



Article

Heavy Metal Pollution Delineation Based on Uncertainty in a Coastal Industrial City in the Yangtze River Delta, China

Bifeng Hu ^{1,2,3,4} , Ruiying Zhao ¹, Songchao Chen ^{3,5}, Yue Zhou ¹, Bin Jin ⁶, Yan Li ^{7,*} and Zhou Shi ¹

¹ Institute of Agricultural Remote Sensing and Information Technology Application, Zhejiang University, Hangzhou 310058, China; hubifeng@zju.edu.cn (B.H.); ruiyingzhao@zju.edu.cn (R.Z.); yuezhou@zju.edu.cn (Y.Z.); shizhou@zju.edu.cn (Z.S.)

² Unité de Recherche en Science du Sol, INRA, Orléans 45075, France

³ InfoSol, INRA, US 1106, Orléans F-4075, France; songchao.chen@inra.fr

⁴ Sciences de la Terre et de l'Univers, Orléans University, Orleans 45067, France

⁵ Unité Mixte de Recherche (UMR) Sol Agro et hydrosystème Spatialisation (SAS), INRA, Agrocampus Ouest, Rennes 35042, France

⁶ Ningbo Agricultural Food Safety Management Station, Ningbo 315000, China; dd923068626@163.com

⁷ Institute of Land Science and Property, School of Public Affairs, Zhejiang University, Hangzhou 310058, China

* Correspondence: liyan522@zju.edu.cn; Tel.: +86-571-5666-2139

Received: 14 March 2018; Accepted: 3 April 2018; Published: 10 April 2018



Abstract: Assessing heavy metal pollution and delineating pollution are the bases for evaluating pollution and determining a cost-effective remediation plan. Most existing studies are based on the spatial distribution of pollutants but ignore related uncertainty. In this study, eight heavy-metal concentrations (Cr, Pb, Cd, Hg, Zn, Cu, Ni, and Zn) were collected at 1040 sampling sites in a coastal industrial city in the Yangtze River Delta, China. The single pollution index (PI) and Nemerow integrated pollution index (NIPI) were calculated for every surface sample (0–20 cm) to assess the degree of heavy metal pollution. Ordinary kriging (OK) was used to map the spatial distribution of heavy metals content and NIPI. Then, we delineated composite heavy metal contamination based on the uncertainty produced by indicator kriging (IK). The results showed that mean values of all PIs and NIPIs were at safe levels. Heavy metals were most accumulated in the central portion of the study area. Based on IK, the spatial probability of composite heavy metal pollution was computed. The probability of composite contamination in the central core urban area was highest. A probability of 0.6 was found as the optimum probability threshold to delineate polluted areas from unpolluted areas for integrative heavy metal contamination. Results of pollution delineation based on uncertainty showed the proportion of false negative error areas was 6.34%, while the proportion of false positive error areas was 0.86%. The accuracy of the classification was 92.80%. This indicated the method we developed is a valuable tool for delineating heavy metal pollution.

Keywords: soil heavy metal pollution; indicator kriging (IK); Nemerow integrated pollution index (NIPI); uncertainty; pollution area definition

1. Introduction

Anthropogenic activities, such as urbanization processes, not only change the inherent properties of affected soils, such as their pH, texture, cation exchange capacity, and bulk density, but also inadvertently cause harmful substances, such as heavy metals, to deposit into the soils [1–3]. Heavy

metals are natural components of the earth's crust, and natural concentrations of heavy metals in soils tend to remain low. However, anthropogenic inputs of several heavy metals into soils greatly exceed natural inputs from pedogenesis [4]. This problem is more prominent in cities, since cities have a high density of anthropogenic activities and populations. Therefore, citizens in cities face greater health threats from heavy metal soil pollution. In China, the Cr, Pb, Cd, Hg, As, Cu, Zn, and Ni concentrations of soils have frequently been found to exceed the national standard value [5–7]. It has been revealed that 16.1% of soil has been polluted by heavy metals according to a national survey of soil pollution released in 2014 by the Ministry of Land and Resources and the Ministry of Environmental Protection of the People's Republic of China [8]. Heavy metal pollution in soil attracts great attention, since it can have a substantially harmful effect on human health by the ways of consumption of infected animals, and the chronic low-level intake of soil metals through ingestion or inhalation [9,10].

Recently, soils have been widely used as a diagnostic tool to determine environmental conditions that influence health [11,12]. However, there is a need for better information on heavy metal contamination in urban soils [13–17]. Accurately delineating pollution is the aim of pollution studies, and is vital to determining cost-effective remediation plans [18]. In addition, most of the existing related studies are based on the spatial distribution of pollutants [19–27], which is usually determined through a number of geostatistical interpolation techniques (generally termed kriging) [28], such as simple kriging (SK) [29], ordinary kriging (OK) [30], cokriging [31], and universal kriging [32]. However, due to frequent, very strong positive skewness of pollution data, the use of these methods is not always optimal and results in uncertainty in the estimations. As a simple but useful method, indicator kriging (IK) has been widely used to map the probabilities of estimates that exceed given threshold values, such as standard values of heavy metal contamination in soil [33–35]. To estimate the probability that the heavy metal concentrations exceed the critical threshold, we need the entire probability distributions of heavy metal contents at points. We may assume that the estimation errors follow normal distributions, but this is unrealistic in many situations. Other solutions are to use disjunctive kriging [36] or indicator kriging [37]. Since indicator kriging can accommodate measurements less than the detection limit [38], in this study, a method based on an uncertainty analysis, which was produced by IK, was employed to delineate composite heavy metal pollution contamination.

The aims of this study were to (1) assess the pollution status of the concentrations of eight heavy metals (Cr, Pb, Cd, Hg, As, Cu, Zn and Ni) in surface soils; (2) investigate the spatial distribution of heavy metals in the surface soils in study area; and (3) explore spatial distribution of uncertainty of composite heavy metals pollution and then delineate pollution sites based on this.

2. Methods and Materials

2.1. Description of the Case Study Area

The studied area (28°51'–30°33' N, 120°55'–122°16' E) is located in the Yangtze River Delta (YRD), which is the most developed economic district in China. It covers an area of 9816 km² and has a population of 7.81 million. It is an important and typical coastal industrial city and is one of 14 cities that implemented the reform and opening policy early in 1984. After that, this area has been undergoing rapid and intense industrialization and urbanization over the past three decades. Nowadays, it is well-known for its industries; chemical, textile and garment, petrochemical engineering, and machinery, and has become an important chemical industrial base in China.

