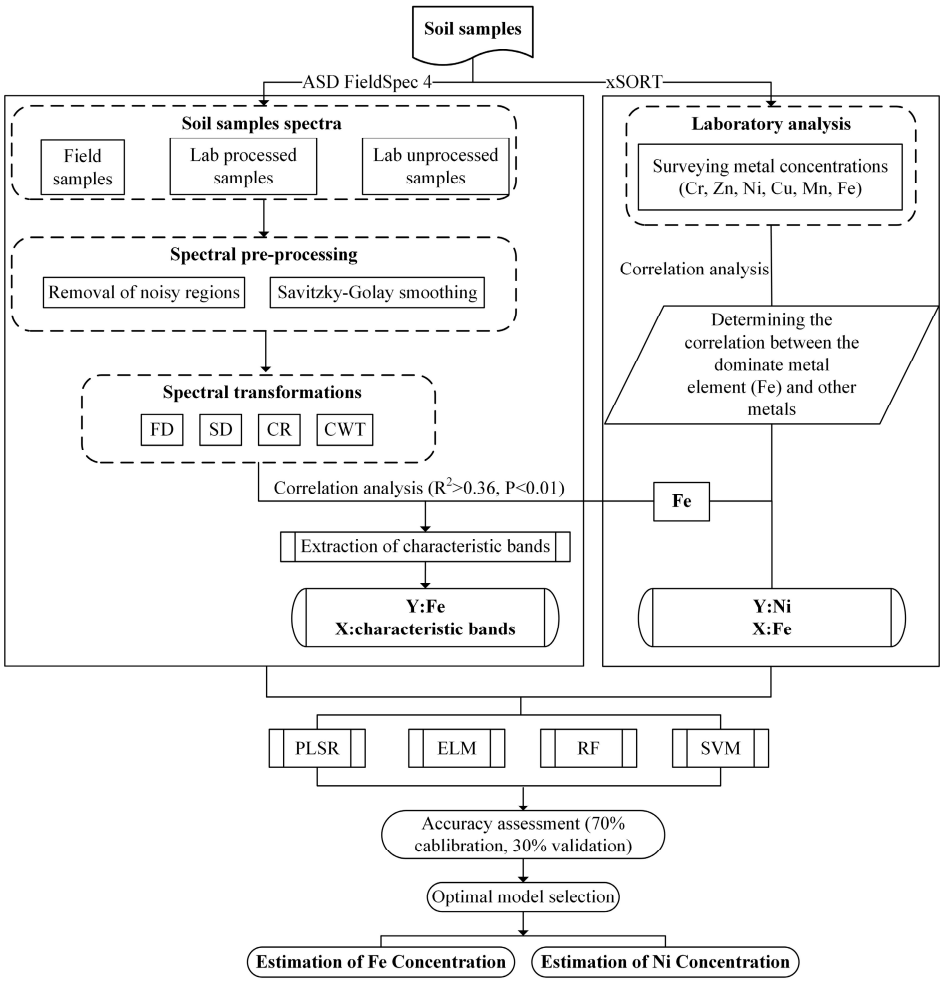

Figure 1. Workflow of the present study.

2.1. Study Area

An abandoned reclamation coal mine in Baishui County, located in Weinan City, Shaanxi Province, northwest of China, was selected as the study area (Figure 2). The elevation of the sampling region with a total area of approximately 12.55 km^2 is decreased from northwest to southeast. The transport route of coals is located in the south of the study area. It belongs to a warm temperate continental monsoon climate, with northeast and northwest prevalent wind directions, an average annual temperature of 11.6 °C, and average annual precipitation of 598.2 mm [40]. In this area, the land is mainly devoted to planting wheat, fruit trees, and other economic crops. The county has an agricultural acreage of 48,000 hectares and a cultivated area of 34,596 hectares [41]. The ecosystem of this area is fragile due to long-time mining resource development. The abandoned coal mine with waste slag accumulation easily leads to heavy metal pollution through rain and wind. Furthermore, a large number of chemical fertilizers are frequently used in the soil for improving the amount of production, which possibly leads to heavy metal accumulation in the soil [42,43].

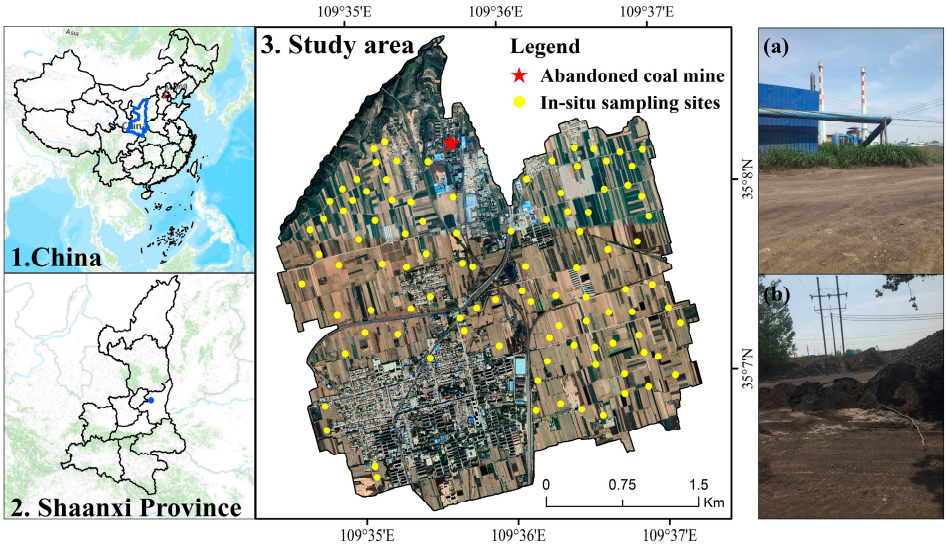

Figure 2. Location of the study area and in-situ sampling sites. (**a**): An abandoned cement factory, (**b**): An empty field heaped with cinders.

2.2. Soil Sampling and Pretreatment

One hundred in-situ topsoils (0–20 cm) samples were collected with a post-hole digger in a reclamation coal mine in Baishui County on 29 April 2020. The sample collecting rule has been adopted to ensure that the interval between the two sites was suitable. Each representative in-situ sample (about 300 g) contained five subsamples that were positioned as wintersweet (HJ/T166-2004). We used a portable receiver of the Global Navigation Satellite System (GNSS) to capture the signals from Continuously Operating Reference Stations (CORS) of Qianxun Corporation and to survey the sampling site's actual position via the Real-time kinematic (RTK) method. The China Geodetic Coordinate System 2000 (CGCS2000) was obtained for each sample site with reliable precision. Furthermore, samples were put into polyethylene bags and were transported to the laboratory for the next pretreatment.

In the lab, samples were airdried at room temperature, and impurities containing gravel and other foreign matters were removed with tweezers, then crushed and grounded using a plastic rod. Then, samples were placed into an oven to dry until the weight did not change. Next, a 0.7 mm nylon aperture sieve was used to sieve samples and put them into clean polyethylene bags for analysis. Each sample with 4 g weight was put into a 32 mm mold to squeeze a tablet under 30-ton pressure for concentrations measuring via SPECTRO xSORT X-ray fluorescence made in German. The average concentration was used as the final record for eliminating the errors. Finally, Sample Result Manager for X-ray fluorescence (SPECTRO xSORT) was introduced to handle metal concentrations data. To ensure data quality and precision, we introduced quality assurance and quality control (QA/QC) to evaluate the data. The GSS-series and GSD-series geochemical reference standard materials (Institute of Geophysical and Geochemical Prospecting, Lang fang, China) were adopted to calibrate the X-ray fluorescence (SPECTRO xSORT) instrument and to ensure the relative standard deviation ranged from 3% to 5%, In this study, six heavy metals were detected, and other heavy metals were below the detection limit.

2.3. Spectral Measurement and Preparation

2.3.1. Spectral Measurement

The spectral reflectance of three different kinds of samples includes field spectra (FS), laboratory-processed spectra (LPS), and laboratory unprocessed spectra (LUS), respectively. In the field, sampling sites were always located in relatively flat and open areas to decrease adjacent reflectance and eliminate shadows during the field spectral survey. To keep the

original state of soil, vegetation, pebbles, roots, and voids were excluded carefully from the surface before measuring soil spectra. Soil spectra measurement was conducted from 10 AM to 2 PM on a sunny day. In the laboratory, a 1000 W halogen lamp was utilized as the light source. Spectral measurements were conducted in a dark room and all surveyors were required to put on black clothes to eliminate unnecessary spectral noise.

Soil spectra were measured in both laboratory and field using a portable object spectrometer manufactured by Analytical Spectral Devices (ASD FieldSpec 4 Standard Res) of America. The ASD spectrometer with a spectral range of 350 to 2500 nm has a 1 nm resolution. To minimize bias, a warm-up with 30 min duration was carried out for the spectrometer. Before surveying, a whiteboard with 99% reflectance was introduced to calibrate the spectrometer [33]. The soil sample was filled with a black dish with a diameter of 10 cm and a depth of approximately 1 cm. The soil part beyond the dish height was scraped with a knife blade to ensure a smooth surface [39]. The observation angle between the light and the vertical direction was set as 15°. The 50 cm distance was adjusted between the light source and the soil samples. In addition, the distance and angle between the probe and the samples were 15 cm and 90°, respectively [39]. Furthermore, every sample was surveyed 10 times to minimize observation errors, and the average spectrum was recorded as the final spectrum after 10 times of scanning.

