sustainability

MDPI

Article

Potential Risk Recognition of Agricultural Land Based on Agglomeration Characteristics of Pollution-Related Enterprises: A Case Study on the Black Soil Region in Northeast China

Xiaofeng Zhao [1,2,3,†], Changhe Wei [4,*,†], Jiufen Liu [1,2,3,*], Xiaohuang Liu [1,3], Xiaoming Wan [5,6], Mei Lei [5,6] and Shaobin Wang [5,6]

1 Comprehensive Survey Command Center for Natural Resources, China Geological Survey, Beijing 100055, China; zhaoxf2011@163.com (X.Z.)
2 School of Earth Science and Resources, China University of Geosciences (Beijing), Beijing 100083, China
3 Key Laboratory of Coupling Process and Effect of Natural Resources Elements, Beijing 100055, China
4 School of Mining and Geomatics Engineering, Hebei University of Engineering, Handan 056038, China
5 Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
6 University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: weichanghe@hebeu.edu.cn (C.W.); liujiufen@mail.cgs.gov.cn (J.L.)
† These authors contributed equally to this work.

Abstract: The black soil region in Northeast China serves as a ballast for food security. However, the presence of scattered polluting enterprises poses a threat to the safety of the surrounding soil and agricultural products. In this study, the distribution patterns and agglomeration features of key industrial enterprises in Northeast China were elucidated through multi-source geographical big data and geographic information system (GIS) spatial analysis. Subsequently, the risk areas were extracted based on their potential impact on the soil environmental quality of the surrounding agricultural lands. The results revealed that pollution-related enterprises were widely distributed but locally clustered in the black soil area. The dominant industries were chemical manufacturing, petroleum processing, coking, and non-ferrous metal mining. The study found that the agricultural land area affected by polluting enterprises was 43,396.13 km^2, with the majority being at a low-risk level (83.42%). High-risk areas (1646.62 km^2) were mostly aggregated west of Hulunbuir, east of Xilingol, and in most of Chifeng. These areas were primarily affected by the non-ferrous metal mining industry. Other high-risk hotspots were mainly influenced by the chemical manufacturing and metal processing industries. The emissions from industrial and mining enterprises are important heavy metals in the agricultural lands in this region. However, it is important to note that there are other sources of pollution as well. These results may contribute to future investigations on soil environmental quality and pollution source control in the black soil region in Northeast China.

Keywords: black soil region; potential pollution risk; pollution-related enterprises; multi-source geographical big data

Citation: Zhao, X.; Wei, C.; Liu, J.; Liu, X.; Wan, X.; Lei, M.; Wang, S. Potential Risk Recognition of Agricultural Land Based on Agglomeration Characteristics of Pollution-Related Enterprises: A Case Study on the Black Soil Region in Northeast China. *Sustainability* **2024**, *16*, 417. https://doi.org/10.3390/su16010417

Academic Editor: Marko Vinceković

Received: 29 October 2023
Revised: 1 December 2023
Accepted: 21 December 2023
Published: 3 January 2024

1. Introduction

The black soil region in Northeast China is an important commodity, grain-producing, and animal husbandry area, with an annual grain yield accounting for approximately one quarter of the entire country [1]. Soil environmental quality and agricultural product safety in the black soil region directly impact strategic food security stability in China [2]. Black soil, "a panda in cultivated land", is subject to various problems like unclear base numbers and the current situation during protection and development work, which seriously restricts the precise implementation of protective strategies for soil and land [3,4]. Cultivated land resources in China are currently insufficient. Protective utilization of cultivated land and the enhancement of its quality have become urgent problems in agricultural development [5].

Therefore, it is necessary to propose higher requirements for soil environmental quality in black soil regions [6].

It Is noteworthy that, when including the eastern part of Inner Mongolia, the overall soil environmental quality in the black soil region in Northeast China is favorable; however, there are still many pollution-related industrial and mining enterprises in such regions with a lengthy heavy industry development history. Pollutants from such enterprises may enter the soil through atmospheric deposition, sewage discharge, and surface runoff, thus leading to pollutant accumulation in Northeast China soils [7,8]. Numerous studies on soil pollution source analyses have shown that industrial and mining enterprise sewage disposal activity has become the main source of soil pollution and risk in China [9–12]. According to the 2014 national survey report on soil contamination, the soil exceedance rate around industrial and mining enterprises was 36.3%, which was much higher than that of the total samples throughout China (16.1%) [13–15]. Soil contamination was particularly prominent in some regions, such as the northeastern old industrial base [16,17]. In a typical black soil region in Heilongjiang Province, an analysis of 450 samples showed a 0.9% exceedance rate of heavy metals in the soil, which was mainly caused by mineral resource exploitation in the region [18].

When a suburban black soil region in Jilin Province was considered as the study object, the analysis data of 183 crop samples indicated that Cr, Ni, and Pb accumulated to different extents at 6.55%, 36.78%, and 53.01%, respectively, and the degree of heavy metal accumulation of Ni and Pb was higher [19]. The survey analysis results of 99 samples from Liaoning showed that the Cd and Pb concentrations at some sampling points exceeded the limiting values. Consequently, the heavy metals in the planted vegetables also exceeded the limiting values [20]. In eastern Inner Mongolia (including the Xilingol and Hinggan Leagues, as well as Chifeng and Tongliao), which has a large-scale coal and electricity base in China, the analysis results of 95 soil samples that were obtained from mining areas of grassland indicated a slight pollution of Cd, Pb, and Ni, while Zr and Zn presented partially moderate pollution [21]. It can be concluded that the pollutants discharged from industrial and mining activity in Northeast China, including eastern Inner Mongolia (which is the primary black soil distribution region), have posed threats to local black soil environmental quality and crop security to different degrees. However, restricted by the difficulties in obtaining large amounts of industrial and mining enterprise data in real time, most existing studies were based on smaller areas (industrial districts and counties) or limited sampling sites [22,23]. In addition, existing soil pollution surveys in the study area are subject to problems such as local concentration, small scale, and insufficient sample size due to the high manpower and material resource costs incurred in sampling and detection [24,25]. Moreover, agricultural land, including cultivated land, forest land, and grassland, is widely distributed in the black soil region in Northeast China. Hence, it is difficult to conduct timely and effective surveys to assess the contamination effects of different enterprises using traditional means of sampling analysis in such a large region [12]. In addition, given the strong spatial variability of soil pollution distribution in the target region, the scope and degree of soil contamination caused by industrial and mining activity on a large regional scale in the black soil region have not been fully understood [21]. Under the increasingly urgent need to protect black lands, a rapid and efficient assessment of agricultural land pollution risks induced by industrial and mining activity is necessary in the black soil region in Northeast China [1].

In recent years, big data have undergone rapid development when applied to research on ecological environmental protection [26–28].In particular, soil pollution risk identification and control based on big data have attracted increasing attention from researchers [29–31]. With the advantages of full data samples, coverage and process analysis, and panoramic decision making, big data technology provides an effective approach for soil pollution risk identification and control [30]. However, the scope of soil pollution in the black soil region in Northeast China has not yet been identified using big data. In recent years, China has emphasized pollution reduction from enterprises to cope with soil pollution [32–35]. Accordingly, studies on soil pollution macroscopic effects around industrial and mining enterprises at regional and national scales have recently been conducted [36,37]. Some studies have indicated that, on a large regional scale, the spatial agglomeration of industrial and mining enterprises shows a significant positive spatial correlation with the degree of soil pollution [12]. This spatial correlation was further enhanced after enterprise pollutant discharge intensity was introduced to weight the enterprise spatial hotspot distribution [38]. This was also consistent with studies showing that the spatial pollution emission intensity differentiation of industrial and mining enterprises directly determines the level and accumulation of soil pollution in the surrounding area [33,34]. Therefore, the degree of enterprise spatial agglomeration, as well as attributes related to pollutant discharge intensity, such as enterprise distribution density, enterprise size, and productive life, have become key indicators for agricultural land pollution risk assessment around enterprises [38]. These results lay a theoretical foundation for exploring the impact of the distribution and agglomeration of polluting industrial and mining enterprises on the soil environmental quality of surrounding agricultural lands and risk zone division based on big data.

In this study, the information on key industrial enterprises related to soil pollution in the black soil region in Northeast China was collected and integrated using multi-source geographic big data methodology. Subsequently, the spatial distribution pattern and composition of high-polluting enterprises in the black soil region were clarified through multiple data cleaning, interpretation, and spatialization. The potential impact of these enterprises on the soil quality of neighboring agricultural land was then described based on the source dispersion pathway approach. Finally, potentially high-risk areas were identified. This study aimed to provide effective guidance for further detailed investigations of soil pollution distribution and pollution source reduction in the black soil region in Northeast China.

2. Materials and Methods

2.1. Spatial Scope Delineation of the Black Soil Region in Northeast China and the Extraction of Agricultural Sensitive Areas

According to the definition given by Liu et al. [39], the concentrated distribution regions of black, black calcium, chestnut, and gray wooded soils in Northeast China, as identified in China's soil genetic classification, were delimited as the scope of the black soil region in Northeast China. Geographically, the Northeast China region includes the Heilongjiang, Liaoning, and Jilin Provinces; Hulunbeier; and the Hinggan, Tongliao, Chifeng, and Xilingol Leagues in the eastern Inner Mongolia Autonomous Region.

The specific spatial distribution data of the black soil region were collected from the China spatial soil-type distribution-related 1 km raster data map layers that were fabricated after digitalization of the 1:1,000,000 soil map obtained by the Second National Soil survey [40]. Since the initial soil types presented dispersed spatial distribution and as different soil types were staggered, after buffering and merging discontinuous patches of relevant soil types and removing small patches through the "central gravity gathering method" [39], we obtained the scope and boundary of the black soil region in Northeast China for the follow-up analysis, as shown in Figure 1a. It can be observed that the target region can be largely divided into three main sections: west–eastern Inner Mongolia, central Songnen, and the eastern black soil subregions.

Within the black soil region scope, in accordance with the Soil Environment Quality Risk Control Standard for Soil Contamination of Agricultural Land (Trial), cultivated land (paddy fields and dry land), garden land (other forest lands in the land use spatial dataset), and grassland (high- and medium-coverage grassland in the land use spatial dataset) were extracted as the agricultural sensitive areas of this target region (Figure 1b) [41]. This assessment was based on the land use spatial distribution of the black soil region in Northeast China, which was extracted from China land-use-related remote sensing data with a 1 km resolution in 2020 [42]. Subsequently, all pollution risks were assessed within this spatial scope and its boundaries.