2.2. Sampling, Processing, and Analysis

Sampling was done in the study area based on 2 × 2 km² grids. In total, 1040 top soil samples (0–20 cm) were collected. A stainless steel hand auger was used. Each soil sample was collected at an intersection point and mixed; five subsamples were collected from five locations within 5 m. A differential global positioning system (GPS) was used to record spatial position of sampling locations (Figure 1). The details of soil and plant sample analysis have been described in previous study [10].

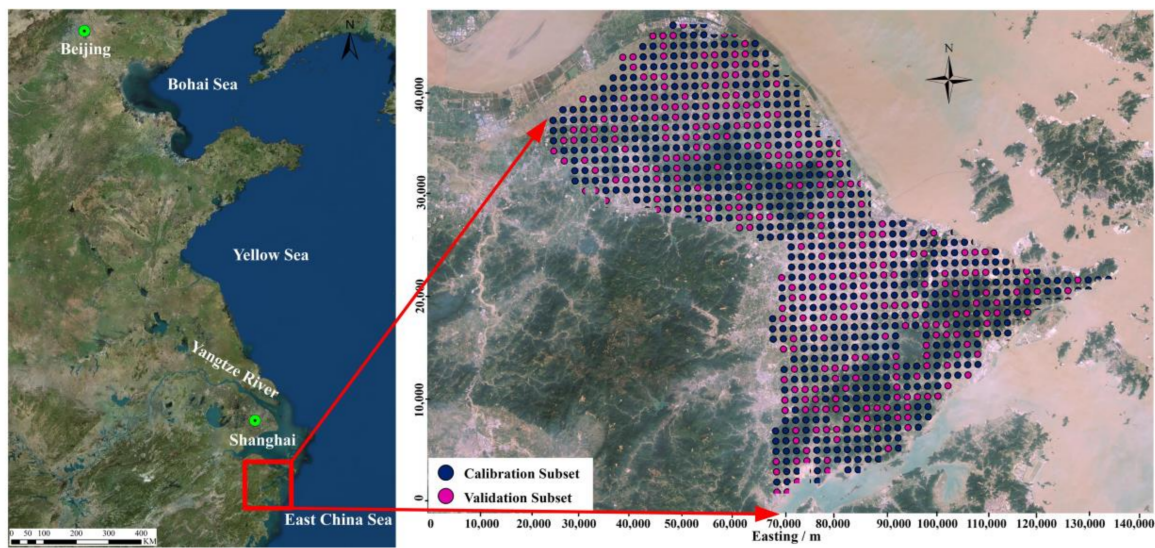


Figure 1. Map of the study area.

2.3. Assessment of Heavy Metal Pollution in Soils

The degree of soil heavy metal pollution was assessed according to related national regulations [39], and a pollution index was used to assess the quality of soil and to estimate the impact of anthropogenic activities [40]. The details are as follows: first, soil pH values were categorized into three classes, <6.5 , $6.5 \leq \text{pH} \leq 7.5$, and >7.5 ; second, the pollution threshold for each soil heavy metal was determined by land use (e.g., paddy fields) and pH class; third, the pollution index (PI) for each heavy metal was determined (Equation (1)); and finally, the Nemerow integrated pollution index (NIPI) was calculated (Equation (2)) [41–44].

$$PI = C_i/S_i \tag{1}$$

where C_i is the concentration of the soil heavy metal i and S_i is the pollution threshold of i heavy metal i in soil.

$$NIPI = ((P_{imax})^2 + (\bar{PI})^2)^{1/2} \tag{2}$$

where P_{imax} is the maximum PI value of each heavy metal and \bar{PI} is the mean PI of each heavy metal. NIPI is a comprehensive index which was used to classify the soils in terms of heavy metal pollution.

PI is divided into four levels from no to high pollution, to indicate the pollution degree and classified as follows: unpolluted ($PI \leq 1$), slightly polluted ($1 < PI \leq 2$), moderately polluted ($2 < PI \leq 3$), and highly polluted ($PI > 3$). However, the classification of NIPI is slightly different from the PI levels, and can be graded as safe ($NIPI \leq 0.7$), precaution ($0.7 < NIPI \leq 1.0$), slight pollution ($1.0 < NIPI \leq 2.0$), moderate pollution ($2.0 < NIPI \leq 3.0$), and heavy pollution ($NIPI > 3.0$)

2.4. Spatial Distribution of Heavy Metals in Soil

Spatial variability in the concentrations of heavy metals was determined using geostatistical methods. Experimental semivariograms were developed to reveal the spatial dependence of soil properties, using the following equation [45]:

$$\gamma^*(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \tag{3}$$

where $\gamma^*(h)$ is the semivariance, $N(h)$ is the number of experimental pairs separated by distance h , $Z(x_i)$ is the measured sample value at point I, and $Z(x_i + h)$ is the measured sample value at

point $i + h$. From the analysis of the experimental variogram, a suitable model is typically fit using weighted least squares and parameters such as range, nugget, and sill prior to the kriging procedure. All geostatistical analyses were carried out using ArcGIS 10.2 (ESRI, ArcGIS 10.2, Redlands, CA, USA). Maps of the spatial distribution of heavy metal concentrations in the study area were generated using ordinary kriging (OK) interpolation of the data in the surface and subsurface soils.

2.5. Delineating Soil Heavy Metal Pollution Based on Uncertainty Analysis

The sample was divided as calibration subset and validation subset with a ratio of 2:1 with a calibration subset of 693 samples and a validation subset of 347 samples. OK and IK was employed on calibration subset to get the spatial distribution of NIPI and spatial map of probability of NIPI > 1. IK is a kriging analysis performed on a binary-transformed sample population. This analysis considers the problem of estimating the probability that the concentration of a pollutant Z exceeds a critical threshold Z_c at an unsampled point u . This approach that was first proposed by Journel (1983) [46] can be used if the spatial correlation of a highly variant parameter is difficult to describe with the raw data. Defining indicators for variables would lead to the following transformation [45]:

$$I(x; z) = \begin{cases} 1 & Z(x) \geq z \\ 0 & Z(x) < z \end{cases} \quad (4)$$

After transforming the observed data to a new set of indicator variables, the experimental semivariogram is calculated for every set of indicators at each cutoff Z_k as

$$\gamma_1^*(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [I(x_i; Z_k) - I(x_i + h; Z_k)]^2 \quad (5)$$

where $\gamma_1^*(h)$ is the indicator experimental semivariogram and $N(h)$ is the number of pairs of indicator transformations $I(x_i; Z_k)$ and $I(x_i + h; Z_k)$ separated by distance vector h .

The conditional cumulative distribution function (ccdf) at each unsampled location, e.g., χ_0 , is then obtained by the indicator kriging estimator:

$$F(\chi_0; Z_k | (n)) = I^*(\chi_0; Z_k) = \sum_{i=1}^n \lambda_i I(x_i; Z_k) \quad (6)$$

where $I^*(\chi_0; Z_k)$ is the estimated indicator transformation at the unsampled location χ_0 and λ_i is the weight assigned to the indicator transformation I at location x_i .