2.3.2. Spectral Smoothing and Spectral Transformation

The preprocessing of original spectra was implemented to decrease spectral noise, including noise area removal, spectral smoothing, and transformation. (1) A detector shift was corrected using ViewSpecPro software based on original reflectance (OR) in LPS, LUS, and FS; (2) Due to the influence of the instrument noise, water vapor, and complex field environment, the spectra at an interval of the edge bands of 350–399 nm and 2401–2500 nm were deleted in LPS and LUS, and the edge bands of 350–399 nm, 1291–1429 nm, and 1731–2500 nm were removed from each soil sample in FS [14,44]; (3) The Savitzky-Golay filter (second order and 21 points) was used to smooth spectra for denoising and eliminating errors produced by baseline drift, tilt, etc. [31,45]; and (4) The smoothed spectral reflectance was transformed based on FDR, SDR, and CR (Equations (1)–(3)) methods to enhance the spectral characteristics [14]. Meanwhile, the CWT method (Equations (4)–(5)) was chosen to compare with the above three spectral transform methods [46–50]. The Gaussian 4 function was chosen as the basis function according to previous studies [12]. Moreover, decomposition scales were divided into ten scales, including 2^1, 2^2, 2^3, . . . , 2^9, and 2^{10}.

$$FDR = \frac{\rho_{\lambda_{i+1}} - \rho_{\lambda_{i-1}}}{2\Delta\lambda} \quad (1)$$

$$SDR = \frac{\rho_{\lambda_{i+1}} - 2\rho_{\lambda_i} + \rho_{\lambda_{i-1}}}{\Delta\lambda^2} \quad (2)$$

$$CR = \frac{\rho_{\lambda_i}}{\rho_{850} + \frac{\rho_{2500} - \rho_{850}}{2500 - 850} \times \left(\rho_{\lambda_i} - \rho_{850}\right)} \quad (3)$$

where $\rho_{\lambda i-1}$, $\rho_{\lambda i}$, and $\rho_{\lambda i+1}$ represent reflectances of the bands λ_{i-1}, λ_i, λ_{i+1}, respectively; and $\Delta\lambda$ is the band interval.

$$Wf(a,b) \le f; \varphi_{a,b} \ge \int_{-\infty}^{+\infty} f(t)\varphi_{a,b}(t)dt \quad (4)$$

$$\varphi_{a,b}(t) = \frac{1}{\sqrt{a}}\varphi\left(\frac{t-b}{a}\right) \quad (5)$$

where $f(t)$ is the measured reflectance from a sample; t is the band interval; $\varphi_{a,b}(t)$ is the basic function; a is the scaling factor; and b is the shift factor.

2.4. Correlation Analysis for Determining the Characteristic Bands

The correlation analysis method was introduced to determine the relationship between spectral reflectance and heavy metal concentrations. Furthermore, the relation between the content of the dominant element in the soil (such as Fe element) and other heavy metal concentrations were obtained for evaluating the feasibility of indirect inversion. The correlation coefficient (r) was used to quantitatively describe the correlation [51]. The computational formula [52] is as follows:

$$r = \frac{\sum_{i=0}^{n}(y_i - \overline{y})(x_i - \overline{x})}{\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2 \sum_{i=0}^{n}(x_i - \overline{x})^2}} \quad (6)$$

where y_i is the content of heavy metal of the ith sample; x_i is the spectral reflectance of the ith band; $\overline{y}$ is the average value of the soil's heavy metal content; $\overline{x}$ is the average value of each band; and n is the number of samples.

2.5. Calibration and Validation

2.5.1. Partial Least Squares Regression (PLSR)

PLSR proposed by Herman O.A. Wold includes multiple linear regression, canonical correlation analysis, and principal factor analysis [53]. PLSR can eliminate the influence of multiple correlations and allow the number of samples to be less than the number of variables. PLSR is good at hyperspectral bands with collinearity and spectral noise [54]. Recently, PLSR has been widely utilized in estimating heavy metal concentrations using hyperspectral remote sensing technology [55].

2.5.2. Extreme Learning Machine (ELM)

The ELM is a new algorithm for single-hidden-layer feedforward networks. For the network, the number of hidden layer nodes needs to be set and the input weight, as well as the offset of hidden elements in the process of execution, does not need to be adjusted. Hence, the ELM has speedy training speeds and higher generalization performance. No parameters need to be manually adjusted except predefined network architecture [56–58].

2.5.3. Random Forest (RF)

The RF algorithm is a bagging method based on a classification and regression tree (CART) [25,59]. The potential of RF is the significance of each feature that can be evaluated with unbiased estimation during the classification process, and the problems with numerous missing data can be tackled. In addition, the RF can be used to process big data without any dimensionality reduction that outperforms traditional models. The basic theory of RF bagging is to select the outcomes of several weak classifiers and form a strong classifier [60,61].

2.5.4. Support Vector Machine (SVM)

Plenty of scholars are paying increasing attention to SVM owing to its excellent generalization performance. SVM minimizes structural risk as a regression aim [62]. SVM uses a kernel function to transform the nonlinear regression into linear regression and solve the problem of small samples and high dimensions. SVM has been successfully applied in the remote sensing field. The prediction performance of the SVM was assessed via the penalty coefficient C and the kernel parameter Gamma. Therefore, the generalization ability and prediction accuracy can be improved by choosing appropriate parameters [63].

2.5.5. Validation

Leave 30% out cross-validation, a widely used method, was introduced to evaluate the reliability of models. The statistics used in the study were the coefficient of determination R^2 and root mean square error (RMSE) [64].

2.5.6. Software

Spectral smoothing, spectral transformation, correlation analysis, PLSR, ELM, RF, and SVM were executed via MATLAB version 2016b. Moreover, Sample Result Manager (a specific software for SPECTRO xSORT, XHH03, SEEGER GMBH, Boschstrasse, Kleve, Germany) was used to export the heavy metal contents, and ArcGIS10.0 were used for analyzing and mapping in this paper.

3. Results

3.1. Statistical Analysis of Heavy Metals in Soil

Table 1 demonstrated the surveyed statistical value of 100 samples in the study area. Statistical analysis showed that the mean contents of Fe (28,787.56 mg/kg) and Mn (580.34 mg/kg) in soil were higher than in others. The mean Ni, Zn, and Cu contents were close to the Shaanxi soil background value. The maximum content of the Cu (135 mg/kg) element and the mean concentration of Ni (27.40 mg/kg) exceeded the national pollution thresholds, according to the Chinese Soil Environmental Quality Standard [65]. This result implied that parts samples of Cu may exist in pollution conditions. Furthermore, the mean concentration of Cr (76.42 mg/kg) was higher than the Shaanxi soil background value. In the soil environment, the accumulation of soil heavy metals is accompanied by an increase in variability. The dispersion and variation characteristics of heavy metal concentration were displayed by the standard deviation (SD) and coefficient of variation (CV), respectively. The CV of the Cu element was the largest among all metal elements (63.42%).

Table 1. Descriptive statistics for soil heavy metals contents of the study area. Unit (mg/kg).

Paraments	Cr	Ni	Zn	Cu	Mn	Fe
Mean	76.42	27.40	67.37	19.57	580.34	28,787.56
Standard deviation (SD)	11.93	2.92	16.66	12.41	38.60	1761.97
Minimum	47.70	20.07	41.65	9.60	469.50	23,839.00
Maximum	107.50	34.12	201.50	135.00	723.00	34,223.00
Coefficient of variation (CV, %)	15.61	10.64	24.73	63.42	6.65	6.12
Shaanxi soil background value	62.50	28.80	69.40	21.40	557.00	-
Chinese soil background value	61.00	26.90	72.40	22.60	583.00	-

A histogram and a box plot of the heavy metals (Figure 3) represented that the distribution of the elements except Cu is normal distribution, which showed better prediction results based on the regression model [66]. A significant extent of changes especially Fe, Mn, and Zn was detected for all metal elements. Parts of the samples had high values, such as Zn (190 mg/kg) and Cu (122 mg/kg).

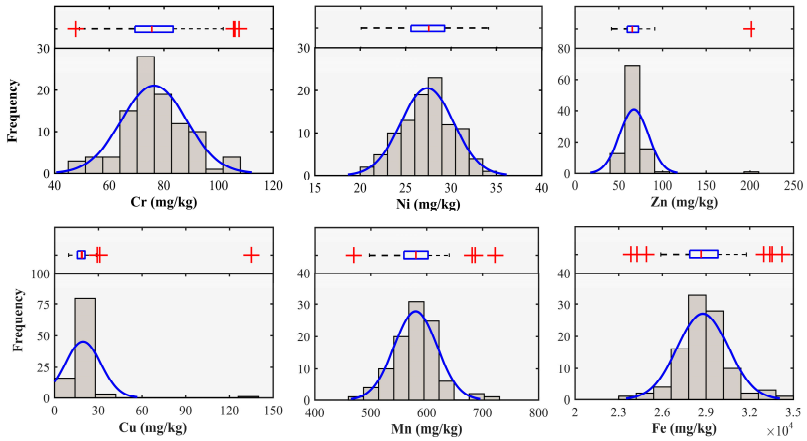

Figure 3. Histograms and box plots of heavy metal contents of the samples in the study area. Note: The red crosses delegate outliers, the red lines inner the blue rectangular represent median value, and the vertical sides of the blue rectangular are quartiles.

3.2. Analysis of Soil Spectral Characteristics

Figure 4 shows the reflectance spectrum of soil samples of the OR and SG smoothed data in the study area. As shown in Figure 4, some differences in terms of the values of spectral reflectance under each experimental method were observed. The overall trend of the spectral reflectance was similar, and the value of spectral reflectance was between 0 and 0.6 for LPS and FS, whereas, the value of spectral reflectance of the LUS was significantly lower than LPS. The FS had higher spectral reflectance than the LUS because FS was measured in the field and the samples contained some alien stuff such as stones and roots. The locations of the spectral absorption were roughly the same, and the peaks could hardly be found from 500 to 800 nm for the gradual increase and fast growth rate of the spectral reflectance, and the spectral reflectance tended to be stable after 800 nm. The spectral absorption peaks and valleys were detected at 800 nm, 1000 nm, 1400 nm, 1900 nm, and 2200 nm. The related experiment indicated that the most significant impact of organic matter is observed from 600 to 800 nm [67]. Published research reported that the absorption peaks between 750 and 1000 nm are owing to the electronic transitions of the Fe^{3+} in oxy/hydroxides [68]. The absorption feature of iron oxides and crystalline Fe was at 900 nm and 850 nm [69]. The absorption valleys occurring at 1400, 1900, and 2200 nm were due to OH^- groups, H_2O molecules, and clay minerals, such as kaolinite, illite, and smectite [70,71].

