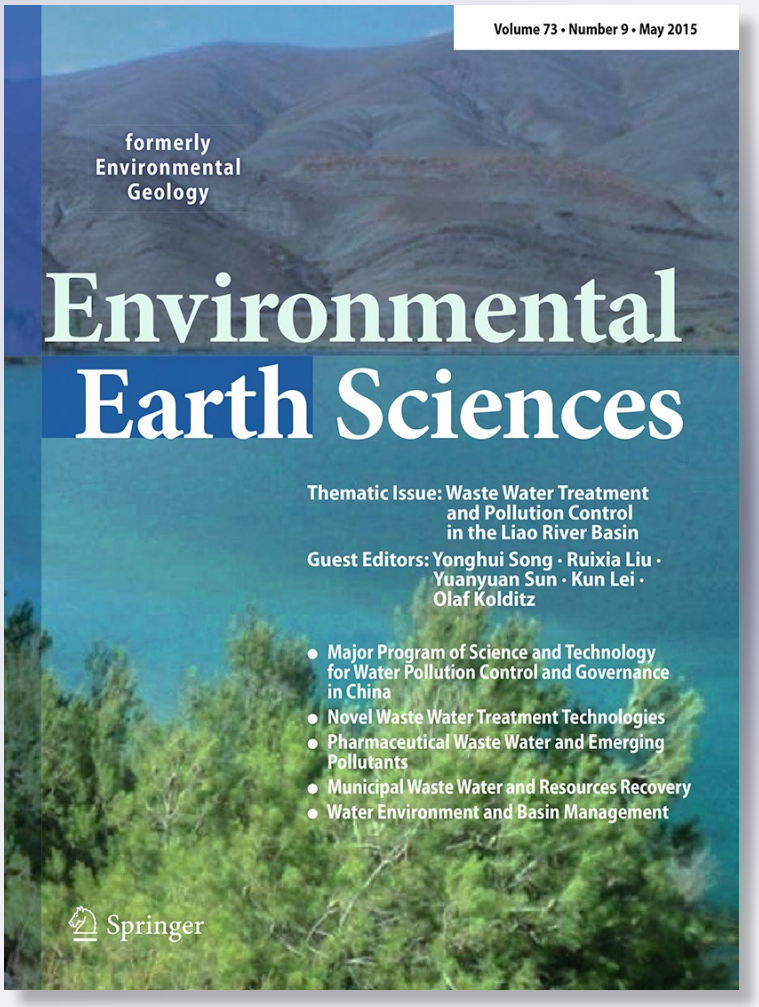
Geostatistical and multivariate analysis of soil heavy metal contamination near coal mining area, Northeastern India

S. K. Reza, Utpal Baruah, S. K. Singh & T. H. Das

Environmental Earth Sciences

ISSN 1866-6280
Volume 73
Number 9

Environ Earth Sci (2015) 73:5425-5433
DOI 10.1007/s12665-014-3797-1



Springer

Environ Earth Sci (2015) 73:5425–5433
DOI 10.1007/s12665-014-3797-1

ORIGINAL ARTICLE

Geostatistical and multivariate analysis of soil heavy metal contamination near coal mining area, Northeastern India

S. K. Reza · Utpal Baruah · S. K. Singh · T. H. Das

Received: 19 March 2014 / Accepted: 11 October 2014 / Published online: 22 October 2014

Abstract The spatial distribution and hazard assessment of heavy metals in the soils of surrounding agricultural fields affected by mine drainage of Ledo coal mining area of Tinsukia district, Assam, India, were investigated using statistics, geostatistics and geographic information system techniques. The amounts of Cr, Cd, Ni and Pb were determined from 83 soil samples collected within the contaminated area. The maps based on ordinary kriging showed that high concentrations of heavy metals were located in the low-lying paddy field and near coal mining site. Indicator kriged probability maps were prepared based on the concentrations to exceed permissible limit (MPL). It was seen that more than 95 % of the studied area has a higher than 50 % probability to exceed the MPL value of Pb and Ni. Multivariate statistical analyses and principal component analysis suggest that Cd and Pb are derived from anthropogenic sources, particularly coal mining activities, whereas Ni and Cr are derived from lithogenic and/or anthropogenic sources.

S. K. Reza (✉) · T. H. Das
National Bureau of Soil Survey and Land Use Planning, Sector-II, DK-Block, Salt Lake, Kolkata, West Bengal, India
e-mail: reza_ssac@yahoo.co.in

T. H. Das
e-mail: thdas@rediffmail.com

U. Baruah
National Bureau of Soil Survey and Land Use Planning, Jamuguri Road, Jorhat, Assam, India
e-mail: ubaruah2@rediffmail.com

S. K. Singh
National Bureau of Soil Survey and Land Use Planning, Nagpur, Maharashtra, India
e-mail: skcssri@gmail.com

Keywords Heavy metals · Geostatistics · Risk assessment · Principal component · Multivariate analysis

Introduction

Coal plays an important role in energy generation, and ~27 % of the world's energy consumption originates from the incineration of coal (Bhuiyan et al. 2010). Underground and open pit coal exploitation includes a phase development in mine and removal of surrounding rocks, which are low in coal content (<30 %) and often contain iron sulphide minerals (Bhuiyan et al. 2010). During the process of opencast and underground coal mining, a variety of rock types with different compositions are exposed to atmospheric conditions and undergo accelerated weathering. These materials are often deposited nearby as mine waste rocks and mine dust. Mining can be a significant source of metal contamination the environment as acid mine drainage (AMD), which usually occurs at coal mining sites in the world, representing serious environmental problems. AMD can occur during the exploitation of coal and coal-bearing minerals and ore bodies containing acid forming metal sulphides such as pyrite (FeS_2) (Adriano 2001; Madejón et al. 2002). The oxidation of such sulphides exposed to atmospheric oxygen during or after mining activities generates acidic waters with high concentrations of sulphate as well as dissolved iron and heavy metals. The low pH may cause further dissolution of local country rock and leaching of additional metals into water (Tabaksblat 2002), thereby adversely impacting on aquatic life and the surrounding vegetation (Cherry et al. 2001).