The main steps of delineating heavy metal soil pollution in this study are as follows:

- (1) IK was employed on calibration subset to evaluate the spatial distribution of the probability of NIPI > 1.0, which is the probability of composite heavy metal pollution in the study region. The higher the pollution probability is, the less the uncertainty is. Therefore, we have a sufficient basis to delineate the area with a high pollution probability as the contaminated zone.
- (2) To obtain the optimal probability for delineating pollution, misclassifications of samples in validation subset as contaminated or clean with different pollution probabilities were plotted. The probability that had the highest accuracy was selected as the optimal threshold probability, meaning that a location with a pollution probability larger than this threshold was regarded as contaminated land; otherwise, the site was classified as clean land.
- (3) The pollution area was delineated according to the optimal pollution probability.
- (4) Misclassification rates of delineating pollution based on composited heavy metal pollution uncertainty based on IK and spatial distribution of NIPI through OK were calculated and compared using a validation subset. Misclassification includes false positive errors, which classifies uncontaminated samples as contaminated sites, resulting in unnecessary

expenditure on site remediation, and false negative errors, which classify polluted sites as unpolluted sites, leading to a potential decline in human health.

2.6. Data Analysis

Microsoft Excel 2010 (Office 2010, Redmond, WA, USA) was deployed to make statistical analyses, and ArcGIS10.2 software (ESRI, ArcGIS 10.2, Redlands, CA, USA) and GS 9.0+ (Gamma Design Software, GS + 9.0, LLC Plainwell, MI, USA) were used to map the sampling sites and perform geostatistical analyses.

3. Results and Discussion

3.1. Exploratory Data Analysis

The mean contents of As, Cd, Cr, Hg, Ni, Pb, and Zn were 6.55, 0.19, 61.84, 33.87, 0.27, 23.85, 39.86, and 99.60 mg/kg, respectively. The mean concentrations of all heavy metals in the samples were less than the mean concentrations in the national secondary standard (CEPA, 1995) [47]. However, the highest levels of Cd, Cu, Hg, Ni, Cu, and Zn were higher than the concentrations in the national secondary standard. The coefficients of variation for heavy metal concentrations decreased in the following order: Cd > Hg > Cu > Pb > Ni > Zn > Cr > As. Among them, As, Cr, Ni, Pb, Cu and Zn exhibited moderate variability, with coefficients of variation of 34.61%, 38.68%, 46.61%, 47.59%, 92.99% and 40.77%, respectively, and Cd and Hg exhibited the greatest variability, with coefficients of variation of 195.13% and 121.82% (as shown in Table 1).

Table 1. Descriptive statistics of heavy metal concentrations in the research area (mg/kg).

Items	As	Cd	Cr	Cu	Hg	Ni	Pb	Zn
Sample numbers	1040	1040	1040	1040	1040	1040	1040	1040
Mean	6.55	0.19	61.84	33.87	0.27	23.85	39.86	99.60
Std	2.27	0.37	23.92	31.50	0.33	11.12	18.97	40.60
Min	1.80	0.01	6.50	1.00	0.02	4.00	16.00	31.50
Max	29.10	11.76	262.70	685.40	3.42	131.99	313.00	749.90
CV (%)	34.61	195.13	38.68	92.99	121.82	46.61	47.59	40.77
Background value [48]	5.75	0.161	56.1	23.1	0.076	20.7	36.2	86.6
Data distribution	Log ND [†]	Log ND	Log ND	Log ND	Log ND	Log ND	Log ND	Log ND

[†] ND is the abbreviation of normal distribution.

3.2. Heavy Metal Pollution Assessment

The mean PIs of studied heavy metals Cr, Pb, Hg, Cd, AS, Cu, Zn and Ni in soil were 0.24, 0.13, 0.54, 0.31, 0.24, 0.33, 0.40 and 0.48, respectively (Table 2). The mean PI of all the heavy metals in the samples was less than 1, which indicated the content of the heavy metals in soil in study are were at safe levels. However, it is important to note that the maximum PIs for Cr and Pb signified that these metals were at low pollution levels; based on the maximum PI for Ni and Zn, these metals were at mild pollution levels; the maximum PI of Hg, Cu, and Cd indicated that they were at a severe level of pollution.

The mean value of the NIPI was 0.59, which was a safe level. However, the maximum NIPI was 14.02, and this level indicates a severe level of pollution. Overall, the soil contamination in the study area was at levels considered safe.

Table 2. Descriptive statistics of the single heavy metal pollution index (PI) of heavy metals in the study area.

Items	Cr	Pb	Hg	Cd	As	Cu	Zn	Ni
Sample numbers	1040	1040	1040	1040	1040	1040	1040	1040
Mean value	0.24	0.13	0.54	0.31	0.24	0.33	0.40	0.48
Std	0.10	0.06	0.66	0.62	0.08	0.31	0.16	0.22
Min	0.03	0.05	0.03	0.02	0.06	0.01	0.13	0.08
Max	1.31	1.04	6.84	19.60	0.97	6.85	3.00	2.64
CV (%)	41.58	47.67	121.86	195.24	34.63	92.91	40.73	46.62

3.3. Spatial Distribution of Soil Heavy Metals

The semivariograms for the four heavy metals are listed in Table 3. The semivariogram of Cr, Hg, As, Cu, and Ni were well fitted with the spherical model, while the semivariogram of Pb, Cd, and Zn were well fitted with the exponential model. The nugget/sill ratio can be used to classify the spatial dependence of heavy metals. If the ratio is less than 25%, then the variable has strong spatial dependence. When the ratio is between 25% and 75%, the variable has moderate spatial dependence, while the variable shows weak spatial dependence when the ratio is greater than 75% [49].

Table 3. Parameters of the semivariogram models of different heavy metals.

Elements	Model Types	C ₀	C	A ₀ (m)	R ²	RSS	C ₀ /(C)	Data Distribution
Cr	Spherical	0.007	0.040	31,100	0.973	3.69×10^{-5}	17.7%	Log ND [†]
Pb	Exponential	0.009	0.003	54,500	0.978	8.24×10^{-6}	26.0%	Log ND [†]
Cd	Exponential	0.014	0.028	11,700	0.703	6.06×10^{-5}	49.8%	Log ND [†]
Hg	Spherical	0.032	0.187	38,200	0.966	1.24×10^{-3}	17.1%	Log ND [†]
As	Spherical	0.011	0.023	42,400	0.987	2.69×10^{-6}	46.7%	Log ND [†]
Cu	Spherical	0.024	0.084	17,000	0.904	2.82×10^{-4}	28.9%	Log ND [†]
Zn	Exponential	0.007	0.020	38,100	0.978	3.03×10^{-6}	36.0%	Log ND [†]
Ni	Spherical	0.005	0.050	26,300	0.900	2.43×10^{-4}	10.6%	Log ND [†]

[†] ND is the abbreviation of normal distribution.