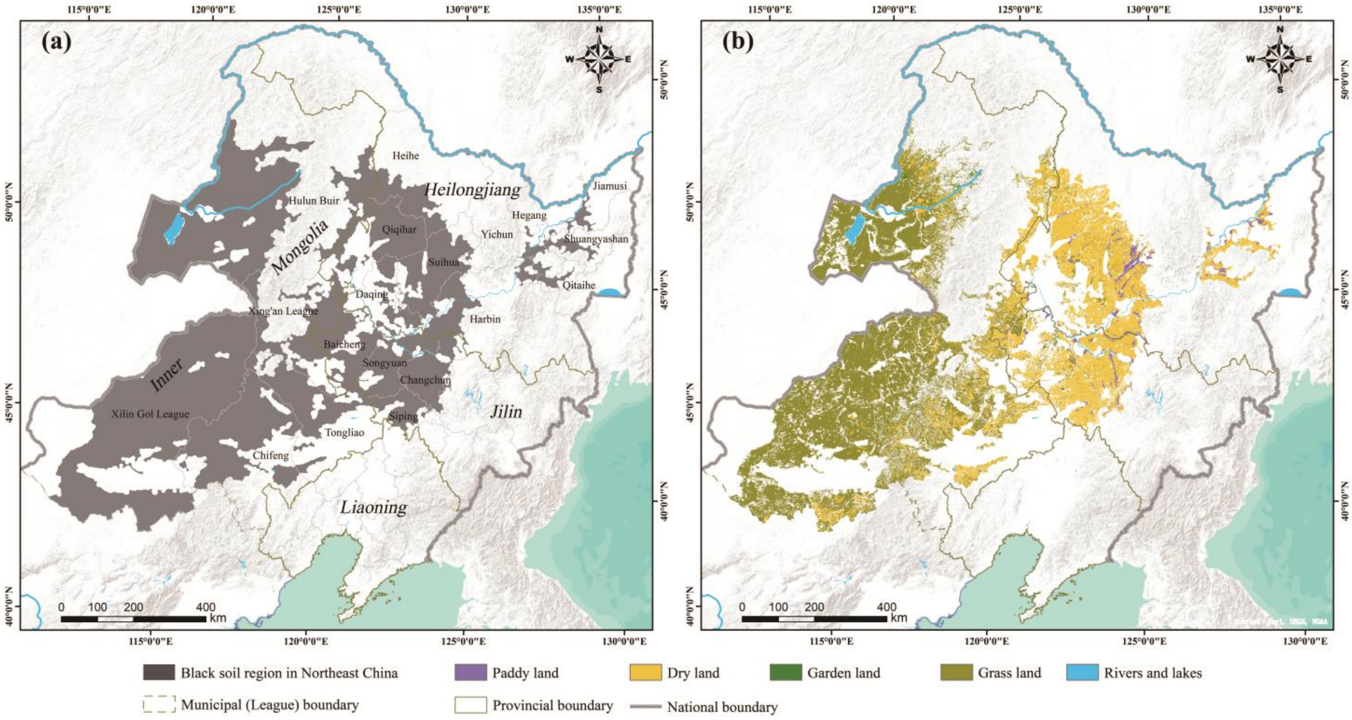

Figure 1. Distribution scope and agricultural sensitive area boundary of the black soil region in Northeast China (data source: Data Center for Resources and Environmental Sciences, the Chinese Academy of Sciences, data acquisition date: 15 August 2023). (**a**): Black soil regional distribution in Northeast China after buffering and merging; (**b**): agricultural sensitive area distribution within the scope of the black soil region in Northeast China.

2.2. Key Soil-Pollution-Related Industry Scope Determination

Defining key soil-pollution-related industries is the premise for collecting relevant enterprise information. In cooperation with a national detailed survey on enterprise land use, China issued the Technical Regulations on Risk Screening and Risk Classification of Enterprise Sites in Production/Closed Enterprises (Trial) in 2017, in which 17 major categories of production and operational industries that might cause soil pollution were screened out based on national industrial classification [43]. Subsequently, the 17 industries were further subdivided into 39 medium types and 73 subtypes in the Guidelines to the Preparation of Detailed Survey Scheme on Provincial Soil Pollution Status and the Guidelines to Soil Environmental Quality Classification for Agricultural Land (Trial), which were used to guide enterprise land use surveys [44,45]. In this study, the enterprise information on key industries related to soil pollution was collected in accordance with the aforementioned scope, with the specific industry types listed in Table 1.

Table 1. Key soil-pollution-related industry scope.

Industry Category	Production and Operational Industries
B Mining industries	07 Petroleum and natural gas exploitation industry 08 Ferrous metal mining and dressing industry 09 Nonferrous metal mining and dressing industry

Table 1. *Cont.*

Industry Category	Production and Operational Industries
C Manufacturing industries	17 Textile industry 19 Leather, fur, feather, as well as their products, and shoemaking industry 22 Papermaking and paper product industry 25 Petroleum refining, coking, and nuclear fuel processing industry 26 Chemical raw materials and products manufacturing industry 27 Pharmaceutical industry 28 Chemical fiber manufacturing industry 31 Ferrous metal smelting and rolling industry 32 Nonferrous metal smelting and rolling industry 33 Metal product industry 38 Electrical machinery and equipment manufacturing industry
G Transportation, warehousing, and postal industries	59 Warehousing industry
N Water conservancy, environment, and public utilities management industries	77 Ecological protection and environmental governance industry 78 Public utilities management industry

2.3. Enterprise Information Collection and Integration in Key Soil-Pollution-Related Industries

As complete enterprise data for key soil-pollution-related industries in various regions of China have not yet been publicized, relevant enterprise information needs to be collected from multiple data sources. In this study, the enterprise data were mainly derived from authoritative public and network data. The former mainly included the National Sewage Discharge Permission Management Information Platform (http://permit.mee.gov.cn/permitExt/defaults/default-index!getInformation.action) (accessed on 8 November 2023), as well as the full-caliber lists of heavy-metal-involved key industries and enterprises, key sewage discharge catalogues, lists of key supervised enterprises related to soil environmental pollution, and lists of hazardous waste management permissions released by provinces in Northeast China within the black soil region. Meanwhile, the data were supplemented by third-party commercial websites, such as the National Enterprise Credit Information Publicity System (https://bt.gsxt.gov.cn/index.html) (accessed on 8 November 2023), the Green Network Environmental Protection Service Center (http://v2.lvwang.org.cn/search) (accessed on 8 November 2023), the Environmental Maps Data Platform (http://www.ipe.org.cn/index.html) (accessed on 8 November 2023), and Tianyancha (https://www.tianyancha.com/) (accessed on 8 November 2023). After the enterprise data were collected according to the industry types and administrative regional scope with black soil, as listed in Table 1, the data types and their attribute fields were unified through document transformation and attribute standardization, and the multi-period data were merged and standardized. Specifically, the dataset attribute fields from different sources (enterprise name, business operation site address, longitude and latitude, industry type, registered capital, plant construction completion time, operating state, and business scope) were unified and were accompanied by information integration, cross-validation, and fusion, thus completing the initial enterprise information dataset integration (Figure 2). To ensure data integration quality, the following principles were followed during attribute information integration. When multiple data sources shared consistent attributes or there was only one data source, integration was performed directly. In the case of multiple data sources and inconsistent attribute content, priority was given to relatively authoritative official public data, particularly records of the national sewage discharge permission information platform. In the face of missing official public data, data from other sources were cross-validated. The attribute fields without data sources were supplemented by combining more network sources (such as third-party website data centers or platforms) during the follow-up data cleaning process [38]. The data were collected using Python based on the PyCharm platform. In addition, the initial data were integrated via Microsoft Excel (Office 2021).

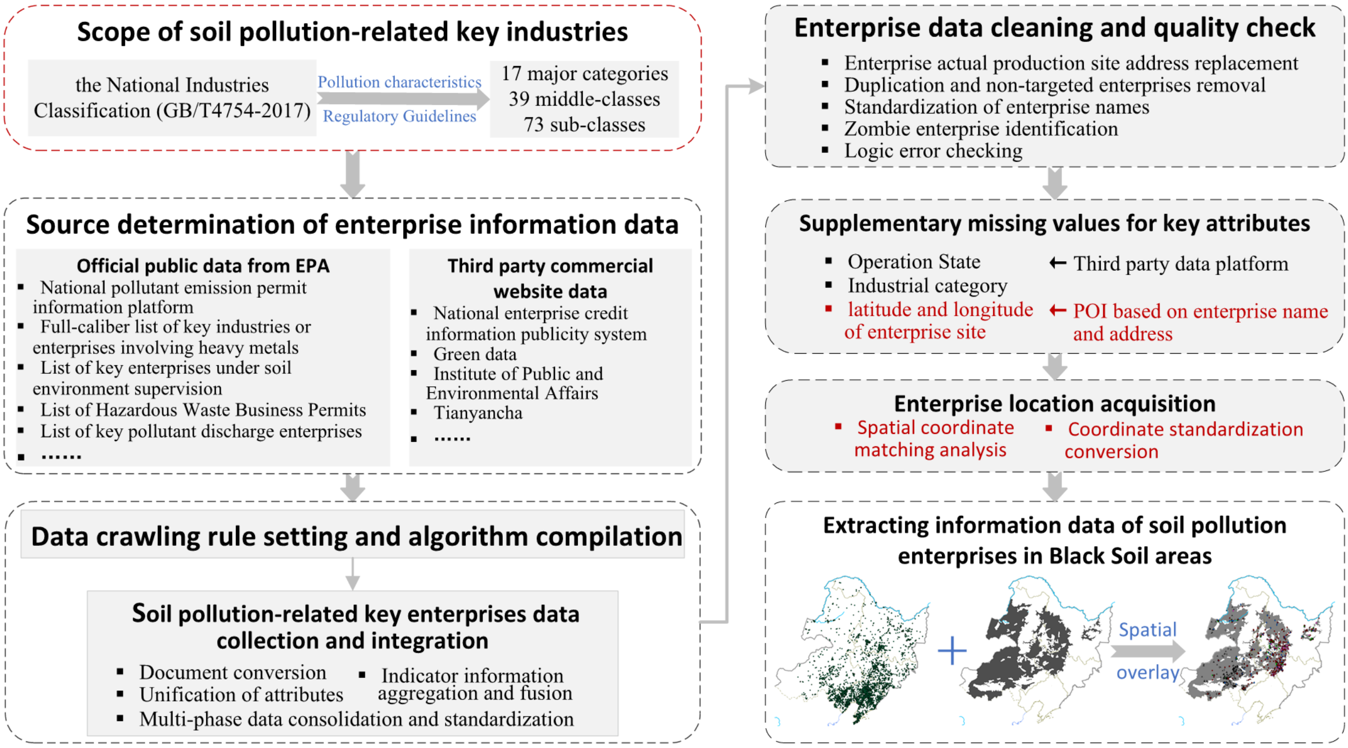

Figure 2. Enterprise data crawling and cleaning flowchart in key soil-pollution-related industries.