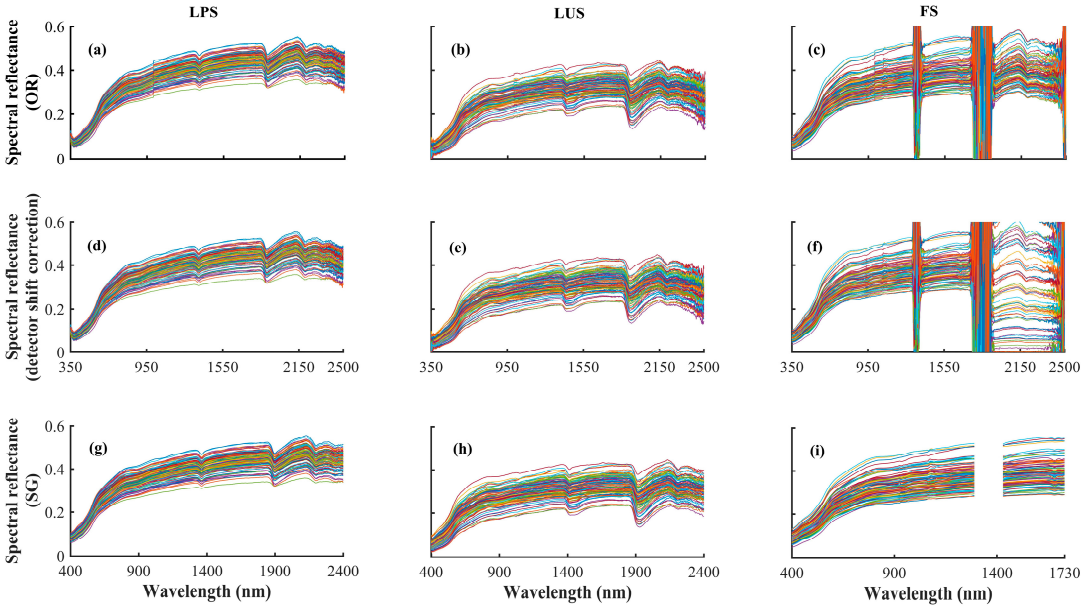

Figure 4. The plots of spectral curves of OR and SG, including laboratory-processed spectra, laboratory unprocessed spectra, and field spectra. (**a**–**c**) represent the original spectra in LPS, LUS, and FS, respectively. (**d**–**f**) represent the corrected detector shift. (**g**–**i**) represent the smoothed spectra. **Note**: Different color lines represent spectra for each soil sample.

First, the OR curve chart under different experimental conditions including FS, LPS, and LUS was transformed using SG smoothed spectral reflectance data. Then, three transformation methods including FDR, SDR, and CR were introduced to further highlight spectral characteristic bands (Figure 5). Clearly, the spectral responses were significantly improved after the spectral transformation, and relatively weak absorption peaks in the original spectral reflectance were highlighted, especially CR. Moreover, the spectral transformation methods used in the present study effectively decreased the impacts of parallel background values and spectral information, and the spectral characteristic bands can be easily captured.

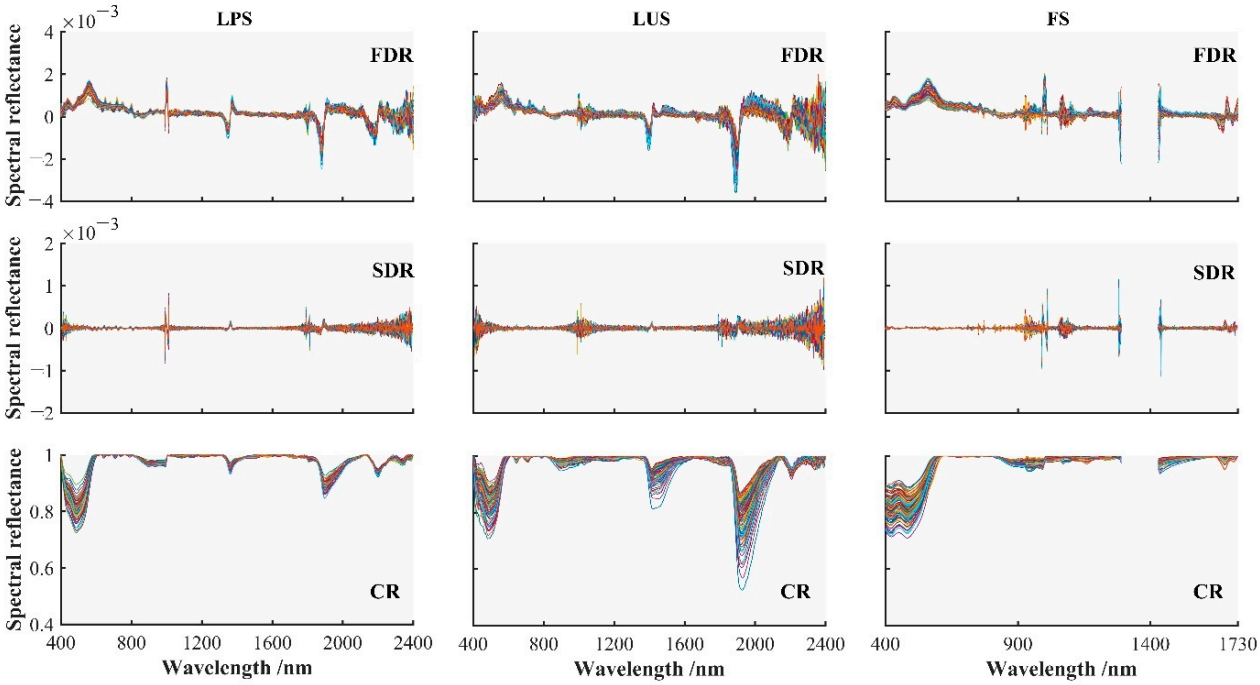

Figure 5. Spectral reflectance curves of soil samples using different spectral transformation methods under three kinds of experimental conditions. Note: Different color lines represent spectra after spectral transformation for each soil sample.

3.3. Correlation Analysis of Heavy Metals Concentration in Soil

The correlation relationship among heavy metals was reflected by R (Table 2). The correlation between each heavy metal and Fe was significantly positive at a level of 0.05. The correlation coefficient between Fe and Mn was 0.851, followed by 0.741 for Fe and Ni. Relatively significant correlations were detected between Ni and Cr (r = 0.296), Ni and Zn (r = 0.353), as well as Cu and Zn (r = 0.876). The relatively high r concerning Cu and Zn demonstrated Zn was a copper-philic element and tend to be compatible with symbiotics. This result demonstrated that the migration and accumulation features of heavy metals were similar in this study. The highest r was 0.851 in terms of Fe and Mn owing to the majority of the Fe and Mn in the soil occurring in the shape of ferromanganese compounds. Furthermore, Mn was more readily aggregated and absorbed by Fe. The previous study confirmed that Mn is one of the major natural sources of soil, and man-made pollution is less [72]. Meanwhile, Ni is mainly distributed on the soil surface and derived from traffic and coal combustion. The study area belonged to the coal mining area and accumulated a lot of waste coal slag. Mining and smelting have also led to Ni contamination of agricultural soils [73]. In addition, some crops accumulated Ni in contaminated soil. Ni and Ni compounds are carcinogens of animals and humans according to the International Agency for Research on Cancer [74,75]. Therefore, Mn was not considered in this work. Ni was mainly studied because of its high risk and higher correlation with Fe.

Table 2. Correlation coefficient matrix for soil metal elements.

Elements	Mn	Cr	Zn	Ni	Cu	Fe
Mn	1					
Cr	0.189	1				
Zn	0.522 **	0.159	1			
Ni	0.646 **	0.296 **	0.353 **	1		
Cu	0.405 **	0.182	0.876 **	0.179	1	
Fe	0.851 **	0.213 *	0.436 **	0.741 **	0.274 **	1

Note: ** Correlation is significant at 0.01 level. * Correlation is significant at the 0.05 level.

3.4. Determining the Characteristic Bands for Fe Element Using R^2 and CWT

The determination coefficients (R^2) between Fe content of soil samples and OR, SG, FDR, SDR, as well as CR under different experimental conditions including FS, LPS, and LUS were determined and exhibited in Figure 6.

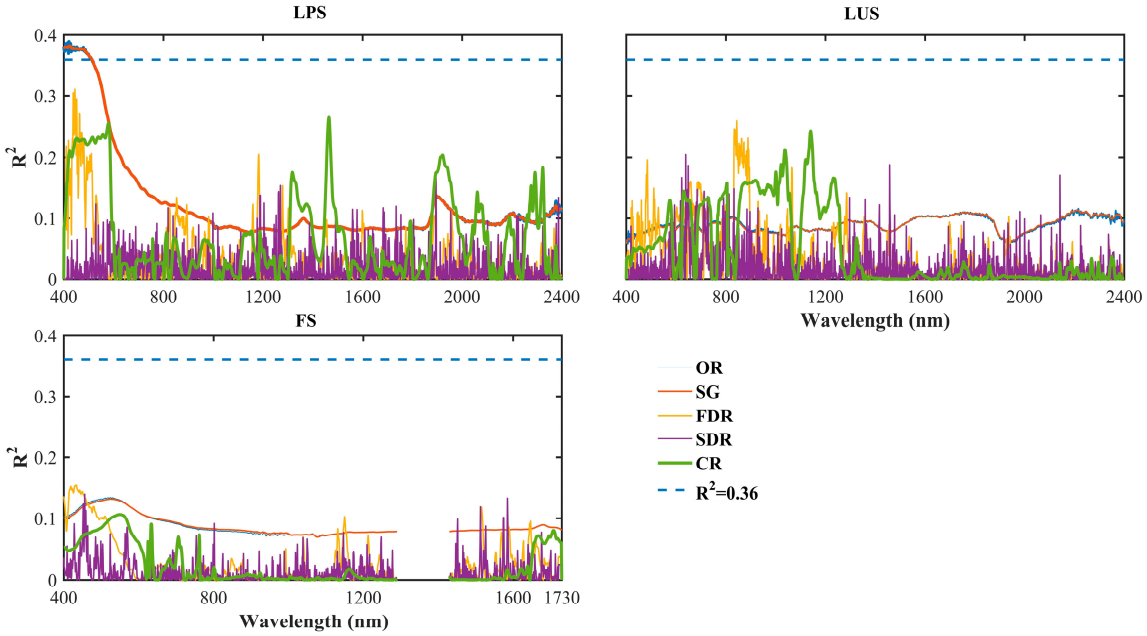

Figure 6. R^2 between Fe content and different spectral reflectance in terms of OR, SG, FDR, SDR, and CR under three experimental conditions.

The previous study indicated that a threshold of 0.36 represents a high correlation between the two variables [76]. The R^2 of OR and SG was larger than 0.36 during 400–511 nm (R^2: 0.36–0.38) and 400–512 nm (R^2: 0.36–0.38) under the LPS. On the contrary, no matter which experimental conditions or spectral transformation we conducted, the R^2 between the Fe concentration and the spectral reflectance obtained by different transformation methods was lower than 0.36. Clearly, the variation range of the visible light curve was more significant compared with near-infrared under the LPS condition. It revealed that the spectral response of Fe in the visible light section was more significant, and this outcome was in line with the findings of some previous studies [77–79]. So, the bands with $R^2 > 0.36$ under LPS conditions were selected as characteristic bands for estimating Fe content using models [76]. Significance tests demonstrated that the correlation coefficient was significant at level $p < 0.01$.

