RESEARCH ARTICLE

Spatial distribution and risk assessment of heavy metals in farmland surface soil in Yongchang, Gansu, China

Yifei Yang[1], Bo Dong[2, 3, *], Dandan Du[2, 3, *], Liqun Cai[2], Yun Li[4], Xiaohong Duan[1]

[1]College of Finance and Economics, [2]College of Resources and Environmental Sciences, Gansu Agricultural University, Lanzhou, Gansu 730070, China. [3]Key Laboratory of Efficient Utilization of Water in Dry Farming of Gansu Province, Dryland Agriculture Institute, Gansu Academy of Agricultural Sciences, Lanzhou, Gansu, China. [4]Liaocheng Jiangbei Shuicheng Tourism Resort Management Committee, Liaocheng, Shandong, China.

Received: March 19, 2025; **accepted:** June 24, 2025.

Soils are prone to the deposition of trace metals, thereby posing potential health risks to humans. As Yongchang county in Gansu province, China is an important mining area, it is necessary to evaluate its heavy metal distribution pattern and pollution characteristics. During environmental governance audits, it was found that construction parties often used point data for evaluation, which was less representative. This study examined eight heavy metals including copper (Cu), zinc (Zn), nickel (Ni), lead (Pb), arsenic (As), chromium (Cr), vanadium (V), and cobalt (Co) in 1,266 cultivated topsoil samples from Yongchang county. The commonly used geographic information system (GIS)-based interpolation techniques including inverse distance weighting (IDW), local polynomial (LP), and radial basis functions (RBF), ordinary Kriging (OK), and empirical Bayesian Kriging (EBK) were employed for spatial distribution mapping of heavy metals. Single-factor index method combined with Nemerow pollution index was used to calculate the extent of heavy metal pollution. The results showed that EBK and LP consistently provided the most accurate predictions of heavy metals concentrations. The concentration of As was greater than the standard values of 25 mg/kg on some towns while other elements all showed lower values than National Soil Environmental Quality Standard. The Nemerow pollution index evaluation results demonstrated that the safety level, alert level, light pollution, and moderate pollution were 85.931%, 12.548%, 1.518%, and 0.004% of the region's arable land, respectively. It is of great significance to improve the efficiency of environmental auditing and make more intuitive judgments on the effectiveness of environmental governance.

Keywords: geographic Information system; soil; heavy metal; interpolation; spatial distribution; pollution assessment.

[*]**Corresponding authors:** Bo Dong and Dandan Du, College of Resources and Environmental Sciences, Gansu Agricultural University, Lanzhou, Gansu 730070, China. Email: ppleyuan@163.com (Dong B), dudandan19@mails.ucas.ac.cn (Du D).

Introduction

While naturally occurring in soils, the levels of heavy metals can rise to hazardous proportions due to both geological processes and human activities, posing risks to humans, flora, and fauna alike [1, 2]. Activities contributing to environmental impact encompass the extraction and processing of metals, combustion of fossil fuels, application of agricultural chemicals like fertilizers and pesticides, manufacture of batteries and various metal goods in industrial settings, as well as the management of sewage sludge and the disposal of urban waste [3]. Soil is

one of the most fundamental materials for human survival and development since environmental quality is directly related to human health and safety [4]. With rapid industrialization and modernization of China's development, heavy metals pollution in the soil's environment is increasingly becoming a prominent issue [5]. Pollution is seriously affecting economic development and biological health of the people [6]. Analyzing the distribution of heavy metals in the region and assessing the resulting pollution in soil can lay a scientific foundation for implementing soil environmental management and conservation strategies.

Evaluation of heavy metal pollution of soil has received a lot of research attention [7]. Geographic information system (GIS) technology has unique advantages in resource and environmental auditing, which can improve audit efficiency, accuracy, ensure data integrity, and facilitate audit evidence collection and verification. GIS has proven to be an effective tool for visualizing environmental contaminants [8]. Due to limitations of time and resources, the number of soil samples obtained from research sites for chemical analysis is usually limited, which leads to poor quality datasets and may undermine experimental results and conclusions [9]. To address any deficiencies in the cognitive design framework, it is possible to employ data interpolation to estimate values in areas that lack sampling by leveraging values from nearby observations [10]. Many more different types of studies have been carried out, in which a variety of deterministic and geostatistical interpolation techniques have been used to extract soil characteristics [11]. Among the common methods employed for mapping soil contamination, there are five key interpolation techniques including inverse distance weighting (IDW), local polynomial (LP), ordinary Kriging (OK), empirical Bayesian Kriging (EBK), and radial basis functions (RBF). Research conducted on trace elements in the soil of Matehuala, Mexico indicated that the most effective interpolation method was IDW [11, 12]. In contrast, Mohammad *et al*. demonstrated that cokriging and ordinary Kriging (OK) were more effective than the IDW approach in estimating the geographic spread of soil properties [13]. Bhunia *et al*. further demonstrated that the method of ordinary Kriging was more effective for estimating the spatial allocation of soil organic carbon (SOC) [14]. Saha *et al*. evaluated the efficiency of IDW, LP, and RBF techniques and reported that all these interpolation techniques had moderate accuracy in predicting the mean concentration of trace metals in soil [2].

Yongchang county in Gansu province, China is an emerging industrial area for heavy metal processing. Heavy metal contamination in Yongchang county is high because of large number of mining and industrial sewage. The soil heavy metal elements have strongly enriched the soil surface layer. Crops generally have excessive heavy metal content, particularly in the case of vegetables [15]. It has been observed that elevated concentrations of metals can directly lead to toxic effects such as the suppression of cytoplasmic enzymes and the impairment of cellular structures, which arise from oxidative stress. Observations have shown that the growth of plants on soils contaminated with heavy metals is adversely affected due to alterations in their physiological and biochemical functions. A persistent decrease in the growth of plants leads to lower yields, ultimately contributing to food shortages. Management and rehabilitation of soil environments are essential to provide a scientific framework for addressing heavy metal pollution, given its considerable effects on local agricultural practices and the ecological development of the region. Consequently, the importance of assessing soil impacted by heavy metals should not be underestimated. Nevertheless, research on the distribution over time and space, as well as the potential ecological risks associated with heavy metals in Yongchang county remains limited. The characteristics of soil in this contaminated region remain largely unexplored. This study conducted a large-scale survey in farmland for heavy metal contamination in topsoil of Yongchang county and characterized

the distribution of heavy metal concentrations by GIS based spatial distribution to estimate the pollution levels of heavy metals in farmland and assess the environmental and ecological risk of the area affected by the mining. The results would provide a scientific foundation for effective soil environmental management and conservation.