2.4. Construction of a Soil Pollution Key Industry Enterprise Dataset

Considering the miscellaneous and diversified data sources and irregular data quality, the initial enterprise datasets required further cleaning and quality checking after integration. Specifically, repeated enterprises were deleted, attributes were fused, logical errors were checked (attribute–address inconsistency and longitude and latitude > 90°), zombie enterprises with government tax filings and recruitment information (without actual production activity) were investigated, and logistics, warehousing, and distribution enterprises that were involved in relevant industries but belonged to non-production-type enterprises were excluded. Meanwhile, because some enterprise registered addresses were office addresses (such as group company headquarters), the production and business operational site addresses were replaced to ensure accurate spatial correspondence between such enterprises and the actual production site. In addition, some key attributes missing after information integration, such as the operating state and industry type, were supplemented by data from third-party websites such as Tianyancha. The most critical longitude and latitude information can be acquired in batches according to enterprise names and addresses, as well as by combining network maps with points of interest (POI, https://lbs.amap.com/api/webservice/guide/api/search/) (accessed on 8 November 2023) to spatialize the enterprise site distribution [46]. As a product of the Internet electronic map service and national basic surveying and mapping results, POI data have been widely applied in the field of big data mining based on spatial positions because of their enormous advantages in data size, coverage, accuracy, and update frequency [12,38,46]. The Gaode Open Platform (https://lbs.amap.com/api/webservice/guide/api/search/) (accessed on 8 November 2023) was used as the POI data source [38]. The data collection, integration, and cleaning processes are shown in Figure 2.

Furthermore, because the enterprise longitude and latitude information are important parameters for subsequent spatial analysis, the precision of such information should be verified through spatial matching after the missing values are supplemented to prevent the incorrect longitude and latitude information caused by the repetition of enterprise name and address registration errors [46]. Specifically, the minimum administrative units (districts and counties) of the falling point of each enterprise were extracted individually from the ArcGIS platform based on the existing longitude and latitude coordinates, and inconsistencies with registered places were corrected through comparisons. In addition, the GCJ-02 national geodetic coordinate system was usually adopted to acquire longitude and latitude information based on POI, which led to unavoidable manual treatment-induced deviations (for confidential reasons); therefore, coordinate transformation was necessary after the spatial matching of longitude and latitude. Details of the spatial matching of longitude and latitude, as well as the coordinate transformation for enterprise sites in key soil-pollution-related industries, can be found in Wei et al. [38]. As such, the entire enterprise data cleaning process was completed, followed by a spatial overlay within the scope of the black soil region in Northeast China. Finally, a high-reliability and high-precision enterprise dataset of key soil-pollution-related industries in the black soil region in Northeast China was obtained [19].

2.5. Spatial Analytical Method for the Agglomeration Characteristics of Polluting Enterprises

2.5.1. Multi-Distance Spatial Clustering Analysis

Spatial distance is a significant parameter that describes the distribution characteristics of spatial elements. The spatial elements typically present different agglomeration patterns based on different spatial distances. In this study, a multi-distance spatial clustering analysis algorithm was used to explore the spatial agglomeration scale of polluting industrial and mining enterprises in the black soil region in Northeast China, as well as to quantify their spatial agglomeration characteristics [47]. Multi-distance spatial clustering analysis was implemented by Ripley's K function in the ArcGIS platform. Ripley's K function is defined as follows:

$$K(d) = A\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{d_{ij}(d)}{n^2}\ (i, j = 1, 2, \cdots, n, i \neq j, d_{ij} \leq d), \tag{1}$$

where n is the number of enterprise points; sd is the distance scale; d_{ij} is the distance between points i and j; and A is the study area. Besag [48] proposed replacing $K(d)$ with $L(d)$ and performing an open-square linear transformation of $K(d)$ to maintain a stable variance. Assuming that random distribution $L(d)$ has an expected value of zero, the $L(d)$ is as follows:

$$L(d) = \sqrt{\frac{K(d)}{\pi}} - d \tag{2}$$

According to the relationship graph, if the $L(d)$ value is greater than the upper packet trace, the enterprise points are significantly aggregated; if the $L(d)$ value is less than the lower packet trace, the enterprises are significantly discrete distributed; and if the $L(d)$ value is located between the upper and lower packet traces, the distribution is significantly random. The statistical significance of spatial clustering or discreteness is measured at the 5% significance level. When the $L(d)$ value is larger than the d value, the further the two curves are apart from each other, that is, the larger the DiffK value, the stronger the aggregated spatial distribution of enterprises is; conversely, the stronger the discrete spatial distribution is. Simultaneously, with the gradual increase in the spatial scale d value, the first peak of the corresponding spatial scale $L(d)$ value of the research object indicates the corresponding d value of spatial aggregation characteristics, and the $L(d)$ peak value can indicate the aggregation degree of the spatial distribution of the object [49].

2.5.2. Spatial Kernel Density Analysis of Polluting Enterprises

Kernel density analysis refers to the continuous simulation of the distribution density of a spatial point or line elements, as well as of the effective mining of the agglomeration region of various elements. Compared with an original scatter diagram that simply expresses the spatial phenomenon distribution, kernel density analysis can construct a smooth distribution surface based on enterprise sites and output the density of each output raster unit in its surrounding neighborhood, thus serving as a more accurate analysis tool to explore deeper spatial distribution characteristic laws [12,50,51]. A greater kernel density value indicates a higher enterprise of agglomeration, and vice versa. The formula for enterprise kernel density analysis can be expressed as follows:

$$f(s) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{s - c_i}{h}\right), \tag{3}$$

where $f(s)$ is the kernel density estimation at spatial position s (number of enterprises/km^2); h is the search radius, also referred to as bandwidth; n is the number of elements with the distance from position s being $\leq h$; c_i is a core element; and K is the spatial weight function.

The kernel density analysis method was adopted to measure the degree of the spatial agglomeration of the polluting enterprise sites at the overall scale in the black soil region. The kernel function was solved using ArcGIS, while the search radius h (also called bandwidth) was comprehensively determined according to the optimal characteristic spatial scale calculated by using Ripley's K function in the multi-distance spatial clustering analysis.

2.6. Delimitation for the Potential Polluted Scope by Polluting Enterprise

According to the wastewater discharge intensity of different industries and atmospheric deposition and diffusion scales, the scope of the potential influence of enterprises belonging to the main polluting industry types in the black soil region in Northeast China was set according to the influence scope of the key soil pollution sources used in the national detailed survey on agricultural land [45]. It is notable that the scope of soil pollution caused by pollutant discharging enterprises mainly depended on atmospheric deposition, surface runoff, and water flow and usage. The pollution scope was impacted by enterprise scale, productive life, terrain, wind velocity, and precipitation. Considering data availability and the complexity of delimiting contaminated areas, as well as that the large-scale study (i.e., higher than provincial level) was less sensitive to the sub-kilometer spatial scope, the process of determining the influence scope of key soil pollution sources specified by the Technical Provisions on Site Layout for Detailed Investigation of Soil Pollution of Agricultural Land and the Guidelines to Soil Environmental Quality Classification for Agricultural Land (Trial) was simplified [52]. The longest diffusion distance provided in the guidelines under the conventional discharge state was used to determine the influence scope of all industry types. From the above processing, the maximum influence scope of the metal mining and dressing (ferrous and nonferrous), nonferrous metal smelting, and environmental governance industries on soil pollution was set at 5 km, and those of the petroleum refining and coking, chemical manufacturing, ferrous metal smelting, and electrical machinery and equipment manufacturing industries were set at 3.5, 2.5, 4.0, and 2.0 km, respectively. The influence scope of the leather-making and metal product manufacturing industries, which mostly existed in the form of park or industrial agglomeration areas, was set at 3 km with reference to industrial parks. According to the above-determined influence scope of enterprises from various industries on soil pollution, caching, fusion, and clipping were performed to obtain grid-based (2 km $\times$ 2 km grids generated per shortest influence distance) potentially polluted areas by high-polluting industrial and mining enterprises in the black soil region in Northeast China.

2.7. Agricultural Soil Potential Pollution Risk Analysis in the Black Soil Region

Agricultural soil pollution risks mainly depend on "source-path", natural meteorological conditions (such as rainfall, wind velocity, and soil parent material), and soil physicochemical properties (such as organic carbon, pH, redox potential, and available content of pollutants) [53–55]. However, for the agricultural soil around polluting enterprises with clear pollution sources, determining whether it is within the effective influence scope of enterprise pollution sources for achieving pollutant transport is a precondition for evaluating pollution risk. Nevertheless, natural meteorological conditions have a significant impact on the potential pollution source influence scope, which has already been demonstrated in the previous delimitation of the potentially polluted scope by enterprises. Meanwhile, existing studies have shown that agricultural soil pollution risk is closely related to the degree of agglomeration of peripheral polluting enterprises and the pollutant discharge intensity [12,33–35], where the latter depends, to a greater extent, on production factors such as enterprise productive life and scale [38]. As this study only focused on black soil, a unique soil type with approximate physicochemical properties in Northeast China, the overall agricultural land spatial variability in the study area was not significant; thus, it was temporarily neglected in the soil pollution risk division. Therefore, based on index data availability, only the potential pollution scope of industrial and mining enterprises, enterprise scale, and the degree of enterprise agglomeration within a unit area were considered in the potential soil pollution risk division for agricultural lands around polluting enterprises in the black soil region. The relevant analysis process was as follows:

(1) First, a spatial overlay analysis of the potential polluting scope of pollution-related enterprises that were acquired in Section 2.6 with the agricultural land in the black soil area of Northeast China was conducted to obtain the potential risk area of agricultural land in the black soil area, and this was followed by a gridding (2 × 2 km) of these risk areas. (2) The position of grid *i* within the potential pollution scope of industrial and mining enterprises or not was identified to determine whether agricultural land was subject to any soil pollution risk. (3) The sum of the registered capital of all enterprises in the agricultural potential risk areas in the black soil region were used as the regional enterprise scale index to reflect the overall pollutant discharge capacity of this region. (4) The number of enterprises in each agricultural land grid was used as the regional industrial activity intensity index to reflect the spatial distribution density of polluting enterprises in grid *i*. (5) The industrial activity intensity and regional enterprise scale indices were graded with the natural breakpoint method (Table 2) [56], and one of the two indices with a higher grade was used and combined with the situation where the agricultural land falls into the potential pollution range of industrial and mining enterprises. This was performed to finally determine the potential risk level of agricultural land in grid *I* (Table 2). Hence, the potential risk level distribution of the agricultural land in the entire black soil region was obtained after identifying the risks in all of the grids of the potential risk areas (Table 2).