The nugget/sill ratios ranged from 26.0% to 49.8% for Pb, Cu, Zn, As and Cd, which indicates that these metals all had moderate spatial dependence. This result indicated that the spatial variation of Pb, Cu, Zn, As, and Cd were affected by both structural and random factors. The nugget/sill ratios for Ni, Cr, and Hg ranged from 10.6% to 17.7%, which indicates that these metals all have strong spatial dependence. This result revealed that the spatial variation of Ni, Cr, and Hg are mainly controlled by structural factors.

The highest Cr concentration in the surface soil was in the central portion of the study area. In the southern part of the study area, the Cr content was very low. The Pb content was high in the central and southern portions of the study area and low in the other areas. The Cd value was high in the mid-western and southeastern parts of the study area, while the value was low mainly in the northern portion of the study area. The concentration of Hg in the central part of the study area was higher than in any other places in the study area. The As and Ni concentrations were low in the southern and central portions of the study area and high in other parts of the study area. The concentrations of Cu and Zn were high in the central region and low in other areas (Figure 2). The central region was the core urban area of the study area. In addition, many anthropogenic sources of metals such as industrial and business production activities, vehicle exhaust and aerial deposition are clustered in the urban area and contribute to the enrichment of heavy metals in the soil.

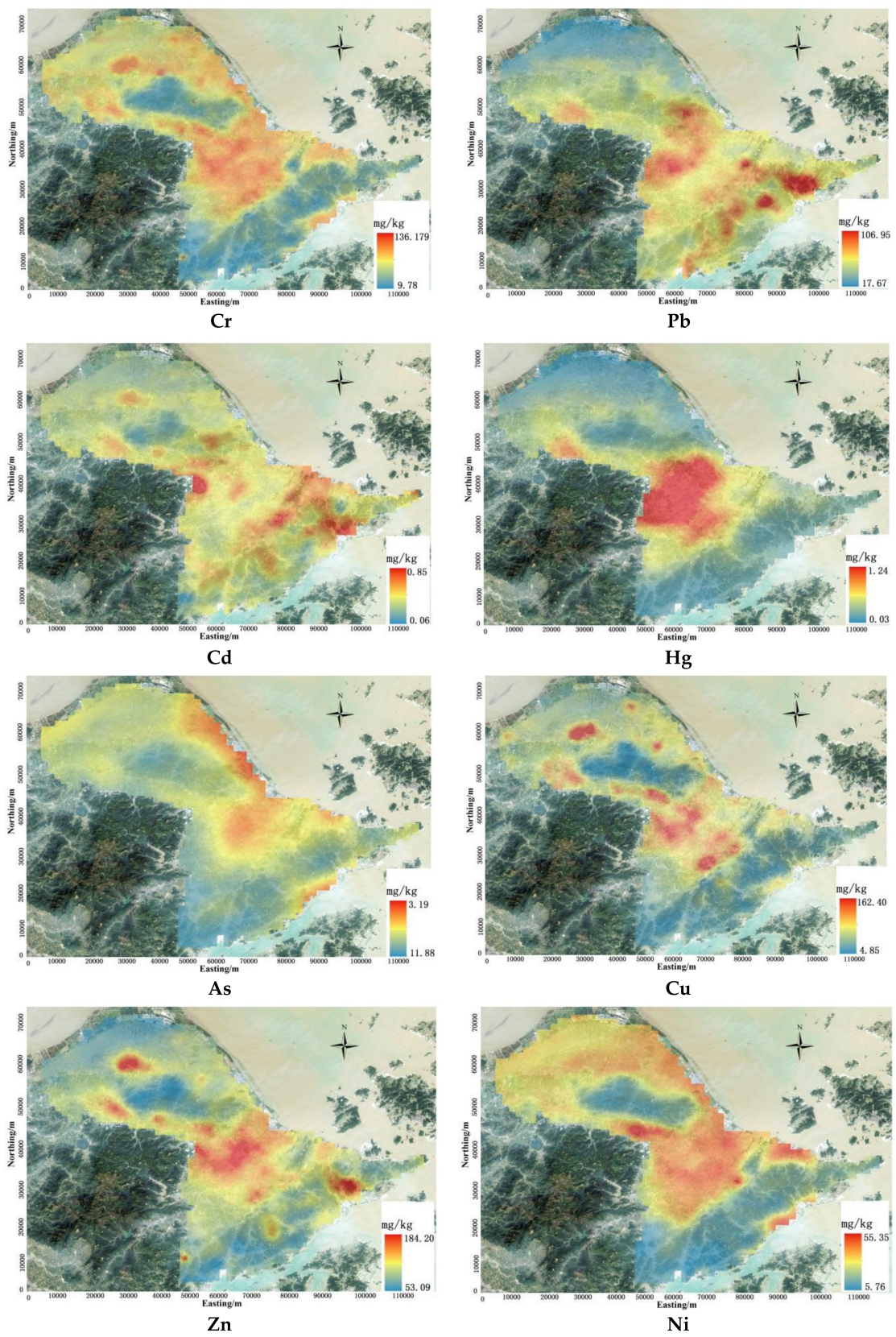


Figure 2. Spatial distribution of heavy metal concentrations.

3.4. Delineating Heavy Metal Soil Pollution Based on Uncertainty Analysis

The spatial distribution of the probability composited heavy metal pollution in the study area is shown in Figure 3.

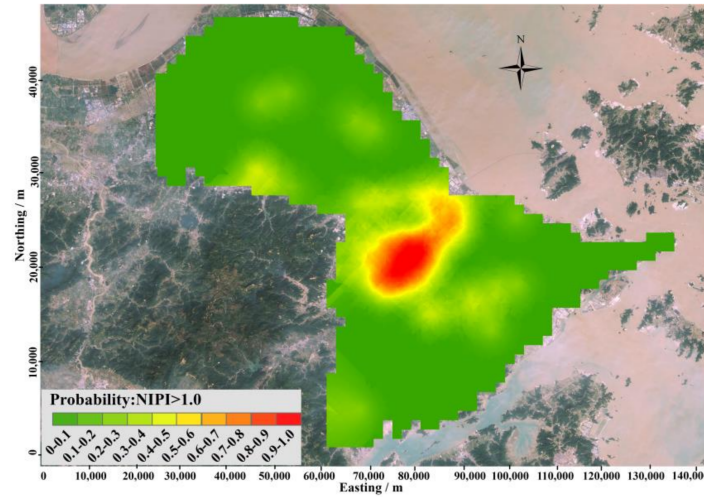


Figure 3. Spatial distribution of the probability of Nemerow integrated pollution index (NIPI) > 1.0.