Materials and methods

Geology and hydrology of the study area

The study was conducted in Yongchang county, Gansu, China (101°04′–102°43′ E, 37°47′21″–38°39′58″ N), which comprises ten administrative towns, covering a total area of 7,439.27 square kilometers with a population of 243,000. The area is characterized by mountainous terrain, plains, Gobi Desert, and desert oasis with a minimum altitude of 1,452 meters, highest elevation of 4,442 meters, and an average altitude of 2,000 meters. Yongchang county has a temperate continental climate with cold winter and hot summer and an average annual temperature of 7℃, average annual rainfall of 183.2 mm, frost-free period of 134 days. It has an average annual sunshine of 2,884.2 hours with a sunshine rate of 65% and annual evaporation of 2,000.6 mm (Figure 1).

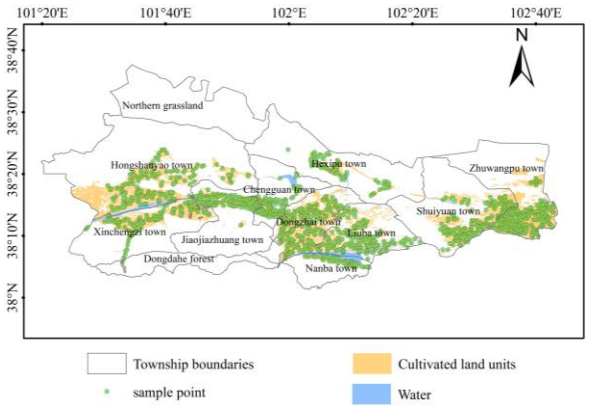

Figure 1. Spatial distribution of sampling points and cultivated land units in the study area.

Sample collection and examination

Sampling units were established at the village level by considering the unique characteristics of each community, variations in land use and soil types, and the distribution patterns of soil classifications relevant to the target species. Based on these factors, the locations of soil sampling points were further optimized. Upon finalizing the sampling units and location counts, geographic coordinates were mapped to reflect current land use patterns with each point representing the centroid of its respective unit. Two field sampling techniques were employed including topographic coordinate positioning using landmark features such as rivers, villages, valleys, and prominent structures for initial approximation followed by global positioning system (GPS) verification and scaled reference mapping involving fixed geographical reference points with measured distances cross-validated by GPS. In mountainous terrain, composite soil samples from 0 - 20 cm depth were collected from 15 points arranged in an S-pattern using vertical drilling, while gentle slopes employed a five-point sampling scheme to minimize spatial bias with all coordinates precisely recorded. Each of the 1,266 sampling sites was assigned a unique identifier documenting location name, coordinates, and collection date. The soil samples were air-dried in the laboratory at room temperature approximately 25℃, then ground by an agate mortar to pass through a 0.075 mm nylon sieve. Approximately 4.0 g of powder sample was squeezed under 40 tons of pressure for 20 seconds, creating a compressed specimen with a thickness of 4 mm and a diameter of 30 mm. The total contents of As, Co, Cr, Cu, Ni, Pb, V, and Zn were determined using X-MET7000 high-accuracy portable X-ray fluorescence spectrometry (Oxford Instruments, High Wycombe, England) with the limits of detection (LODs) of 1.600, 1.100, 1.200, 10.000, 1.000, 1.000, 1.400, and 2.000 mg/kg, respectively [16].

Assessment of agricultural land resource management units

For the identification of management units for crop cultivation, relevant maps of the 2019

administrative map, the land usage map from 2019 with a scale of 1:50,000, along with the 2008 soil map with a scale of 1:50,000 were utilized. The data were provided by Gansu Engineering Research Center for Smart Agriculture, Lanzhou, Gansu, China. A map indicating the primary management unit was then produced by combining the present land use map, the soil maps, and the administrative map.

Interpolation approaches for the spatial distribution of metals

(1) Inverse distance weighted (IDW)

The combination of multivariate statistical analysis with GIS in inverse distance weighting (IDW) is one of the most widely used deterministic interpolation methods in soil studies. IDW is based on the principle that the influence of known points decreases with distance, and the weights assigned are inversely proportional to the power of the distance. This approach provided larger weights to spatially adjacent points compared to those farther away, which aligned conceptually with the idea that closer observations were more relevant and was expressed as follows [17, 18].

$$Z = \sum_{i=1}^{n}(Z_i/d_i^p)/\sum_{i=1}^{n}(1/d_i^p) \quad (1)$$

where Z was the approximate value at the interpolation point. Z_i was the calculated value at point i. n was the total number of values obtained through interpolation. d_i was the distance between the interpolation point Z and the calculated value Z_i. p was the weighting power. In this study, IDW calculations were performed based on adjacent observation points, meaning that each known point adjusted independently of the others [16]. However, the predicted values were constrained within the range of the values used for interpolation. Since IDW was a weighted average method, the average predicted value could not exceed the maximum input value or fall below the minimum input value. Consequently, IDW might not capture extreme features such as ridges or valleys if these extrema were not present in the sampled data.

(2) Local polynomial (LP)

Local polynomial (LP) interpolation techniques have been utilized in meteorological research for over 50 years [19]. LP interpolation involves fitting a unique polynomial equation to localized regions based on observed values, the spatial extent of the region, the type of observation neighborhood, and the selected kernel function [20]. The goal of polynomial interpolation is to determine a polynomial function that best fits a specified set of observation points. While a global polynomial may cover the entire surface, it often fails to capture local variations effectively. The LP method addresses this issue by using localized polynomial fits, which can better accommodate natural variations in the data [21]. Within the local region, a weight is assigned to each observation point, which is typically achieved using a kernel function and determines the weight based on the distance between the observation point and the prediction point. In this study, several kernel functions including constant, exponential, Gaussian, Epanechnikov, quartic, and polynomial of degree 5 were evaluated as below.

$$Z_i = (1 - \frac{d_i}{R})p \quad (2)$$

where, Z_i was the average observation value at the i-th measurement point. d_i was the difference between the observation point and the prediction point. R was the considered neighboring region. p was the order of the polynomial function defined by the operator.

(3) Radial basis functions (RBF)

Radial basis function (RBF) interpolation, also known as spline interpolation, is a precise interpolation technique rooted in the principles of artificial neural networks (ANN) [20]. This technique includes five different basis functions as thin plate splines (TPS), splines with tension (ST), inverse multiquadrics (IMQ), completely regularized splines (CRS), and multiquadrics (MQ).