Table 2. Agricultural land potential risk division method in the black soil region.

Risk Level [1]	Whether Agricultural Land Grid *i* Is Located within the Potential Pollution Scope of Industrial and Mining Enterprises	Industrial Activity Intensity Index in Agricultural Land Grid *i*	Enterprise Scale Index in Grid *i*
No risk	No	0	0
Low risk	Yes	[0, 1]	[0, 106]
Medium risk	Yes	(1, 5]	(106, 1700]
High risk	Yes	(5, 26]	(1700, 1,050,000]

[1] The potential risk level of grid *i* was obtained by comprehensively judging three indices: the industrial activity intensity, enterprise scale, and whether it was within the potential pollution scope. The final risk level was determined to be the highest among the above three indicators.

3. Results

3.1. Spatial Distribution Characteristics of High-Polluting Industrial and Mining Enterprises in the Black Soil Region in Northeast China

3.1.1. Overall Distribution of High-Polluting Industrial and Mining Enterprises

According to the constructed enterprise dataset of key soil-pollution-related industries, the soil-polluting enterprises in Northeast China were mainly concentrated in large and medium-sized cities, such as most parts of Liaoning Province, the middle of Jilin Province, Daqing and Harbin in Heilongjiang Province, and Chifeng in eastern Inner Mongolia. Comparatively, the industrial enterprise proportion was low in the black soil region in Northeast China, as agriculture, forestry, and animal husbandry (cultivated land, forest land, and grassland accounted for 89% of the black soil region area, as shown in Figure 1b) played a dominant role. However, because the base number of polluting enterprises in Northeast China was large (13,439), a total of 2151 enterprises from key soil-pollution-related industries in the black soil region in Northeast China (accounting for 16.01% in the entire Northeast China) were extracted through an overlay analysis with the black soil region spatial distribution scope.

The degree of the spatial agglomeration of polluting enterprise sites in the black soil region was investigated using multi-distance spatial clustering analysis (Ripley's K function). The results showed that, within a spatial scope of 10–550 km, the $L(d)$ value was always greater than the lower envelope line, thus indicating that the polluting enterprises maintained a significantly concentrated distribution, that is, their agglomerative distribution was of a favorable statistical significance. When the characteristic spatial scale reached approximately 210 km, the degree of spatial enterprise agglomeration was at its strongest (Figure 3). Meanwhile, the curve of the predicted values intersected with the expected values at approximately 550 km, indicating agglomerative variability within this distance. When the scale was >550 km, the degree of agglomeration decreased, and the result could be quantified as the maximum distance of the polluting enterprise spatial agglomeration.

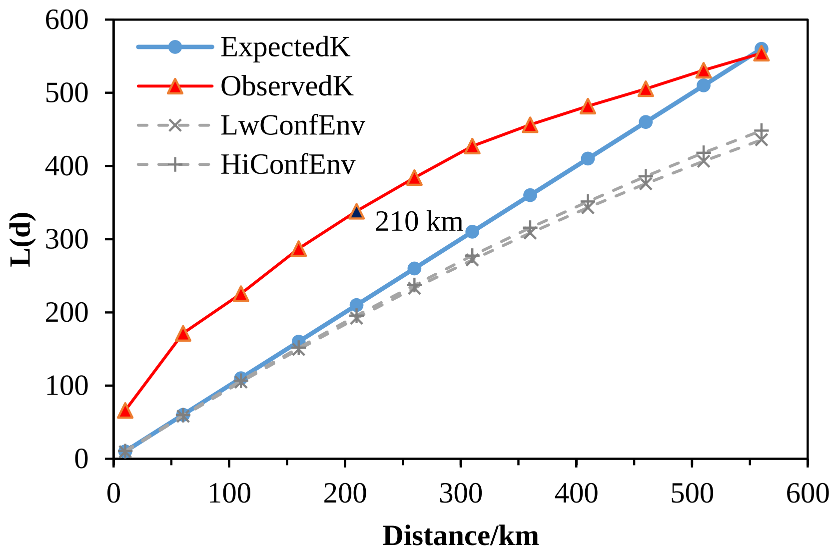

Figure 3. Multi-distance spatial clustering analysis of the high-polluting industrial and mining enterprise sites in the black soil region in Northeast China.

From the kernel density analysis results, it can be further seen that, in the black soil subregion in eastern Inner Mongolia and the Songnen black soil subregion (which cover the largest area with concentrated industrial and mining activity), polluting enterprises accounted for the highest proportion. Considering various industries as an entirety, polluting enterprises were widely distributed throughout the region (Figure 4a), and they presented a dense hotspot distribution in the south of the black soil subregion in eastern

Inner Mongolia (Chifeng) and in the mid-east Songnen black soil subregion (Changchun, Harbin, and Daqing). Moreover, such polluting enterprises were locally agglomerated to some extent in Qitaihe, Hegang, Jiamusi, and Shuangyashan, which are covered by the Sanjiang black soil subregion (Figure 4b).

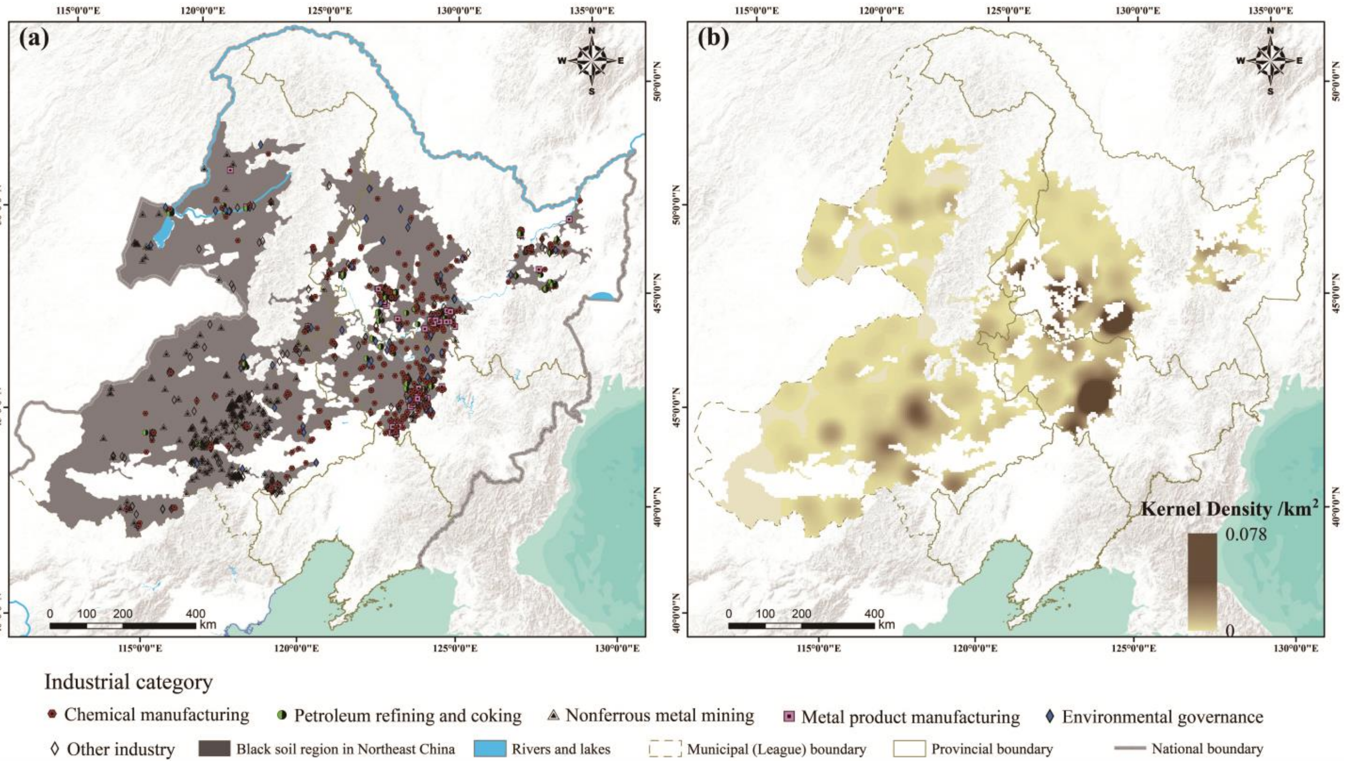

Figure 4. Spatial distribution of the high-polluting industrial and mining enterprises in the black soil region in Northeast China (data acquired on 21 September 2020). (**a**): Enterprise site distribution, (**b**): enterprise kernel density analysis results (cell size = 5 km and search radius: 210 km).

3.1.2. Composition and Agglomeration Characteristics of Polluting Enterprises in Various Industry Types

Through a statistical analysis of the enterprise information dataset of key soil-pollution-related industries, it was found that the black soil region in Northeast China included 10 of the 17 soil-pollution-related key industries (Table 3). Considering that the primary oil-producing area (Daqing) and coal-producing areas (Hegang, Qitaihe) in Northeast China were covered, the related downstream chemical manufacturing industry became the leading polluting industry in the black soil region in Northeast China with 1014 enterprises, accounting for 47.14% of polluting enterprises in this region. The petroleum refining and coking industry was related to energy and chemistry, and the nonferrous metal mining and dressing industry followed successively; however, their quantity and scale were significantly lower than those of the chemical manufacturing enterprises, accounting for only 13.3% and 12.83%, respectively, of the polluting enterprises in the study area. The number of enterprises in other industry types—except for the metal product manufacturing (140 enterprises, including 133 metal surface and heat treatments enterprises) and environmental governance industries (131 enterprises, including 67 sewage treatment plants)—was less than 100.