With growing public concern throughout the world over health hazards caused by polluted agricultural products, many studies have been conducted on metal and metalloid

contamination in soils, water and sediments from metalliferous mines (Lee 2006; Chopin and Alloway 2007; Anawar et al. 2008). It is known that serious systemic health problems can develop as a result of excessive accumulation of dietary heavy metals such as Cd, Cr and Pb in the human body (Petrisor et al. 2004). Heavy metals are of great concern in soil pollutants because they can threaten the health of human beings and animals through the food chain (Ladwani Kiran et al. 2012).

Risk estimation involves calculation of risk in affected areas and provides valuable information regarding feasible rehabilitation options. The methodology is mainly based on the principle "source—pathway—target". Risk is better assessed if quantitative techniques are used to account for spatial and temporal variations. A probabilistic assessment takes into account variability of parameters and uncertainty in measurement (Korre et al. 2002). Geostatistics is extensively used to assess the level of soil contamination and estimation of risk in contaminated sites by preserving the spatial distribution and uncertainty of the estimates. In addition, geostatistics and GIS provide useful tools for the study of spatial uncertainty and hazard assessment (McGrath et al. 2004; Komnitsas and Modis 2006; Reza et al. 2013).

The calculation of risk takes into account generic standards (target and intervention values) that are used to assess soil quality and classify soils as clean, slightly, moderately or heavily contaminated. The target values are protective levels and indicate the desired soil quality while the intervention values are indicative of serious contamination (Swartjes 1999). The assessment of risk for the population is a much more complex procedure and requires the establishment of human toxicological and eco-toxicological intervention values as well as exposure rates over various periods. A generic methodology that combines quantitative probabilistic human health risk assessment and spatial geostatistical methods has been recently proposed (Gay and Korre 2006). This methodology enables the calculation of risk to human health from exposure to contaminated land in a manner that preserves its spatial distribution and provides a measure of uncertainty in the assessment.

In the northeastern India, coalfields are confined particularly in the Tinsukia district of upper Assam. Coal mining activities in these areas have been in operation since 1882. Most of the coal fields now been closed due to declining of production while the collieries of Tikka, Borgolai, Ledo, Tipang and Namdang of Makum coalfield have so far produced more than 25 Mt of coal out of the reserve estimated at 130 Mt up to a depth of 300 m (Nesa and Azad 2008). These fields are still under operation in full swing. Our study investigates the extent of contamination of heavy metals (Cr, Cd, Ni and Pb) in soil by ledo coal mine using geostatistics and GIS techniques to reveal the spatial distribution patterns and provides a basis for hazard assessment.

Materials and methods

Study area

The study was carried out near the Ledo coal mining area of Tinsukia district, Assam, north-eastern India, extended between 27°17′05″ to 27°20′45″ N latitude and 95°39′35″ to 95°44′53″ E longitude covering an area of 1,968 ha (Fig. 1). The climate is humid subtropical. The average annual rainfall ranges between 2,000 and 2,500 mm with maximum rainfall during July–September. The climate is moderately warm during summer but cold in winter. Mean monthly minimum and maximum temperatures were 7 and 36 °C, respectively.

Soil sampling and analysis

A total of 83 surface soil samples were collected from a depth of 0–25 cm (plough layer) using a square 500 × 500 m grid (Fig. 1). The total area sampled was nearly 20 km^2, corresponding to a sampling density of four samples km^{-2} and covering not only the waste disposal site, but also the surrounding cultivated areas with the help of a hand-held global positioning system. Soil samples were air-dried and ground to pass through a 2-mm sieve. A combined glass calomel electrode was used to determine the pH of aqueous suspension (1:2.5 soil:solution ratio) (Richards 1954). Organic carbon was determined by the Walkley and Black (1934) method. Digestion of 0.50 g samples was performed with concentrated HNO_3, HF and $HClO_4$ in a microwave digester (model Start D, Milestone). Subsequently, the total concentration of heavy metals was determined by a Shimadzu AA6300 atomic absorption spectrophotometer.

Geostatistical analysis based on GIS

Spatial interpolation and GIS mapping techniques were employed to produce spatial distribution and risk assessment maps for the four investigated heavy metals, and the software used for this purpose was ArcGIS v.9.3 (ESRI Co, Redlands, USA). In ArcGIS, kriging can express the spatial variation and allow a variety of map outputs and at the same time minimize the errors of predicted values. Moreover, it is very flexible and allows users to investigate graphs of spatial autocorrelation. Kriging, as applied within moving data neighbourhoods, is a non-stationary algorithm which corresponds to a non-stationary random function