RBF interpolation uses these functions to generate predictions based on a specified region and the distances from known data points. Each prediction involves combining the effects of all relevant basis functions according to their influence on the prediction point as follows.

$$Z(x) = \sum_{i=1}^{m} a_i f_i(x) + \sum_{i=1}^{n} b_j \varphi(d_j) \quad (3)$$

$$\varphi(d) = \ln(\frac{cd}{2})^2 + E_1(cd)^2 + \gamma \quad (4)$$

where d was the difference between the estimated point and the observed point. c was the smoothing factor. E_1 was the modified Bessel function. γ was the Euler constant.

(4) Ordinary Kriging (OK)

Advanced geostatistical techniques such as Kriging represent a sophisticated category of interpolation methods. Geostatistical algorithms are not only proficient at generating predictive surfaces but also provide some indication of the reliability or efficiency of the predictions [22]. Kriging is widely recognized for its effectiveness in characterizing spatial variability, particularly in soil studies, and is valued for its ability to produce optimal and constructive estimates for unsampled areas. Ordinary Kriging (OK) is based on a statistical model that includes autocorrelation or statistical correlation between observed points [23]. Considering the spatial orientation relationship between known sample points and predicted points, different ordinary Kriging methods use different semivariogram models that include spherical semivariogram model, circular semivariogram model, exponential semivariogram model, Gaussian or normal distribution semivariogram model, linear semivariogram model with a sill. These models are used to define the spatial correlation structure in ordinary Kriging in this study. OK as a spatial interpolation predictor is expressed as a weighted sum of the following data.

$$Z(x) = \sum_{i=1}^{n} \lambda_i Z(x_i) \quad (5)$$

where Z(x) was the predicted value at location x. Z(xi) was the known value at sampled location x_i. λ_i was the weight assigned to each known value, determined based on the spatial correlation structure.

(5) Empirical Bayesian Kriging (EBK)

Empirical Bayesian Kriging (EBK) enhances traditional Kriging by integrating Bayesian statistical methods to optimize model parameters and improve prediction accuracy. Unlike conventional Kriging, which relies on a predefined semivariogram model, EBK adjusts the model through simulation and subset techniques to account for errors in the predicted semivariogram [24, 25]. It assumes that the predicted semivariogram accurately represents the spatial variability of the interpolation area, allowing for linear predictions with varying spatial dispersions [23]. EBK combines two geostatistical concepts of intrinsic random function Kriging (IRFK) [26], which addresses spatial data randomness and autocorrelation, and the linear mixed model (LMM) [27], which incorporates external trends. By integrating these approaches, EBK provides a unified computational model that improves stability and reliability in predictions. Additionally, EBK employs various semivariogram models such as power, linear, thin plate spline, exponential, and Whittle functions to define the spatial correlation structure, enhancing the accuracy and reliability of the interpolation results. Both IRFK and LMM models are fitted using the same procedure, combining these approaches into a single computational framework as below.

$$Z_i = y(s_i) + \varepsilon_i, i = \overline{1.....K} \quad (6)$$

where Z_i was the measured value at the observation point s_i . $y(s)$ was the Gaussian process at the location s under study. ε_i was the measurement error. K was the total number of measurements.

Accuracy evaluation

To allow for an evaluation of the various model performances, the collected data was

categorized into two sets with 80% of sampling points being used for model training and 20% of sample points for model validation against the predicted values generated using the reference interpolation methods that included IDW, LP, RBF, OK, and EBK to estimate the spatial distribution of heavy metals in agricultural soils. To determine the most suitable method, comparisons were conducted based on mean relative error (MRE), root mean square error (RMSE), and coefficient of variation (CV) according to various criteria. A minimal prediction error signified that the forecasting outcome was deemed both satisfactory and exemplary. In this research, it was deemed that a smaller CV value was regarded as the most favorable. CV serves as a statistical tool to quantify how data points are distributed relative to the average within a dataset. For soil samples, the CV served as an indicator of the extent of variability in data relative to the average concentration of metals present. An increased CV value signified that there was greater variability around the means. A lower value of CV was advantageous, indicating that the range of data points was closely grouped around the average. The following equations were for the metrics utilized to evaluate the precision of the predictions.

$$\mathrm{MRE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{z^*(x_i)-Z(x_i)}{Z(x_i)}\right| \quad (7)$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}[z^*(x_i)-Z(x_i)]^2}{n}} \quad (8)$$

$$\mathrm{CV} = \frac{\text{standard deviation of predicted values}}{\text{mean of predicted values}} \quad (9)$$

where $Z(x_i)$ was the observed value at point i. $z^*(x_i)$ was the predicted value at point i. n was the number of observed sample data points.

Environmental risk assessment methods and grades

This study utilized the single factor index method combined with Nemerow pollution index method to calculate the extent of heavy metal environmental risk as follows.

For single Factor Index method:

$$p_i = c_i / s_i \quad (10)$$

where Pi was single factor index of pollutant i (mg/kg). Ci was measured concentrations of pollutant i (mg/kg). Si was national soil environmental quality standards (mg/kg).

For Nemerow pollution index method:

$$\mathrm{Pc}=\sqrt{\left[\left(\max pi\right)^2+\left(\overline{pi}\right)^2\right]/2} \quad (11)$$

where Pc was Nemerow pollution index. max pi was the max value of single pollution index. pi was the average value of single pollution index. According to the actual situation of Yongchang county topsoil pH, values of National Soil Environmental Quality Standard (GB15618-2008) large than 7.5 were used as evaluation index [28], which were Cu (100 mg/kg), Zn (300 mg/kg), Ni (100 mg/kg), Pb (80 mg/kg), As (25 mg/kg), Cr (250 mg/kg), V (130 mg/kg), and Co (40 mg/kg). The grading criteria for single-factor pollution assessment were established as light pollution level (Pi < 1.0), moderate pollution level (1.0 ≤ Pi < 2.0), strong pollution level (2.0 ≤ Pi < 3.0), and very strong pollution level (Pi ≥ 3.0). For the comprehensive environmental risk assessment, the Nemerow pollution index classification followed the standard methodology with five pollution grades defined as safety level (Pc < 0.7), alert level (0.7 ≤ Pc < 1.0), light pollution (1.0 ≤ Pc < 2.0), moderate pollution (2.0 ≤ Pc < 3.0), and heavy pollution (Pc ≥ 3.0) [29].