According to the spatial distribution of the enterprises associated with the primary industry types in the black soil region (Figure 4a), the nonferrous metal mining and dressing industry was concentrated in the south and west of the black soil subregion in eastern Inner Mongolia, specifically including the areas west of Hulunbuir and east of Xilingol, as well as the majority of Chifeng. The chemical manufacturing and metal product manufacturing industries also presented a concentrated distribution, where the chemical manufacturing, petroleum refining, and coking industries were primarily concentrated in regions with developed petrochemical industries, such as Daqing, Harbin, Changchun, and Siping in the central-south Songnen black soil subregion. Moreover, they were also concentrated in the advantaged coal-producing and coal chemical areas, such as Qitaihe, Hegang, Kiamusa, and Shuangyashan, in the Sanjiang black soil subregion. Other industries have not shown a clear clustering trend and are coupled with a relatively small quantity, resulting in an overall scattered distribution.

Table 3. Statistical enterprise distribution of the key soil-pollution-related industries in the black soil region in Northeast China.

Industry Code	Major Industry Type	Number of Enterprises	Proportion
26	Chemical raw materials and chemical products manufacturing industry	1014	47.14%
25	Petroleum refining and coking industry	286	13.30%
09	Nonferrous metal mining and dressing industry	276	12.83%
33	Metal product manufacturing industry (metal surface and heat treatments)	149	6.93%
77	Ecological protection and environmental governance industry	131	6.09%
08	Ferrous metal mining and dressing industry	83	3.86%
19	Leather, fur, feather and their products and shoemaking industry	79	3.67%
32	Nonferrous metal smelting and rolling industry	73	3.39%
31	Ferrous metal smelting and rolling industry	46	2.14%
38	Electrical machinery and equipment manufacturing industries (lead accumulator manufacturing)	14	0.65%

A recently released national general survey on soil contamination revealed that the chemical raw materials and product manufacturing industry, as well as the petroleum, coal, and other fuel processing industries, have become heavily polluting industries in China, thus leading to the highest atmospheric volatile organic compound, volatile phenol, and cyanide emissions in wastewater. The heavy metal emissions of the nonferrous metal mining and dressing, metal product manufacturing, and nonferrous metal smelting and rolling industries were also ranked in the top three industries [57]. The existence of heavily polluting industrial enterprises in the main grain-producing areas will certainly have adverse impacts on the soil and crop safety in the black soil region in Northeast China.

3.2. Potential Impact Distribution of the High-Polluting Industrial and Mining Enterprises in the Black Soil Region in Northeast China

Industrial and mining enterprises release pollutants primarily by discharging "three wastes" that enter the soil through sewage irrigation, atmospheric deposition, and surface runoff [7,8]. Therefore, industrial and mining enterprises can be regarded as point sources from a large regional scale, and the "three wastes" spatial transmission distance was taken as the radius to delimit the influence scope of a known industrial and mining industry on soil. According to the influence distance of the key soil-pollution-related industry types determined in Section 2.6, the potential influence scope of the heavily polluting industrial and mining enterprises in the black soil region in Northeast China is shown in Figure 5. Based on a calculation of the grid area, the influence area of polluting enterprises in the entire black soil region in Northeast China was approximately 44,053.96 km^2, accounting for 8.42% of the black soil region. As shown in Figure 4, because the chemical manufacturing and nonferrous metal mining and dressing industries were aggregated in the south, the area of potential pollution distribution in this region accounted for a high proportion.

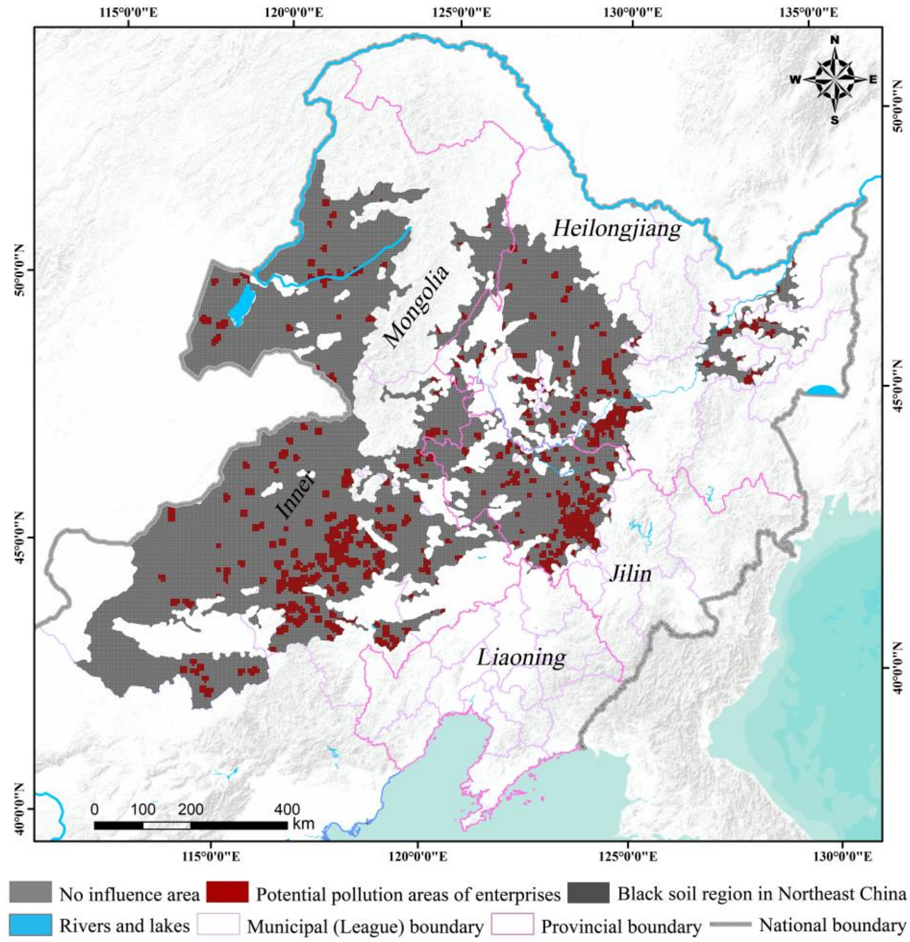

Figure 5. Potential soil pollution distribution of the polluting enterprises in the black soil region in Northeast China based on the influence scope of enterprises associated with different industries.

3.3. Distribution of High-Risk Agricultural Lands Affected by Industrial and Mining Enterprises

As an important grain-producing and livestock breeding area, the agricultural land and crop safety in the black soil region in Northeast China are key issues requiring attention [1,58]. Soil and crop safety in the surrounding agricultural lands have been seriously threatened historically, from low early-stage environmental management levels to the disorderly pollutant discharge from enterprises belonging to various industries across the country, including Northeast China [9,59,60]. To explore the potential risks of the agricultural land affected by industrial and mining enterprises in the black soil region in Northeast China, a spatial overlay analysis was performed to obtain the agricultural land distribution in the black soil region with areas potentially polluted by polluting enterprises (Figure 5), and this was conducted according to the method and procedures mentioned in Section 2.7. On this basis, the potential agricultural land risk areas in the black soil region were obtained and comprehensively graded (specific methods are detailed in Table 2) according to the industrial activity intensity and enterprise scale indices within each grid. Therefore, the potential risk division of agricultural land in the black soil region was achieved. The detailed distribution is shown in Figure 6.

Based on grid statistics, the agricultural land area potentially affected by polluting enterprises in the black soil region in Northeast China was 43,396.13 km^2, accounting for 11.3% of the total agricultural land area in the black soil region. From the influence area of enterprises associated with various industries, the influence area of the nonferrous metal mining and dressing industry was the largest (11,655.45 km^2), followed by that of the chemical raw materials and product manufacturing industry (8448.30 km^2) and that of the ecological protection and environmental governance industry (5524.32 km^2) (Table 4) successively. It is notable that despite the large quantity, enterprises in the chemical manufacturing industry showed a relatively small influence area. Thus, the total agricultural land area affected by such enterprises was less than that of the nonferrous metal mining and dressing industry.

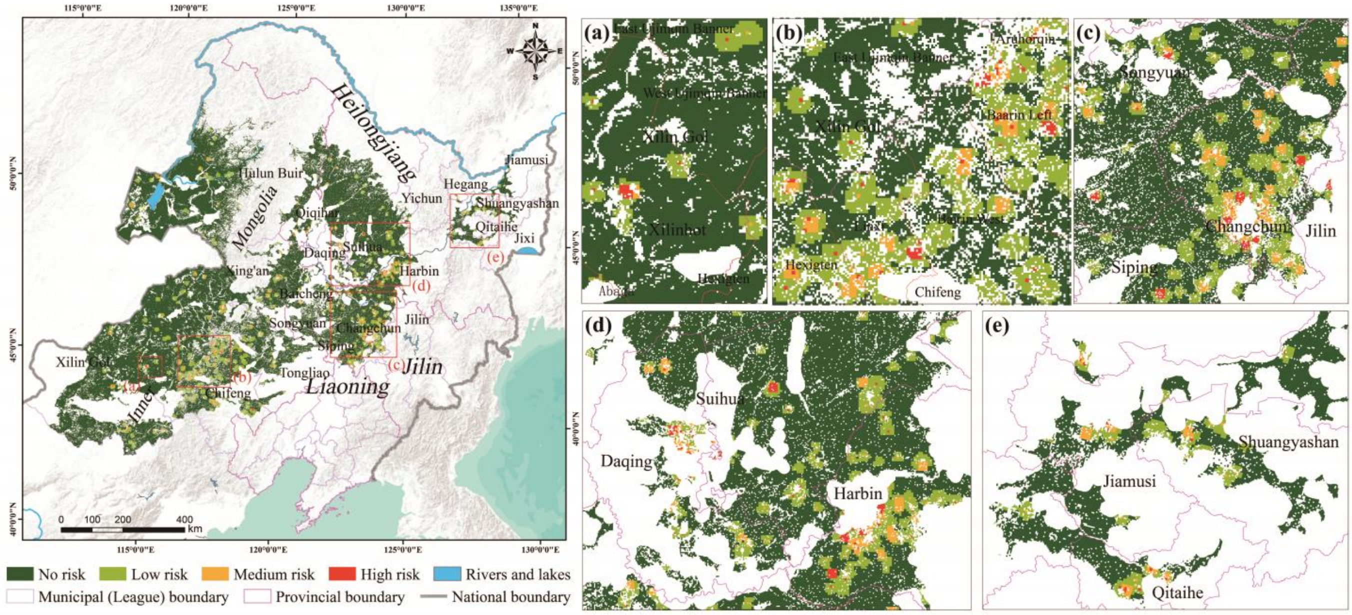

Figure 6. Potential soil pollution risk level distribution of agricultural lands in the black soil region. (**a**): Peripheral Xilinhot region, (**b**): banner counties in the north of Chifeng, (**c**): peripheral Changchun region, (**d**): Harbin–Daqing–Suihua region, (**e**): northwest Kiamusa–Qitaihe–Shuangyashan region.