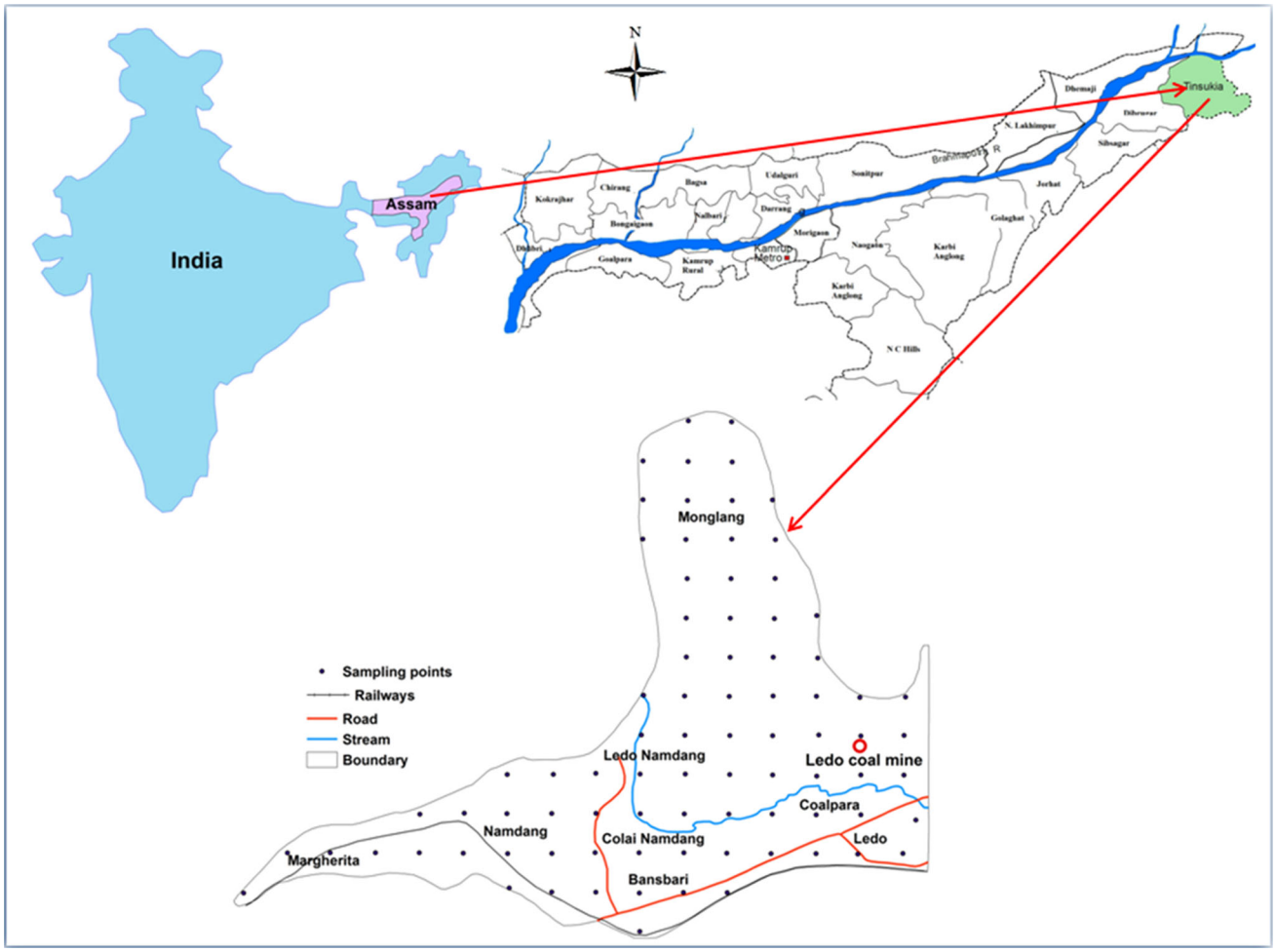


Fig. 1 Location and grid map of the study area

model with varying mean but stationary covariance (Deutsch and Journal 1992). In kriging, a semivariogram model was used to define the weights of the function (Webster and Oliver 2001), and the semivariance is an autocorrelation statistic defined as follows (Mabit and Bernard 2007):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i + h)]^2,$$

where $z(x_i)$ is the value of the variable z at location of x_i, h the lag and $N(h)$ the number of pairs of sample points separated by h. For irregular sampling, it is rare for the distance between the sample pairs to be exactly equal to h. That is, h is often represented by a distance band.

During pair calculation for computing the semivariogram, maximum lag distance was taken as half of the minimum extent of sampling area. In this study, omnidirectional semivariogram was computed for the heavy metals because no significant directional trend was observed. Best-fit model with minimum root mean square error (RMSE) was selected for each heavy metal. Using the model semivariogram, basic spatial parameters such as nugget (C_0), sill ($C + C_0$) and range (A) were calculated from the fitting curves which provide information about the structure as well as the input parameters for the kriging interpolation. Nugget is the variance at zero distance, sill is the lag distance between measurements at which one value for a variable does not influence neighbouring values and range is the distance at which values of one variable become spatially independent of another (Lopez-Granados et al. 2002).

Indicator kriging

The probability maps of soil Cr, Cd, Ni and Pb concentration that exceeds the respective FAO (2000) maximum permissible limit value (MPL) of 100, 3, 30 and 50 mg kg^{-1}, respectively, were prepared using indicator kriging. The indicator kriging estimates the probability that the concentration of a pollutant exceeds a specific threshold value at a given location (Deutsch and Journal 1992; Lin

et al. 2002). It is a nonlinear geostatistics where the conventional linear kriging estimators are applied to the data after a nonlinear transformation. Here the nonlinear transform is to a discrete (binary) indicator variable. These techniques have been widely applied (Van Meirvenne and Goovaerts 2001).