Results

Statistical analysis of heavy metal content

The results showed that the average concentrations of 8 heavy metals in the arable layer soil of Yongchang county didn't exceed the value of National Soil Environmental Quality Standard (GB15618-2008) (Table 1). The single factor index of them was less than 1, indicating

Table 1. Descriptive statistics of cultivated land arable layer soil Cu, Zn, Ni, Pb, As, Cr, V, and Co contents in Yongchang county.

Element	Mean (mg/kg)	SD	Skewness	Kurtosis	Single factor index	Variation coefficient	Environ standards (mg/kg)	Samples
Cu	18.53	8.92	0.41	0.28	0.19	0.48	100	1,226
Zn	59.91	14.8	2.9	33.28	0.20	2.05	300	1,266
Ni	32.47	11.81	0.2	0.03	0.32	0.36	100	1,263
Pb	18.53	19.2	2.76	14.94	0.23	1.03	80	600
As	20.7	9.75	0.23	0.78	0.83	0.47	25	1,101
Cr	61.46	31.88	0.26	-0.29	0.24	0.52	250	1,288
V	80.82	49.9	0.61	0.001	0.62	0.61	130	1,122
Co	13.86	8.47	0.75	0.41	0.34	0.61	40	1,106

light pollution level. The descending order of the degree of pollution was As, V, Co, Ni, Cr, Pb, Zn and Cu, while the average values of the degree of contamination were 0.83, 0.62, 0.34, 0.32, 0.24, 0.23, 0.20 and 0.19, respectively. Among them, the highest average pollution levels were obtained with As and the lowest with Cu. The minimum and maximum values reflected the concentration range of heavy metals, which presented an order of V, Zn, Pb, Cr, As, Ni, Cu and Co. The trend of median concentrations of heavy metals was toward minimum direction with less than half of the median value of the sample. The coefficient of variation of the heavy metals followed a descending order of Zn, Pb, V, Co, Cr, Cu, As and Ni with the CV values as 2.05, 1.03, 0.61, 0.61, 0.52, 0.48, 0.47 and 0.36, respectively. Zinc and lead had their CVs exceeding 1, and the lowest value of 0.36 was observed with Ni. The high CV found in the heavy metals could be explained by the presence of the metals in the soils due to outside interference, which was related to the chemical industry in Yongchang county. The heavy metals showed a higher degree of movement towards the right side with Zn and Pb having the highest skewness, which reflected the industrial development or other causes that made heavy metal content in some areas more than the original level. From the peak values, Zn and Pb had deeper distribution than normal. As and Co had slightly more prominent than normal distribution. Zinc and lead contents experienced dramatic high changes, mainly because Yongchang county is an important non-ferrous metal smelting and processing area. Drought, water shortage, and conventional sewage irrigation increased the accumulation of heavy metals in the soil. The cultivated land, industrial and mining areas were found at different distances in Yongchang county, which might be responsible for the dramatic changes in the phenomenon of heavy metal contents.

Comparison of five methods for interpolation

Overall, the various interpolation methods estimated the average levels of trace metals in soil with a fair degree of precision (Figure 2). An evaluation of five interpolation methods across various functions revealed the optimal technique for the eight selected trace metals. The criteria established favored the MRE and CV values that were the closest to 0 with a strong emphasis on achieving the lowest RMSE. An evaluation was conducted on three interpolation methods based on deterministic approaches (IDW, LP, and RBF) alongside two methods rooted in geostatistics (OK and EBK) to determine the most effective technique for interpolation. In the area studied, the concentrations of eight distinct soil metals were estimated using various interpolation methods with EBK employing the Whittle semivariogram for chromium, LP utilizing an exponential semivariogram for vanadium, EBK with an exponential semivariogram for cobalt and copper, OK with a circular semivariogram for nickel, LP applying a disjoint parabolic function for zinc and lead, and RBF implementing a spline with tension function for arsenic. The results indicated that EBK and LP emerged as the most effective techniques in this research.

