Journal Pre-proof

Cadmium source identification in soils and high-risk regions predicted by geographical detector method

Yinjun Zhao, Qiyu Deng, Qing Lin, Changyu Zeng, Cong Zhong

PII: S0269-7491(19)37684-5

DOI: https://doi.org/10.1016/j.envpol.2020.114338

Reference: ENPO 114338

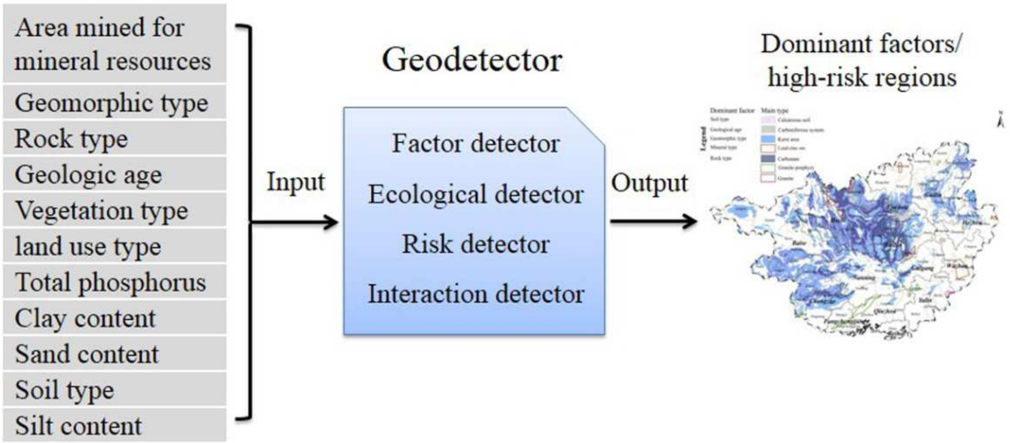
To appear in: *Environmental Pollution*

Received Date: 20 December 2019

Revised Date: 26 February 2020

Accepted Date: 4 March 2020

Please cite this article as: Zhao, Y., Deng, Q., Lin, Q., Zeng, C., Zhong, C., Cadmium source identification in soils and high-risk regions predicted by geographical detector method, *Environmental Pollution* (2020), doi: https://doi.org/10.1016/j.envpol.2020.114338.



Cadmium (Cd) contamination in soils has become a serious and widespread environmental problem, especially in areas with high natural background Cd values, but the mechanism of Cd enrichment in these areas is still unclear. We have Quantitatively identified main natural sources of high background values of Cd in the soils in karst areas and main anthropogenic sources of Cd in the soils in mining areas. We concluded that the natural sources of Cd in soils in karst areas mainly derived from the weathering and deposition processes of carbonatite from the Carboniferous system and the anthropogenic sources of Cd in soils mainly derived from the mining of lead-zinc ores. We have also identified the high-risk regions for Cd in the soils, where should be pay more attention. This study provides some resources for other studies and for establishing pollution control strategies in Guangxi.

Cadmium source identification in soils and high-risk regions predicted by geographical detector method

Yinjun Zhao[a,b], Qiyu Deng[c], Qing Lin[a,b,*], Changyu Zeng[c], Cong Zhong[a,b]

[a] Key Laboratory of Environment Change and Resources Use in Beibu Gulf, Ministry of Education, Nanning Normal University, Nanning 530001, PR China

[b] Key Laboratory of Earth Surface Process and Intelligent Simulation (Guangxi), Nanning Normal University, Nanning 530001, PR China

[c] School of Geography and Planning, Nanning Normal University, Nanning 530001, PR China

Corresponding author: 765649477@qq.com

Abstract

Cadmium (Cd) contamination in soils has become a serious and widespread environmental problem, especially in areas with high natural background Cd values, but the mechanism of Cd enrichment in these areas is still unclear. This study uses the Guangxi Zhuang Autonomous Region (Guangxi), a typical area with a high background Cd level and Cd pollution related to mining activities, as an example to explore the source and predict areas with high Cd risk in soils based on the geographical detector method. The areas with high Cd in Guangxi soils were classified into non-mining areas and mining areas according to their potential Cd sources. The results show that the rich Cd content in the soils from the non-mining area of Guangxi was mainly derived from the soil type (q = 0.34), geological age (q = 0.27), rock type (q = 0.26) and geomorphic type (q = 0.20). Specifically, the Cd content was derived from the weathering and deposition processes of carbonatite from the Carboniferous system in the karst area. The high Cd content in the soils of the mining area of Guangxi was mainly derived from the area mined for mineral resources (q = 0.08) and rock type (q =0.05). Specifically, the Cd content was derived from the mining of lead-zinc ores. The areas in Guangxi with a high risk of Cd soil pollution are mostly concentrated in karst areas, such as Hechi, Laibin, Chongzuo, southern Liuzhou and Baise, northern Nanning city and northeastern Guilin city, and some mining areas. These results indicated that the high Cd concentration in the soils of large areas of Guangxi is probably due to natural sources, while the high Cd concentration around mining areas is due to anthropogenic sources. The results will be useful for soil restoration and locating and controlling contaminated agricultural land.

Key words: Cadmium; Source; High-risk region; Geographical detector; Guangxi

Capsule abstract

We quantitatively identified the main natural and anthropogenic sources of Cd in the soils and predicted the high-risk regions.

Corresponding author: Qing Lin, 765649477@qq.com

1 Introduction

The heavy metal cadmium (Cd) is a highly toxic element with a wide variety of adverse effects on humans, animals and plants (Waalkes, 2000; Berglund et al., 2015; Benavides et al., 2005; Chou et al., 2005). Cd exerts toxic effects on the lungs, kidneys, bones, prostate, heart, skin, placenta and gastrointestinal tract (Pinot et al., 2000) and has been clearly shown to be a carcinogen for humans and experimental animals (Waalkes, 2000; Żukowska and Biziuk, 2008). Cd is a naturally occurring element and a natural component of rocks; the Cd concentration is 0 ~ 11 mg kg^{-1} in Earth's crust (Bowen, 1979). Cd is transferred into the environment (e.g., soils) by both natural and anthropogenic processes and is generally present in the environment at low levels.

However, the mean concentrations of Cd in soils with no obvious anthropogenic sources vary between 0.06 and 1.10 mg kg^{-1} (Pinot et al., 2000; Kabata-Pendias, 2010). This difference means that the weathering of parent rocks may produce natural Cd enrichment in soils (Kabata-Pendias and Dudka, 1991; Naidu et al., 1997; Liu et al., 2013). High background values of Cd in soils have been reported in many regions worldwide, especially in karst areas mostly covered by black shake (NEPA, 1990; Pinot et al., 2000; Kabata-Pendias, 2010; Liu et al., 2013; Luo, 2018; Zhang et al., 2018). Compared to other shales or the average crust, black shales are commonly highly enriched in Cd because Cd generally occurs in sulfides in black shale and is easily released when exposed to oxygen and water (Kabata-Pendias and Dudka, 1991; Liu et al., 2017).

On the other hand, anthropogenic Cd found in atmospheric emissions, which is mostly due to mining and smelting of nonferrous metals that account for 73% of all anthropogenic sources (Pacyna and Pacyna, 2001), is thought to be a primary source of environmental Cd because it accounts for 70% of the total cadmium (Pinot et al., 2000; Cheng et al., 2014). Approximately 80% ~ 90% of anthropogenic Cd ultimately enters soils (Zhang et al., 2015; Liao et al., 2015). According to an assessment of Cd in China based on 486 studies (Zhang et al., 2015), the average cadmium concentration in soils was ranked as follows: mining and smelting areas (8.27 mg kg^{-1}) > wastewater irrigation areas (1.91 mg kg^{-1}) > urban and suburban areas (0.46 mg kg^{-1}) > remote areas (0.16 mg kg^{-1}). Specifically, the high soil Cd concentration surrounding mining areas was mainly due to mineral excavation, ore transportation, smelting and refining, and the disposal of tailings and wastewater, which can allow Cd to enter and accumulate in the soil (Salomons 1995; Shao et al., 2013). The soils surrounding mining areas are more severely polluted by Cd than those far from mining activities (Li et al., 2014). In general, the total Cd concentration in soils comes from volcanic activity and geological weathering of parent rocks

(Liu et al., 2013), together with inputs from anthropogenic sources, such as mining, smelting, sewage irrigation, fertilizers (phosphate fertilizers), manures and other agricultural materials (Alloway and Steinnes, 1999; Razo et al., 2004; Luo et al., 2009).