It is assumed that a soil property z at location x take value $z(x)$. In geostatistics, this value is treated as a realization of the random function $Z(x)$. An indicator transformation of $z(x)$ can be defined by

$$\omega_c(x) = 1 \quad \text{if } z(x) \le z_c, \quad 0 \text{ otherwise},$$

where z_c is a threshold value of the property. In indicator geostatistics, $\omega_c(x)$ is regarded as a realization of the random $\Omega_c(x)$,

$$\Omega_c(x) = 1 \quad \text{if } z(x) \le z_c, \quad \text{else } 0.$$

It can be seen that

$$\text{Prob}[Z(x) \le z_c] = E[\Omega_c(x)] = G[Z(x); z_c],$$

where Prob[], E[] denote, respectively, the probability and the expectation of the terms within the square brackets, and $G[Z(x); z_c]$ is the cumulative distribution function of $Z(x)$ at value z_c. The principal of IK is to estimate the conditional probability that $z(x)$ is smaller than or equal to a threshold value z_c, conditional on a set of observations of z at neighbouring sites, by kriging $\Omega_c(x)$ from a set of indicator-transformed data.

A set of data on z is transformed to the indicator variable $\omega_c(x)$. The variogram of the underlying random function $\Omega_c(x)$ is then estimated by

$$\gamma_c(h) = \frac{1}{2M_h} \sum_{i=1}^{M_h} [\omega_c(x_i) - \omega_c(x_i + h)]^2,$$

where M_h represents pairs of observations that are separated by the lag interval h. A set of estimates of this indicator variogram at different lags may then be modelled by one of the authorized continuous functions used to describe variograms (Webster and Oliver 2001).

An estimate of the indicator random function may then be obtained for a location x by kriging from the neighbouring indicator-transformed data. IK is equivalent to simple kriging of the indicator variables $\omega_c(x)$ using the mean within the kriging neighbourhood as the expectation.

Accuracy assessment

Accuracy of the spatial distribution maps was evaluated through cross-validation approach (Davis 1987). Among three evaluation indices used in this study, mean absolute error (MAE) and mean squared error (MSE) measure the accuracy of prediction, whereas goodness of prediction (G) measures the effectiveness of prediction. MAE is a measure of the sum of the residuals (e.g. predicted minus observed) (Voltz and Webster 1990).

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} z(x_i) - \hat{z}(x_i)$$

where $\hat{z}(x_i)$, is the predicted value at location i. Small MAE values indicate less error. The MAE measure, however, does not reveal the magnitude of error that might occur at any point and hence MSE will be calculated,

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} [z(x_i) - \hat{z}(x_i)]^2.$$

Squaring the difference at any point gives an indication of the magnitude, e.g. small MSE values indicate more accurate estimation, point-by-point. The G measure gives an indication of how effective a prediction might be relative to that which could have been derived from using the sample mean alone (Schloeder et al. 2001).

$$G = \left[1 - \frac{\sum_{i=1}^{N} [z(x_i) - \hat{z}(x_i)]^2}{\sum_{i=1}^{N} [z(x_i) - \bar{z}]^2}\right] \times 100,$$

where z is the sample mean. If $G = 100$, it indicates perfect prediction, while negative values indicate that the predictions are less reliable than using sample mean as the predictors. The comparison of performance between interpolations was achieved using mean absolute error (MAE).

Multivariate statistical analysis

The identification of pollutants sources is conducted with the aid of multivariate statistical analyses, such as principal component analysis (PCA) and correlation analysis. Multivariate analyses of the data in this work were carried out by SPSS v.16.0 software (SPSS Inc., Chicago, USA). Bartlett sphericity test and Kaiser-Mayer-Olkin test indicated that the normalized data were suitable for PCA. Varimax with Kaiser Normalization rotation was applied to maximize the variances of the factor loadings across variances for each factor. In addition, the correlations between the original variables are presented in the form of non-parametric Pearson correlation coefficients.

Results and discussion

Descriptive statistics of heavy metals and other soil properties

The statistical characteristics of soil Cr, Cd, Ni and Pb are listed in Table 1. The mean value of each heavy metal was much higher than the median, indicating the central tendency dominated by the outliers in the distribution. The

Table 1 Summary statistics of heavy metal concentrations and selected soil properties

	pH	Organic carbon (%)	Cr mg kg^{-1}	Cd	Ni	Pb
Mean	4.7	1.60	112.3	2.60	87.5	183.1
Median	4.7	1.28	1.13	2.32	83.7	1.83
SD	0.44	0.96	27.0	1.53	31.9	70.2
CV (%)	9.4	60.0	24.0	58.8	36.5	38.3
Minimum	3.7	0.13	55.0	0.10	28.9	4.20
Maximum	5.7	6.76	225.4	6.20	176.6	326.8
Skewness	0.10	2.30	1.21	0.46	0.46	0.02
Kurtosis	−0.39	9.22	2.64	−0.77	−0.19	−0.66
Distribution pattern			Normal	Normal	Normal	Normal

SD standard deviation, *CV* coefficient of variation

difference in mean and median values and the high coefficient of variation in this study would be attributed to the extremely high values of heavy metals. In the present investigation, among the heavy metals studied (Cr, Cd, Ni and Pb), the mean concentration of Pb was high (183.1 mg kg^{-1}). A high mean concentration of Pb (433 mg kg^{-1}) has also been reported in coal mine-affected agricultural soils (Bhuiyan et al. 2010). Pb is generally associated with mineral matter in coal, primarily with sulphides such as galena (PbS), clausthalite (PbSe) (Hower and Robertson 2003) and pyrite, as well as aluminosilicates and carbonates (Wang et al. 2003). Oxidation of sulphide-bearing minerals exposed to weathering has resulted in acid mine drainage, which is characterized by extreme acidity and salinity and a high level of dissolved metals (e.g., Cd and Pb). The greatest and the smallest standard deviation were observed in the Pb (70.2) and pH (0.44), respectively. Organic carbon, Cd and Pb exhibit a high variation (>50 %) according to guidelines provided by Warrick (1998). Skewness is the most common form of departure from normality. If a variable has positive skewness, the confidence limits on the variogram are wider than they would otherwise be and consequently, the variances are less reliable.