As shown in Figure 3, the central portion of the study area had the highest probability of composite heavy metal pollution. The probability of composite heavy metal contamination in the soil in this region was greater than 80%. This result indicated that there was little uncertainty around delineating this area as a contaminated area.

The spatial distribution of NIPI in study area was shown in Figure 4 which was got by OK. The red areas represent areas with NIPIs larger than 1.0, and are regarded as polluted areas. The green areas represent areas with NIPIs less than 1.0, and are regarded as unpolluted areas. As revealed by Figure 4, the polluted areas are mainly located in the central region of the study area, and almost overlapped with an area with high probability of being an integrative polluted area, shown in Figure 3. It indicated the central region of study area was polluted by heavy metals. Since this region was the core urban area, anthropogenic activities made main contribution to heavy metals pollution in these areas.

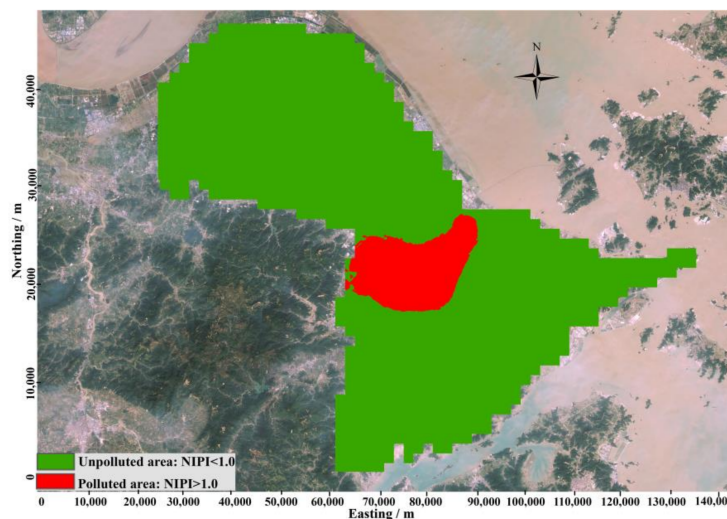


Figure 4. Delineation of composite heavy metal pollution according to the spatial distribution of Nemerow integrated pollution index (NIPI).

The threshold probability that was used to define polluted and unpolluted areas is critical for delineating contamination. To obtain the optimal probability for delineating pollution, a plot of the percentage of locations misclassified versus the probability threshold was generated (as shown in Figure 5). When the threshold probability is close to 1.00, the misclassification rate will be close to the proportion of integrative polluted samples with NIPI larger than 1.0, and more and more potential polluted samples would be classified as unpolluted samples. Therefore, when threshold probability increased from 0.95 to 1.00, the misclassification rate was sharply increased from 8.36% to 23.34%, and vice versa. The probability value at which a misclassification was minimal was 0.60. In addition, the misclassification rate was 7.20%. This threshold was finally chosen to classify areas as polluted and unpolluted lands as shown in Figure 6. According to the threshold level, polluted areas are mainly distributed in the central part of study area, which is consisted with the result revealed by Figures 3 and 4.

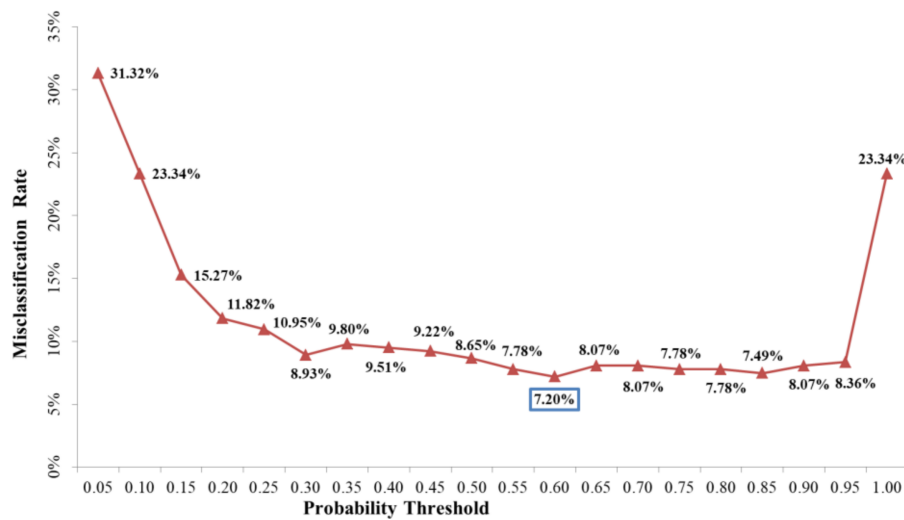


Figure 5. Misclassification rate vs the probability thresholds of composite heavy metal pollution risk.

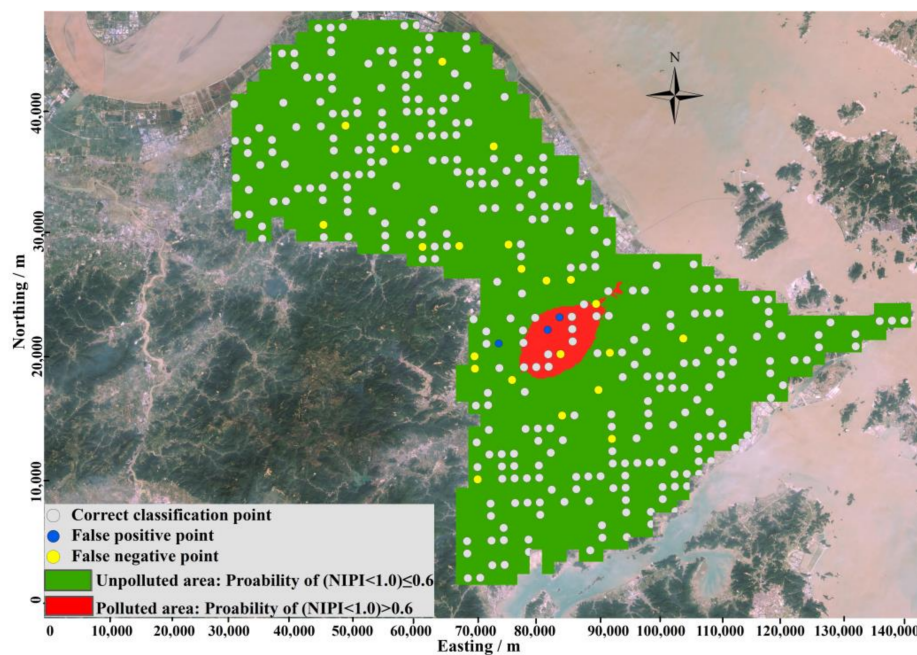


Figure 6. Delineation of composite heavy metal pollution according to pollution probability.