High background Cd levels in soils are mainly controlled by high Cd values in regional rocks and the process of rock weathering (Naidu et al., 1997; Liu et al., 2013; Wen et al., 2020). To explore the leading factors of Cd accumulation is very important to understand the anomalies in the high levels of Cd in soils. The impacts of rocks with high Cd on the background values of Cd in soils are well known. For example, the level of Cd in soils is very high and exhibits large anomalies in black shale, phosphorite (Liu et al., 2017) and cadmiferous Pb-Zn ore areas. However, the anomalies in the levels of high Cd in the soils of karst areas most likely involve many factors.

Intensive geochemical studies can explore the sources of Cd anomalies in soils and can help reveal the mechanisms of Cd enrichment in soils (Wang et al., 2019). However, it is important to determine the factors that play a key role in the enrichment of Cd in soils before carrying out geochemical studies. Spatial statistical analysis may be a beneficial complement to traditional experimental methods for identifying natural sources and the mechanisms of Cd enrichment in soils (Liu et al., 2013; Shi et al., 2020). The traditional experiments used to study Cd in soils would be greatly improved by including pertinent factors, such as rock type (Luo, 2018; Zhang et al., 2018; Ministry of Land and Resources, 2014), total phosphorus (McDowell et al., 2013; Jia et al., 2019) and mining-related activities (Zhang et al., 2015; Liu et al., 2013), which can be determined based on the driving factors and high-risk areas identified by spatial statistical models. In addition, spatial statistical analyses can be used to determine the overall spatial configuration and general condition of large areas.

Compared with traditional statistical methods, such as correlation analysis, geostatistical techniques (Lv, 2019), multiple regression analysis (Qiu et al., 2016), principal component analysis (Ordonez et al., 2003) and geographically weighted regression (Liu et al., 2013), a growing number of studies have applied the geographical detector method (GDM) in recent years to quantify the influence of the factors controlling the spatial patterns of geographical phenomena. The GDM has been applied to many topics, such as heavy metal pollution (Shi et al., 2018; Luo et al., 2019; Qiao et al., 2018), precipitation changes (Zhao et al., 2017), and endemic disease patterns (Wang et al., 2010). The GDM can capture the spatial relationship between the driving factors and the concentration of heavy metals and the interactions among the driving factors, and it does not require the data to meet strict assumptions. Based on the spatially stratified heterogeneity of geographical phenomena, the GDM assumes that if a geographical phenomenon (Y; e.g., the spatial distribution of Cd) is consistent with the spatial distribution of another geographical factor (X; e.g., rock type), then factor X will have a definite determinant power on the occurrence and development of phenomenon Y (Wang et al., 2010; Wang et al., 2016; Zhao et al., 2017). The free software for executing the GDM was downloaded from http://www.geodetector.org.

The Guangxi Zhuang Autonomous Region (Guangxi) of China has a high geochemical background (Wen et al., 2019) and a background Cd value in the soils of 0.27 mg kg^{-1}, which is 2.75 times the national average of China (NEPA, 1990). In the karst regions of Guangxi, the background values of Cd in limestone and limestone soils are 0.22 mg kg^{-1} and 1.12 mg kg^{-1}, respectively, which are the highest in China (NEPA, 1990). Although previous studies have explored some findings based on the analysis of limited samples from some small karst areas

(Ministry of Land and Resources, 2014; Wang et al., 2019; Wen et al., 2020), the sources and mechanisms of Cd enrichment in soils are still unclear. In contrast, high Cd concentrations in soils have been reported in mining-related areas in Guangxi (Zhang et al., 2008; Zhang et al., 2009; Guo et al., 2018; Liu et al., 2016; Wang et al., 2018).

Based on the high Cd background values and the major anthropogenic sources of Cd (mining-related activities) identified in Guangxi, we assumed that the high background value of Cd in the soils comes from a natural source (i.e., the geological weathering of parent rocks in the karst area) and that the anthropogenic source of the high Cd levels in Guangxi soils is mainly related to mining. This study was designed to test the above assumptions. The high Cd levels in Guangxi soils can be categorized into two types: non-mining areas (corresponding to the natural sources of Cd in the soils) and mining areas (corresponding to the anthropogenic sources of Cd in soils). Longzhou County, a karst area with few mining-related activities, and the Diaojiang River basin, a karst area highlighted by mining activities, were selected as the non-mining area and mining area, respectively, to identify the driving factors behind the high Cd concentrations in the soils using GDM. Mashan County, which has relatively limited mining-related activity, and Huanjiang County, a lead-zinc (Pb-Zn) ore area, were selected as test areas to validate the findings. Finally, we identified and predicted the regions with a high risk of a high Cd concentration in the soils in Guangxi. This study provides resources for other studies and for establishing pollution control strategies in Guangxi.

2 **Materials and Methodology**

2.1 Study area

Guangxi is located in Southwest China, sharing its southwestern border with Vietnam (Fig.

1). The karst area in Guangxi, which accounts for 37.80% of the whole region (236,700 km^2), belongs to the central karst geomorphology of eastern Asia, which is one of the largest karst areas in the world (Liu et al., 2008). The plain accounts for only 26.90% of the region. Guangxi has diverse and large-scale mineral resources, such as Pb-Zn ores and manganese ores and is one of the ten nonferrous metal-producing areas in China. The leading reserves of some mineral resources in China, and even the world, are found in Guangxi. Based on the development of the mining industry, Guangxi was identified as one of fourteen key provincial regions for the integrated control of heavy metal pollution in the "Twelfth Five-Year Plan".

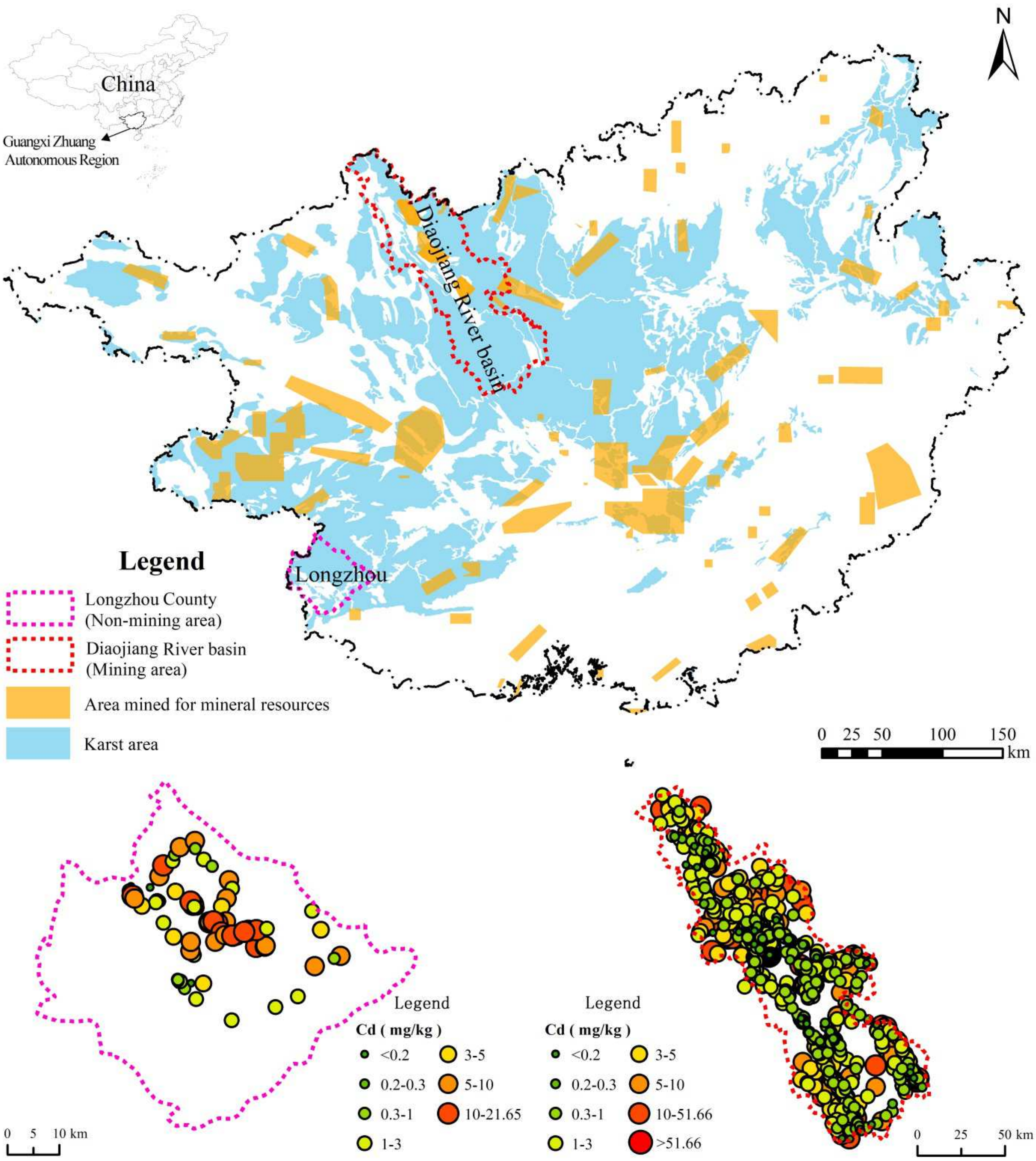


Fig. 1. Basic information related to the study area.