Semivariogram analysis of heavy metals

Semivariogram parameters (nugget, sill and range) for each heavy metal with best-fitted modal was identified based on minimum root mean square error (RMSE). Analysis of the isotropic variogram indicated that the Cd and Pb semivariograms were well described with the exponential model, with the distance of spatial dependence being 3,949 and 7,021 vm, respectively, while the Cr semivariogram was well described with the spherical model, with the distance of spatial dependence being 833 m; similarly, Ni was well described with the circular model, with the distance of spatial dependence being 993 m (Table 2).

Table 2 Semivariogram model and parameters of heavy metals

Heavy metals	Fitted model	Nugget (C_0)	Sill ($C + C_0$)	Range (A) (m)	Nugget/sill
Cr	Spherical	0.213	0.439	833	0.483
Cd	Exponential	0.012	0.147	3,949	0.082
Ni	Circular	0.357	0.760	993	0.470
Pb	Exponential	0.066	0.735	7,021	0.090

In the semivariogarm analysis, the nugget values represent the variability of measured heavy metals level at zero distance, which are positive in this study for all the heavy metals. This spatial random variance is caused by the artificial nature of heavy metal pollution in soil; meaning that anthropogenic input like coal mining is a significant source of heavy metals in the study area. Sill is the lag distance between measurements in semivariogram at which one value for a variable does not influence neighbouring values. The range (distance of spatial dependence) expressed as distance could be interpreted as the diameter of the zone of influence that represented the average maximum distance over which a heavy metal of two samples was related. At distances less than the range, measured properties of two samples became similar with decreasing distance between the two points. Thus, the range provided an estimate of areas of spatial dependence. The ratio of nugget and sill is commonly used to express the spatial autocorrelation of regional variables, which also indicates the predominant factors among all natural and anthropogenic factors (Robertson et al. 1997). The ratios of nugget and sill between 0.25 and 0.75 represented moderate spatial dependence; those below 0.25 represented strong spatial dependence; and all others represented weak dependence. Cd and Pb were strongly spatially dependent suggesting that they are affected by anthropogenic factors only while Cr and Ni were moderately spatially dependent suggesting that they are affected by either anthropogenic or natural factors or both.

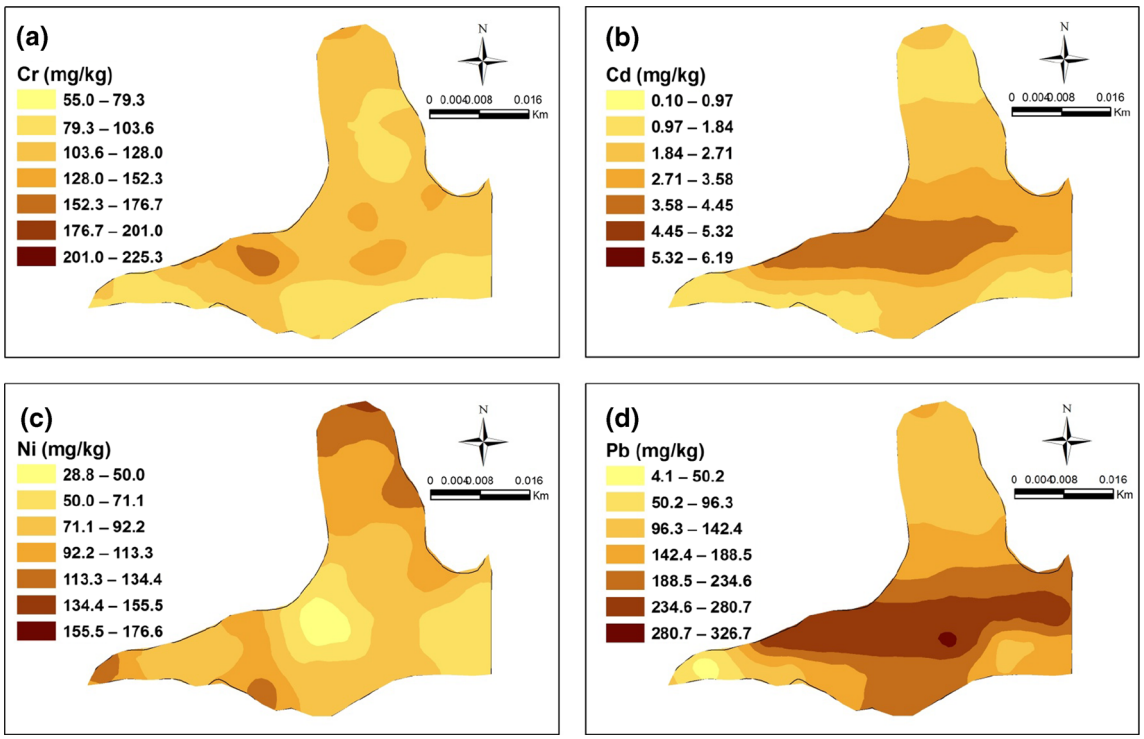


Fig. 2 Spatial distribution maps of **a** chromium, **b** cadmium, **c** nickel, **d** lead