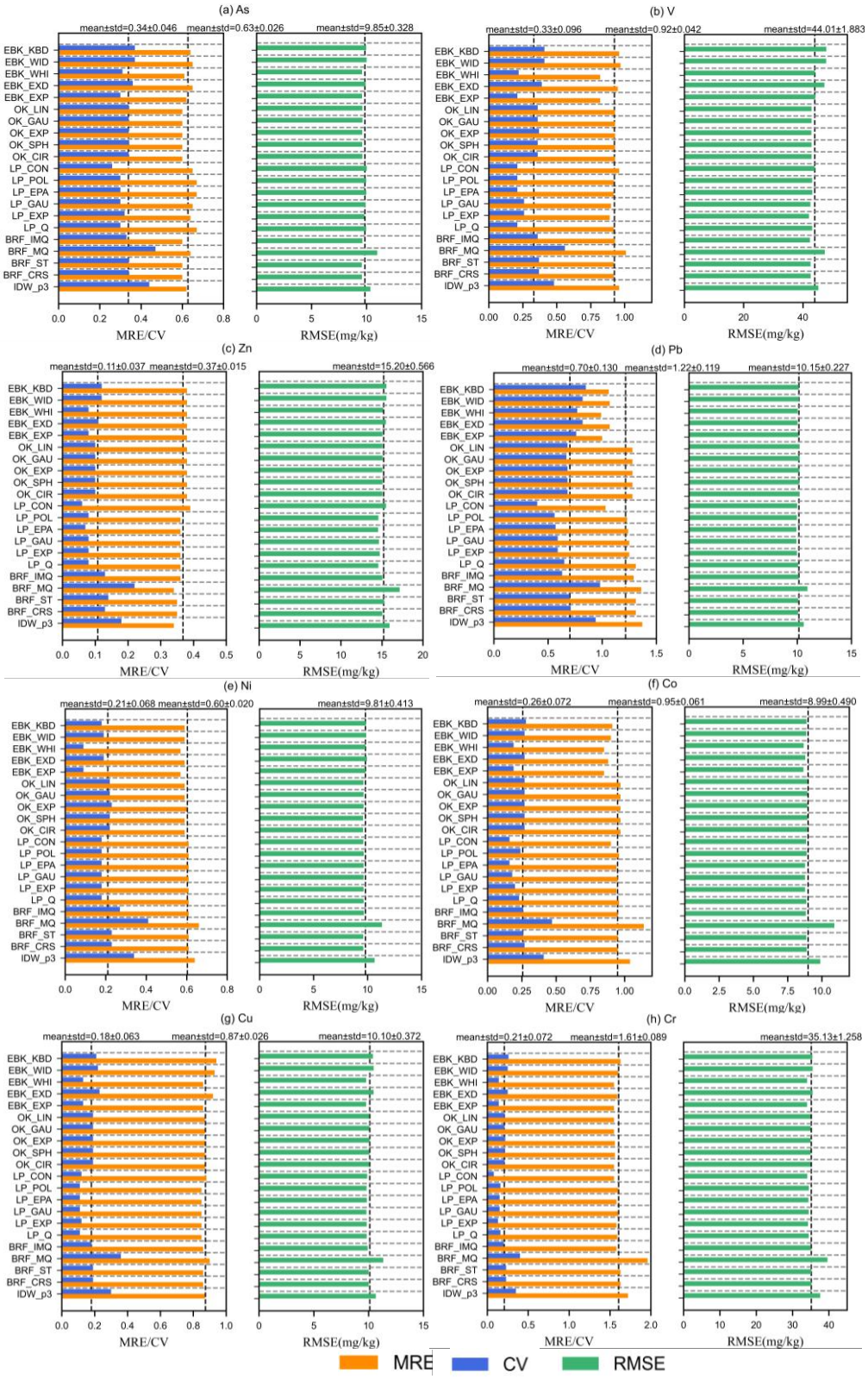

Figure 2. Performance of the five interpolation techniques under different functions. EBK_KBD, EBK_WID, EBK_WHI, EBK_EXD, EBK_EXP were EBK with K_Bessel_Detrended, Whittle_Detrended, Whittle_Detrended, Exponential_Detrended, Exponential function. OK_LIN, OK_GAU, OK_EXP, OK_SPH, OK_CIR were OK with linear, Gaussian, exponential, spherical, circular function. LP_CON, LP_POL, LP_EPA, LP_GAU, LP_EXP were LP with constant, polynomial of degree 5, Epanechnikov, Gaussian, exponential function. BRF_IMQ, BRF_MQ, BRF_ST, BRF_CRS were BRF with Inverse_Multiquadric_Function, Multiquadrics, Spline_With_Tension, Completely_Regularized_Spline. IDW_p3 was IDW with power 3.

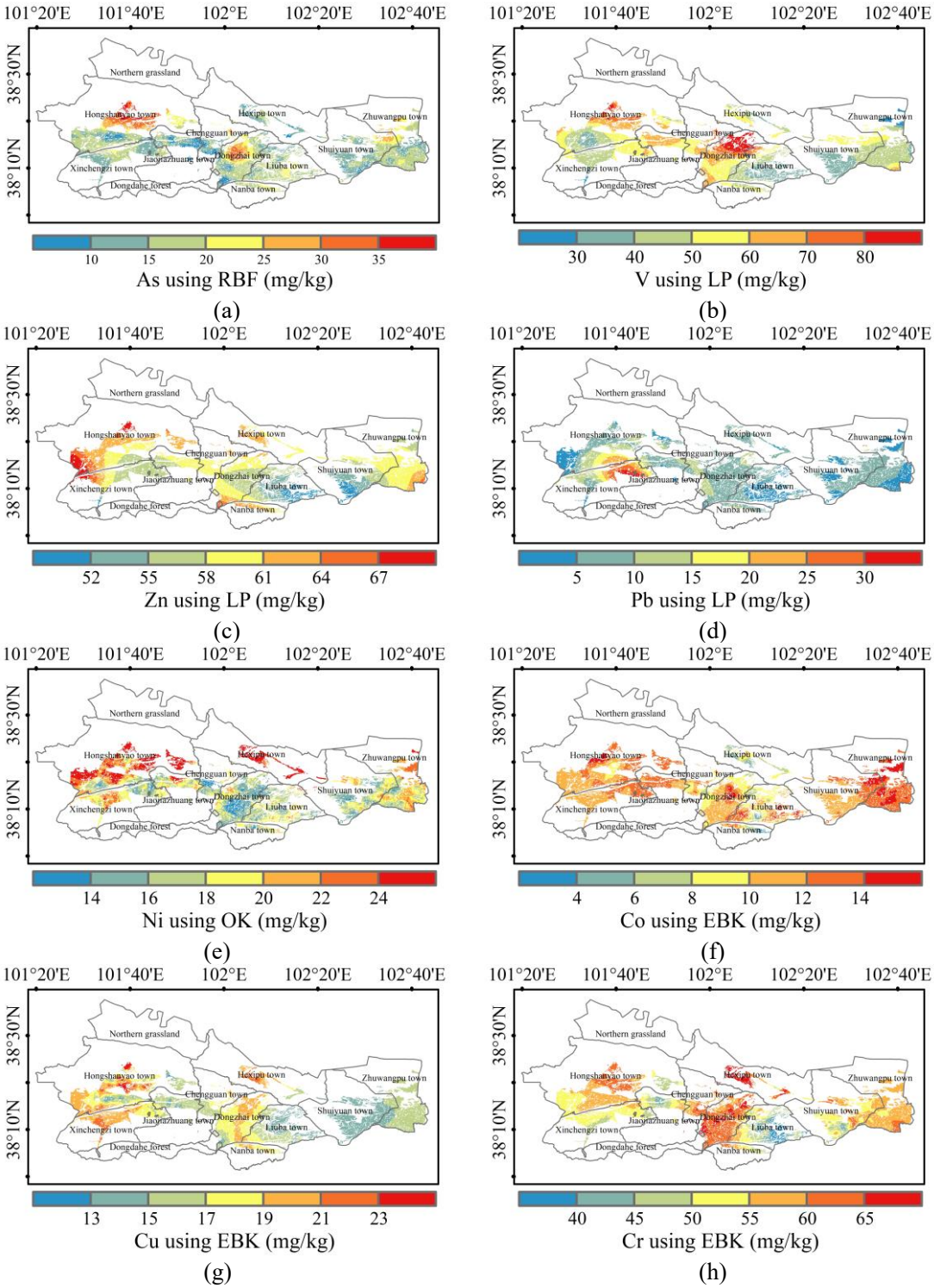

Figure 3. Spatial distribution of 8 elements contents of arable layer soil in Yongchang county. (**a**). As using RBF. (**b**). V using LP. (**c**). Zn using LP. (**d**). Pb using LP. (**e**). Ni using OK. (**f**) Co using EBK. (**g**). Cu using EBK. (**h**). Cr using EBK.