The Diaojiang River basin (approximately 3,580 km^2) is located in the northwest of Guangxi, and its main branch is 229 km. This basin has the largest Zn-Sn mine in China. From upstream to downstream, there is a tin-polymetallic ore field near the town of Dachang, a Zn-Pb-Sn mine near the town of Wuxu and a Mn mine near the town of Laren (Liu et al., 2016; Wang et al., 2018). Mining activities have occurred for several decades and have released a large amount of heavy metals (e.g., Cd) into the soils (Wang et al., 2018). Longzhou County is mostly covered by karst areas and has few mining activities. The Diaojiang River basin and Longzhou County were selected as typical examples of mining and non-mining areas, respectively, to build two models.

2.2 Sample collection and measurement

A total of 100 soil samples were randomly collected in Longzhou County in July 2016 (Fig. 1), and 17 soil samples were collected along a road in Mashan County in April 2017. All the soil samples were air-dried, ground in a stainless-steel grinder chamber, and then sieved through a 100-mesh grid sieve (0.15 mm) to remove impurities (Shi et al., 2018). The soil Cd content was measured using inductively coupled plasma-mass spectrometry (ICP-MS). Another 837 records of the Cd contents in soil samples from the Diaojiang River basin (from September 2013) were provided by the Chinese Research Academy of Environmental Sciences and processed by standard procedures (Bai et al., 2017).

2.3 GDM

The GDM proposed by Wang et al. (2010) was reviewed to measure the spatially stratified heterogeneity of the Cd concentration (*H*) and explore how a factor (*D*) explains the spatial pattern of *H* in this research context. The four components of the GDM (the factor detector, ecological detector, risk detector and interaction detector) are detailed below.

We assume that the Cd concentrations have a spatial distribution similar to that of an impact factor if the impact factor leads to the observed distribution of Cd concentrations based on the factor detector. All the impact factors are quantified by power values as follows:

$$q = 1 - \frac{1}{n\sigma_H^2} \sum_{i=1}^{m} (n_{D,i} \cdot \sigma_{H_{D,i}}^2) \quad (1)$$

where D represents a driving factor layer (e.g., soil type), which must already be categorized; m is the number of zones (categories) of factor D (D={D_1, D_2, D_3, …, D_m}); H layer represents the spatial distribution of the Cd concentration; q represents the power of determinant D on H; n and σ_H^2 represent the total number of samples and the global variance in H over the entire study area, respectively; and $n_{D,i}$ and $\sigma_{H_{D,i}}^2$ represent the number of samples in the i-th subregion of factor D and the variance in H over the i-th subregion of factor D, respectively. In general, the value range of q is [0, 1] (Wang et al., 2010). A high q means that the driving factor category explains more of the spatial variation in the Cd concentration.

The ecological detector compares the potential driving factors (e.g., C factor) and determines the one that is more significant than the others (e.g., D factor) in causing the spatial pattern of the Cd concentration. This is measured using the F-test:

$$F = \frac{n_{C,p}(n_{C,p}-1)\sigma_{C,m}^2}{n_{D,p}(n_{D,p}-1)\sigma_{D,m}^2} \quad (2)$$

where F is the test value of F; $n_{C,p}$ and $n_{D,p}$ denote the number of samples of impact factors C and D in sample unit p, respectively; and $\sigma_{C,m}^2$ and $\sigma_{D,m}^2$ are the dispersion variances of the impact factors C and D, respectively.

The risk detector compares the difference of the average values of different types or ranges of a driving factor on the Cd accumulation in the soils through a t-test. The larger the difference is, the greater the contribution to Cd accumulation in soils.

$$t_{ij} = \frac{R_i - R_j}{\left[\sigma_i^2/n_i + \sigma_j^2/n_j\right]^{1/2}} \quad (3)$$

where t_{ij} is the test value of t; $\mathrm{R_i}$ and R_j are the average values of the spatial pattern of the Cd concentration over property i and property j of the driving factor R, respectively; σ_i^2 and σ_j^2 are the variances in the spatial pattern of the Cd concentration from property i and property j, respectively; and n_i and n_j are the sample sizes of the two properties.

The interaction detector shows that when two different factors (x1 and x2) are combined, they weaken, enhance, or are independent of each other in their impact on Cd accumulation. The effect is determined by comparing $q(x1 \cap x2)$ with the values of $q(x1)$ and $q(x2)$, where the '∩' symbol denotes a new layer created by overlaying the x1 layer and the y layer (Wang et al., 2010). If $q(x1 \cap x2) < \min(q(x1), q(x2))$, the variables nonlinearly weaken each other; if $\min(q(x1), q(x2)) < q(x1 \cap x2) < \max(q(x1), q(x2))$, the variables weaken each other uniformly; if $q(x1 \cap x2) > \max(q(x1), q(x2))$, the variables enhance each other bidirectionally; and if $q(x1 \cap x2) > q(x1) + q(x2)$, the variables nonlinearly enhance each other. If $q(x1 \cap x_2) = q(x_1) + q(x_2)$, then the variables are independent of each other (Wang et al., 2010; Zhao et al., 2017).

2.4 Data processing of the driving factors

According to a literature review, expert consultation and access to research data, all the potential driving factors related to the spatial pattern of Cd concentrations, such as the geomorphic type (GT), area mined for mineral resources (AMMR), soil type (ST), vegetation type (VT), total phosphorus in the soil (TP), land use type (LUT), rock type (RT), geologic age (GA) and soil context, including silt content (STC), sand content (SDC) and clay content (CYC) in soils, were considered. RT and GA were derived from a geologic map of Guangxi at a scale of 1:1750,000. GT was derived from a geomorphic map of Guangxi at a scale of 1:1000,000. ST, VT, LUT and soil texture (SDC, STC and CYC) were provided by the Data Center for Resources and Environmental Sciences (RESDC), Chinese Academy of Sciences (http://www.resdc.cn). AMMR was derived from the General Mineral Resources Planning of Guangxi (2008~2015), which was released by the Bureau of Land and Resources of Guangxi (Fig. 1). All the driving factors were mapped and categorized using original categories, expert knowledge or a categorization algorithm (Fig. 1, 2), such as equal interval, quantile and k-means, which depend on the improvement of the *q* value (Wang and Xu, 2017).