Table 3 Evaluation performance of ordinary kriged map of heavy metals through cross-validation

Heavy metals	Mean absolute error (MAE)	Mean square error (MSE)	Goodness of prediction (G)
Cr	−0.443	664.5	8.0
Cd	−0.007	1.57	32.1
Ni	−0.376	723.2	28.0
Pb	0.647	2,368.5	51.3

Spatial distribution and risk assessment of heavy metals pollution

Using the available measurements for Cr, Cd, Ni and Pb concentration as well as the aforementioned structural models, spatial maps of these pollutants were produced using the ordinary kriging procedure (Journel and Huijbregts 1978). The spatial distribution maps of Cr, Cd, Ni and Pb (Fig. 2a–d, respectively) showed that high concentration of heavy metals was located in the low-lying paddy field and near coal mining site. Evaluation indices resulting from cross-validation of spatial maps of soil properties (Table 3) for all the soil heavy metals the prediction of goodness (G) value was greater than zero, which indicates that spatial prediction using semivariogram parameters is better than assuming mean of observed value as the values for any unsampled location. This also shows that semivariogram parameters obtained from fitting of experimental semivariogram values were fairly reasonable to describe the spatial variation.

In order to obtain data that may be used in the future for the assessment of the health risk due to elevated soil heavy metal concentration in cultivated areas, spatial maps of the probability that these pollutants exceed the corresponding maximum permissible limits (MPL) are produced. Figure 3a–d shows the indicator kriged probability maps of soil Cr, Cd, Ni and Pb based on the concentrations to exceed the respective FAO (2000) MPL value of 100, 3, 30 and 50 mg kg^{-1}, respectively. It was seen that more than 95 % area has higher than 50 % probability to exceed this MPL value of Pb and Ni. About 40 % area of the study site was having higher concentration than MPL value of Cd and concentrated at the centre of the study area. While, more than 80 % of the study area having higher concentration than MPL value of Cr and irregularly distributed.

Source identification based on multivariate statistics

For further evaluation of extent of metal contamination in the study area and source identification, principal component analysis was used following standard procedure reported in literature (Dragović et al. 2008; Yu et al. 2008; Franco-Uría et al. 2009; Alexakis 2011; Zheng et al. 2013).

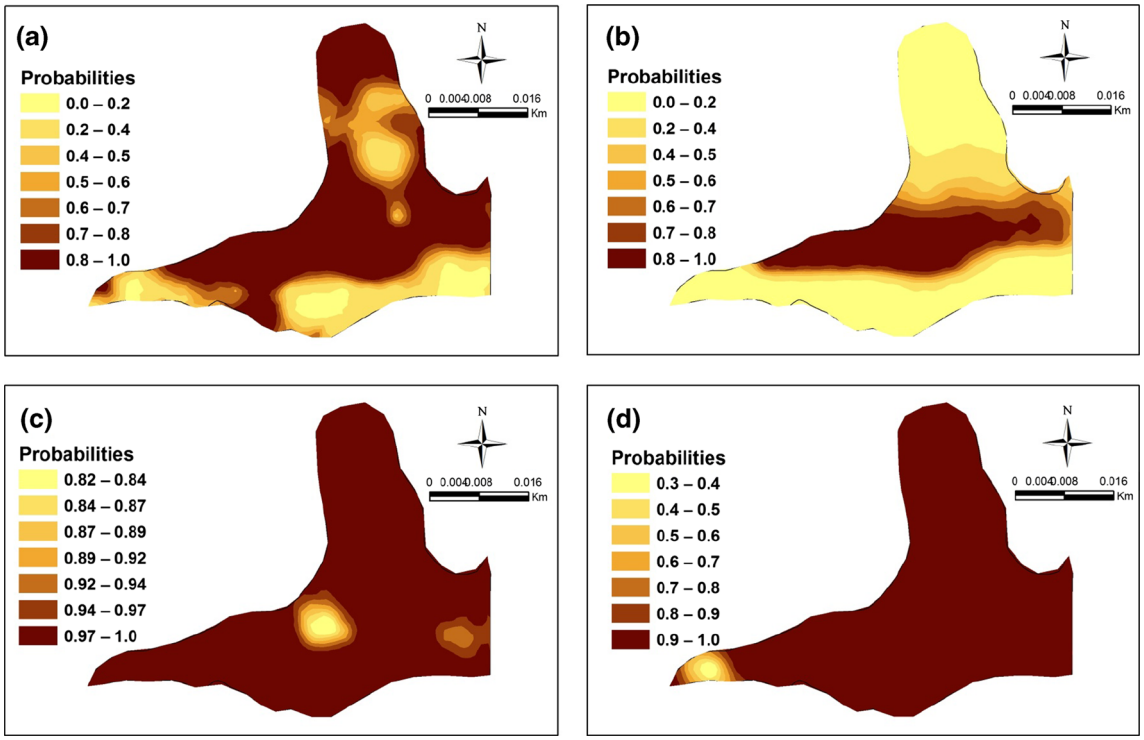


Fig. 3 Risk assessment maps of **a** chromium, **b** cadmium, **c** nickel, **d** lead