Table 4. Statistics of the agricultural land area affected by potential pollution enterprises from different industry categories (based on actual buffer area).

Industry Code	Major Industry	Agricultural Land Area Affected (km^2) [1]
09	Nonferrous metal mining and dressing industry	11,655.45
26	Chemical raw materials and product manufacturing industry	8448.30
77	Ecological protection and environmental governance industry	5524.32
25	Petroleum refining and coking industry	3883.85
08	Ferrous metal mining and dressing industry	3874.79
32	Nonferrous metal smelting and rolling industry	3294.83
33	Metal product manufacturing industry (metal surface and heat treatments)	1713.27
31	Ferrous metal smelting and rolling industry	1263.80
19	Leather, fur, feather and their products and shoemaking industry	851.84
38	Electrical machinery and equipment manufacturing industry (lead accumulator manufacturing)	178.28

[1] Agricultural land areas affected by different industries may overlap.

In terms of risk level, the agricultural land in the black soil region remains generally free from risks (Figure 6 and Table 5) and accounts for 88.71%; the low-risk agricultural lands account for 9.42%; and the high-risk lands account for 0.43%. High-risk areas are aggregated in the south and west of the black soil subregion in eastern Inner Mongolia, namely the west of Hulunbuir, east of Xilingol, and most of Chifeng (Figure 6a,b), which are mainly affected by the nonferrous metal mining and dressing industry. In addition, high-risk areas are also agglomerated in Daqing, Harbin, and Changchun in the central-south Songnen black soil subregion (Figure 6c,d), which are mainly affected by the chemical and metal product manufacturing industries. Given the weaker industrial activity and pollutant discharge intensity, the medium-risk areas are mainly distributed around high-risk areas and diffused toward the surrounding areas (Figure 6).

Table 5. Summary of the affected areas and risk levels of different agricultural land types (unit: km^2).

Risk Level	Paddy Field	Dry Land	Garden Land	Pastureland [1]	Total
No risk	16,874.87	104,061.71	1572.92	218,373.38	340,882.88
Low risk	1826.94	17,337.24	117.57	16,918.85	36,200.61
Medium risk	474.86	2810.54	3.00	2260.50	5548.90
High risk	91.11	913.84	4.00	637.67	1646.62
Subtotal	2392.91	21,061.62	124.57	19,817.02	43,396.13
Total	19,267.78	125,123.33	1697.49	238,190.41	384,279.00

[1] Pastureland only includes high- and medium-coverage grasslands.

On a smaller scale, several regions have agglomerated high-risk areas, including the Xilinhot peripheral region, the banner counties north of Chifeng, the Changchun peripheral region, the Harbin–Daqing–Suihua region, and the northeast Kiamusa–Qitaihe–Shuangyashan region. In addition, the high-risk areas (those given priority to survey areas) were located at city perimeters. With an increase in the distance from the city areas, agricultural land risk level declined (the risk gradually weakened from the city center to the outskirts) (Figure 6).

In addition, the risk level of the agricultural land types differed (Table 5). Agricultural land areas at potential risk were sorted in descending order as follows: dry land > pastureland > paddy fields > garden land. The proportions of the three risk levels for dry land were significantly higher than those for the other types of agricultural land, with high, medium, and low risk ratios of 0.73%, 2.25%, and 13.86%, respectively. The proportions of high-, medium-, and low-risk areas in all four types of the agricultural land affected by industrial and mining enterprises were 0.43%, 1.44%, and 9.42%, respectively.

4. Discussion

Many samples must be collected in the traditional division of the scope of regional soil pollution and risk assessment, thus consuming enormous manpower and material resources. Therefore, the current soil environmental quality surveys on the black soil region in Northeast China mostly concentrate on small local regional scales with very limited sampling, thus leading to a lack of understanding of agricultural soil pollution risk levels and their spatial distribution in the entire black soil region in Northeast China [18–21]. Despite the overall good soil environmental quality in Northeast China, serious soil pollution in local regions impedes the green, ecological, and organic development of black soil grain barns to some extent. To protect the black soil region in Northeast China, there is an urgent need to delimit the soil pollution risk distribution rapidly and economically, particularly to effectively identify high-risk areas for prior surveys.

In this study, a high-precision dataset of 2151 polluting industrial and mining enterprises and their spatial distribution in the black soil region in Northeast China was acquired by crawling and cleaning multi-source geographical big data. The spatial agglomeration characteristic analysis showed that, within a scale of 550 km, all polluting industrial and mining enterprises were agglomerated in the black soil region in Northeast China. Specifically, they were mainly concentrated in Chifeng in the black soil subregion of eastern Inner Mongolia and Changchun, Harbin, and Daqing of the mid-east Songnen black soil subregion. Moreover, they also agglomerated to some extent in the local regions of Qitaihe, Hegang, Kiamusa, and Shuangyashan in the Sanjiang black soil subregion (Figure 4a,b). Previous research has been conducted on the spatiotemporal emission characteristics of atmospheric pollutant sources in Northeast China, and it showed that the pollutant discharge in Northeast China is mainly concentrated in the south and northeast of Heilongjiang Province, the middle and west of Jilin Province, the middle and southwest of Liaoning Province, and in the mid-east of the four eastern league cities [61]. The discharge intensities in Harbin, Changchun, Jilin, Shenyang, Jinzhou, Anshan, and Yingkou were higher than those in surrounding cities. In addition, the spatial distribution pattern of the atmospheric pollutants SO_2 and non-methane volatile organic compounds, mainly derived from industrial sources, were consistent with polluting enterprise spatial agglomeration trends in this study. It revealed that enterprise distribution information based on big data crawling is reliable. As for enterprise type, the two types of polluting industrial and mining enterprises with the largest quantities throughout Northeast China were chemical manufacturing and its associated petroleum refining and coking industry, as well as nonferrous metal mining and dressing, in which the latter showed the broadest influence scope. Spatial agglomeration characteristics of the nonferrous metal mining and dressing industry, which is associated with a resource-based industry, are mainly affected by resource endowments [62]. Spatial agglomeration characteristics of the chemical manufacturing industry are mainly associated with location factors such as traffic conditions and urbanization level. The research results on the degree of the spatial agglomeration mentioned above provide a basis for identifying high-risk areas.

After collecting multi-source geographical big data, the requisite spatial distribution and potential influence scope of polluting industrial and mining enterprises in the study area were obtained. Subsequently, a complete system for assessing and grading pollution risks induced by enterprise pollution sources in agricultural lands on a large regional scale was established by combining the soil pollution risk identification around industrial and mining enterprises on a large scale with the relevant influencing factors based on remote sensing land-use-type distribution data and geographic information system (GIS) spatial analyses. A refined pollution risk division of agricultural land in the entire black soil region in Northeast China was realized using the above method at a 2 km resolution. According to the division results, agricultural lands in the black soil region in Northeast China presented a distribution pattern featuring an overall low pollution risk with aggravated local pollution risks. High-risk areas were mainly concentrated around the Xilingol and Chifeng city areas in the black soil subregion in eastern Inner Mongolia, as well as near the main industrial and mining areas. These areas were mainly affected by the large pollutant transport scope and high pollutant discharge intensity of the industrial and mining enterprises that were mainly engaged in nonferrous metal mining and dressing activity [38,63,64]. The agricultural lands distributed in these areas are subject to high soil pollution risks. Meanwhile, due to the presence of plenty of pollutant discharge sources, i.e., mainly due to the enterprises that belong to the chemical and metal product manufacturing industries, multiple high-risk area hotspots were also distributed around industrial lands that showed more activity, such as Daqing, Harbin, and Changchun in the central south Songnen black soil subregion. Although the aforementioned industry types had a relatively similar pollutant transport distance and influence scope, high-density discharge source agglomeration still resulted in extreme pollution risks to agricultural lands in such areas [45]. Meanwhile, although the intensity of industrial activities and pollution emissions has weakened in the periphery of the high-risk areas mentioned above, due to spillover effects, there are still multiple medium risk areas that are scattered, and appropriate attention should be paid to such areas in the process of agricultural land risk control [45]. Overall, the abovementioned high-risk areas, including the southeastern black soil region in eastern Inner Mongolia and the south-central Songnen black soil subregion, are key regions for soil pollution surveys and pollution source control. Moreover, the scope delimitation of agricultural land with high soil pollution risks in the black soil region will substantially reduce the regional area for a detailed soil survey on agricultural lands. Furthermore, they will provide data support for further improving soil environmental quality in the black soil region, as well as for implementing emission reductions from the enterprise pollution sources in key industries.

This study provides a basis for identifying and dividing the black soil regions that may be subject to soil pollution. However, this method has several limitations. For instance, the soil pollution risk was graded according to the regional industrial activity intensity and general enterprise scale indices that reflect the number of enterprises and their registered capital. However, the risk level might also be significantly influenced by other factors, such as the actual enterprise scale, productive life, and industry type [65–67]. Moreover, due to being limited by data availability and insufficient relevant quantitative studies, the influence scope of polluting enterprises was conservatively estimated and simplified. Therefore, this framework is currently applicable for identifying potential risks on the agricultural land affected by pollutant-related enterprises on a large regional scale (provincial level and above). This can be used to guide subsequent detailed surveys of soil contamination in key regions and the implementation of source control measures for enterprises in key industries. In future research, the influence scope of polluting enterprises and the subsequent risk division method will be optimized by fully collecting pollutant discharge, enterprise scale, and related natural geographical data.