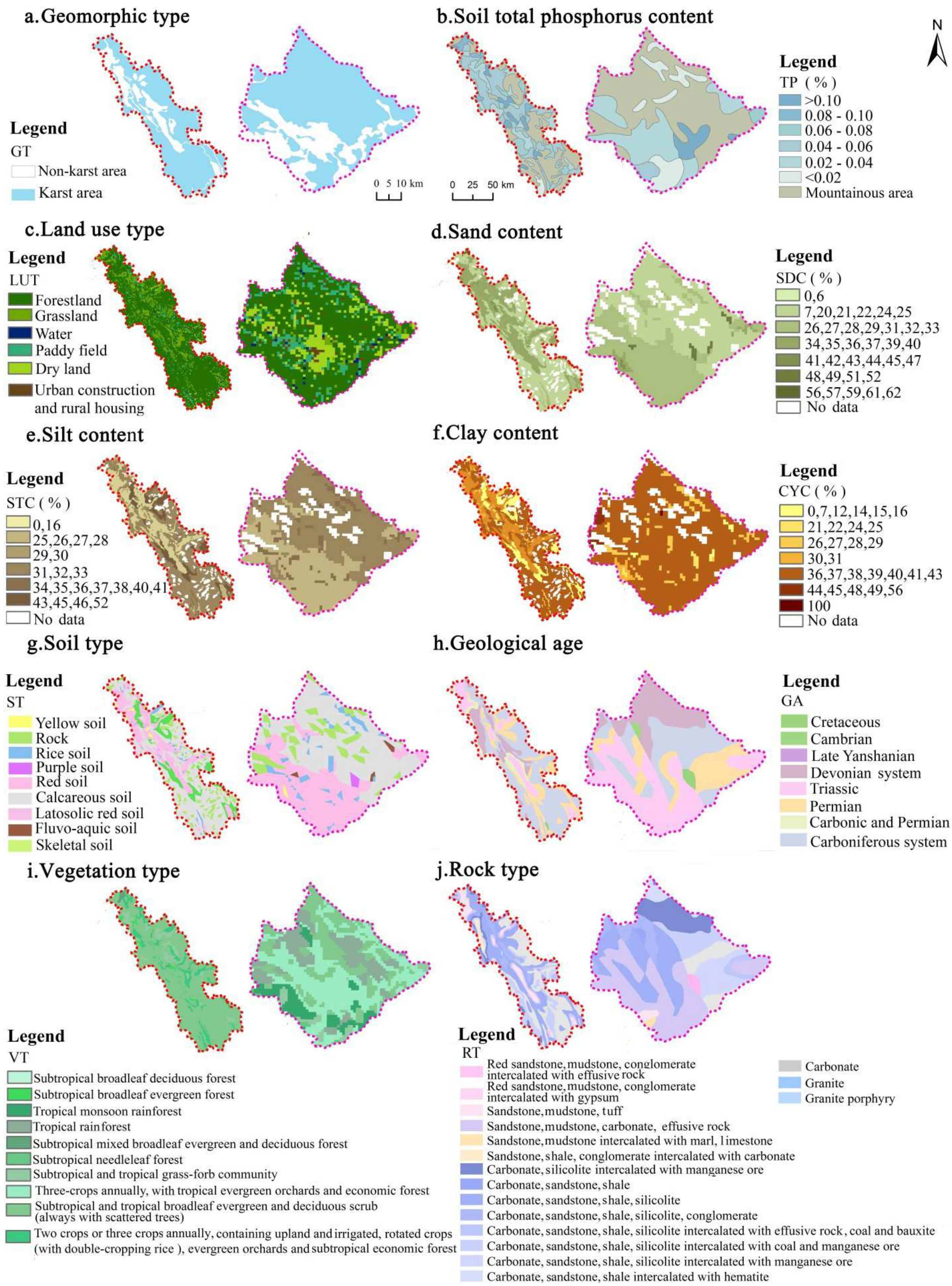


Fig. 2. Spatial distribution of the paired factors driving Cd enrichment in the Diaojiang River basin (left) and Longzhou County (right) soils.

3 Results and Analysis

3.1 Cd concentration in soils

Table 1 shows the descriptive statistics for the Cd concentrations in Longzhou County and Diaojiang River basin soils. The Cd concentrations in the rocks and soils varied, with ranges of 0.02~2.10 and 0.03~49.29 mg kg^{-1}, respectively. Rocks from the Carboniferous system had the highest average Cd content (0.89 mg kg^{-1}). The Cd contents in the rocks in Longzhou County (average values varying from 0.29-0.89 mg kg^{-1}) largely exceeded the Cd content in the Earth's crust (0.11 mg kg^{-1}) (Bowen, 1979). The Cd contents in 92% of the soil samples from Longzhou County exceeded the background value of Cd in the soils from Guangxi (0.267 mg kg^{-1}), and the maximum was 184.59 times the background value. Similarly, the Cd contents in the soils in the Diaojiang River basin varied greatly, ranging from 0.01 ~ 178.30 mg kg^{-1}.

Table 1. Cd concentrations in rocks with different geological ages and soils from Longzhou County and the Diaojiang River basin (mg kg^{-1}).

Study area	Type		Average value	Standard deviation	Maximum	Minimum
Longzhou	Rock (n=86)	Devonian system	0.29	0.32	1.37	0.02
		Carboniferous system	0.89	0.44	1.81	0.20
		Permian system	0.79	0.57	2.10	0.25
		Cambrian system	0.38	0.21	0.52	0.23
	Soil (n =100)		6.54	8.15	49.29	0.03
Diaojiang River basin	Soil (n = 837)		5.33	13.48	178.30	0.01

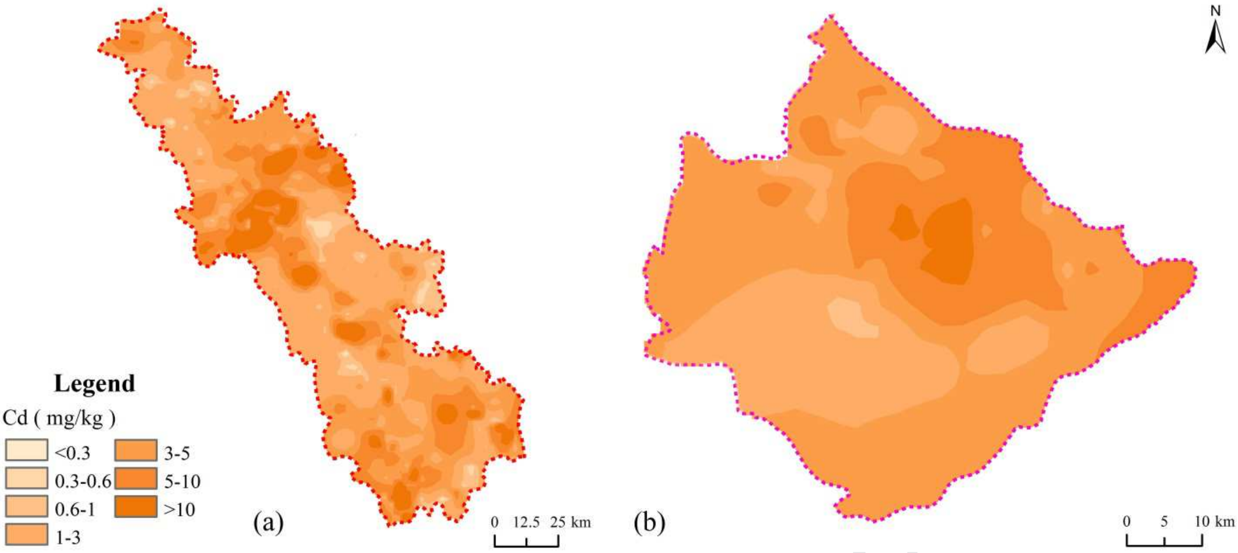


Fig. 3. Spatial distribution of Cd concentrations in the Diaojiang River basin (a) and Longzhou County (b).

Fig. 3 shows the spatial distribution of Cd in the soils from the two study areas based on an inverse distance weighted interpolation map. Most of the study area had a high Cd content (> 0.3 mg kg^{-1}), which means that Cd pollution is commonly found in the soils throughout the study regions. The high Cd contents were relatively organized. These findings suggest that the entire study area has a high risk of Cd pollution, especially in some clustered areas with high Cd concentrations.

3.2 The dominant factors driving high Cd concentrations

Based on the factor detector model, all the driving factors were ranked by their impact on the Cd concentrations in the soils of Longzhou County and the Diaojiang River basin (Table 2). In Longzhou County (the non-mining area), all the driving factors were statistically significant at the 0.05 level, which means that all the driving factors have significant impacts on the spatial distribution of Cd in the soils. Among the driving factors, the *q* value of ST was the highest (0.34), and the *q* values of ST, SDC, CYC, GA, RT and TP were larger than 0.20. This result indicates that these factors could be regarded as leading factors (Li et al., 2013; Liu et al., 2013) that strongly explain the spatial pattern of the Cd concentration, i.e., ST, SDC, CYC, GA, RT, TP and GT are potential leading factors that may have a relatively strong influence on the patterns of high Cd concentrations in this non-mining area, while the *q* values of VT, LUT and STC are comparatively small, and these factors likely have a relatively weak influence on the spatial pattern of high Cd concentrations. According to the ecological detector, ST, a potential leading factor, was statistically significant when paired with SDC and CYC (Table S1). That is, ST clearly had a greater impact than SDC and CYC on the enrichment of the Cd concentration. Thus, SDC and CYC were eliminated from the list of potential dominant factors. The same analysis was conducted for the other potential leading factors (Table S1). We concluded that ST, GA, RT and GT were the leading factors that had the greatest impact on the enrichment of the Cd concentration, and the remaining factors may have had a relatively weak influence on the Cd concentration.