Then, the result of pollution scope delineation based on probability threshold was used to assess the validation subset. After that, the assessment based on probability threshold delineation was compared with the assessment based on NIPI to define the accuracy of pollution scope delineation. A false positive error occurs when unpolluted samples are classified as polluted, and a false negative error occurs when polluted samples are classified as unpolluted.

As represented by Figure 6, false positive errors occurred rarely, and are mainly located on polluted regions. These samples are mainly located at the contaminated area or the boundary between the contaminated area and the uncontaminated area. When using IK to obtain spatial pollution probabilities, the probabilities are affected by the surrounding polluted samples, and contamination probabilities in these areas will be higher than their actual values. However, false negative errors mainly occurred in isolation in many areas of the study area. This result mainly occurred because the NIPIs in these locations were high compared with surrounding locations, and when using IK to obtain spatial pollution probabilities that are affected by surrounding unpolluted areas, contamination probabilities in these areas will be lower than their actual values. These locations were also easily ignored, and thus, the threats to the health of citizens were also ignored. So, enough attention is needed to focus on these locations to provide potential and hidden health caused by soil heavy metals pollution in these locations.

The misclassification rate of contamination delineation method is based on spatial distribution of NIPI (Figure 4), and the contamination delineation method based on uncertainty probability (Figure 6) were calculated and compared, as represented in Table 4. When pollution is defined based on the spatial pollution of NIPI (as shown in Figure 4), the validation samples with false negative errors totaled 17, which is 4.90% of the total study area. In addition, samples with false positive errors totaled 17, which accounted for 4.90% of the total study area. Overall, the accuracy of classification was 90.20%.

Table 4. Statistics of the delineated area (total validation sample number = 347).

Items	Classification Based on Uncertainty Probability of NIPI > 1		Classification Based on Spatial Distribution of NIPI	
	Sample Number	Proportion	Sample Number	Proportion
False positive errors	3	0.86%	17	4.90%
False negative errors	22	6.34%	17	4.90%
Correct	322	92.80%	313	90.20%

When delineate pollution based on uncertainty of heavy metals pollution as shown in Figure 6, the validation samples with false negative errors totaled 22, which is 6.34% of the total study area. In addition, samples with false positive errors totaled 3, which accounted for 0.86% of the total study area. Overall, the accuracy of classification was 92.8%. This result indicated that this method can delineate pollution with very high accuracy, and the accuracy is higher than the pollution delineation method based on spatial distribution of NIPI obtained by OK. In addition, this method was also very easy to implement.

4. Conclusions

Surface soil contamination by eight studied heavy metals was evaluated using PI and NIPI with a set of 1040 soil samples in a coastal industrial city in the YRD of China. The spatial distribution of heavy metals was assessed by OK. In addition, IK was used to investigate the spatial uncertainty of the probability of pollution and to delineate soil heavy metal contamination. The results showed that the mean concentrations of all the heavy metals in the samples were less than the mean concentrations in the national secondary standard. The mean PI of Cr, Pb, Hg, Cd, AS, Cu, Zn and Ni were 0.24, 0.13, 0.54, 0.31, 0.24, 0.33, 0.40 and 0.48, respectively, all of which were at safe levels. The mean value of the NIPI was 0.59, which was also at a safe level.

The central region was the core urban area in the study area. Heavy metals were most accumulated in this area. This region also had large probabilities of composite heavy metal pollution, with pollution probabilities even higher than 0.8. A concern is that this region also has the highest population density. Thus, measures should be taken to hinder the heavy metal accumulation that occurs there.

After validation with the validation subset samples, 0.60 was selected as the optimal probability to use to define contamination locations. According to the spatial distribution of uncertainty probability of composite heavy metals contamination (Figure 3), polluted zones were mainly distributed in the central part of the study area. The sample proportion of false negative errors was 6.34% compared with 4.90% calculated from the pollution delineation method based on OK, while the percentage of samples with false positive errors was 0.86% compared with 4.90% when using pollution delineation method based on OK. The accuracy of classification was 92.80% compared with 90.2% using pollution delineation method based on OK, which indicated that this method could delineate pollution regions with very higher and satisfied accuracy.

However, to more precisely delineate the polluted areas, more attention needs to be paid to the following issues:

- (1) The available content of heavy metals should be used to replace the total concentrations of heavy metals to get a conclusion which is closer to reality.
- (2) Other factors that control metal bioavailability, such as chemical partitioning, which in turn is affected by soil chemical properties, should also be considered.
- (3) In this study, sample density was still sparse, and sampling density needs to be improved to obtain a higher resolution map.
- (4) In this study, the ratio of the sample number of validation subset and calibration subset is 1:2, and the validation subset was randomly extracted from samples. However, the proportion of sample number of validation subset and calibration subset and the spatial pattern of validation subset may have a certain effect on the choice of optimum threshold probability which is used to define pollution sites.
- (5) Contamination in soils cannot be adequate, and thresholds based on local variability should be used for properly assessing heavy metals contamination, which cannot be archived by IK.

Acknowledgments: This work was supported by the Key Research and Development Project of Zhejiang Province (2015C02011), the National Key Research and Development Program of China (2016YFD0201200), the National Natural Science Foundation of China (41771244) and the Fundamental Research Funds for the Central Universities. We also acknowledge the support received by Bifeng Hu from the Chinese Scholarship Council (under grant agreement No. 201706320317) for 3 years' Ph.D. study in INRA and Orléans University.