The CWT and correlation analysis were implemented on the SG for selecting characteristic bands. The R^2 was shown in Figure 7. Clearly, the R^2 changed significantly among different decomposition scales (Figure 7). Furthermore, R^2 values achieved stabilized with increasing in the wavelet decomposition scales. The effect of decomposition scales from 4 to 9 in the LPS, 2 to 8 in the LUS, and 4 to 8 in the FS were better than other scales

according to the values. For LPS (Figure 7a), the more significant red region with large coefficient values can be found in Figure 7. The characteristic bands for the Fe element were determined according to the bands with larger R^2 ($R^2 > 0.36$) [76] at a different wavelength and eight decomposition scales including 1 (401–406 nm), 2 (401–404 nm, and 406–408 nm), 3 (401–408 nm and 412–417 nm), 4 (402–418 nm and 431–438 nm), 5 (403–434 nm and 444–459 nm), 6 (400 nm and 437–440 nm), 8 (541 nm, 616 nm, 899 nm, 913 and 914 nm), and 9 (557–659 nm and 1176–1298 nm) in LPS. The corresponding R^2 value were 0.374–0.378, 0.37–0.379, 0.362–0.388, 0.362–0.382, 0.361–0.384, 0.36–0.379, 0.361–0.434, and 0.36–0.429. The results showed that the important bands of Fe belong to the visible light bands. The maximum R^2 of LUS and FS are 0.24 and 0.21, respectively, both less than 0.36 and insignificant. Therefore, soil spectra were not screened out for characteristic bands and estimated Fe content in LUS and FS.

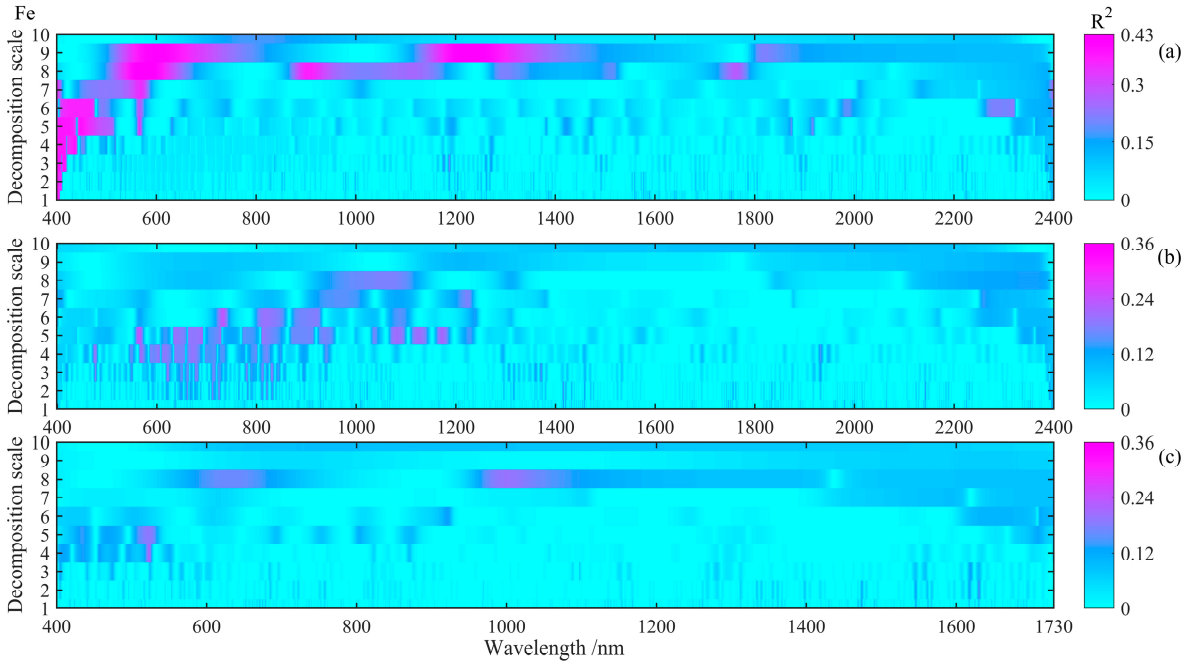

Figure 7. R^2 between Fe content and spectral reflectance transformed by CWT using ten decomposition scales under three experimental conditions. (**a**): LPS, (**b**): LUS, (**c**): FS.

3.5. Estimating Metals Concentration Using RF, PLSR, and ELM Methods

3.5.1. Fe Concentration Estimation

The characteristic bands chosen by OR, SG and CWT in LPS were used to construct inversion models (Tables 3 and 4). The RF methods were introduced to determine the quantitative relationship between Fe content and OR, SG as well as spectral reflectance transformed by CWT at different scales. Meanwhile, the PLSR, SVM, and ELM were selected as comparison models with RF for choosing the optimum estimation model. Clearly, for OR (Table 3), the results of RF (ntree = 100) had the highest R_v^2 (0.64) and the lowest $RMSE_v$ (899.9), followed by ELM (R_v^2 = 0.55, $RMSE_v$ = 1345.5), SVM (R_v^2 = 0.49, $RMSE_v$ = 1575.5), and PLSR (R_v^2 = 0.30, $RMSE_v$ = 1674.8). For SG (Table 3), the results of RF (ntree = 100) outperformed other methods with the highest R_v^2 (0.64) and the lowest $RMSE_v$ (969.7), followed by ELM (the number of hidden neurons = 21) (R_v^2 = 0.41, $RMSE_v$ = 1676.7), PLSR (R_v^2 = 0.36, $RMSE_v$ = 1591.5), and SVM (−c = 4) (R_v^2 = 0.36, $RMSE_v$ = 1702.1) through adjusting the parameters circularly (Table 3). For the CWT method, the best outcomes occurred in the decomposition scales of 9, and RF (ntree = 100) demonstrated better performance than other methods with the highest R_v^2 (0.71) and the lowest $RMSE_v$ (1019.1), followed by ELM (the number of hidden neurons = 35) (R_v^2 = 0.47, $RMSE_v$ =1765.9), PLSR (R_v^2 = 0.41, $RMSE_v$ = 1143.1) and SVM (−c = 2) (R_v^2 = 0.41, $RMSE_v$ = 1539.2). It was obvious that RF was superior to the other three models. Therefore, the model constructed by RF

at the decomposition scales of 9 was chosen as the estimation model for inferring soil Fe content.

Table 3. Accuracy evaluation for inferring Fe concentration using PLSR, ELM, RF, and SVM models based on OR and SG transformation.

Transform	Model	Calibration Dataset (n = 70)		Validation Dataset (n = 30)	
		R_c^2	$RMSE_c$	R_v^2	$RMSE_v$
OR	PLSR	0.44	1214.3	0.30	1674.8
	ELM	0.56	1177.6	0.55	1345.5
	RF	0.84	864.6	0.64	899.9
	SVM	0.62	977.1	0.49	1575.5
SG	PLSR	0.39	1286.7	0.36	1591.5
	ELM	0.554	1035.3	0.548	1468.0
	RF	0.85	863.4	0.64	969.7
	SVM	0.54	1143.7	0.36	1702.1

Table 4. Accuracy evaluation for inferring Fe concentration using PLSR, ELM, RF, and SVM models based on CWT.

Decomposition Scales	Model	Calibration Dataset (n = 70)		Validation Dataset (n = 30)	
		R_c^2	$RMSE_c$	R_v^2	$RMSE_v$
1	PLSR	0.40	1269.6	0.36	1631.9
	ELM	0.59	1172.5	0.40	1084.7
	RF	0.83	824.1	0.60	1186.3
	SVM	0.44	1348.6	0.39	1419.0
2	PLSR	0.39	1374.4	0.33	1411.2
	ELM	0.54	1170.5	0.42	1591.4
	RF	0.86	806.2	0.59	1084.6
	SVM	0.42	1435.2	0.36	1258.3
3	PLSR	0.39	1291.8	0.38	1569.2
	ELM	0.49	1207.9	0.37	1622.6
	RF	0.83	774.4	0.53	1376.3
	SVM	0.43	1349.3	0.38	1504.1
4	PLSR	0.39	1477.1	0.32	1137.7
	ELM	0.58	1211.0	0.42	1672.5
	RF	0.86	837.0	0.68	695.0
	SVM	0.52	1245.5	0.32	1460.0
5	PLSR	0.38	1383.8	0.37	1418.0
	ELM	0.64	1017.3	0.46	1580.4
	RF	0.81	880.7	0.59	1096.3
	SVM	0.47	1115.0	0.39	1776.0
6	PLSR	0.40	1226.6	0.32	1750.4
	ELM	0.62	1149.2	0.41	1348.6
	RF	0.81	862.7	0.58	1099.2
	SVM	0.46	1257.0	0.30	1575.0
8	PLSR	0.46	1397.5	0.42	1145.9
	ELM	0.62	1149.2	0.41	1348.6
	RF	0.83	758.7	0.61	1276.0
	SVM	0.61	1074.0	0.4	1498.9
9	PLSR	0.45	1433.3	0.41	1143.1
	ELM	0.64	1095.1	0.47	1765.9
	RF	0.85	768.9	0.71	1019.1
	SVM	0.66	955.1	0.41	1539.2

3.5.2. Ni Concentration Estimation

Table 2 demonstrated that the concentration of Fe and Ni has close relations with a relatively higher correlation coefficient (R = 0.741). Therefore, the present study used in-situ

measured Fe content to infer Ni content, and the inversion models were established with Fe as the dependent variable and Ni as the independent variable, including PLSR, ELM, RF, and SVM (Table 5). Clearly, the outcomes of RF outperformed other methods with the highest R_v^2 (0.69) and the lowest $RMSE_v$ (1.94), followed by SVM (R_v^2 = 0.57, $RMSE_v$ = 2.13), ELM (R_v^2 = 0.52, $RMSE_v$ = 2.53), and PLSR (R_v^2 = 0.51, $RMSE_v$ = 2.55).

Table 5. Accuracy evaluation for inferring Ni concentration using indirect inversion models including PLSR, ELM, RF, and SVM.

Model	Calibration Dataset (n = 70)		Validation Dataset (n = 30)	
	R_c^2	$RMSE_c$	R_v^2	$RMSE_v$
PLSR	0.60	1.63	0.51	2.55
ELM	0.59	1.92	0.52	2.53
RF	0.83	1.24	0.69	1.94
SVM	0.59	1.89	0.57	2.13

To exhibit the capability of the estimation method of Fe and the indirect inversion model of Ni, the scatterplots concerning the measured content and predicted content for Fe and Ni based on the RF model was illustrated in Figure 8. For the testing samples of the Fe element, the majority of sample contents were underestimated and the estimated Fe content was lower than the measured Fe content. For the testing samples of Ni element, most scatter points were underestimated when the measured Ni content was larger than 25 mg/kg.