Spatial distribution characters of heavy metals

The spatial variability of heavy metals leads to the formation of a regular spatial distribution pattern, which is a specific manifestation of the spatial variability of heavy metals [30]. Geostatistical methods are usually used to insert

spatial data to form maps, which can more intuitively and accurately describe the spatial distribution of heavy metals in soil, such as shape, size and location [31]. It has benefits to well understand the situation of spatial distribution pattern of heavy metals. The most suitable interpolation methods were used to determine the spatial distribution characters of heavy metals of the cultivated land based on sampling points and the actual measurement data of Yongchang county. By using function of ArcGIS Geostatistical Analyses, interpolation combined with soil environmental quality standards (GB15618-2008) [28] and the single factor index raster data classification was applied to obtained As, V, Co, Ni, Pb, Cu, Zn and Cr concentration levels map of Yongchang county (Figure 3). The result illustrated that the levels of As exceeded the permissible limit of 25 mg/kg in Hongshanyao Town, Dongzhai Town, as well as in smaller regions of Zhuwangpu Town and Shuiyuan Town (Figure 3a), which indicated that human activities in the region had negatively influenced the concentrations of this element. Activities such as ore processing, the generation of slag, and the accumulation of metal particulates in the soil around mining areas might contribute to the elevated levels of arsenic found in certain locations. The levels detected for other elements remained below the thresholds established by the National Soil Environmental Quality Standard (GB15618-2008). The distribution pattern of Pb resembled an island configuration, featuring two segments that declined outward from a central peak concentration area located near Xinchengzi Town (Figure 3d). V exhibited two relatively high concentration areas located in the vicinity of Hongshanyao Town and Dongzhai Town (Figure 3b). Co exhibited analogous patterns in its spatial distribution (Figure 3f). The distribution of Zn and Ni reflected a comparable trend, exhibiting elevated levels in Hongshanyao Town, Xinchengzi Town, Hexipu, and Zhuwangpu Town (Figure 3c and 3e). Three notable areas of elevated concentrations for Co and Cr were identified in Hongshanyao, Dongzhai, and Zhuwangpu Town (Figure 3f and 3h).

Assessment of environmental risk

Furthermore, 8,411 patches were obtained by assigning heavy metals concentration raster data value on arable land. Nemerow pollution index method was used to calculate the value of each patch, and the patch was classified according to grades of environmental risk assessment. The Nemerow pollution index of 7,453 patches of arable land was less than 0.7, consisting of an area of 8,0274.19 hm^2 with safety level. There were 907 patches' Nemerow pollution index ranging from 0.7 to 1.0, consisting of an area of 11,721.7 hm^2 as the pollution risk alert level, while the light pollution level consisted of 50 patches covering an area of 1,417.83 hm^2. The cultivated lands safety level, alert level, light pollution, and moderate pollution levels accounted for 85.931%, 12.548%, 1.518%, and 0.004% of the region's arable land, respectively. The area of pollution accounted for 1.522% of total arable land in Yongchang county, which constituted an area of 1,421.47 hm^2. The environmental risk level and space distribution of cultivated land arable layer soil in Yongchang county were illustrated in Figure 4. Most of cultivated land belonged to safety level. Alert level of cultivated land distributed around the urban periphery of Dongzhai Town, Shuiyuan Town, and Zhuwangbu Town, where the construction land and industrial and mining land were located. Light pollution and moderate pollution levels were mainly distributed in the urban periphery of Hongshanyao Town. The pollution of these regions might be caused by human activities including discarded garbage, fertilizer, and transportation. Land irrigation of wastewater and application of sludge to land and city dust precipitation were the main causes for the heavy metal pollution [15]. Zonal and spatial correlation could be indicative of the main reasons of As, V, Pb, and Co pollution. The light pollution and moderate pollution level land accounted for the proportion in the research area to be small, while large alert level arable land area was dominated responsible for the high environmental risk. Pollution existed in almost all areas of human activity in Yongchang county. Heavy traffic, vigorous human activities, and low

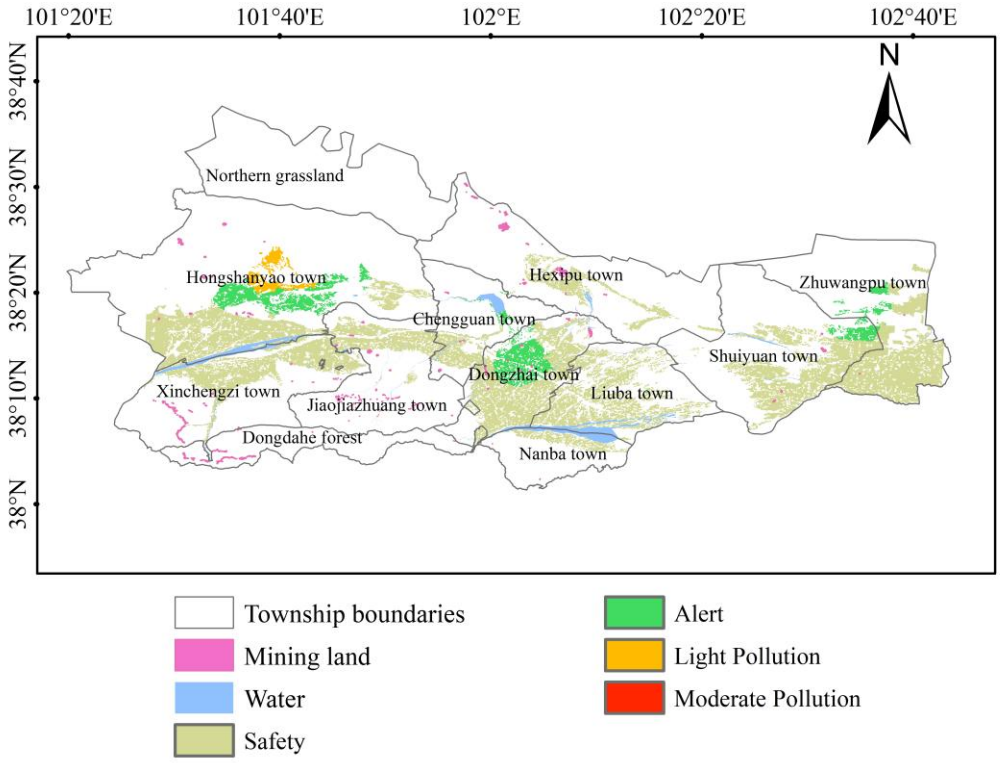

Figure 4. Environmental risk level and space distribution of cultivated land arable layer soil in Yongchang county.

vegetation cover might be responsible for high levels of soil As, V, Pb, and Co contamination [32].