In the Diaojiang River basin (the mining area), all the selected factors except SDC were statistically significantly associated with the Cd concentration at the 0.05 level. All the driving factors can be easily categorized into a group with high *q* values (AMMR and RT) and a group with low *q* values (the remainder of the factors) because of clear clustering in the *q* values (Table 2). AMMR and RT were identified as potential leading factors. There was not a statistically significant difference between AMMR and RT according to the ecological factors, which means that they both have a great impact on the enrichment of the Cd concentrations in the soil of the Diaojiang River basin. We concluded that AMMR and RT were the leading factors for the enrichment of the Cd concentration in the soil of the Diaojiang River basin.

Table 2. *q* values of the factor detector in the mining area and non-mining area

Study area	Driving factor	*q* value	Study area	Driving factor	*q* value
Longzhou County (Non-mining area)	ST	0.34***	**Diaojiang River basin** (Mining area)	AMMR	0.08***
	SDC	0.30***		RT	0.05***
	CYC	0.30***		GA	0.03***
	GA	0.27***		TP	0.03***
	RT	0.26***		VT	0.02***
	TP	0.23***		GT	0.01***
	GT	0.20***		LUT	0.01**
	VT	0.20***		ST	0.00***
	LUT	0.13***		STC	0.00***
	STC	0.02***		CYC	0.00***
				SDC	0.00

Note: *** and ** indicate statistical significance at 1% and 5%, respectively.

3.3 The effect of the interaction of the driving factors

Each pair of driving factors had a greater impact than single factors according to the interaction detector (Table S2). The results clearly show that the *q* values of the two dominant factors (ST and GA) enhanced all the other factors. The *q* values of 29 pairs exceeded the largest *q* value found for a single factor (q = 0.34). For example, GA and TP alone each had low *q* values (0.27 and 0.23, respectively), but their interaction explained 49% (i.e., q (GA∩TP) = 0.49) of the spatial distribution of the Cd concentration in Longzhou County, and their combined *q* value was much greater than the largest *q* value for a single factor (q = 0.34). In addition, the *q* values between ST and other factors were enhanced (Table S2). The same interaction effects were found in the Diaojiang River basin. The *q* values of 11 pairs exceeded the largest *q* value, which was found for AMMR. AMMR enhanced the impact of all the other factors on the accumulation of Cd in soils (Table S3).

3.4 Prediction of areas with a high risk of Cd in Guangxi soils

According to the risk detector, we compared the effect of the type or range of each dominant factor on the high Cd concentrations. For the non-mining area, limestone soils (of the ST factors), the Carboniferous system (of the GA factors), carbonatite (of the RT factors) and the karst landform (of the GT factors) had the highest mean Cd concentration among all types or ranges of each dominant factor and had a significantly greater impact than the other types or ranges of each dominant factor on the enrichment of the Cd concentration (Table 3). Thus, those types or ranges (of each dominant factor) were used to determine the threshold values of high risk (Table 3). In general, the highest Cd concentrations in the soils of the non-mining area mostly originated in limestone soils, which were possibly derived from the geological weathering of carbonatite of the Carboniferous system located in the karst area; therefore, areas with these characteristics had the highest risk of Cd contamination. In the mining area, we concluded that Pb-Zn ore contributed the most to the input of Cd into the surrounding soils, and granite porphyry and granite are ore-forming materials that are the sources of Pb-Zn ore (Table 3). In addition, the areas with the characteristics of the main types or ranges of the dominant driving factors have very high Cd contents, ranging from 4.84 ~ 9.75 mg kg^{-1} (Table 3).

Table 3. Main influential types (ranges) of the dominant factors and their impacts on the enrichment of the Cd concentration

Study area	Dominant factor	Main type (range) of the dominant factor	Mean Cd concentration (mg kg^{-1})
Longzhou County (Non-mining area)	ST	Limestone soil	4.84
	GA	Carboniferous system	5.27
	RT	Carbonatite	5.98
	GT	Karst landform	4.25
Diaojiang River basin (Mining area)	AMMR	Metal ores related to Pb-Zn	9.75
	RT	Granite porphyry	7.89
		Granite	5.38

We extrapolated the conclusions of the GDM in Longzhou County and the Diaojiang River basin, two typical experimental areas, to all of Guangxi. Fig. 4 shows the main types (ranges) of the dominant driving factors (Table 3) for all of Guangxi; there is a higher risk of Cd in the soils where these factors occur than in the soils without influential factors. These high-risk areas for Guangxi soils are mostly concentrated in Hechi, Laibin, Chongzuo, southern Liuzhou and Baise, northern Nanning city and northeastern Guilin city and some mining areas. Fig. 4 also indicates where many of the abovementioned driving factor types overlapped (i.e., where the risk was high). The cities of Hechi, Laibin and Chongzuo should be the highest priority areas for soil restoration because these areas had the highest number of influential driving factor types or ranges.

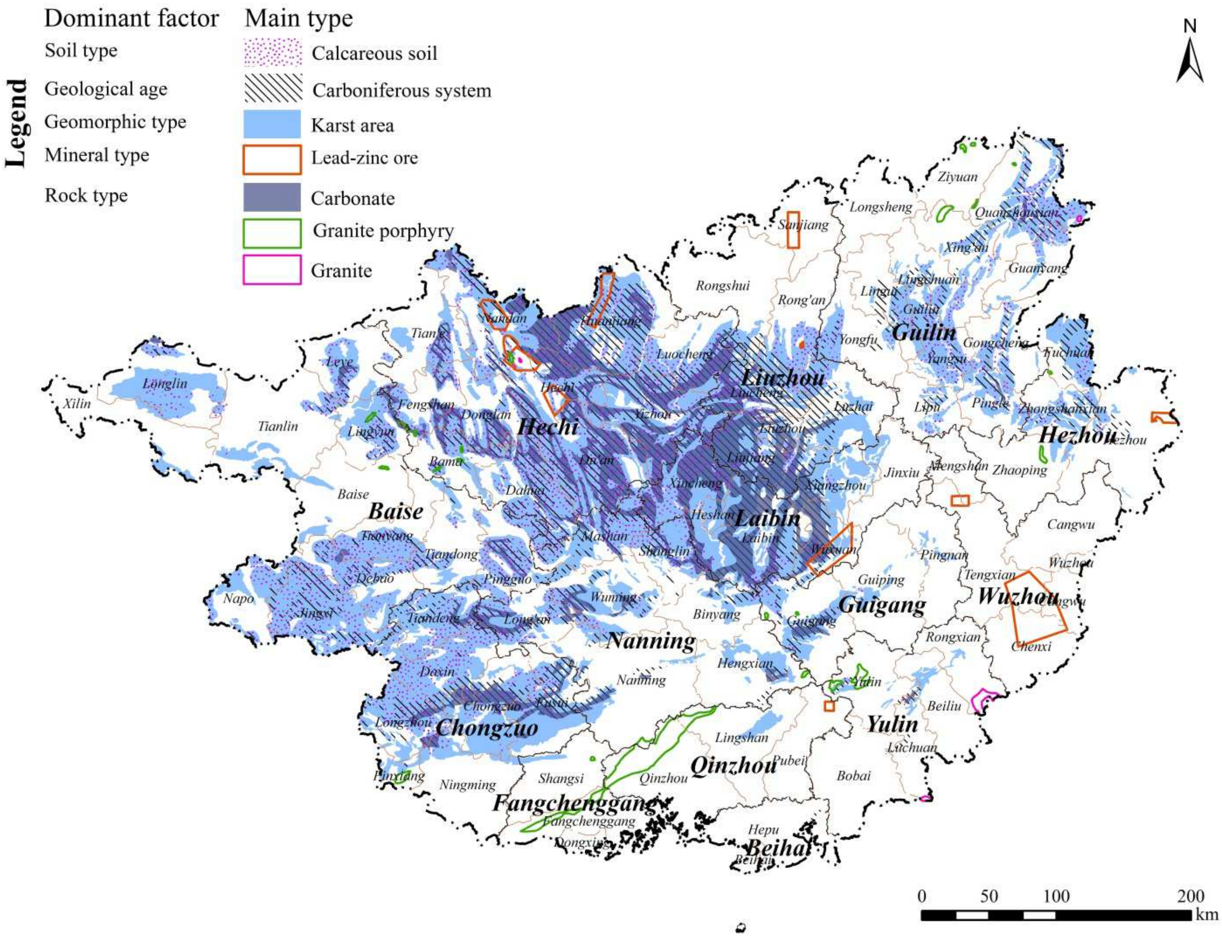


Fig. 4. Areas of Guangxi with a high risk of Cd in the soils.

4 Discussion

In this study, GDM was employed to measure the contribution of various driving factors to the enrichment of Cd in Guangxi soils and then predict the high-risk areas. The GDM performed well and had advantages; for example, (1) it revealed the causal relationship between the Cd concentration and driving factors; (2) it easily processed categorical variables, which is rare for statistical analysis methods; and (3) it worked with input data that were either relatively easy to obtain or free, such as remote sensing data, geological maps and geomorphological maps (Shi et al., 2018).