Table 4 Matrix of the three principal component (PC) accounting for most of the total variance

Heavy metals	PC1	PC2	PC3	Communities
Cr	0.37	0.26	0.89	0.99
Cd	0.86	−0.17	0.38	0.91
Ni	−0.25	0.95	0.20	1.00
Pb	0.92	−0.23	0.20	0.94
Percentage of variance	58.17	33.21	4.76	
Cumulative percent	58.17	91.38	96.14	

Table 5 Pearson correlation coefficient of heavy metal contents

Heavy metals	Cr	Cd	Ni	Pb
Cr	1.00			
Cd	0.583**	1.00		
Ni	0.335**	−0.291**	1.00	
Pb	0.468**	0.832**	−0.411**	1.00

** Correlation is significant at $p < 0.01$ level (2-tailed)

The loadings of measured heavy metal concentrations in the coordinate system of three principal components (PC) were obtained by analysing the correlation matrix. For the purposes of determining inter-element relationships, factor loadings between 0.50 and 0.75 were defined as ''moderate'', loadings >0.75 were defined as ''high'', and loading <0.5 were defined as ''low'' (Stamatis et al. 2011).

In detail, principal component 1(PC1) has the high positive loadings of Cd and Pb (> +0.86) and accounts for 58.17 % of variance (Table 4) and is the most important component. PC1 could be better explained as anthropogenic source, specifically derived from coal mine effluents. Lead may be released from galena and clausthite minerals that are associated with coal seams. Studies elsewhere have also reported on similar observations (Hower and Robertson 2003; Sakurovs et al. 2007). Geochemical weathering of sulphide minerals derived from mine drainage leads to accumulation of heavy metals in the soil (Bhuiyan et al. 2010). For example, galena (PbS) will release Pb into the environment by oxidation through the reaction described below: $PbS(s) + 2O_2\ (aq) \rightarrow Pb^{2+} + SO_4^{2-}$

Similarly, Greenokite (CdS) and associated minerals sphalerite (ZnS) and galena (PbS) will release Cd into the environment by oxidation reaction (Evanko and Dzombak 1997).

Meanwhile there are significant correlations between their levels in the soils of the study area (Table 5), which imply that these two heavy metals in soils may have originated from similar source, specifically from coal mine effluents. Considering the above reasons, the components loading of PC1 may have been derived from coal mine drainage sources, and PC1 may be defined as a coal mine drainage component. PC2, which has high positive loading of Ni (+0.95), low positive loading of Cr (+0.26) and low

negative loadings of Cd and Pb (<−0.23), accounts for 33.21 % of variance. PC2 can be considered as a measure of leaching of crustal materials because an important fraction of all the metals is lithogenic. PC3 has high positive loading of Cr (+0.89), low positive loadings of Cd, Ni and Pb (>+0.20) and accounts for 4.76 % of variance, indicating a mixed source from both lithogenic and anthropogenic inputs.

Conclusions

Variogram structure of heavy metals showing that there are clear spatial patterns of heavy metals on the distribution map and also that the current sampling density is sufficient enough to indicate such spatial patterns. The kriging interpolated map showed that high concentration of heavy metals was located in the low-lying paddy field and near coal mining site. The probability map produced based on kriging interpolation provided useful information for hazard assessment. Multivariate analysis (principal component analysis, cluster analysis) and correlation matrix used in this study provide important tools for the source identification. However, further information is need for more details about possible and real sources.

Acknowledgments The authors would like to thank D. P. Dutta, Senior Technical Officer, for assistance rendered during the field and laboratory work of this study. They also gratefully acknowledge Dr. G. Doerhoefer for editorial help and anonymous reviewers for their valuable comments, which greatly improved this manuscript.

References

Adriano DC (2001) Trace elements in terrestrial environments-biogeochemistry, bioavailability and risks of metals. Springer, New York

Alexakis D (2011) Diagnosis of stream sediment quality and assessment of toxic element contamination sources in East Attica, Greece. Environ Earth Sci 63:1369–1383

Anawar HM, Garcia-Sanchez A, Santa Regina I (2008) Evaluation of various chemical extraction methods to estimate plant-available arsenic in mine soils. Chemosphere 70:459–1467

Bhuiyan MAH, Parvez L, Islam MA, Dampare SB, Suzuki S (2010) Heavy metal pollution of coal mine-affected agricultural soils in the northern part of Bangladesh. J Hazard Mater 173:384–392

Cherry DS, Currie RJ, Souek DJ, Latimer HA, Trent GC (2001) An integrative assessment of a watershed impacted by abandoned mined land discharges. Environ Pollut 111:377–388

Chopin EIB, Alloway BJ (2007) Distribution and mobility of trace elements in soils and vegetation around the mining and smelting areas of Tharsis, Ríotinto and Huelva, Iberian pyrite belt, SW Spain. Water Air Soil Pollut 182:245–261

Davis BM (1987) Uses and abuses of cross-validation in geostatistics. Math Geol 19:241–248

Deutsch CV, Journal AG (1992) Geostatistical software library and user's guide. Oxford University Press, New York

Dragović S, Mihailović N, Gajić B (2008) Heavy metals in soils: distribution, relationship with soil characteristics and radionuclides and multivariate assessment of contamination sources. Chemosphere 72:491–549

Evanko C, Dzombak D (1997) Remediation of metals-contaminated soils and groundwater. Technology evaluation report te-97-01. Carnegie Mellon University, Pittsburg