Discussion

With the help of GIS technology for point data obtained during agri-environmental auditing, spatial interpolation technique was used in this research to predict the unsampled arable land and infer the spatial distribution of each element. Spatial interpolation technique as a fast and intuitive graphical traceability method can provide accurate spatial prediction of element content, matrix factor assignment, and so on to infer the potential sources of the elements [33]. GIS-based interpolation techniques including IDW, LP, and RBF, OK, EBK were used to map the spatial distributions of the elements. Each method had its own advantages and limitations, so choosing the best interpolation technique was crucial for accurate and reliable spatial mapping. The accuracy of interpolation depended on the precision of the definition of boundaries and contaminated areas [34]. Several studies have been conducted on the effectiveness of spatial interpolation techniques. However, the conclusions were not consistent. Some of them found that the Kriging method was superior to the IDW method [13], while others demonstrated that the Kriging method was comparable to other techniques [35]. To determine the best method for describing the distribution patterns of heavy metals in various environmental situations, several interpolation methods including their parameterization and validation were compared and evaluated in this study. The results showed that EBK and LP consistently provided the most accurate prediction of heavy metal concentrations in Yongchang county, which was consistent with the results of Kravchenko *et al* that the OK prediction accuracy was higher than that of the IDW in most cases [36]. Interpolation methods are different for different types of landforms. In general, simple Kriging interpolation is more accurate for areas dominated by mountains. Kriging or tensile spline

interpolation is more accurate for areas dominated by plains and terraces, while pan-Kriging interpolation is more accurate for plains or basins of comparable size distribution, hills, and mountains [37]. In China, the scientists mainly adopt methods such as the single-factor index method, Nemero comprehensive pollution index method, pollution load index method, potential ecological risk index method, and land accumulation index method for the evaluation of heavy metal enrichment in agricultural soils [33]. The core principle of these diverse evaluation methods lies in comparing the concentration of pollutants in the soil with the soil environmental quality standards or the background value of the soil environment, thereby reflecting the degree of pollutant enrichment in the soil. Nemero index method is a common means of comprehensive assessment of environmental pollutant status, which considers the extreme value of the multi-factor environmental quality index and can comprehensively reflect the degree of contamination of the factors in the environment of the study area and the overall environmental quality changes [38]. In this study, Nemero integrated pollution index method was employed to assess the heavy metal enrichment condition of the farmland soil within the study area. The evaluation outcomes consistently demonstrated that a minor portion of the regional soil presented a relatively high level of soil pollution risk. Specifically, the soil with a medium pollution level accounted for an area proportion of 0.004%. Subsequently, an analysis of the pollution causes in the more severely polluted areas was conducted based on the spatial distribution map of the evaluation indices and discovered that there existed over 60 mineral deposits in Yongchang county, predominantly consisting of iron, copper, nickel, lead, zinc, cobalt, and other minerals. Among them, nickel reserves accounted for 80% of China's nickel reserves, while copper reserves were more than 3 million tons, ranking 2nd in the country. Hongshanyao Town is located under Yanzhi Mountain with high content of zinc, nickel, cobalt, and copper. Agricultural soils with moderate and light pollution risk were mainly distributed around the mining site, and there was an enrichment condition of heavy metals in soil in the dense mining area. By applying GIS technology to study the distribution and pollution status of heavy metals in Yongchang county, strong technical support has been provided in key audit steps such as data collection, audit suspicion mining, and audit suspicion verification. Meanwhile, geographic information technology can accurately display the characteristics and spatial distribution of geographic information, and its powerful visualization analysis function has obvious advantages over traditional auditing methods. It also has good adaptability to various resource and environmental audits. In the future, according to the characteristics of resources and environment audit, further integration of geographic information technology combined with artificial intelligence, the internet, and other technologies can provide strong technical support for achieving full audit coverage.

Acknowledgements

This research was funded by the National Key Research and Development Plan of China (Grant No. 2021YFD190070406), Gansu Provincial Agricultural Science and Technology Special Project (Grant No. GNKJ-2021-32), Gansu Agricultural University's Project and Lanzhou Science and Technology Plan Project (Grant No. 2024-8-39).

References

1. Tchounwou PB, Yedjou CG, Patlolla AK, Sutton DJ. 2012. Heavy metal toxicity and the environment. In: Luch A (ed) Molecular, Clinical and Environmental Toxicology, Volume 3: Environmental Toxicology. Springer Basel, New York, pp 133-164.
2. Saha A, Gupta BS, Patidar S, Martínez-Villegas N. 2023. Optimal GIS interpolation techniques and multivariate statistical approach to study the soil-trace metal(loid)s distribution patterns in the agricultural surface soil of Matehuala, Mexico. J Hazard Mater Adv. 9:100243.
3. Tong Y, Yue T, Gao J, Wang K, Wang C, Zuo P, *et al*. 2020. Partitioning and emission characteristics of Hg, Cr, Pb, and As