5. Conclusions

(1) This study provides a methodology that integrates the spatialization of the information data of key pollutant-related industries on a regional scale and potential risk assessments of surrounding agricultural land using multi-source geographical big data technology and "source-diffusion pathway" risk elements. The case was conducted in the black soil area in Northeast China. A total of 2151 enterprises from key soil-pollution-related industries in the black soil region in Northeast China were extracted. These polluting enterprises exhibited significant spatial clustering characteristics and presented a dense hotspot distribution in the south (Chifeng) of the black soil subregion in eastern Inner Mongolia and the east-central region (Changchun, Harbin, and Daqing) of the Songnen black soil subregion. The chemical manufacturing industry is the leading polluting industry type in the study area, accounting for nearly half of the polluting enterprises in this region (47.14%), followed by the energy and chemistry-related petroleum refining and coking industries (13.3%) and the nonferrous metal mining and dressing industry (12.83%).

(2) Considering the "source-diffusion pathway" risk factor, the area of agricultural lands affected by polluting enterprises in the black soil region in Northeast China reached 43,396.13 km^2. The area affected by the potential impacts of non-ferrous metal mining and processing was the greatest, amounting to 11,655.45 m^2. The potential risk level of agricultural land contamination in the black soil zone was mainly categorized as no risk (88.71%), followed by low risk (9.42%). And the overall risk was controllable. A few high-risk areas (0.43%) were relatively agglomerated in the south and west of the black soil subregions in eastern Inner Mongolia. These areas are primarily affected by the non-ferrous metal extraction industry, which has high pollution emission intensity. Moreover, the high-risk areas are also aggregated in the central-south Songnen black soil subregions and are mainly influenced by high-density chemical and metal product manufacturing industries. The potential high-risk areas mentioned above should be prioritized for future investigations into agricultural land pollution in the Northeast black soil region, as well as for the prevention and control of pollution sources.

(3) Through the construction of a regional scale information data mining method for polluting-related enterprises and a methodology for identifying agricultural land potential pollution risks around enterprises, the aforementioned methodological framework is expected to effectively guide regional soil pollution risk level classification and pollution source reduction for key industry enterprises. Furthermore, it can be expanded to a national scale to provide data and technical support for detailed soil pollution investigations in the future.

Author Contributions: Conceptualization, C.W.; methodology, C.W. and X.W.; software, C.W. and X.Z.; formal analysis, X.Z. and M.L.; data curation, C.W.; writing—original draft preparation, C.W. and X.Z.; writing—review and editing, C.W. and X.Z.; visualization, C.W. and X.Z.; supervision, X.L. and S.W.; project administration, X.L. and J.L.; funding acquisition, C.W., X.L. and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (grant no. 42377462), the project of China Geological Survey (grant no. DD20230514), and the National Natural Science Foundation of China Youth Fund (grant no. 42307576).

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 authors.

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

References

1. Ministry of Agriculture; National Development and Reform Commission; Ministry of Finance; Ministry of Land and Resources; Ministry of Environmental Protection; Ministry of Water Resources. Outline of Northeast Black Land Protection Planning (2017–2030). 2017. Available online: https://www.moa.gov.cn/nybgb/2017/dqq/201801/t20180103_6133926.htm (accessed on 28 November 2023).
2. Liu, H.; Li, S.; Wu, M.; Sun, F.; Wang, Q.; Dong, X. Current situation and perspectives of black soil protection from the integrated angle of quantity, quality, and ecology in Northeast China. *Chin. J. Soil Sci.* **2021**, *52*, 544–552. [CrossRef]
3. Yang, Z.; Gao, C.; Zang, S.; Yang, X. Applicability evaluation of the SWIM at river basins of the black soil region in Northeast China: A case study of the upper and middle Wuyuer River basin. *J. Geogr. Sci.* **2017**, *27*, 817–834. [CrossRef]
4. Duan, X.; Xie, Y.; Liu, B.; Liu, G.; Feng, Y.; Gao, X. Soil loss tolerance in the black soil region of Northeast China. *J. Geogr. Sci.* **2012**, *22*, 737–751. [CrossRef]
5. Xu, M.; Lu, C.; Zhang, W.; Li, L.; Duan, Y. Situation of the quality of arable land in China and improvement strategy. *Chin. J. Agric. Resour. Region. Plan.* **2016**, *37*, 8–14.
6. China State Farm & Tropical Agriculture. Focus on the Development of Organic Industry and Strengthen the Protection and Utilization of Black Soil to Lead the Green Development Trend of China's Agriculture. 2021. Available online: http://www.farmchina.org.cn/ShowArticles.php?url=Um4BZFw8UGUCMwRgVWNWOgNg (accessed on 28 November 2023).
7. Yang, Y.; Liu, C.; Pan, W.; Zhang, G.; Zhu, W. Heavy metal accumulation from zinc smelters in a carbonate rock region in Hezhang County, Guizhou Province, China. *Water Air Soil Pollut.* **2006**, *174*, 321–339. [CrossRef]
8. Ettler, V.; Šebek, O.; Grygar, T.; Klementová, M.; Bezdička, P.; Slavíková, H. Controls on metal leaching from secondary Pb smelter air-pollution-control residues. *Environ. Sci. Technol.* **2008**, *42*, 7878–7884. [CrossRef]
9. Zhang, X.; Zhong, T.; Liu, L.; Ouyang, X. Impact of soil heavy metal pollution on food safety in China. *PLoS ONE* **2015**, *10*, e0135182. [CrossRef]
10. Liu, X.; Tian, G.; Jiang, D.; Zhang, C.; Kong, L. Cadmium (Cd) distribution and contamination in Chinese paddy soils on national scale. *Environ. Sci. Pollut. Res.* **2016**, *23*, 17941–17952. [CrossRef]
11. Wang, P.; Chen, H.; Kopittke, P.M.; Zhao, F. Cadmium contamination in agricultural soils of China and the impact on food safety. *Environ. Pollut.* **2019**, *249*, 1038–1048. [CrossRef]
12. Jia, X.; Hu, B.; Marchant, B.P.; Zhou, L.; Zhou, S.; Zhu, Y. A methodological framework for identifying potential sources of soil heavy metal pollution based on machine learning: A case study in the Yangtze Delta, China. *Environ. Pollut.* **2019**, *250*, 601–609. [CrossRef]
13. Wang, Y.; Wang, S.; Jiang, L.; Ma, L.; Li, X.; Zhong, M.; Zhang, W. Does the geographic difference of soil properties matter for setting up the soil screening levels in large countries like China. *Environ. Sci. Technol.* **2022**, *56*, 5684–5693. [CrossRef] [PubMed]
14. Ministry of Environmental Protection; Ministry of Land and Resources. National Soil Pollution Survey Bulletin. 2014. Available online: http://www.mee.gov.cn/gkml/sthjbgw/qt/201404/t20140417_270670.htm (accessed on 28 November 2023).
15. Liu, M.; Gao, F.; Han, J.; Wang, X.; Guo, X. Ecological harm and source analysis of heavy metals in small watershed soil of black soil region. *J. Chin. Agric. Univ.* **2020**, *25*, 12–21.
16. Gao, F.; Wang, X.; Han, J.; Liu, M.; Guo, X. Heavy metal pollution characteristics and health risk assessment of arable soil in a small watershed in the black soil region of Northeast China: A case study of Haigou River watershed. *J. Chin. Agric. Univ.* **2020**, *25*, 73–83.
17. Zhu, Y.; Wang, D.; Li, W.; Yang, Y.; Shi, P. Spatial distribution of soil trace element concentrations along an urban-rural transition zone in the black soil region of Northeastern China. *J. Soils Sediments* **2019**, *19*, 2946–2956. [CrossRef]
18. Wang, S.; Sun, B.; Wang, C.; Pei, Z.; Li, S.; Zuo, X.; Zhang, X.; Liu, J. Evaluation and analysis of soil heavy metal pollution status in typical black soil region of Northeast China. *J. Anhui Agric. Sci.* **2013**, *41*, 4350–4352.
19. Wu, S.; Wang, D.; Li, W.; Wang, X.; Yan, Z. Risk zoning of heavy metals in a peri-urban area in the Black soil farmland based on agricultural products. *Environ. Sci.* **2022**, *43*, 454–462. [CrossRef]
20. Shi, R.; Lv, J.; Liu, Y. Spatial distribution and bio-accumulation of cadmium and lead in soil, rice and vegetables in typical pollution areas, China. *Bull. Environ. Contam. Toxicol.* **2020**, *104*, 307–313. [CrossRef]
21. Fang, A. Study on Ecological Cumulative Effects of Mining Areas in Arid and Semi-Arid Grassland in Eastern Inner Mongolia. Ph.D. Thesis, China University of Mining and Technology, Xuzhou, China, 2020. [CrossRef]