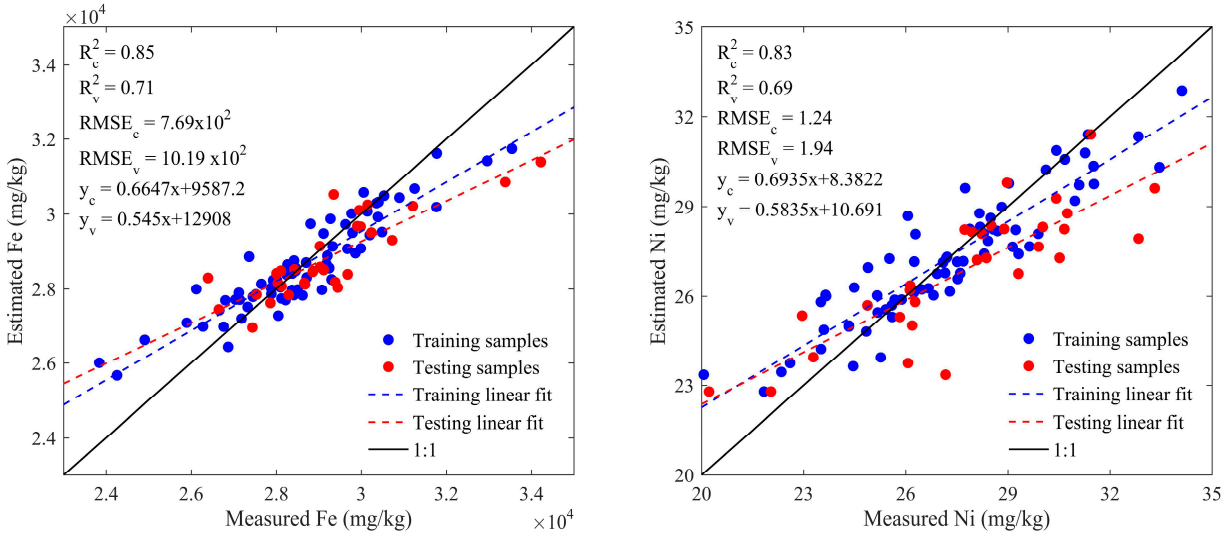

Figure 8. The scatterplots for Fe and Ni content estimation under LPS conditions based on the RF model.

3.6. The Spatial Distribution of Soil Ni Element

The distributions of Fe and Ni in the study area were mapped via the Kriging interpolation method in the ArcGIS toolbox based on the optimal model (Figure 9). The interpolation results of concentrations illustrated that some differences can be found between the predicted and measured concentration values of the two elements, but the spatial distribution trend and the heavy metals concentrations range were close. A little difference between the measured and predicted Fe distributed northwest of the study area, and the difference of Ni in the north of the study area was observed. From the distribution of the predicted Fe and Ni elements content, a significant positive correlation was obtained between the Ni content and Fe content, and the spatial distribution trend was similar, indicating that significant absorption and aggregation relations existed between Ni and Fe elements. The predicted Ni content in the study area decreases first and then increases from southeast to northwest. The predicted Ni element was mainly distributed in the southeast of the study area with the highest concentrations than other positions based on Figure 9, and the enrichment of

the Ni element exceeded the background value of Shaanxi Province (28.80 mg/kg). On the contrary, the lowest contents of Ni occurred in the northeast and southwest part of the study area and were lower than the background value of Shaanxi and China (28.80 mg/kg, 26.90 mg/kg).

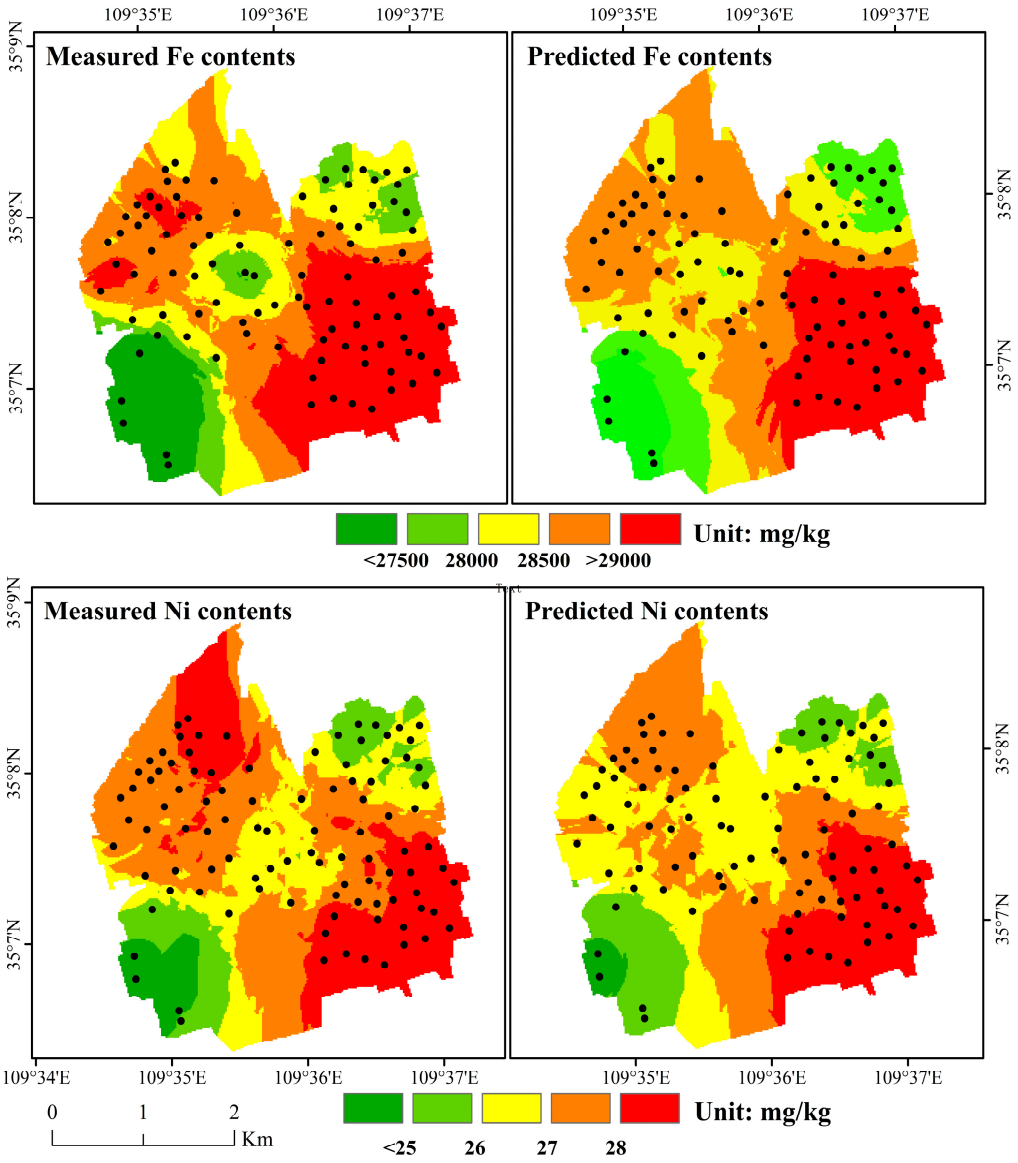

Figure 9. The spatial distribution map of Fe and Ni elements in the study area.

4. Discussion

4.1. The Possible Reason for the Accumulation of Heavy Metals in the Study Area

Soil heavy metals contamination is becoming a worldwide environmental issue, especially in developing countries due to rapid urbanization and industrialization [80,81]. For the study area, Cr mean concentration surpassed both the background value of China and Shaanxi. Although Ni content was lower than the threshold of Shaanxi, it was higher than the background value of China. Mn average concentration surpassed the background value of Shaanxi but was lower than the threshold of China. Most previous studies revealed that the source of Mn is a natural source [82]. The mean content of Fe and Mn in soil was higher than others for Fe and Mn compounds infiltrating into the soil during mining. In general, though the mining area has been remediated, three (Cr, Ni, and Mn) heavy metals in soil still have slight accumulation and the heavy metal contaminations still have not

been eliminated. Though the Cu and Zn mean contents of all samples were lower than both the threshold of China and Shaanxi (Table 1), parts of the soil samples' mean concentration exceeded the national and local thresholds. We inferred that Cu and Zn may be affected by anthropogenic activities [83]. In addition, parts of metals in the study area with relatively large CV implied human inputs may be the main source of contamination. The field survey conducted by our team found that the pesticides, fertilizers, and other agrochemicals used for promoting products and killing insects of the crop may be the reason for increasing the concentration of Zn, Cu, and Cr. Given that the research area was near the downtown of Baishui County and the coal mine, dense population distribution, coal mine chemical industries, and air pollution deposition, possibly concentrated Cr, Ni, and Zn elements were present [75].

4.2. The Characteristics of Spectral Response for Heavy Metals and the Efficiency of the Spectral Transformation Methods

To our knowledge, as the concentration of metals increases the spectral reflectance tended to decrease. Previous studies showed that polluted soil samples revealed a significant spectral absorbance than clean soil samples, especially in the range of 700–2500 nm [27,84–86]. However, the spectral response of heavy metals in the soils was relatively weak because the concentration of the heavy metals was very low. Moreover, the chemical and physical attributes of the soil were complicated that led to the spectral reflectance of soil samples being a mixture of spectra but not a pure one [28,87]. The suitable spectral transformation methods are meant for enhancing the spectral response-ability in estimating heavy metal concentrations based on VNIR spectral technology [57,88]. Therefore, the Savitzky-Golay filter (second order and 21 points) and the FDR, SDR, CR, and CWT methods were used to smooth the spectra and enhance the spectral characteristics, respectively. Multiple points were tried in this experiment, including 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, and 25. The 21 points with the highest accuracy were adopted in this study. Clearly, the spectral responses were significantly improved after the spectral transformation (Figure 5), and relatively weak absorption peaks in the original spectral reflectance were highlighted. Moreover, the spectral transformation methods used in the present study effectively decreased the impacts of parallel background values and spectral information, and the spectral characteristic bands can be easily captured. However, according to our experiment results, FDR and SDR were unsuitable to establish an inversion model in the study. It proved that the CWT and correlation analysis can be efficiently used to choose characteristic bands.

4.3. The Precision Comparison of Estimation Models Combined with Different Transformation Methods in Inferring Metal Content

For Fe simulation, the current study demonstrated that the precision of the RF combined with CWT outperforms the other methods. Some previous studies also obtained a similar outcome [46,51]. To ensure global transformation, sufficient training samples were required by previous spectral transformation methods. CWT can be implemented from the time domain due to a time-scale analysis of signals. Hence, the spectral features can be extracted more efficiently for the time domain as well as frequency domains were used to characterize the local features of the signal. The RF method [89], a data-mining technique originating from a classification and regression tree (CART) model, is a tree-based ensemble learning method [38]. Hence, many regression trees are contained by RF. Complex data pretreatment is not mandatory by RF, and the calculation ability is relatively stronger than the other methods such as artificial neural networks and gradient-boosted machines [90]. In addition, the relative significance of each spectral band can be detected using the RF algorithm.