among air pollution control devices in Chinese coal-fired industrial boilers. Energy Fuels. 34(6):7067-7075.
4. Chai L, Wang Y, Wang X, Ma L, Cheng Z, Su L, *et al*. 2021. Quantitative source apportionment of heavy metals in cultivated soil and associated model uncertainty. Ecotoxicol Environ Saf. 215:112150.
5. Zhao FJ, Ma Y, Zhu YG, Tang Z, McGrath SP. 2015. Soil contamination in China: Current status and mitigation strategies. Environ Sci Technol. 49(2):750-759.
6. Wen M, Ma Z, Gingerich DB, Zhao X, Zhao D. 2022. Heavy metals in agricultural soil in China: A systematic review and meta-analysis. Eco-Environ Health. 1(4):219-228.
7. Lv J, Liu Y, Zhang Z, Dai J, Dai B, Zhu Y. 2015. Identifying the origins and spatial distributions of heavy metals in soils of Ju country (Eastern China) using multivariate and geostatistical approach. J Soils Sediments. 15(1):163-178.
8. Fischer A, Lee MK, Ojeda AS, Rogers SR. 2021. GIS interpolation is key in assessing spatial and temporal bioremediation of groundwater arsenic contamination. J Environ Manage. 280:111683.
9. Liu CW, Jang CS, Liao CM. 2004. Evaluation of arsenic contamination potential using indicator kriging in the Yun-Lin aquifer (Taiwan). Sci Total Environ. 321(1-3):173-188.
10. Tobler WR. 1970. A computer movie simulating urban growth in the Detroit region. Econ Geogr. 46(2):234-240.
11. Saha A, Gupta BS, Patidar S, Martínez-Villegas N. 2022. Spatial distribution based on optimal interpolation techniques and assessment of contamination risk for toxic metals in the surface soil. J South Am Earth Sci. 115:103763.
12. Saha A, Gupta BS, Patidar S, Martínez-Villegas N. 2022. Spatial distribution and source identification of metal contaminants in the surface soil of Matehuala, Mexico based on positive matrix factorization model and GIS techniques. Front Soil Sci. 2:1041377.
13. Mohammad ZM, Taghizadeh R, Akbarzadeh A. 2010. Evaluation of geostatistical techniques for mapping spatial distribution of soil pH, salinity and plant cover affected by environmental factors in Southern Iran. Not Sci Biol. 2(4):92-102.
14. Bhunia GS, Shit PK, Maiti R. 2018. Comparison of GIS-based interpolation methods for spatial distribution of soil organic carbon (SOC). J Saudi Soc Agric Sci. 17(2):114-126.
15. Rai PK, Lee SS, Zhang M, Tsang YF, Kim KH. 2019. Heavy metals in food crops: Health risks, fate, mechanisms, and management. Environ Int. 125:365-385.
16. Dai LH, Liu XW, Wang D, Liu Y, Fang K, Jiang M. 2011. Application of polarized energy dispersive X ray fluorescence spectrometry in soil environmental determination. Chin J Spectrosc Lab. 28(3):836-841.
17. Qi Z, Gao X, Qi Y, Li J. 2020. Spatial distribution of heavy metal contamination in mollisol dairy farm. Environ Pollut. 263:114621.
18. Hou D, O'Connor D, Nathanail P, Tian L, Ma Y. 2017. Integrated GIS and multivariate statistical analysis for regional scale assessment of heavy metal soil contamination: A critical review. Environ Pollut. 231:1188-1200.
19. Gilchrist B, Cressman GP. 1954. An experiment in objective analysis. Tellus. 6(3):309-318.
20. Johnston K, Hoef JMV, Krivoruchko K, Lucas N. 2001. Using ArcGIS geostatistical analyst. ESRI Press, Redlands, pp 1-306.
21. Liao Y, Li D, Zhang N. 2018. Comparison of interpolation models for estimating heavy metals in soils under various spatial characteristics and sampling methods. Trans GIS. 22(2):409-434.
22. Oliver MA, Webster R. 1990. Kriging: a method of interpolation for geographical information systems. Int J Geogr Inf Syst. 4(3):313-332.
23. Ghosh M, Pal DK, Santra SC. 2020. Spatial mapping and modeling of arsenic contamination of groundwater and risk assessment through geospatial interpolation technique. Environ Dev Sustain. 22(4):2861-2880.
24. Krivoruchko K, Gribov A. 2019. Evaluation of empirical Bayesian kriging. Spat Stat. 32:100368.
25. Gribov A, Krivoruchko K. 2020. Empirical Bayesian kriging implementation and usage. Sci Total Environ. 722:137290.
26. Allard D. 2012. Book review: Geostatistics: Modeling spatial uncertainty by Jean-Paul Chilès and Pierre Delfiner. Math Geosci. 45(3):377-380.
27. Coburn TC. 2006. Book review: Statistical methods for spatial data analysis by Oliver Schabenberger and Carol A. Gotway. Math Geol. 38(4):511-513.
28. Liu YL, Zhang LJ, Han XF, Zhuang TF, Shi ZX, Lu XZ, *et al*. 2012. Spatial variability and evaluation of soil heavy metal contamination in the urban-transect of Shanghai. Environ Sci. 33(2):599-605.
29. Guo Y, Huang M, You W, Cai L, Hong Y, Xiao Q, *et al*. 2022. Spatial analysis and risk assessment of heavy metal pollution in rice in Fujian Province, China. Front Environ Sci. 10:1082340.
30. Sulieman MM, Kaya F, Keshavarzi A, Hussein AM, Al-Farraj AS, Brevik EC, *et al*. 2024. Spatial variability of some heavy metals in arid harrats soils: Combining machine learning algorithms and synthetic indexes based-multitemporal Landsat 8/9 to establish background levels. Catena. 234:107579.
31. Khodoli Zangeneh D, Amanipoor H, Battaleb-Looie S. 2023. Evaluation of heavy metal contamination using cokriging geostatistical method (case study of Abteymour oilfield in southern Iran). Appl Water Sci. 13(8):200.
32. Huang HN, Nan ZR, Hu XN, Liu XW, Li Y, Ding HX. 2009. Spatial distributions of heavy metals and assessment of potential ecological risk in Jinchang urban area. Adm Tech Environ Monit. 21(1):30-34.
33. Dong B, Zhang R, Gan Y, Cai L, Freidenreich A, Wang K, *et al*. 2019. Multiple methods for the identification of heavy metal sources in cropland soils from a resource-based region. Sci Total Environ. 651:3127-3138.
34. Saha A, Sen Gupta B, Patidar S, Hernández-Martínez JL, Martín-Romero F, Meza-Figueroa D, *et al*. 2024. A comprehensive study of source apportionment, spatial distribution, and health risks assessment of heavy metal(loid)s in the surface soils of a semi-arid mining region in Matehuala, Mexico. Environ Res. 260:119619.
35. Yanto, Apriyono A, Santoso PB, Sumiyanto. 2022. Landslide susceptible areas identification using IDW and Ordinary Kriging interpolation techniques from hard soil depth at middle

western Central Java, Indonesia. Nat Hazards. 110(3):1405-1416.

36. Kravchenko A, Robertson GP. 2007. Can topographical and yield data substantially improve total soil carbon mapping by regression kriging? Agron J. 99(1):12-17.
37. Bogunovic I, Mesic M, Zgorelec Z, Jurisic A, Bilandzija D. 2014. Spatial variation of soil nutrients on sandy-loam soil. Soil Tillage Res. 144:174-183.
38. Tomczyk P, Wdowczyk A, Wiatkowska B, Szymańska-Pulikowska A. 2023. Assessment of heavy metal contamination of agricultural soils in Poland using contamination indicators. Ecol Indic. 156:111161.