22. Wong, S.; Li, X.; Zhang, G.; Qi, S.; Min, Y. Heavy metals in agricultural soils of the Pearl River Delta, South China. *Environ. Pollut.* **2002**, *119*, 33–44. [CrossRef]
23. Lu, A.; Wang, J.; Qin, X.; Wang, K.; Han, P.; Zhang, S. Multivariate and geostatistical analyses of the spatial distribution and origin of heavy metals in the agricultural soils in Shunyi, Beijing, China. *Sci. Total Environ.* **2012**, *425*, 66–74. [CrossRef]
24. Niu, L.; Yang, F.; Xu, C.; Yang, H.; Liu, W. Status of metal accumulation in farmland soils across China: From distribution to risk assessment. *Environ. Pollut.* **2013**, *176*, 55–62. [CrossRef]
25. Huang, Y.; Wang, L.; Wang, W.; Li, T.; He, Z.; Yang, X. Current status of agricultural soil pollution by heavy metals in China: A meta-analysis. *Sci. Total Environ.* **2019**, *651*, 3034–3042. [CrossRef] [PubMed]
26. Cheng, C.; Li, W.; Song, X. Thinking on big data construction of ecological environment. *Chin. Environ. Manag.* **2015**, *7*, 9–13. [CrossRef]
27. Liu, L.; Zhang, L.; Zhao, F.; Zhao, M.; Zhao, H.; Shao, R.; Xu, M. Opportunities and challenges faced by ecological big data. *Acta Ecol. Sin.* **2017**, *37*, 4896–4904.
28. Qi, Y.; Niu, C.; He, H. Political connection and environmental protection investment: A study based on ownership difference. *Sustainability* **2023**, *15*, 15982. [CrossRef]
29. Huang, G.; Zhu, S.; Wang, X.; Tian, Z.; Ji, G.; Lu, R.; Cui, X.; Chen, X. Natural language processing and machine learning-based suspected soil contamination enterprise identification. *Chin. J. Environ. Eng.* **2020**, *14*, 3234–3242. [CrossRef]
30. Wang, X.; Huang, G.; Zhu, W.; Ji, G. The overall technical strategy of big data driven site pollution risk management and control. *Environ. Prot.* **2020**, *48*, 64–66. [CrossRef]
31. Wang, X. A major task related to the goal of beautiful China—Some thoughts on promoting the battle of pure land. *Chin. Ecol. Civiliz.* **2019**, *1*, 37–38.
32. Liu, B.; Ai, S.; Zhang, W.; Huang, D.; Zhang, Y. Assessment of the bioavailability, bioaccessibility and transfer of heavy metals in the soil-grain-human systems near a mining and smelting area in NW China. *Sci. Total Environ.* **2017**, *609*, 822–829. [CrossRef]
33. Zhou, Y.; Aamir, M.; Liu, K.; Yang, F.; Liu, W. Status of mercury accumulation in agricultural soil across China: Spatial distribution, temporal trend, influencing factor and risk assessment. *Environ. Pollut.* **2018**, *240*, 116–124. [CrossRef]
34. Zhou, Y.; Niu, L.; Liu, K.; Yin, S.; Liu, W. Arsenic in agricultural soils across China: Distribution pattern, accumulation trend, influencing factors, and risk assessment. *Sci. Total Environ.* **2018**, *616*, 156–163. [CrossRef]
35. Qiao, P.; Yang, S.; Lei, M.; Chen, T.; Dong, N. Quantitative analysis of the factors influencing spatial distribution of soil heavy metals based on geographical detector. *Sci. Total Environ.* **2019**, *664*, 392–413. [CrossRef] [PubMed]
36. Ministry of Environmental Protection. China's 12th 5-Year Planning of Comprehensive Prevention and Control of Heavy Metal Pollution. 2011. Available online: https://www.mee.gov.cn/ywdt/hjnews/201409/W020140903603840681475.pdf (accessed on 28 November 2023).
37. Ministry of Ecology and Environment. Opinions on Strengthening Pollution Prevention and Control of Heavy Metal Related Industries. 2018. Available online: http://www.jsxishan.gov.cn/uploadfiles/201902/15/2019021515231726066606.pdf (accessed on 28 November 2023).
38. Wei, C.; Lei, M.; Chen, T.; Zhou, C.; Gu, R. Method on site-specific source apportionment of domestic soil pollution across China through public data mining: A case study on cadmium from non-ferrous industries. *Environ. Pollut.* **2022**, *295*, 118605. [CrossRef] [PubMed]
39. Liu, B.; Zhang, G.; Xie, Y.; Shen, B.; Gu, Z.; Ding, Y. Delineating the black soil region and typical black soil region of northeastern China. *Chin. Sci. Bull.* **2021**, *66*, 96–106. [CrossRef]
40. Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences. China's Spatial Distribution of Soil Type 1 km Raster Data Layer. 2017. Available online: https://www.resdc.cn/data.aspx?DATAID=145 (accessed on 28 November 2023).
41. Ministry of Ecology and Environment of the People's Republic of China. Soil Environment Quality Risk Control Standard for Soil Contamination of Agricultural Land (Trial). 2018. Available online: https://www.mee.gov.cn/ywgz/fgbz/bz/bzwb/trhj/201807/t20180703_446029.shtml (accessed on 28 November 2023).
42. Resources and Environment Science Data Center, Chinese Academy of Sciences. 1 km Grid Remote Sensing Monitoring Data of Land Use in 2020 in China. 2020. Available online: https://www.resdc.cn/data.aspx?DATAID=335 (accessed on 28 November 2023).
43. Ministry of Environmental Protection. Technical Regulations on Risk Screening and Risk Classification of Enterprise Sites in Production/Closed Enterprises. 2017. Available online: http://www.mee.gov.cn/gkml/hbb/bgt/201708/t20170818_420021.htm (accessed on 28 November 2023).
44. Ministry of Environmental Protection. Guidelines to the Preparation of Detailed Survey Scheme on Provincial Soil Pollution Status. 2017. Available online: https://www.mee.gov.cn/gkml/hbb/bgt/201708/W020170818581370519351.pdf (accessed on 28 November 2023).
45. Ministry of Environmental Protection. Guidelines to Soil Environmental Quality Classification for Agricultural Land (Trial). 2017. Available online: http://www.mee.gov.cn/xxgk2018/xxgk/xxgk13/202012/t20201201_810765.html (accessed on 28 November 2023).
46. Jiang, Y.; Wang, H.; Lei, M.; Hou, D.; Chen, S.; Hu, B.; Huang, M.; Song, W.; Shi, Z. An integrated assessment methodology for management of potentially contaminated sites based on public data. *Sci. Total Environ.* **2021**, *783*, 146913. [CrossRef] [PubMed]

47. Luo, J.; Sun, Z.; Zhang, X. Analysis of oasis landscape pattern evolution based on Ripley's *K* function: A case study of Ganzhou District, Zhangye City. *Res. Soil Water Conserv.* **2019**, *26*, 224–231. [CrossRef]
48. Besag, J.E. Discussion on Dr Ripley's Paper. *J. R. Stat. Soc. Ser. B* **1977**, *39*, 192–212. [CrossRef]
49. O'Driscoll, R.L. Description of spatial pattern in seabird distributions along line transects using neighbour K statistics. *Mar. Ecol. Prog. Ser.* **1998**, *165*, 81–94. [CrossRef]
50. Jayasinghe, D.Y.; Jayasinghe, C.L. An investigation into adult human height distributions using kernel density estimation. *Sankhya Ser. B.* **2022**, *84*, 79–105. [CrossRef]
51. Yang, J.; Zhu, J.; Sun, Y.; Zhao, J. Delimitating urban commercial central districts by combining kernel density estimation and road intersections: A case study in Nanjing City, China. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 93. [CrossRef]
52. Department of Ecology and Environment of Henan Province. Technical Provisions on Site Layout for Detailed Investigation of Soil Pollution of Agricultural Land. 2020. Available online: https://sthjt.henan.gov.cn/2020/03-27/1374863.html (accessed on 28 November 2023).
53. Zhao, F.; Ma, Y.; Zhu, Y.; Tang, Z.; McGrath, S.P. Soil contamination in China: Current status and mitigation strategies. *Environ. Sci. Technol.* **2015**, *49*, 750–759. [CrossRef]
54. Yuan, Z.; Luo, T.; Liu, X.; Hua, H.; Zhuang, Y.; Zhang, X.; Zhang, L.; Zhang, Y.; Xu, W.; Ren, J. Tracing anthropogenic cadmium emissions: From sources to pollution. *Sci. Total Environ.* **2019**, *676*, 87–96. [CrossRef] [PubMed]
55. Zhao, X.; Lei, M.; Chen, T. Species, habitat characteristics, and screening suggestions of cadmium hyperaccumulators in China. *Environ. Sci.* **2023**, *44*, 2786–2798. [CrossRef]
56. Chen, J.; Yang, S.; Li, H.; Zhang, B.; Lv, J. Research on geographical environment unit division based on the method of natural breaks (Jenks). *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *40*, 47–50. [CrossRef]
57. Ministry of Ecological Environment. The Second National Pollution Census Bulletin. 2020. Available online: https://www.mee.gov.cn/xxgk2018/xxgk/xxgk01/202006/t20200610_783547.html (accessed on 28 November 2023).
58. Jiang, F.; Qian, Y.; Guo, Z.; Gao, L.; Zhang, Z.; Cao, Z.; Guo, J.; Liu, F.; Peng, X. Evaluating the regional suitability of conservation tillage and deep tillage based on crop yield in the black soil of Northeast China: A Meta-analysis. *Acta Pedol. Sin.* **2022**, *59*, 935–952.
59. 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]
60. Shifaw, E. Review of heavy metals pollution in China in agricultural and urban soils. *J. Health Pollut.* **2018**, *8*, 180607. [CrossRef] [PubMed]
61. Zhang, M.; Chen, W.; Gao, C.; Zhang, X.; Jiang, Y.; Xiu, A. Spatio-temporality of anthropogenic air pollutants in Northeast China based on current emission inventories. *Sci. Geogr. Sin.* **2020**, *40*, 1940–1948. [CrossRef]
62. Qin, Z.; Ni, Y.; Jiang, X.; Wang, R. Spatial agglomeration and regional transformation of industrial and mining enterprises in resource-based areas: A case study of Shanxi province. *J. Nat. Resour.* **2021**, *36*, 948–964. [CrossRef]
63. Cheng, G.; Li, Y.; Zhang, M. Research progress on desulfurization technology of high-sulfur bauxite. *Trans. Nonferrous Met. Soc. China* **2022**, *32*, 3374–3387. [CrossRef]
64. Cheng, G.; Zhang, J.; Su, H.; Zhang, Z. Synthesis and characterization of a novel collector for the desulfurization of fine high-sulfur bauxite via reverse flotation. *Particuology* **2023**, *79*, 64–77. [CrossRef]
65. Zabeo, A.; Pizzol, L.; Agostini, P.; Critto, A.; Giove, S.; Marcominr, A. Regional risk assessment for contaminated sites Part 1: Vulnerability assessment by multicriteria decision analysis. *Environ. Int.* **2011**, *37*, 1295–1306. [CrossRef] [PubMed]
66. Lisa, P.; Critto, A.; Agostini, P.; Marcomini, A. Regional risk assessment for contaminated sites Part 2: Ranking of potentially contaminated sites. *Environ. Int.* **2011**, *37*, 1307–1320. [CrossRef]
67. Agostini, P.; Pizzol, L.; Critto, A.; D'Alessandro, M.; Zabeo, A.; Marcomini, A. Regional risk assessment for contaminated sites Part 3: Spatial decision support system. *Environ. Int.* **2012**, *48*, 121–132. [CrossRef] [PubMed]