Furthermore, the present study developed an indirect inversion method using spectral data for inferring heavy metal concentrations in the soil. Although the direct quantitative relations between spectral reflectance and metal contents have been determined by published research, the accuracy was not satisfactory due to the low content of heavy metals in the soils and the weak spectral response. Fortunately, the spectral response of the organic matter, iron, and other constant elements is relatively significant. Though spectral charac-

teristic bands of heavy metals were hardly captured due to their relatively low contents during the VNIR region, the indirect relations between heavy metals and organic matter, clay minerals, and iron oxide can be determined for inferring heavy metal contents [18,39]. This study revealed that the relationship between Ni and Fe was obvious. Although the correlation between Mn and Ni is significant, Mn was not selected as a proxy to inverse soil heavy metals indirectly because soil heavy metals were mainly adsorbed on iron oxide, organic matter, and clay minerals instead of Mn. The indirect inversion model of soil Ni was constructed based on the relation between Fe and Ni. The estimation capability of the indirect inversion method was evaluated by the current. The indirect inversion method was reliable and feasible.

4.4. Limitations of This Paper

Though some findings were obtained by the present study, some limitations still exist. Firstly, the indirect inversion method using spectral data was restricted by the relationship between heavy metals and iron. Only the significant correlations between the heavy metals and iron can be assured that the precision of the indirect inversion model in estimating heavy metals content can be acceptable. Secondly, the current study was carried out locally not globally. Therefore, the suitability of the methods used in this study needs to be verified in other areas owing to the influence of uncertainties such as different spectral noise. Thirdly, the difference between field spectra and laboratory spectra needs to be further studied. If the inversion model can be established by field spectra, the practical value will be enlarged. Therefore, we plan to correct the field spectra using laboratory spectra to eliminate spectral differences caused by environmental factors and improve model accuracy in the future. The results obtained from current research should be treated cautiously because the spectral reflectance of the soil samples was surveyed under laboratory conditions that are completely different from the outdoor environment. Up to now, retrieving heavy metal contents based on remotely sensed imagery and spectral reflectance of in-situ soil samples is still a big challenge due to the complicated chemical and physical attributes of the soil, atmospheric absorption and scattering, soil moisture content, particle size, plant cover, and litter. In addition, heavy metals' spectral characteristics are very complicated due to their coexistence in the soil. We plan to extract the relevant spectral bands as modeling variables based on the adsorption and retention of heavy metals by various soil components to improve the accuracy of the inversion model, including iron oxide, organic matter, and clay. In addition, we plan to map the heavy metals contents distribution using hyperspectral images collected by satellite or UAV (unmanned aerial vehicle) and to verify the feasibility of the indirect inversion method in other areas in the future.

5. Conclusions

The present study proposed an indirect inversion method based on the significant relationship between Fe and Ni for inferring Ni content using hyperspectral reflectance data. The precision of different inversion models combined with spectral transformation methods was compared. Some findings were obtained as follows: Two heavy metals including Cr (76.42 mg/kg) and Ni (27.40 mg/kg) in the study area have slight accumulation and the concentrations exceeded the national pollution thresholds according to the Chinese Soil Environmental Quality Standard. The correlation between each heavy metal and Fe was positively significant at a level of 0.05. The correlation coefficient was 0.741 concerning Fe and Ni. The characteristic bands for the Fe element were determined based on correlation analysis and CWT in LPS. Namely, the characteristic bands for the Fe element were determined on eight decomposition scales, including 1 (401–406 nm), 2 (401–404 nm and 406–408 nm), 3 (401–408 nm and 412–417 nm), 4 (402–418 nm and 431–438 nm), 5 (403–434 nm and 444–459 nm), 6 (400 nm and 437–440 nm), 8 (541 nm, 616 nm, 899 nm, 913 and 914 nm), and 9 (557–659 nm and 1176–1298 nm) in LPS. For the CWT method, the best outcomes occurred in the decomposition scales of 9, and RF demonstrated better performance than other methods with the highest R_v^2 (0.71) and the

lowest $RMSE_v$ (1019.1), followed by ELM (R_v^2 = 0.47, $RMSE_v$ = 1765.9), PLSR (R_v^2 = 0.41, $RMSE_v$ = 1143.1) and SVM (R_v^2 =0.41, $RMSE_v$ = 1539.2) through adjusting the parameters circularly. The outcomes of RF outperformed other methods with the highest R_v^2 (0.69) and the lowest $RMSE_v$ (1.94), followed by SVM (R_v^2 = 0.57, $RMSE_v$ = 2.13), ELM (R_v^2 = 0.52, $RMSE_v$ = 2.53), and PLSR (R_v^2 = 0.51, $RMSE_v$ = 2.55). The feasibility of the indirect inversion method was proven in this study. We will continue to explore the sensitive spectral bands for retrieving toxic metals and conduct experiments in mapping heavy metals' spatial distribution using hyperspectral remote sensing images.

Author Contributions: All authors contributed to the study conception and design. Conceptualization, Y.S. and B.G.; data curation, Y.S., D.Z. and X.G.; investigation, Y.S., Y.L. and L.S.; methodology, B.G.; software, Y.S.; supervision, B.G.; validation, Y.B.; visualization, D.Z. and L.S.; writing—original draft, Y.S.; writing—review and editing, B.G. and Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Natural Science Foundation of Shaanxi Province (2021JM-388) and the Fund Project of Shaanxi Key Laboratory of Land Consolidation (300102352505).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy restrictions.

Acknowledgments: We are grateful to Haorui Bai for checking and editing.

Conflicts of Interest: The authors declared that there is no conflict of interest.

References

1. Proshad, R.; Kormoker, T.; Mursheed, N.; Islam, M.M.; Bhuyan, M.I.; Islam, M.S.; Mithu, T.N. Heavy metal toxicity in agricultural soil due to rapid industrialization in Bangladesh: A review. *Int. J. Adv. Geosci.* **2018**, *6*, 83–88. [CrossRef]
2. Tan, W.-F.; Zhang, R.; Cao, H.; Huang, C.-Q.; Yang, Q.-K.; Wang, M.-K.; Koopal, L.K. Soil inorganic carbon stock under different soil types and land uses on the Loess Plateau region of China. *Catena* **2014**, *121*, 22–30. [CrossRef]
3. Lamine, S.; Petropoulos, G.P.; Brewer, P.A.; Bachari, N.-E.-I.; Srivastava, P.K.; Manevski, K.; Kalaitzidis, C.; Macklin, M.G. Heavy Metal Soil Contamination Detection Using Combined Geochemistry and Field Spectroradiometry in the United Kingdom. *Sensors* **2019**, *19*, 762. [CrossRef]
4. Wang, F.; Gao, J.; Zha, Y. Hyperspectral sensing of heavy metals in soil and vegetation: Feasibility and challenges. *ISPRS J. Photogramm. Remote Sens.* **2018**, *136*, 73–84. [CrossRef]
5. Guo, B.; Su, Y.; Pei, L.; Wang, X.; Zhang, B.; Zhang, D.; Wang, X. Ecological risk evaluation and source apportionment of heavy metals in park playgrounds: A case study in Xi'an, Shaanxi Province, a northwest city of China. *Environ. Sci. Pollut. Res. Int.* **2020**, *27*, 24400–24412. [CrossRef] [PubMed]
6. Alshahri, F.; El-Taher, A. Assessment of Heavy and Trace Metals in Surface Soil Nearby an Oil Refinery, Saudi Arabia, Using Geoaccumulation and Pollution Indices. *Arch. Environ. Contam. Toxicol.* **2018**, *75*, 390–401. [CrossRef]
7. Guo, B.; Wang, Y.; Pei, L.; Yu, Y.; Liu, F.; Zhang, D.; Wang, X.; Su, Y.; Zhang, D.; Zhang, B.; et al. Determining the effects of socioeconomic and environmental determinants on chronic obstructive pulmonary disease (COPD) mortality using geographically and temporally weighted regression model across Xi'an during 2014–2016. *Sci. Total Environ.* **2021**, *756*, 143869. [CrossRef]
8. Guo, B.; Su, Y.; Pei, L.; Wang, X.; Wei, X.; Zhang, B.; Zhang, D.; Wang, X. Contamination, Distribution and Health Risk Assessment of Risk Elements in Topsoil for Amusement Parks in Xi'an, China. *Pol. J. Environ. Stud.* **2021**, *30*, 601–617. [CrossRef]
9. Naz, A.; Chowdhury, A.; Mishra, B.K.; Karthikeyan, K. Distribution of heavy metals and associated human health risk in mine, agricultural and roadside soils at the largest chromite mine of India. *Environ. Geochem. Health* **2018**, *40*, 2155–2175. [CrossRef]
10. Li, Z.; Ma, Z.; van der Kuijp, T.J.; Yuan, Z.; Huang, L. A review of soil heavy metal pollution from mines in China: Pollution and health risk assessment. *Sci. Total Environ.* **2014**, *468–469*, 843–853. [CrossRef]
11. Hu, Z.; Wang, C.; Li, K.; Zhu, X. Distribution characteristics and pollution assessment of soil heavy metals over a typical nonferrous metal mine area in Chifeng, Inner Mongolia, China. *Environ. Earth Sci.* **2018**, *77*, 638. [CrossRef]
12. Zhang, S.; Shen, Q.; Nie, C.; Huang, Y.; Wang, J.; Hu, Q.; Ding, X.; Zhou, Y.; Chen, Y. Hyperspectral inversion of heavy metal content in reclaimed soil from a mining wasteland based on different spectral transformation and modeling methods. *Spectrochim. Acta Part A: Mol. Biomol. Spectrosc.* **2019**, *211*, 393–400. [CrossRef] [PubMed]
13. Pandit, C.M.; Filippelli, G.M.; Li, L. Estimation of heavy-metal contamination in soil using reflectance spectroscopy and partial least-squares regression. *Int. J. Remote Sens.* **2010**, *31*, 4111–4123. [CrossRef]