Our findings showed that ST, GA, RT and GT were the dominant factors driving Cd accumulation in soils through the natural weathering of parent rocks. This finding is consistent with the findings of other studies (e.g., Wen et al., 2020; Zhang et al., 2018). Zhang et al. (2018) reported that in an area of southwestern China (e.g., Guangxi) with a high regional geochemical background value of Cd, the excessive accumulation of Cd in soils is substantially attributable to natural pedogenic processes, such as the biogeochemical weathering of metal-enriched carbonate rocks (e.g., limestone). Carbonate rock is the main parent rock in karst areas. During the process of soil formation, calcium carbonate, the main chemical constituent of the rock, is dissolved and leached, while Cd is retained and gradually accumulates in the topsoil, which makes the Cd concentration in the topsoil 10-20-fold higher than that in the parent rocks (Ministry of Land and Resources, 2014; Zhang and Song, 2018).

Our findings also show that most areas with a high risk of Cd accumulation in the soils were distributed in northwestern Guangxi (Fig. 4). As shown in Fig. 4, Mashan County, with few mining areas, and Huanjiang County, with Pb-Zn ore mining, were selected as test areas to validate the accuracy of the high-risk areas for the non-mining and mining areas, respectively. A total of 17 soil samples were collected from the high-risk areas in Mashan County (Fig. 5), and their Cd concentrations varied from 0.49 to 2.68 mg kg^{-1} (Table S4). All samples were higher than the background value of Cd in Guangxi soils (0.27 mg kg^{-1}). This result indicates the accuracy of the high-risk areas that were delimited based on the main influential types of the dominant factors.

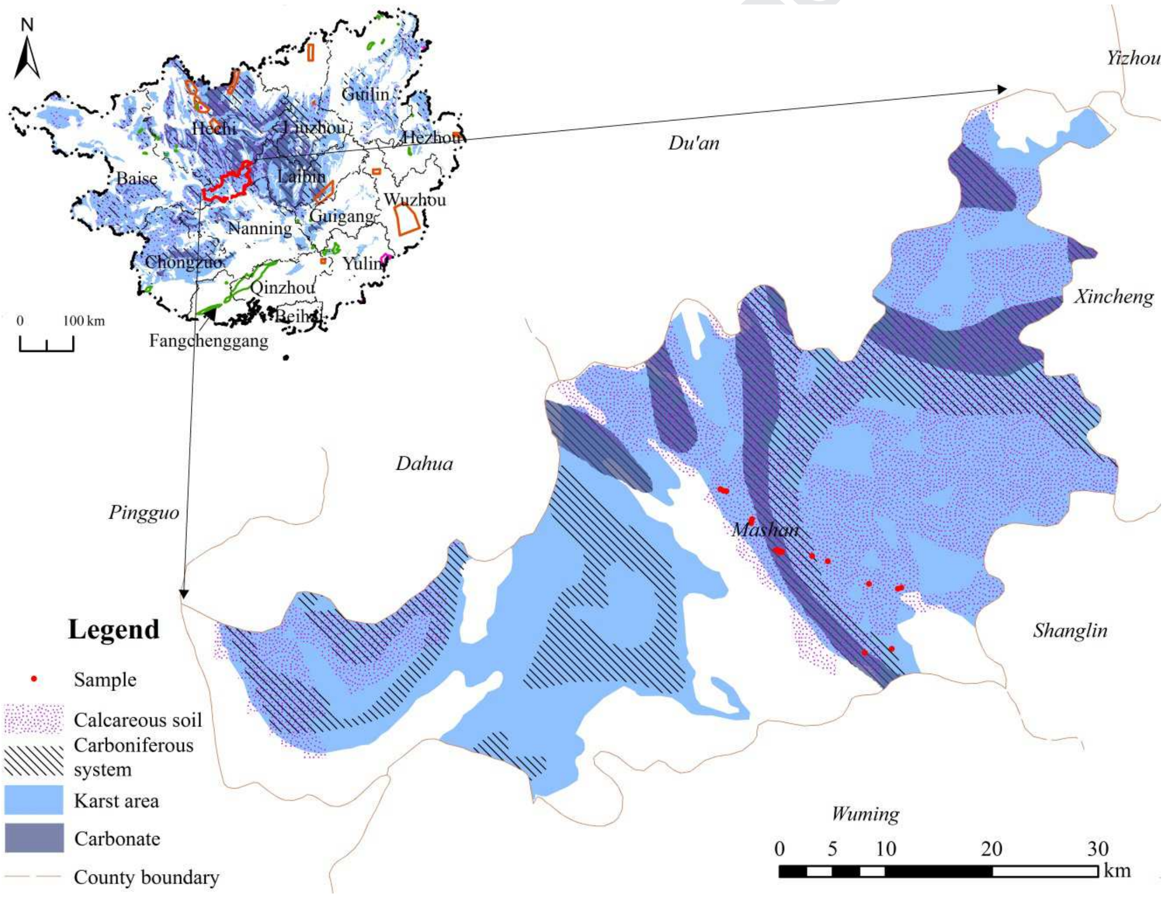


Fig. 5. High-risk regions in Mashan County (i.e., non-mining validation area) and the locations of the sample points

The area planned for mining Pb-Zn ore in Fig. 6 (approximately 256.34 km^2), including the Beichuan and Yamai mining points, was identified as an area with high risk of Cd pollution in Huanjiang County. This region was derived from Fig. 4 and overlaid with the spatial distribution of Cd in soils reported by Liu (2017) to validate the accuracy of the high-risk regions. Fig. 6 clearly shows that the minimum Cd concentration in soils is greater than the background value of Cd in Guangxi (0.27 mg kg-1). The Cd concentrations in the soils of the mining area in this region are significantly higher than those in farmland and natural soils (Liu, 2017). Qiao et al. (2017) reported that the mining activities at the Beichuan and Yamai mining points have polluted nearby and downstream soils with heavy metals (e.g., Cd), and mining activities were the primary influencing factor for Cd pollution (Qiao et al., 2019). Therefore, Cd concentrations in soils in this region were dominated by mining-related activities even if we considered the contribution of natural sources of Cd in soils because this region is located in a karst area. Cd pollution in soils due to Pb-Zn exploitation has been previously demonstrated and is widely distributed (Asami, 1984; Zhang et al., 2002; Pérez-Sirvent, 2009) because Cd principally occurs as a byproduct of Zn refinement (Pinot et al., 2000). Fluvial transportation and atmospheric dispersion are the key pathways of the movement of Cd from mining sources to the environment (Razo et al., 2004). Soils adjacent to mining areas are heavily polluted by Cd (Asami, 1984; Zhang et al., 2002). This study also confirmed that mining activities, especially Pb-Zn exploitation, are the dominant reason for the high Cd concentrations in soils of mining areas.

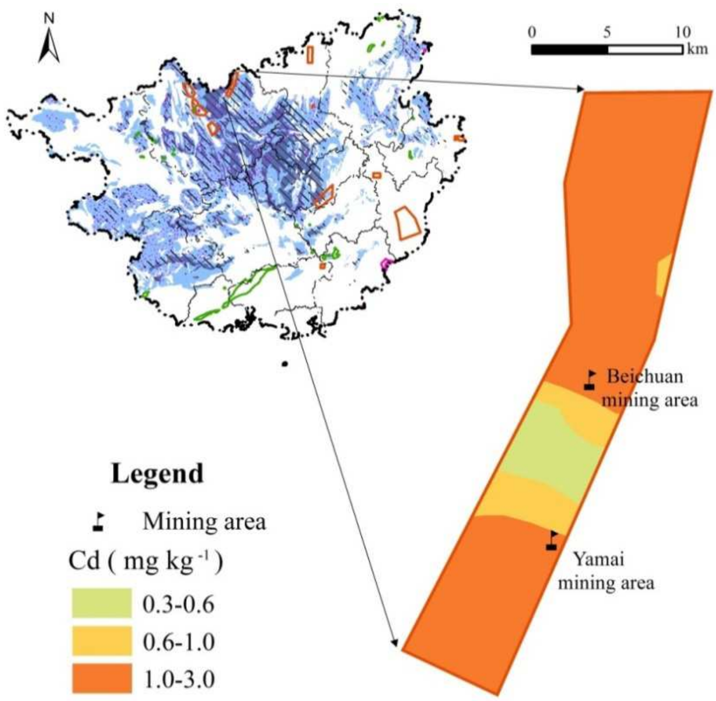


Fig. 6. High-risk region planned for mining Pb-Zn ore in Huanjiang County and its spatial distribution of Cd in soils mapped by Liu (2017)

The results shown in Fig. 4 are important to managers, as these results indicate that soil restoration should be conducted in Guangxi as soon as possible. Moreover, the areas with the highest Cd concentrations, such as Hechi, Laibin, Chongzuo, southern Liuzhou and Baise, northern Nanning, northeastern Guilin and some mining areas, should receive the most attention. The risk area is large; thus, from the perspective of food safety, locating and controlling contaminated agricultural land may be urgent and necessary.