FAO (2000) Land and water digital media series. Soil and terrain database, land degradation status and soil vulnerability assessment for Central and Eastern Europe. Food and Agriculture Organization (FAO). Available: http://www.FAO.org/docrep/TO551E/TO551eob.htm

Franco-Uría A, López-Mateo C, Roca E, Fernández-Marcos ML (2009) Source identification of heavy metals in pastureland by multivariate analysis in NW Spain. J Hazard Mater 165:1008–1015

Gay JR, Korre A (2006) Aspatially-evaluatedmethodology for assessing risk to a population fromcontaminated land. Environ Pollut 142:227–234

Hower J, Robertson JD (2003) Clausthalite in coal. Int J Coal Geol 53:219–225

Journel AG, Huijbregts CJ (1978) Mining geostatistics. Academic, London

Komnitsas K, Modis K (2006) Soil risk assessment of As and Zn contamination in a coal mining region using geostatisretics. Sci Total Environ 371:190–196

Korre A, Durucan S, Koutroumani A (2002) Quantitative-spatial assessment of the risks associated with high Pb loads in soils around Lavrio, Greece. Appl Geochem 17:1029–1045

Ladwani Kiran D, Ladwani Krishna D, Manik VS, Ramteke DS (2012) Assessment of heavy metal contaminated soil near coal mining area in Gujarat by toxicity characteristics leaching procedure. Int J Life Sci Biotechnol Pharm Res 4:73–80

Lee S (2006) Geochemistry and partitioning of trace metals in paddy soils affected by metal mine tailings in Korea. Geoderma 135:26–37

Lin YP, Chang TK, Shih CW, Tseng CH (2002) Factorial and indicator kriging methods using a geographic information system to delineate spatial variation and pollution sources of soil heavy metals. Environ Geol 42:900–909

Lopez-Granados F, Jurado-Exposito M, Atenciano S, Gracia-Ferrer A, De La Orden MS, Gracia-Toreres L (2002) Spatial variability of agricultural soil parameters in Southern Spain. Plant Soil 246:97–105

Mabit L, Bernard C (2007) Assessment of spatial distribution of fallout radionuclides through geostatistics concept. J Environ Radioactivity 97:206–219

Madejón P, Murillo JM, Maranon T, Cabrera FR, Lopez R (2002) Bioaccumulation of As, Cd, Cu, Fe and Pb in wild grasses affected by the Aznalcollar mine spill (SW Spain). Sci Total Environ 290:105–120

McGrath D, Zhang C, Carton OT (2004) Geostatistical analyses and hazard assessment on soil lead in Silvermines area, Ireland. Environ Pollut 127:239–248

Nesa N, Azad P (2008) Studies on trace metal levels in soil and water of Tipong, Tirap and Kikak collieries of Makum coal field, Tinsukia, Assam. Pollut Res 27:237–239

Petrisor I, Dobrota S, Komnitsas K, Lazar I, Kuperberg CM, Serban M (2004) Artificial inoculation—perspectives in tailings phytostabilization. Int J Phytoremediat 6:1–15

Reza SK, Baruah U, Sarkar D (2013) Hazard assessment of heavy metal contamination by the paper industry, north-eastern India. Int J Environ Studies 70:23–32

Richards LA (1954) Diagnosis and improvement of saline and alkali soils. USDA Handbook, p 60

Robertson GP, Klingensmith KM, Klug MJ, Paul EA, Crum JC, Ellis BG (1997) Soil resource, microbial activity, and primary production across an agricultural ecosystem. Ecol Appl 7:158–170

Sakurovs R, French D, Grigore M (2007) Quantification of mineral matter in commercial cokes and their parent coals. Int J Coal Geol 72:81–88

Schloeder CA, Zimmermen NE, Jacobs MJ (2001) Comparison of methods for interpolating soil properties using limited data. Soil Sci Soc Am J 65:470–479

Stamatis G, Alexakis D, Gamvroula D, Migiros G (2011) Groundwater quality assessment in Oropos–Kalamos basin, Attica, Greece. Environ Earth Sci 64:973–988

Swartjes F (1999) Risk-based assessment of soil and groundwater quality in the Netherlands: standards and remediation urgency. Risk Anal 19:1235–1249

Tabaksblat LS (2002) Specific features in the formation of the mine water microelement composition during ore mining. Water Res 29:333–345

Van Meirvenne M, Goovaerts P (2001) Evaluating the probability of exceeding a site-specific soil cadmium contamination threshold. Geoderma 102:75–100

Voltz M, Webster R (1990) A comparison of kriging, cubic splines and classification for redicting soil properties from sample information. J Soil Sci 41:473–490

Walkley A, Black IA (1934) An examination of the digtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. Soil Sci 37:29–38

Wang J, Sharma A, Tomita A (2003) Determination of the modes of occurrence of trace elements in coal by leaching coal and coal ashes. Energy Fuels 17:29–37

Warrick AW (1998) Spatial variability. In: Hillel D (ed) Environmental soil physics. Academic, New York, pp 655–675

Webster R, Oliver MA (2001) Geostatistics for environmental scientists. Wiley, Chichester

Yu L, Xin G, Gang W, Zhang Q, Su Q, Xiao G (2008) Heavy metal contamination and source in arid agricultural soil in central Gansu Province, China. J Environ Sci 20:607–612

Zheng Y, Gao Q, Wen X, Yang M, Chen HWuZ, Lin X (2013) Multivariate statistical analysis of heavy metals in foliage dust near pedestrian bridges in Guangzhou, South China in 2009. Environ Earth Sci 70:107–113