14. Cheng, H.; Shen, R.; Chen, Y.; Wan, Q.; Shi, T.; Wang, J.; Wan, Y.; Hong, Y.; Li, X. Estimating heavy metal concentrations in suburban soils with reflectance spectroscopy. *Geoderma* **2019**, *336*, 59–67. [CrossRef]
15. Han, L.; Chen, R.; Zhu, H.; Zhao, Y.; Liu, Z.; Huo, H. Estimating Soil Arsenic Content with Visible and Near-Infrared Hyperspectral Reflectance. *Sustainability* **2020**, *12*, 1476. [CrossRef]
16. Tan, K.; Ma, W.; Chen, L.; Wang, H.; Du, Q.; Du, P.; Yan, B.; Liu, R.; Li, H. Estimating the distribution trend of soil heavy metals in mining area from HyMap airborne hyperspectral imagery based on ensemble learning. *J. Hazard. Mater.* **2021**, *401*, 123288. [CrossRef]
17. Shen, Q.; Xia, K.; Zhang, S.; Kong, C.; Hu, Q.; Yang, S. Hyperspectral indirect inversion of heavy-metal copper in reclaimed soil of iron ore area. *Spectrochim. Acta Part A: Mol. Biomol. Spectrosc.* **2019**, *222*, 117191.
18. Lu, Q.; Wang, S.; Bai, X.; Liu, F.; Wang, M.; Wang, J.; Tian, S. Rapid inversion of heavy metal concentration in karst grain producing areas based on hyperspectral bands associated with soil components. *Microchem. J.* **2019**, *148*, 404–411.
19. Sun, W.; Zhang, X.; Sun, X.; Sun, Y.; Cen, Y. Predicting nickel concentration in soil using reflectance spectroscopy associated with organic matter and clay minerals. *Geoderma* **2018**, *327*, 25–35. [CrossRef]
20. Shi, T.; Chen, Y.; Liu, Y.; Wu, G. Visible and near-infrared reflectance spectroscopy—An alternative for monitoring soil contamination by heavy metals. *J. Hazard. Mater.* **2014**, *265*, 166–176. [CrossRef]
21. Alsbou, E.M.E.; Al-Khashman, O.A. Heavy metal concentrations in roadside soil and street dust from Petra region, Jordan. *Environ. Monit. Assess.* **2017**, *190*, 48. [CrossRef]
22. Zhang, S.; Fei, T.; You, X.; Wan, Y.; Wang, Y.; Bian, M. Two hyperspectral indices for detecting cadmium and lead contamination from arice canopy spectrum. *Land Degrad. Dev.* **2020**, *32*, 66–78. [CrossRef]
23. Wang, J.; Zou, L.; Li, Z.; Mu, H.; Zhou, P.; Yang, J.; Zhao, Y.; Qin, K. A detection method of trace metal elements in black soil based on hyperspectral technol-ogy: Geological implications. *J. Geomech.* **2021**, *27*, 418–429.
24. Chen, T.; Chang, Q.; Clevers, J.G.P.W.; Kooistra, L. Rapid identification of soil cadmium pollution risk at regional scale based on visible and near-infrared spectroscopy. *Environ. Pollut.* **2015**, *206*, 217–226. [CrossRef]
25. Douglas, R.K.; Nawar, S.; Cipullo, S.; Alamar, M.C.; Coulon, F.; Mouazen, A.M. Evaluation of vis-NIR reflectance spectroscopy sensitivity to weathering for enhanced assessment of oil contaminated soils. *Sci. Total Environ.* **2018**, *626*, 1108–1120. [CrossRef] [PubMed]
26. Khosravi, V.; Doulati Ardejani, F.; Yousefi, S.; Aryafar, A. Monitoring soil lead and zinc contents via combination of spectroscopy with extreme learning machine and other data mining methods. *Geoderma* **2018**, *318*, 29–41. [CrossRef]
27. Chakraborty, S.; Weindorf, D.C.; Li, B.; Ali Aldabaa, A.A.; Ghosh, R.K.; Paul, S.; Ali, M.N. Development of a hybrid proximal sensing method for rapid identification of petroleum contaminated soils. *Sci. Total Environ.* **2015**, *514*, 399–408. [CrossRef] [PubMed]
28. Liu, J.; Zhang, Y.; Wang, H.; Du, Y. Study on the prediction of soil heavy metal elements content based on visible near-infrared spectroscopy. *Spectrochim. Acta Part A: Mol. Biomol. Spectrosc.* **2018**, *199*, 43–49. [CrossRef] [PubMed]
29. Hong, Y.; Shen, R.; Cheng, H.; Chen, Y.; Zhang, Y.; Liu, Y.; Zhou, M.; Yu, L.; Liu, Y.; Liu, Y. Estimating lead and zinc concentrations in peri-urban agricultural soils through reflectance spectroscopy: Effects of fractional-order derivative and random forest. *Sci. Total Environ.* **2018**, *651*, 1969–1982. [CrossRef] [PubMed]
30. Kemper, T.; Sommer, S. Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy. *Environ. Sci. Technol.* **2002**, *36*, 2742–2747. [CrossRef] [PubMed]
31. Hou, L.; Li, X.; Li, F. Hyperspectral-based Inversion of Heavy Metal Content in the Soil of Coal Mining Areas. *J. Environ. Qual.* **2019**, *48*, 57–63. [CrossRef] [PubMed]
32. Sawut, R.; Kasim, N.; Abliz, A.; Hu, L.; Yalkun, A.; Maihemuti, B.; Qingdong, S. Possibility of optimized indices for the assessment of heavy metal contents in soil around an open pit coal mine area. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *73*, 14–25. [CrossRef]
33. Wei, L.; Yuan, Z.; Zhong, Y.; Yang, L.; Hu, X.; Zhang, Y. An Improved Gradient Boosting Regression Tree Estimation Model for Soil Heavy Metal (Arsenic) Pollution Monitoring Using Hyperspectral Remote Sensing. *Appl. Sci.* **2019**, *9*, 1943. [CrossRef]
34. Guo, B.; Zhang, B.; Su, Y.; Zhang, D.; Wang, Y.; Bian, Y.; Suo, L.; Guo, X.; Bai, H. Retrieving zinc concentrations in topsoil with reflectance spectroscopy at Opencast Coal Mine sites. *Sci. Rep.* **2021**, *11*, 19909. [CrossRef]
35. Zhang, B.; Guo, B.; Zou, B.; Wei, W.; Lei, Y.; Li, T. Retrieving soil heavy metals concentrations based on GaoFen-5 hyperspectral satellite image at an opencast coal mine, Inner Mongolia, China. *Environ. Pollut.* **2022**, *300*, 118981. [CrossRef]
36. Meng, X.; Bao, Y.; Ye, Q.; Liu, H.; Zhang, X.; Tang, H.; Zhang, X. Soil Organic Matter Prediction Model with Satellite Hyperspectral Image Based on Optimized Denoising Method. *Remote Sens.* **2021**, *13*, 2273. [CrossRef]
37. Taghizadeh-Mehrjardi, R.; Fathizad, H.; Ali Hakimzadeh Ardakani, M.; Sodaiezadeh, H.; Kerry, R.; Heung, B.; Scholten, T. Spatio-Temporal Analysis of Heavy Metals in Arid Soils at the Catchment Scale Using Digital Soil Assessment and a Random Forest Model. *Remote Sens.* **2021**, *13*, 1698. [CrossRef]
38. Tan, K.; Wang, H.; Chen, L.; Dud, Q.; Due, P.; Pan, C. Estimation of the spatial distribution of heavy metal in agricultural soils using airborne hyperspectral imaging and random forest. *J. Hazard. Mater.* **2020**, *382*, 120987. [CrossRef]
39. Zhou, W.; Yang, H.; Xie, L.; Li, H.; Huang, L.; Zhao, Y.; Yue, T. Hyperspectral inversion of soil heavy metals in Three-River Source Region based on random forest model. *Catena* **2021**, *202*, 105222. [CrossRef]
40. Dong, H.; Zhao, J.; Xie, M. Heavy Metal Concentrations in Orchard Soils with Different Cultivation Durations and Their Potential Ecological Risks in Shaanxi Province, Northwest China. *Sustainability* **2021**, *13*, 4741. [CrossRef]

41. Wang, X.; Rahman, Z.U.; Lv, Z.; Zhu, Y.; Ruan, R.; Deng, S.; Zhang, L.; Tan, H. Experimental Study and Design of Biomass Co-Firing in a Full-Scale Coal-Fired Furnace with Storage Pulverizing System. *Agronomy* **2021**, *11*, 810. [CrossRef]
42. Wei, X.; Wang, N.; Luo, P.; Yang, J.; Zhang, J.; Lin, K. Spatiotemporal Assessment of Land Marketization and Its Driving Forces for Sustainable Urban–Rural Development in Shaanxi Province in China. *Sustainability* **2021**, *13*, 7755. [CrossRef]
43. Zha, X.; Luo, P.; Zhu, W.; Wang, S.; Lyu, J.; Zhou, M.; Huo, A.; Wang, Z. A bibliometric analysis of the research on Sponge City: Current situation and future development direction. *Ecohydrology* **2021**, *14*, e2328. [CrossRef]
44. Sun, W.; Zhang, X. Estimating soil zinc concentrations using reflectance spectroscopy. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *58*, 126–133. [CrossRef]
45. Zhang, X.; Sun, W.; Cen, Y.; Zhang, L.; Wang, N. Predicting cadmium concentration in soils using laboratory and field reflectance spectroscopy. *Sci. Total Environ.* **2019**, *650*, 321–334. [CrossRef]
46. Liu, W.; Li, M.; Zhang, M.; Long, S.; Guo, Z.; Wang, H.; Li, W.; Wang, D.; Hu, Y.; Wei, Y.; et al. Hyperspectral inversion of mercury in reed leaves under different levels of soil mercury contamination. *Environ. Sci. Pollut. Res. Int.* **2020**, *27*, 22935–22945. [CrossRef]
47. Cheng, T.; Rivard, B.; Sánchez-Azofeifa, G.A.; Feng, J.; Calvo-Polanco, M. Continuous wavelet analysis for the detection of green attack damage due to mountain pine beetle infestation. *Remote Sens. Environ.* **2010**, *114*, 899–910. [CrossRef]
48. Clark, R.N.; Roush, T.L. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *J. Geophys. Res. Solid Earth* **1984**, *89*, 6329–6340. [CrossRef]
49. Dehaan, R.L.; Taylor, G.R. Field-derived spectra of salinized soils and vegetation as indicators of irrigation-induced soil salinization. *Remote Sens. Environ.* **2002**, *80*, 406–417. [CrossRef]
50. Hong, Y.; Shen, R.; Cheng, H.; Chen, S.; Chen, Y.; Guo, L.; He, J.; Liu, Y.; Yu, L.; Liu, Y. Cadmium concentration estimation in peri-urban agricultural soils: Using reflectance spectroscopy, soil auxiliary information, or a combination of both? *Geoderma* **2019**, *354*, 113875. [CrossRef]
51. Hong, Y.; Chen, S.; Chen, Y.; Linderman, M.; Mouazen, A.M.; Liu, Y.; Guo, L.; Yu, L.; Liu, Y.; Cheng, H.; et al. Comparing laboratory and airborne hyperspectral data for the estimation and mapping of topsoil organic carbon: Feature selection coupled with random forest. *Soil Tillage Res.* **2020**, *199*, 104589. [CrossRef]
52. Liu, W.; Yu, Q.; Niu, T.; Yang, L.; Liu, H. Inversion of Soil Heavy Metal Content Based on Spectral Characteristics of Peach Trees. *Forests* **2021**, *12*, 1208. [CrossRef]
53. Martens, M.; Martens, H.; Wold, S. Preference of cauliflower related to sensory descriptive variables by partial least squares (PLS) regression. *J. Sci. Food Agric.* **1983**, *34*, 715–724. [CrossRef]
54. Shi, T.; Wang, J.; Chen, Y.; Wu, G. Improving the prediction of arsenic contents in agricultural soils by combining the reflectance spectroscopy of soils and rice plants. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 95–103. [CrossRef]
55. Dotto, A.C.; Dalmolin, R.S.D.; Ten Caten, A.; Grunwald, S. A systematic study on the application of scatter-corrective and spectral-derivative preprocessing for multivariate prediction of soil organic carbon by Vis-NIR spectra. *Geoderma* **2018**, *314*, 262–274. [CrossRef]
56. Guo, W.; Shang, L.; Zhu, X.; Nelson, S.O. Nondestructive Detection of Soluble Solids Content of Apples from Dielectric Spectra with ANN and Chemometric Methods. *Food Bioprocess Technol.* **2015**, *8*, 1126–1138. [CrossRef]
57. Hong, Y.; Chen, S.; Zhang, Y.; Chen, Y.; Yu, L.; Liu, Y.; Liu, Y.; Cheng, H.; Liu, Y. Rapid identification of soil organic matter level via visible and near-infrared spectroscopy: Effects of two-dimensional correlation coefficient and extreme learning machine. *Sci. Total Environ.* **2018**, *644*, 1232–1243. [CrossRef]
58. Lu, Q.; Wang, S.; Bai, X.; Liu, F.; Tian, S.; Wang, M.; Wang, J. Rapid estimation of soil heavy metal nickel content based on optimized screening of near-infrared spectral bands. *Acta Geochim.* **2020**, *39*, 116–126. [CrossRef]
59. Ham, J.; Chen, Y.; Crawford, M.M.; Ghosh, J. Investigation of the random forest framework for classification of hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 492–501. [CrossRef]
60. Guo, B.; Bian, Y.; Zhang, D.; Su, Y.; Wang, X.; Zhang, B.; Wang, Y.; Chen, Q.; Wu, Y.; Luo, P. Estimating Socio-Economic Parameters via Machine Learning Methods Using Luojia1-01 Nighttime Light Remotely Sensed Images at Multiple Scales of China in 2018. *IEEE Access* **2021**, *9*, 34352–34365. [CrossRef]
61. Krogh, A.; Vedelsby, J. Neural Network Ensembles, Cross Validation, and Active Learning. In Proceedings of the International Conference on Neural Information Processing Systems, Denver, CO, USA, 1 January 1994.
62. Du, Y.; Cui, X.; Xu, Q.; Han, D.; Guo, X.; Cao, G. Spatial Characteristics of Soil Organic Matter and As Content in Source Regions of Yangtze River and Yellow River. *Chin. J. Grassl.* **2012**, *34*, 24–29.
63. Melgani, F.; Bruzzone, L. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1778–1790. [CrossRef]
64. Wei, L.; Yuan, Z.; Wang, Z.; Zhao, L.; Zhanng, Y.; Lu, X.; Cao, L. Hyperspectral Inversion of Soil Organic Matter Content Based on a Combined Spectral Index Model. *Sensors* **2020**, *20*, 2777. [CrossRef] [PubMed]
65. CNEMC. *The Background Values of Elements in Chinese Soils*; China national Environmental Monitoring Center: Beijing, China; China Environmental Science Press: Beijing, China, 1990; pp. 15–505.
66. Yi, P.; Kheir, R.B.; Adhikari, K.; Malinowski, R.; Greve, M.B.; Knadel, M.; Greve, M.H. Digital Mapping of Toxic Metals in Qatari Soils Using Remote Sensing and Ancillary Data. *Remote Sens.* **2016**, *8*, 1003.
67. Xu, B.B.; Ji, G.S. A Preliminary Research of Geographic Regionalization of China Land Background and Spectral Reflectance Characteristics of Soil. *Remote Sens. Environ. China* **1991**, *6*, 142–151.

68. Pyo, J.; Hong, S.M.; Kwon, Y.S.; Kim, M.S.; Cho, K.H. Estimation of heavy metals using deep neural network with visible and infrared spectroscopy of soil. *Sci. Total Environ.* **2020**, *741*, 140162. [CrossRef] [PubMed]
69. Stevens, A.; Nocita, M.; Tóth, G.; Montanarella, L.; van Wesemael, B. Prediction of Soil Organic Carbon at the European Scale by Visible and Near InfraRed Reflectance Spectroscopy. *PLoS ONE* **2013**, *8*, e66409.
70. Clark, R.N.; King, T.V.V.; Klejwa, M.; Swayze, G.A.; Vergo, N. High spectral resolution reflectance spectroscopy of minerals. *J. Geophys. Res.* **1990**, *95*, 12653–12680. [CrossRef]
71. Madejova, J.; Komadel, P. Baseline studies of the clay minerals society source clays: Infrared methods. *Clays Clay Miner.* **2001**, *49*, 410–432. [CrossRef]
72. Jie, L.; Feng, C.; Zeng, G.; Gao, X.; Zhong, M.; Li, X.; Li, X.; He, X.; Fang, Y. Spatial distribution and source identification of heavy metals in surface soils in a typical coal mine city, Lianyuan, China. *Environ. Pollut.* **2017**, *225*, 681.
73. Guan, Q.; Wang, F.; Xu, C.; Pan, N.; Lin, J.; Zhao, R.; Yang, Y.; Luo, H. Source apportionment of heavy metals in agricultural soil based on PMF: A case study in Hexi Corridor, northwest China. *Chemosphere* **2017**, *193*, 189–197. [CrossRef]
74. Kelepertzis, E. Accumulation of heavy metals in agricultural soils of Mediterranean: Insights from Argolida basin, Peloponnese, Greece. *Geoderma* **2014**, *221–222*, 82–90. [CrossRef]
75. Chen, X.; Kumari, D.; Cao, C.J.; Plaza, G.; Achal, V. A review on remediation technologies for nickel-contaminated soil. *Hum. Ecol. Risk Assess. Int. J.* **2019**, *26*, 571–585. [CrossRef]
76. Boker, A.; Brownell, L.; Donen, N. The Amsterdam preoperative anxiety and information scale provides a simple and reliable measure of preoperative anxiety. *Can. J. Anesth.* **2002**, *49*, 792–798. [CrossRef] [PubMed]
77. Kayande, K.S.; Deshmukh, R.R.; Janse, P.V.; Kayte, J.N. Hyper spectral Analysis of Soil Iron Oxide using Fieldspec4 Spectroradiometer. *Int. J. Comput. Sci. Eng.* **2018**, *6*, 395–399. [CrossRef]
78. Rathod, P.H.; Müller, I.; Van der Meer, F.D.; de Smeth, B. Analysis of visible and near infrared spectral reflectance for assessing metals in soil. *Environ. Monit. Assess.* **2016**, *188*, 588. [CrossRef] [PubMed]
79. Hunt, G.R. Spectral signatures of particulate minerals in the visible and near infrared. *Geophysics* **1977**, *44*, 501–513. [CrossRef]
80. Jin, G.; Fang, W.; Shafi, M.; Wu, D.; Li, Y.; Zhong, B.; Ma, J.; Liu, D. Source apportionment of heavy metals in farmland soil with application of APCS-MLR model: A pilot study for restoration of farmland in Shaoxing City Zhejiang, China. *Ecotoxicol. Environ. Saf.* **2019**, *184*, 109495. [CrossRef] [PubMed]
81. Liu, K.; Wang, F.; Li, J.; Tiwari, S.; Chen, B. Assessment of trends and emission sources of heavy metals from the soil sediments near the Bohai Bay. *Environ. Sci. Pollut. Res.* **2019**, *26*, 29095–29109. [CrossRef]
82. Huang, Y.; Li, T.; Wu, C.; He, Z.; Japenga, J.; Deng, M.; Yang, X. An integrated approach to assess heavy metal source apportionment in peri-urban agricultural soils. *J. Hazard. Mater.* **2015**, *299*, 540–549. [CrossRef] [PubMed]
83. Chen, H.; Lu, X.; Chang, Y.; Xue, W. Heavy metal contamination in dust from kindergartens and elementary schools in Xi'an, China. *Environ. Earth Sci.* **2014**, *71*, 2701–2709. [CrossRef]
84. Douglas, R.K.; Nawar, S.; Alamar, M.C.; Mouazen, A.M.; Coulon, F. Rapid prediction of total petroleum hydrocarbons concentration in contaminated soil using vis-NIR spectroscopy and regression techniques. *Sci. Total Environ.* **2018**, *616–617*, 147–155. [CrossRef] [PubMed]
85. Okparanma, R.N.; Mouazen, A.M. Combined Effects of Oil Concentration, Clay and Moisture Contents on Diffuse Reflectance Spectra of Diesel-Contaminated Soils. *Water Air Soil Pollut.* **2013**, *224*, 1539. [CrossRef]
86. Zou, B.; Jiang, X.; Feng, H.; Tu, Y.; Tao, C. Multisource spectral-integrated estimation of cadmium concentrations in soil using a direct standardization and Spiking algorithm. *Sci. Total Environ.* **2020**, *701*, 134890. [CrossRef]
87. Xue, Y.; Zou, B.; Wen, Y.; Tu, Y.; Xiong, L. Hyperspectral Inversion of Chromium Content in Soil Using Support Vector Machine Combined with Lab and Field Spectra. *Sustainability* **2020**, *12*, 4441. [CrossRef]
88. Gholizadeh, A.; Saberioon, M.; Carmon, N.; Boruvka, L.; Ben-Dor, E. Examining the Performance of PARACUDA-II Data-Mining Engine versus Selected Techniques to Model Soil Carbon from Reflectance Spectra. *Remote Sens.* **2018**, *10*, 1172. [CrossRef]
89. Schapire, R.E. Random Forests. *Mach. Learn* **2001**, *45*, 5–32.
90. Sachdeva, S.; Kumar, B. Comparison of gradient boosted decision trees and random forest for groundwater potential mapping in Dholpur (Rajasthan), India. *Stoch. Environ. Res. Risk Assess.* **2021**, *35*, 287–306. [CrossRef]