There are three limitations in this study. First, we cannot be sure that the parent rocks of the soils in the study area were near rocks. Second, we extrapolated the results from Longzhou County and the Diaojiang River basin to all of Guangxi, which may have introduced uncertainty. Third, the mechanisms behind the natural sources of high Cd in the soils of the Guangxi karst area require further study.

5 Conclusions

This study explored the leading factors driving Cd accumulation in Guangxi soils and identified high-risk areas using GDM. The results show that the GDM performs well and provides a new framework for exploring the driving factors that control the spatial patterns of high Cd concentrations in soils and identifying high-risk areas. The high Cd concentrations in Guangxi soils were divided into two types (non-mining area and mining area) to explore their respective enrichment mechanisms. For the soils from the non-mining area of Guangxi, the enrichment of Cd was mostly driven by ST, GA, RT and GT, and Cd was derived from the weathering processes of carbonatite from the Carboniferous system in the karst area. Regarding the soils from the mining area in Guangxi, the high concentrations of Cd were mostly driven by AMMR and RT and derived mainly from the mining of Pb-Zn ores. The combined action of multiple factors on the enrichment of Cd in the soils of Guangxi was greater than the effect of single factors. The areas at a high risk for Cd soil pollution in Guangxi are mostly concentrated in Hechi, Laibin, Chongzuo, southern Liuzhou and Baise, northern Nanning city, northeastern Guilin city and some mining areas. The high risk of Cd soil pollution in these relatively large areas was mainly due to the widespread natural sources of Cd.

Acknowledgments

The authors would like to thank the anonymous reviewers for their constructive comments and suggestions for the revision of this paper. This study is supported by the Natural Science Foundation of China (41661085, 41763004, 41461021), the Guangxi Scientific Project (No. AD19110140)， the Guangxi Scholarship Fund of Guangxi Education Department, the Guangxi

Natural Science Foundation (2016GXNSFAA380094) and the Innovation Project of Guangxi Graduate Education (YCSW2018177) and.

Appendix A. Supplementary data

The following are the Supplementary data for this article:

CRediT authorship contribution statement

Yinjun Zhao: Formal analysis, writing - review & editing. **Qiyu Deng**: Investigation, formal analysis, writing - original draft. **Qing Lin**: Writing - review. **Changyu Zeng**: Writing - review. **Cong Zhong**: Writing - review.

Declaration of competing interest

The authors declare no conflicts of interest.

References

1. Alloway, B. J., Steinnes, E., 1999. Anthropogenic additions of cadmium to soils. In: Cadmium in soils and plants. Springer Dordrecht, pp. 97-123.
2. Asami, T., 1984. Pollution of soils by cadmium. In: Changing metal cycles and human health. Springer, Berlin, Heidelberg, pp. 95-111.
3. Bai, Y. L., Sun, W. B.，Li, N.，Gu, Q. B., 2017. Spatial distribution and pollution characteristics of heavy metals in soil of Diaojiang River basin based on geostatistic. Journal of Mining Science and Technology,，2, 409-415 (in Chinese).
4. Benavides, M. P., Gallego, S. M., Tomaro, M. L., 2005. Cadmium toxicity in plants. Braz. J. Plant Physiol. 17, 21-34.
5. Berglund, M., Larsson, K., Grandér, M., Casteleyn, L., Kolossa-Gehring, M., Schwedler, G., Castaño, A., Esteban, M., Angerer, J., Koch, H. M., Schindler, B. K., Schoeters, G., Smolders, R., Exley, K., Sepai, O., Blumen, L., Horvat, M., Knudsen, L. E., Mørck, T. A., Joas, A., Joas, R., Biot, P., Aerts, D., Cremer, K. D., Overmeire., I. V., Katsonouri, A., Hadjipanayis, A., Cerna, M., Krskova, A., Nielsen, J. K. S., Jensen, J. F., Rudnai, P., Kozepesy, S., Griffin, C., Nesbitt, I., Gutleb, A. C., Fischer, M. E., Ligocka, D., Jakubowski, M., Reis, M. F., Namorado, S., Lupsa, I., Gurzau, A. E., Halzlova, K., Jajcaj, M., Mazej, D., Tratnik, J. S., Lopez, A., Cañas, A., Lehmann, A., Crettaz, P., Hond, E. D., Govarts, E., 2015. Exposure determinants of cadmium in European mothers and their children. Environ. Res. 141, 69-76.

6. Bowen, H. J. M., 1979. Environmental chemistry of the elements. Academic Press.

7. Cheng, K., Tian, H. Z., Zhao, D., Lu, L., Wang, Y., Chen, J., Liu, X. G., Huang, Z., 2014. Atmospheric emission inventory of cadmium from anthropogenic sources. Int. J. Environ. Sci. Technol. 11, 605-616.8.

8. Cho, U. H., Seo, N. H., 2005. Oxidative stress in Arabidopsis thaliana exposed to cadmium is due to hydrogen peroxide accumulation. Plant Science 168, 113-120.

9. Guo, G., Song, B., Lei, M., Wang, Y., 2018. Rare earth elements (REEs) in PM10 and associated health risk from the polymetallic mining region of Nandan County, China. Human and Ecological Risk Assessment: An International Journal, 1-16.

10. Jia, X. L., Hu, B. F., Marchant, B. P., Zhou, L. Q., Shi, Z., Zhu, Y. W., 2019. 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. 250, 601-609.

11. Kabata-Pendias, A., 2010. *Trace elements in soils and plants*. CRC press.

12. Kabata-Pendias, A., Dudka, S., 1991. Baseline data for cadmium and lead in soils and some cereals of Poland. Water Air Soil Pollut. 57, 723–731.

13. Kabata-Pendias, A., Mukherjee, A. B., *2007. Trace elements from soil to human. Springer Science & Business Media.*

14. Li, X. W., Xie, Y. F, Wang, J. F., Christakos, G., Si, J. L., Zhao, H. N., Ding, Y. Q., Li, J., 2013. Influence of planting patterns on fluoroquinolone residues in the soil of an intensive vegetable cultivation area in northern China. Sci. Total Environ. *458, 63-69.*

15. Li, Z. Y., Ma, Z. W., van der Kuijp, T. J., Yuan, Z. W., Huang, L., 2014. A review of soil heavy metal pollution from mines in China: pollution and health risk assessment. Sci. Total Environ. 468, 843-853.

16. Liu, C., 2017. Study on the spatial variability of heavy metals in soil and its influencing factors in Xijiang River draining of Guangxi. Guilin University of Technology. (in Chinese).

17. Liu, G. N., Wang, J., Zhang, E., Hou, J., Liu, X. H., 2016. Heavy metal speciation and risk assessment in dry land and paddy soils near mining areas at Southern China. Environ Sci Pollut Res 23, 8709–8720.

18. Liu, Q. F., Li, M. D., Duan, J. N., Wu, H. Y., Hong, X., 2013. Analysis on influence factors of soil Pb and Cd in agricultural soil of Changsha suburb based on geographically weighted regression model. Transactions of the Chinese Society of Agricultural Engineering *29*, 225-234.

19. Liu, Y. S., Wang, J. Y., Deng, X. Z., 2008. Rocky land desertification and its driving forces in the karst areas of rural Guangxi, southwest China. J. Mt. Sci. 5, 350-357.

20. Liu, Y. Z, Xiao, T. F., Ning, Z. P., Li, H. J., Tang, J., Zhou, G. Z., 2013. High cadmium concentration in soil in the Three Gorges region: Geogenic source and potential bioavailability. Applied geochemistry *37*, 149-156.

21. Liu, Y. Z, Xiao, T. F, Perkins, R. B., Zhu, J. M., Zhu, Z. J., Xiong, Y., Ning, Z. P., 2017. Geogenic cadmium pollution and potential health risks, with emphasis on black shale. Journal of Geochemical Exploration 176, 42-49.