Author Contributions: Yan Li and Bifeng Hu conceived and designed the research; Bifeng Hu, Ruiying Zhao and Songchao Chen performed the research; Bifeng Hu, Ruiying Zhao, Songchao Chen and Yue Zhou analysed the data; Bifeng Hu, Songchao Chen, Bin Jin, and Yan Li contributed reagents/materials/analysis tools; Bifeng Hu, Yan Li, Ruiying Zhao and Zhou Shi wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Vega, F.A.; Covelo, E.F.; Andrade, M.L.; Marcet, P. Relationships between heavy metals content and soil properties in minesoils. *Anal. Chim. Acta* **2004**, *524*, 141–150. [[CrossRef](#)]
2. World Reference Base (WRB). *World Reference Based for Soil Resources 2006*; World Soil Resources Report 103; Food and Agriculture Organization of the United Nations: Rome, Italy, 2006.
3. Obrador, A.; Alvarez, J.M.; Lopez-Valdivia, L.M.; Gonzalez, D.; Novillo, J.; Rico, M.I. Relationships of soil properties with Mn and Zn distribution in acidic soils and their uptake by a barley crop. *Geoderma* **2007**, *137*, 432–443. [[CrossRef](#)]
4. Facchinelli, A.; Sacchi, E.; Mallen, L. Multivariate statistical and GIS-based approach to identify heavy metal sources in soils. *Environ. Pollut.* **2001**, *114*, 313–324. [[CrossRef](#)]
5. Wei, B.G.; Yang, L.S. A review of heavy metal contaminations in urban soils, urban road dusts and agricultural soils from China. *Microchem. J.* **2010**, *94*, 99–107. [[CrossRef](#)]

6. Hu, B.F.; Wang, J.Y.; Jin, B.; Li, Y.; Shi, Z. Assessment of the potential health risks of heavy metals in soils in a coastal industrial region of the Yangtze River Delta. *Environ. Sci. Pollut. Res.* **2017**, *24*, 19816–19826. [[CrossRef](#)] [[PubMed](#)]
7. Hu, B.F.; Chen, S.C.; Hu, J.; Xia, F.; Xu, J.F.; Li, Y.; Shi, Z. Application of portable XRF and VNIR sensors for rapid assessment of soil heavy metal pollution. *PLoS ONE* **2017**, *12*, 1–13. [[CrossRef](#)] [[PubMed](#)]
8. National Council of SPCAs (NSPCIR); Ministry of Environmental Protection; Ministry of Land and Resources. The National Soil Pollution Condition Investigation Report [EB/OL]. NSPCIR: Beijing, China, 2014. Available online: <http://www.gov.cn/foot/site1/20140417/782bcb88840814ba158d01.pdf> (accessed on 17 April 2014).
9. Qu, C.S.; Ma, Z.W.; Yang, J.; Liu, Y.; Bi, J.; Huang, L. Human exposure pathways of heavy metals in a lead-zinc mining area, Jiangsu Province, China. *PLoS ONE* **2012**, *7*, 1–11. [[CrossRef](#)] [[PubMed](#)]
10. Hu, B.F.; Jia, X.L.; Hu, J.; Xu, D.Y.; Xia, F.; Li, Y. Assessment of Heavy Metal Pollution and Health Risks in the Soil-Plant-Human System in the Yangtze River Delta, China. *Int. J. Environ. Res. Public Health* **2017**, *14*, 1042. [[CrossRef](#)] [[PubMed](#)]
11. Cheng, J.L.; Shi, Z.; Zhu, Y.W. Assessment and mapping of environmental quality in agricultural soils of Zhejiang Province, China. *J. Environ. Sci.* **2007**, *19*, 50–54. [[CrossRef](#)]
12. Xia, F.; Peng, J.; Wang, Q.L.; Zhou, L.Q.; Shi, Z. Prediction of heavy metal content in soil of cultivated land: Hyperspectral technology at provincial scale. *J. Infrared Millim. Waves* **2015**, *34*, 593–598.
13. Li, X.D.; Lee, S.L.; Wong, S.C.; Shi, W.Z.; Thornton, I. The study of metal contamination in urban soils of Hong Kong using a GIS-based approach. *Environ. Pollut.* **2004**, *129*, 113–124. [[CrossRef](#)] [[PubMed](#)]
14. Saby, N.P.A.; Arrouays, D.; Boulonne, L.; Pochot, J.C. Geostatistical assessment of Pb in soil around Paris, France. *Sci. Total Environ.* **2006**, *367*, 212–221. [[CrossRef](#)] [[PubMed](#)]
15. Wong, C.S.C.; Li, X.D.; Thornton, I. Urban environmental geochemistry of trace metals. *Environ. Pollut.* **2006**, *142*, 1–16. [[CrossRef](#)] [[PubMed](#)]
16. Jiang, Y.X.; Chao, S.H.; Liu, J.W.; Yang, Y.; Chen, Y.J.; Zhang, A.C.; Cao, H.B. Source apportionment and health risk assessment of heavy metals in soil for a township in Jiangsu Province, China. *Chemosphere* **2017**, *168*, 1658–1668. [[CrossRef](#)] [[PubMed](#)]
17. Lequy, E.; Saby, N.P.A.; Ilyin, I.; Bourin, A.; Sauvage, S.; Leblond, S. Spatial analysis of trace elements in a moss bio-monitoring data over France by accounting for source, protocol and environmental parameters. *Sci. Total Environ.* **2017**, *590–591*, 602–610. [[CrossRef](#)] [[PubMed](#)]
18. Verstraete, S.; Van, M.M. A multi-stage sampling strategy for the delineation of soil pollution in a contaminated brownfield. *Environ. Pollut.* **2008**, *154*, 184–189. [[CrossRef](#)] [[PubMed](#)]
19. Zawadzki, J.; Fabijańczyk, P. Use of variograms for field magnetometry analysis in Upper Silesia Industrial Region. *Studia Geophys. Geod.* **2007**, *51*, 535–550. [[CrossRef](#)]
20. Zawadzki, J.; Fabijańczyk, P. The geostatistical reassessment of soil contamination with lead in metropolitan Warsaw and its vicinity. *Int. J. Environ. Pollut.* **2008**, *35*, 1–12. [[CrossRef](#)]
21. Saby, N.P.A.; Thioulouse, J.; Jolivet, C.; Ratié, C.; Boulonne, L.; Bispo, A.; Arrouays, D. Multivariate analysis of the spatial patterns of 8 trace elements using the French soil monitoring network data. *Sci. Total Environ.* **2009**, *407*, 5644–5652. [[CrossRef](#)] [[PubMed](#)]
22. Marchant, B.; Saby, N.P.A.; Lark, R.; Bellamy, P.; Jolivet, C.; Arrouays, D. Robust analysis of soil properties at the national scale: Cadmium content of French soils. *Eur. J. Soil Sci.* **2010**, *61*, 144–152. [[CrossRef](#)]
23. Arrouays, D.; Saby, N.P.A.; Thioulouse, J.; Jolivet, C.; Boulonne, L.; Ratié, C. Large trends in french topsoil characteristics are revealed by spatially constrained multivariate analysis. *Geoderma* **2011**, *161*, 107–114. [[CrossRef](#)]
24. Saby, N.P.A.; Marchant, B.; Lark, R.; Jolivet, C.; Arrouays, D. Robust geostatistical prediction of trace elements across France. *Geoderma* **2011**, *162*, 303–311. [[CrossRef](#)]
25. Lacarce, E.; Saby, N.P.A.; Martin, M.P.; Marchant, B.P.; Boulonne, L.; Meersmans, J.; Jolivet, C.; Arrouays, D. Mapping soil Pb stocks and availability in mainland France combining regression trees with robust geostatistics. *Geoderma* **2012**, *170*, 359–368. [[CrossRef](#)]
26. Zawadzki, J.; Fabijańczyk, P. Geostatistical evaluation of lead and zinc concentration in soils of an old mining area with complex land management. *Int. J. Environ. Sci. Technol.* **2013**, *10*, 729–742. [[CrossRef](#)]
27. Huang, J.Y.; Malone, B.P.; Minasny, B.; Mcbratney, A.B.; Triantafyllis, J. Evaluating a bayesian modelling approach (INLA-SPDE) for environmental mapping. *Sci. Total Environ.* **2017**, *609*, 621–632. [[CrossRef](#)] [[PubMed](#)]
28. Hu, B.F.; Wang, J.Y.; Fu, T.T.; Li, Y.; Shi, Z. Application of Spatial Analysis on Soil Heavy Metal Contamination: A Review. *Chi. J. Soil Sci.* **2017**, *48*, 1014–1022. (In Chinese)

29. Vicente-Serrano, S.; Saz-Sánchez, M.; Cuadrat, J. Comparative analysis of interpolation methods in the middle Ebro valley (Spain): Application to annual precipitation and temperature. *Clim. Res.* **2003**, *24*, 161–180. [CrossRef]
30. Wu, C.F.; Huang, J.Y.; Minasny, B.; Zhu, H. Two-dimensional empirical mode decomposition of heavy metal spatial variation in agricultural soils, southeast china. *Environ. Sci. Pollut. Res.* **2017**, *24*, 8302–8314. [CrossRef] [PubMed]
31. Qiu, M.; Li, F.; Wang, Q.; Chen, J.; Yang, G.; Liu, L. Driving forces of heavy metal changes in agricultural soils in a typical manufacturing center. *Environ. Monit. Assess.* **2015**, *187*, 239. [CrossRef] [PubMed]
32. Yang, Y.; Mei, Y.; Zhang, C.; Zhang, R.; Liao, X.; Liu, Y. Heavy metal contamination in surface soils of the industrial district of Wuhan, China. *Hum. Ecol. Risk Assess. Int. J.* **2016**, *22*, 126–140. [CrossRef]
33. Cattle, J.A.; McBratney, A.B.; Minasny, B. Kriging method evaluation for assessing the spatial distribution of urban soil lead contamination. *J. Environ. Qual.* **2002**, *31*, 1576–1588. [CrossRef] [PubMed]
34. Zhang, C.; Wu, L.; Luo, Y.; Zhang, H.; Christie, P. Identifying sources of soil inorganic pollutants on a regional scale using a multivariate statistical approach: Role of pollutant migration and soil physicochemical properties. *Environ. Pollut.* **2008**, *151*, 470–476. [CrossRef] [PubMed]
35. Ungaro, F.; Ragazzi, F.; Cappellin, R.; Giandon, P. Arsenic concentration in the soils of the Brenta Plain (Northern Italy): Mapping the probability of exceeding contamination thresholds. *J. Geochem. Explor.* **2008**, *96*, 117–131. [CrossRef]
36. Von Steiger, B.; Webster, R.; Schulin, R.; Lehmann, R. Mapping heavy metals in polluted soil by disjunctive kriging. *Environ. Pollut.* **1996**, *94*, 205–215. [CrossRef]
37. Goovaerts, P.; Webster, R.; Dubois, J.P. Assessing the risk of soil contamination in the Swiss Jura using indicator geostatistics. *Environ. Ecol. Stat.* **1997**, *4*, 31–48. [CrossRef]
38. Brus, D.J.; de Gruijter, J.J.; Walvoort, D.J.J.; de Vries, F.; Bronswijk, J.J.B.; Romkens, P.F.A.M.; de Vries, W. Mapping the probability of exceeding critical thresholds for cadmium concentrations in soils in the Netherlands. *J. Environ. Qual.* **2002**, *31*, 1875–1884. [CrossRef] [PubMed]
39. State Environment Protection Administration of China (SEPA). *Technical Guidelines for Risk Assessment of Contaminated Sites*; SEPA: Beijing, China, 2009. Available online: www.mep.gov.cn/gkml/hbb/bgth/200910/W020091009550671751947.pdf (accessed on 14 March 2018).
40. Loska, K.; Wiechula, D.; Korus, I. Metal contamination of farming soils affected by industry. *Environ. Int.* **2004**, *30*, 159–165. [CrossRef]
41. China National Environmental Protection Agency (CEPA). *The Technical Specification for Soil Environmental Monitoring*; Report No. HJ/T 166-2004; China National Environmental Protection Agency: Beijing, China, 2004.
42. Chen, H.; Teng, Y.; Lu, S.; Wang, Y.; Wang, J. Contamination features and health risk of soil heavy metals in China. *Sci. Total Environ.* **2015**, *512*, 143–153. [CrossRef] [PubMed]
43. Singh, S.; Raju, N.J.; Nazneen, S. Environmental risk of heavy metal pollution and contamination sources using multivariate analysis in the soils of Varanasi environs, India. *Environ. Monit. Assess.* **2015**, *187*, 345. [CrossRef] [PubMed]
44. Zang, F.; Wang, S.; Nan, Z.; Ma, J.; Zhang, Q.; Chen, Y.; Li, Y. Accumulation, spatio-temporal distribution, and risk assessment of heavy metals in the soil-corn system around a polymetallic mining area from the Loess Plateau, northwest China. *Geoderma* **2017**, *305*, 188–196. [CrossRef]
45. Webster, R.; Oliver, M.A. *Geostatistics for Environmental Scientists*; John Wiley & Sons Ltd.: Chichester, UK, 2001.
46. Journel, A.G. Nonparametric estimation of spatial distributions. *Math. Geol.* **1983**, *15*, 445–468. [CrossRef]
47. China National Environmental Protection Agency (CEPA). *Environmental Quality Standard for Soils*; Report No. GB15618-1995; China National Environmental Protection Agency: Beijing, China, 1995. (In Chinese)
48. Dong, Y.X.; Zheng, W.; Zhou, G.H. *Soil Geochemical Background value of Zhejiang Province*; The Geological Publishing House: Beijing, China, 2008.
49. Liu, X.M.; Wu, J.J.; Xu, J.M. Characterizing the risk assessment of heavy metals and sampling uncertainty analysis in paddy field by geostatistics and GIS. *Environ. Pollut.* **2006**, *141*, 257–264. [CrossRef] [PubMed]