22. Luo, L., Ma, Y. B., Zhang, S. Z., Wei, D. P., Zhu, Y. G., 2009. An inventory of trace element inputs to agricultural soils in China. Journal of Environmental Management 90, 2524-2530.

23. Luo, L. L., Mei, K., Qu, L. Y., Zhang, C., Chen, H., Wang, S. Y., Di, D., Huang, H., Wang, Z. F., Xia, F., Dahlgren, R. A., Zhang, M. H., 2019. Assessment of the Geographical Detector Method for investigating heavy metal source apportionment in an urban watershed of Eastern China. Sci. Total Environ. *653*, 714-722.

24. Luo, Y. M., 2018. Regional difference in soil pollution and strategy of soil zonal governace and remediation in China. Chinese Science Bulletin 2, 145-152.

25. Lv, J., 2019. Multivariate receptor models and robust geostatistics to estimate source apportionment of heavy metals in soils. Environ. Pollut. 244, 72-83.

26. Ministry of Land and Resources, P, 2014. In: Ministry of Land and Resources (Ed.), The Report on the National General Survey of Soil Contamination. Ministry of Land and Resources, PRC Ministry of Land and Resources, PRC.

27. Naidu, R., Kookana, R. S., Sumner, M. E., Harter, R. D., Tiller, K. G., 1997. Cadmium sorption and transport in variable charge soils: a review. Journal of Environmental Quality 26, 602-617.

28. NEPA (National Environmental Protection Agency), 1990. The background values of soil elements of China. Chinese Environmental Science Press, Beijing (in Chinese).

29. Ordonez, A., Loredo, J., De Miguel, E., Charlesworth, S., 2003. Distribution of heavy metals in the street dusts and soils of an industrial city in Northern Spain. Archives of Environmental Contamination and Toxicology 44, 0160-0170.

30. Pacyna, J. M., Pacyna, E. G., 2001. An assessment of global and regional emissions of trace metals to the atmosphere from anthropogenic sources worldwide. Environmental Reviews 9, 269-298.

31. Pérez-Sirvent, C., Martínez-Sánchez, M. J., García-Lorenzo, M. L., Molina, J., Tudela, M. L., 2009. Geochemical background levels of zinc, cadmium and mercury in anthropically influenced soils located in a semi-arid zone (SE, Spain). Geoderma 148, 307-317.

32. Pinot, F., Kreps, S. E., Bachelet, M., Hainaut, P., Bakonyi, M., Polla, B. S., 2000. Cadmium in the environment: sources, mechanisms of biotoxicity, and biomarkers. Reviews on environmental health 15, 299-324.

33. Qiao, P. W., Lei, M., Guo, G. H., Yang, J., Zhou, X. Y., Chen, T. B., 2017. Quantitative analysis of the factors influencing soil heavy metal lateral migration in rainfalls based on Geographical detector software: A case study in Huanjiang county, China. Sustainability 9, 1227.

34. Qiao, P. W., Yang, S. C, Lei, M., Chen, T. B., Dong, N., 2019. Quantitative analysis of the factors influencing spatial distribution of soil heavy metals based on geographical detector. Sci. Total Environ. 664, 392-413.

35. Qiu, L. F., Wang, K., Long, W. L., Wang, K., Hu, W., Amable, G. S., 2016. A comparative assessment of the influences of human impacts on soil Cd concentrations based on stepwise linear regression, classification and regression tree, and random forest models. PloS one 11, e0151131.

36. Shao, X., Cheng, H. G., Duan, X. L., Lin, C. Y., 2013. Concentrations and chemical forms of heavy metals in agricultural soil near the world's largest and oldest tungsten mine located in China. *Chemical Speciation & Bioavailability 25*, 125-132.

37. Shi, R. G., Lv, J. G., Liu, Y. Z., 2020. Spatial Distribution and Bio-accumulation of Cadmium and Lead in

Soil, Rice and Vegetables in Typical Pollution Areas, China. Bull Environ Contam Toxicol 104, 307–313.

38. Shi, T. Z., Hu, Z. W., Shi, Z., Guo, L., Chen, Y. Y., Li, Q. Q., Wu, G. F., 2018. Geo-detection of factors controlling spatial patterns of heavy metals in urban topsoil using multi-source data. Sci. Total Environ. 643, 451-459.

39. Waalkes, M. P., 2000. Cadmium carcinogenesis in review. Journal of inorganic biochemistry 79, 241-244.

40. Wang, J., Liu, G. N., Wu, H., Zhang, T., Liu, X. H., Li, W. Q., 2018. Temporal-spatial variation and partitioning of dissolved and particulate heavy metal (loid)s in a river affected by mining activities in Southern China. Environ Sci Pollut Res 25, 9828–9839.

41. Wang, J. F., Bai, X. Y., Liu, F., Zhang, J., Chen, F., Lu, Q., 2019. Enrichments of Cadmium and Arsenic and Their Effects on the Karst Forest Area. Int. J. Environ. Res. Public Health 16, 4665.

42. Wang, J. F., Li, X. H., Christakos, G., Liao, Y. L., Zhang, T., Gu, X., Zheng, X. Y., 2010. Geographical detectors - based health risk assessment and its application in the neural tube defects study of the Heshun Region, China. International Journal of Geographical Information Science 24, 107-127.

43. Wang, J. F., Zhang, T. L., Fu, B. J., 2016. A measure of spatial stratified heterogeneity. Ecological Indicators 67, 250-256.

44. Wang, L., Cui, X. F., Cheng, H. G., Chen, F., Wang, J. T., Zhao, X. Y., Lin, C. Y., Pu, X., 2015. A review of soil cadmium contamination in China including a health risk assessment. Environ Sci Pollut Res 22, 16441–16452.

45. Wen, Y. B., Li, W., Yang, Z. F., Zhang, Q. Z, Ji, J. f., 2020. Enrichment and source identification of Cd and other heavy metals in soils with high geochemical background in the karst region, Southwestern China. Chemosphere 245, 125620.

46. Wen, Y. B., Li, W., Yang, Z. F., Zhuo, X. X., Guan, D. X., Song, Y. X., Guo, C., Ji, J. F., 2019. Evaluation of various approaches to predict cadmium bioavailability to rice grown in soils with high geochemical background in the karst region, Southwestern China. Environ. Pollut. 113645.

47. Zhang, X. Y., Chen, D. M., Zhong, T. Y., Zhang, X. M., Cheng, M., Li, X., 2015. Assessment of cadmium (Cd) concentration in arable soil in China. Environ Sci Pollut Res 22, 4932-4941.

48. Zhang, S., Song, J., 2018. Geochemical cadmium anomaly and bioaccumulation of cadmium and lead by rapeseed (Brassica napus L.) from noncalcareous soils in the Guizhou Plateau. Sci. Total Environ. 644, 624-634.

49. Zhang, S., Song, J., Cheng, Y., McBride, M. B., 2018. Derivation of regional risk screening values and intervention values for cadmium−contaminated agricultural land in the Guizhou Plateau. Land Degrad Dev. 29, 2366-2377.

50. Zhang, X. W., Yang, L. S, Li, Y. H., Li, H. R., Wang, W. Y., Ye, B. X., 2012. Impacts of lead/zinc mining and smelting on the environment and human health in China. Environ Monit Assess 184, 2261-2273.

51. Zhang, X. Y., Tang, L. S., Zhang, G., Wu, H. D., 2009. Heavy metal contamination in a typical mining town of a minority and mountain area, South China. Bull Environ Contam Toxicol 82, 31.

52. Zhang, X. Y., Tang, X. L., Zhao, C. L., Zhang, G., Hu, H. S., Wu, H. D., Hu, B., Mo, L. P., Huang, Li, Wei, J. G., 2008. Health risk evaluation for the inhabitants of a typical mining town in a mountain area, South

China. Annals of the New York Academy of Sciences 1140, 263-273.

53. Zhao, Y. J., Deng, Q. Y., Lin, Q., Cai, C. T., 2017. Quantitative analysis of the impacts of terrestrial environmental factors on precipitation variation over the Beibu Gulf Economic Zone in Coastal Southwest China. Sci. Rep. 7, 44412.

54. Żukowska, J., Biziuk, M., 2008. Methodological evaluation of method for dietary heavy metal intake. J. Food Sci. 73, R21-R29.

- The geographical detector method provides a new integrated analysis framework.
- The natural and anthropogenic Cd and Cd sources in Guangxi soils were detected.
- The natural Cd came from the weathering of carbonatite of the Carboniferous system.
- The main anthropogenic Cd was derived from the mining of lead-zinc ores.
- The high-risk regions for Cd in the soils were mostly concentrated in karst areas.

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: