International Research Journal of Applied and Basic Sciences. Vol., 3 (5), 991-999, 2012
Available online at http:// www. irjabs.com
ISSN 2251-838X ©2012

Accuracy Assessment of Geostatistical Methods for Zoning of heavy metals in soils of urban-industrial areas

E. Mahmoudabadi[1*], F. Sarmadian[2], GH. Savaghebi[3], Z. Alijani[1]

1- M.Sc. student, 2- Associate Prof, and 3- Prof Department of Soil Science, College of Agriculture & Natural Resources, University of Tehran, Iran.

*_**Coressponding Author Email:**_ Ebrahim1985@gmail.com

Abstract

According to the IPI (integrated pollution index) report, almost 65 percent of all cities and dust of their roads were polluted with heavy metals. These areas threaten the health of human. Identifying, mapping and monitoring the pollution are the first step in contamination management. Hence, in recent decades many methods are presented to identify and mapping based on computer software that among these methods can be named Geostatistical methods. Geostatistical methods for considering the spatial correlation of data have a great importance in geochemical study. This study evaluates the efficiency of three methods Kriging, Cokriging and IDW, in addition compares them together. The best method was used for zoning and mapping metals lead, cadmium, zinc, copper, nickel, iron and manganese in Cheetgar park in west of Tehran. Results showed that Kriging was the best interpolation method which showed the higher accuracy in estimating of data compared to other methods. IDW was also as a suitable method to estimate the heavy metals while Cokriging due to low correlation with covariates was detected as an inappropriate method.

Keywords: Cokriging, Geostatistics, Heavy Metals, IDW, Kriging, Urban Soils

Introduction

Soil is an important part of urban ecosystem that directly or indirectly effects on the general quality of life. (Van Kamp et al 2003). Human activities such as industry, transportation, fossil fuels and waste consumption cause to soil pollution with heavy metals (Ajmone-marson and Biasioli, 2010). In the urban areas due to proximity of soil to human, metals can show their toxic effects through skin contact with soil, ingestion, entering to the stomath and inhalation of particles containing trace elements (Abrahams, 2002; Poggio et al., 2009).

Study of spatial distribution and identify the contaminant urban soils resources to recognize the extremely contaminated areas and determine the potential of contamination source is very important (Imperato et al., 2003; Zang, 2006; Wu et al., 2010). Recent researches suggest that soils in the urban areas have high variability even in the small ranges (Madrid et al 2006, Wei & Yang, 2009). This subject not only has been seen just about the usual factor of soil such as pH and cation exchangeable capacity but also be about the soil contamination that is due to the nature spatial variability of the soil and environment in which the wide heterogenic of human activities is appeared. Land use change usually occurs quickly in the urban areas, therefore specified points are lost and the new locations appear that need to the new sampling. Thus, the use of Rs and Gis and new computer software are necessary and inevitable (Ajmone-Marsan and Biasioli, 2010). One of the strategies in spatial analysis of environmental data is interpolation methods for the study of spatial distribution pattern of these data and making the favorable map. Geostatistics as a branch of statistic science is able to prepare the data processing and their spatial description. In addition, Geostatistics can describe and feature the pattern of spatial data and provide an estimate and quantitative map of the distribution of pollution with a minimum variance (Isaak and Srivastava, 1989). In this regard Geostatistical methods such as non-parametric estimator like weighted moving average or parametric geostatistical methods such as Kriging and Cokriging is considered (Hajrasuliha et al., 1980).

Dayani et al (2009); Hooker and Nathanail (2006) used the simple Kriging estimator for mapping the pollution of heavy metal to estimate concentration of lead in unsampled areas, and used the Kriging error maps as a credit controller of obtained maps. As was noted by Dayani et al (2009) in the margin of area and distance between samples in which sampling density has been decreased, there was highest estimation variance. Spatial Variability of copper, zinc, lead, chromium and cadmium was examined in Hangzong by Xing Mei et al (2006). They used of Kriging and log normal Kriging for making the pollution maps of these elements. In the study of three heavy metals iron, cobalt and vanadium in the soils of Hamedan, simple Kriging was used for cobalt, discrete Kriging for iron and vanadium was examined with exponential model for interpolation of metals concentration (Madani et al., 2009). The purpose of this study is to evaluate the efficiency and compare some Geostatistical methods such as Kriging, Cokriging, IDW for assessing and spatial variations of heavy metals iron, manganese, copper, zinc, nickel, cadmium and lead in the soils of Chitgar park and near the industrial area.

Materials and methods

The study area is located between 517581 m and 519831 m E, 3952982 m and 3955982 m N, approximately 700 ha, in part of forest park Chitgar and around land in west of Tehran and near the Tehran-Karaj highway. The maximum altitude of sea is 1313 m, the minimum altitude is 1225 m and the mean altitude of sea is 1269 m. the mean annual precipitation based on fourteen-year (1996-2010) report by Chitgar synoptic station is 267 mm. stratified grid sampling method was performed in the sampling area to select the sample locations. One hundred sixteen samples were taken with a distance 250 meters intervals from 0-20 cm soil depth.

After transferring the samples to the laboratory and preparation of them (air drying, squash and passing through 2 mm sieve), extraction was carried out by nitric acid 4 normal. In the obtained extraction the amount of heavy metals lead(pb), cadmium(cd), nickel(Ni), zinc(zn), copper(cu), manganese(Mn) and iron(Fe) were studied by using of atomic absorption system in shimadzu model (AA.670) (Richards et al., 1998).
Statistical description of data was performed by SPSS 16.0. As well spatial structure analyzing of data was done by GS+ software version 5.1.

Variography

Variogram calculation is the first step in examining the spatial variable. Variogram is estimated by Goovaerts (1998) as follows:

$$\gamma(h)=\frac{1}{2N(h)}\sum_{i=1}^{N(h)}\{Z(x_i)-Z(x_i+h)\} \qquad (1)$$

Where: $\gamma(h)$ is semivariogram, N(h) is number of pairs, $z(x_i)$ is observed value of the variable x at location i and $Z(x_i+h)$ is observed value of variable at distance h from x_i . Next step is to fit the best theoretical model to empirical variogram that choice of this model is done by calculating the minimum sum of square deviations and maximum coefficients of determination (R^2). The most common technique is to calculate the weight of the minimum sum of square deviations (Webster and Oliver, 2002). The fitted model gives information about the spatial structure of data for zoning by the estimator.

Geostatistical estimators

Geostatistical estimate involves two stages. First stage is the knowledge and modeling the spatial structure of variable that is determined by the analyzing of semivariogram. Second stage estimates the required variable that is related to the first stage (Mohammadi, 2007).

Kriging method

Kriging method similar the weighted moving average method, for estimating the unknown point, rates the weight for each measured samples:

$$Z^* = \sum_{i=1}^{n} \lambda_i Z(x_i) \quad (2)$$

Where: Z^* is estimated value of spatial variable, Z (x_i) is observed value of spatial variation at point x_i that indicates the significance of point i in estimating.

Use of this estimator is possible if the variable x is normally distributed. Otherwise should be eighter used of non linear Kriging or made the distribution of variables as normal in such a way.

Kriging method is based on weighted moving average and it can be called as the best non diagonal linear estimator with the least variance (Fazeli, 2010). Kriging with any estimate gives the amount of error that by using of this unique Kriging's characteristic can specify the parts that error is height (Mohammadi, 2007).

Cokriging method

Sometimes it may be a variable for reasons such as difficult of sampling or expensive laboratory measurements, does'nt enough sampled and can't be estimated carefully based on them. In such cases can modify the estimate and increase its accuracy by considering correlation between these and other variables which is sampled well. This can be done using Cokriging. Cokriging equation is expressed as follows:

$$Z^*(x_i) = \sum_{e'=1}^{n} \lambda_{e'i} . x_i \sum_{k=1}^{n} \lambda_{k.y} (x_k) \quad (3)$$

Where: $Z^*(x_i)$ is estimated value for point x_i, λ_i is the weight related to the variable z, λ_k is the weight related to the auxiliary variable, $Z(x_i)$ is observed value of main variable, $y(x_k)$ is observed value of auxiliary variable.

Inverse Distance Weighting (IDW) method:

Inverse distance estimators are the most important methods of classical statistical interpolation. In this method the weight of the sample point on the unknown point is calculated based on distance between the known and unknown points. These weights are controlled by weighting powers, so that greater powers reduce the effect of farther points of estimated point and smaller powers distribute the weights more uniform among the neighbor's points. The following equation shows how is calculating the weights in this method:

$$y_i = \left(\frac{D_0}{D_i} \right)^{\alpha} - 1 \quad (4)$$

Where: λ_i is weight of i sample point, D_i is distance between i sample point and unknown point, α is weighting power, D_0 is neighborhood radius.

Selection the power in the IDW method can be done by using cross validation method. There is a point in this way which the points that have the same distance from the estimation points give the equal weight and have not been considered their location and arrangement.

Efficiency evaluation of estimator

There are several methods for validating the interpolation methods that one of the most important of them is cross validation method. In this method one point is removed temporarily and estimated using other points and interpolation methods. Then this point is returned to the its place and next point is removed and so the estimate is done for all points. Finally, the observed estimated value columns are compared in the form of different criteria.

Criteria that were used to compare the observed and estimated values including root mean square error, mean bias error and mean absolute error which are calculated by the equations 5, 6 and 7 respectively. Whatever RMSE, MBA and MAE criteria closer to zero indicate a more accurate and less error of method (Webster & Oliver,200).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\hat{Z}(x_i) - Z(x_i)\right]^2} \qquad (5)$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}\left[\hat{Z}(x_i) - Z(x_i)\right] \qquad (6)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{Z}(x_i) - Z(x_i)\right| \qquad (7)$$

Where: n is number of observed points, $\hat{Z}(x_i)$ is estimated value for i point, $Z(x_i)$ is observed value for i point, $\overline{Z}\ (x_i)$ is mean of observed values.

Result and discussion

Statistical summary of data for seven heavy metals has been studied in the table 1. Kolmogorov Smirnov test was used in order to study of normalizing the data. Normality test results showed that the metals zinc, manganese, lead, iron and cadmium has not a normal distribution while nickel and copper are followed the normal. Since Geostatistical analysis need to the data with normal distribution, data transformation was done logarithmically.

Table 1. statistical summary of heavy metals values

variable	N	Range	Minimum	Maximum	Mean	Std. Deviation	Skewness	Kurtosis	CV%
Zn	116	74.75	45.25	120.00	76.54	17.37	0.87	0.55	22.69
Pb	116	28.30	11.50	39.80	21.33	5.48	1.18	1.96	25.68
Fe	116	14125.00	17800.00	31925.00	25558.04	3668.06	-0.41	-0.48	14.35
Ni	116	32.73	4.69	37.42	21.20	7.32	0.02	-0.30	34.55
Mn	116	642.50	631.25	1273.75	905.43	139.50	0.73	0.29	15.41
Cu	116	20.20	22.80	43.00	34.52	4.52	-0.45	-0.29	13.10
Cd	116	0.27	0.03	0.30	0.13	0.07	0.53	-0.75	53.31

The first step in the utilization of Kriging method is investigation the spatial structure in the data by Variogram analysis. Table 2 shows the variogram characteristic for studied metals. Fitted model for metals zinc, iron, lead, copper and cadmium is spherical while for nickel and manganese is exponential.

Table 2. variogram analysis for studied heavy metals

Parameter	Model	Nugget Effect (C_0)	Sill (C0+C)	Range (A)	Proportion C0/(C0+C)	R^2	RSS
Zn	Spherical	0.03	0.052	1350	0.57	0.90	2.6E-05
Fe	Spherical	0.014	0.028	2638	0.5	0.94	1.01E-05
Ni	Exponential	21.1	62.3	2247	0.34	0.97	25.6

Mn	Exponential	1.3	2.96	2553	0.43	0.96	0.34
Pb	Spherical	0.027	0.057	775	0.47	0.94	1.66E-5
Cu	Spherical	13	21	1200	0.62	0.92	2.63
Cd	Spherical	3.7E-03	5.2E-03	1100	0.71	0.71	3.8E-07

Another basic condition for using Geostatistical methods is being Stationary data. This means that features of a random distribution function does not change from one point to another. Variogram analysis is one of the detection methods for the Stationary data. If semivariogram does'nt reach to the certain limit and does'nt fixed and if also pair of points show sinus modes is indicating lack of reliability (Sarmadian et al., 2009). After modeling the variogram; Kriging, Cokriging and IDW methods was used to predict the spatial changes of soil characteristics. In order to use of Cokriging estimator being of significant correlation is not the main condition to use the auxiliary variable for estimating the main variable but also both of these variables should be have a spatial structure as well its common variogram should be fallow a Structured pattern (Fazeli, 2010). thus for the elements cadmium, copper, lead and zinc due to lack of high correlation with studied factors of soil and lack of common variogram was not used of Cokriging method (Table 3).

Table 3. Pearson correlation coefficient values between soil properties and heavy metals in the whole study area

	pH	CaCO3	Clay	OC	C/N	Zn	Pb	Fe	Ni	Mn	Cu	Cd
pH	1											
CaCO3	0.184ns	1										
Clay	0.117ns	-0.049ns	1									
OC	0.063ns	.232*	-.195*	1								
C/N	0.127ns	.239*	-.223*	.762**	1							
Zn	.312**	-.219*	0.132ns	.219*	0.119ns	1						
Pb	0.072ns	-0.145ns	0.044ns	.229*	.188*	.397**	1					
Fe	0.184ns	-.641**	0.112ns	-.375**	-.337**	.276**	0.061ns	1				
Ni	0.1ns	-.297**	.510**	-0.03ns	-0.03ns	.204*	.272**	0.069ns	1			
Mn	0.122ns	-.441**	.545**	-.350**	-.213*	.254**	0.144ns	.398**	.300**	1		
Cu	0.149ns	-.469**	0.031ns	-0.03ns	-0.1ns	.378**	.269**	.643**	0.06ns	0.137ns	1	
Cd	0.102ns	-0.095ns	.258**	0.083ns	0.069ns	.290**	.214*	0.033ns	.301**	0.12ns	.203*	1

OC: Organic Carbon, C/N: Carbon per Nitrogen ratio
** Significant at 0.01 levels, *Significant at 0.05 level and ns: non-significant

MBA, MAE and RMSE characteristics was used to compare and evaluate the three interpolation methods of Kriging, Cokriging and IDW (table 4).

By getting further from zero value RMSE, MAE and MBE factor showing the result in accuracy reduction or a growth in deviation (Vakernagel, 2002). Accordingly Kriging estimator showed the higher accuracy or less deviation among the methods.

Among the studied metals There was no differences between Kriging and IDW for cadmium and MBA, MAE and RMSE factors were similar for these two methods but for other studied heavy metals, Kriging estimator showed the higher accuracy (table 4). Also Cokriging showed the least accuracy among three methods that this is due to lack of high correlation coefficient between the soil factors with studied heavy metals.

Kriging is an unbiased estimator with the lowest estimation variance. It gives the error with any estimate that using this unique property of Kriging can specify the areas where error is high and more data needed. Minimizing the error variance is the most important characteristic of Kriging that distinguishes it from other estimators (Mohammadi, 2007). For studied variables it showed the less value of MBE compared to IDW method.

Sarmadian et al (2009) also showed that Kriging was the best estimator for zoning the DTPA form of iron, manganese, zinc, lead and copper. As well Xing mei et al (2006) used ordinary and log normal Kriging to prepare pollution maps of copper, zinc, lead, chromium and cadmium in Hangzong, China. Atteia and Dubois et al (1994), Carlos et al (2003), Hooker and Nathanail (2006), Jiachun et al (2006), Baghaie et al (2007), Luoe et al (2008), Dayani et al (2009) and many other researchers said that Kriging model is more accurate than other interpolation methods. After selecting the most appropriate method, interpolation was done for each of heavy metals and then distribution map of heavy metals manganese, nickel, lead, cadmium, copper, zinc and iron was produced for area (figure1).

Table4. comparing the accuracy of studied estimators in estimating the studied heavy metals value.

variable	Estimation method	RMSE	MAE	MBA
	Ordinary Criging	0.071	0.058	-0.008
	Co Criging	-	-	-
Cd	IDW WP* 1	0.070	0.058	-0.007
	IDW WP 2	0.071	0.058	-0.008
	IDW WP 3	0.073	0.059	-0.010
	Ordinary Criging	4.162	3.284	-0.115
	CoCriging	-	-	-
Cu	IDW WP 1	4.234	3.321	-0.202
	IDW WP 2	4.215	3.311	-0.202
	IDW WP 3	4.277	3.340	-0.187
	Ordinary Criging	3100.051	2535.88	-247.139
	CoCriging	4737.227	3301.302	-743.233
Fe	IDW WP 1	3156.022	2580.463	-330.891
	IDW WP 2	3127.571	2534.271	-305.379
	IDW WP 3	3126.453	2501.572	-266.797
	Ordinary Criging	103.161	80.849	-0.222
	CoCriging	104.875	83.198	-0.528
Mn	IDW WP 1	103.594	79.446	-4.407
	IDW WP 2	103.578	80.588	-3.305
	IDW WP 3	104.648	81.774	-2.748
	Ordinary Criging	6.553	5.291	-0.088
	CoCriging	6.729	5.451	-0.041
Ni	IDW WP 1	6.567	5.196	-0.246
	IDW WP 2	6.502	5.182	-0.205
	IDW WP 3	6.500	5.230	-0.179
	Ordinary Criging	4.828	3.533	-0.587
	CoCriging	-	-	-
Pb	IDW WP 1	5.062	3.550	-0.784
	IDW WP 2	4.938	3.516	-0.696
	IDW WP 3	4.872	3.561	-0.606
	Ordinary Criging	16.864	12.169	-1.888
	CoCriging	-	-	-
Zn	IDW WP 1	16.975	12.391	-2.413
	IDW WP 2	16.826	12.147	-2.251
	IDW WP 3	16.845	12.255	-2.022

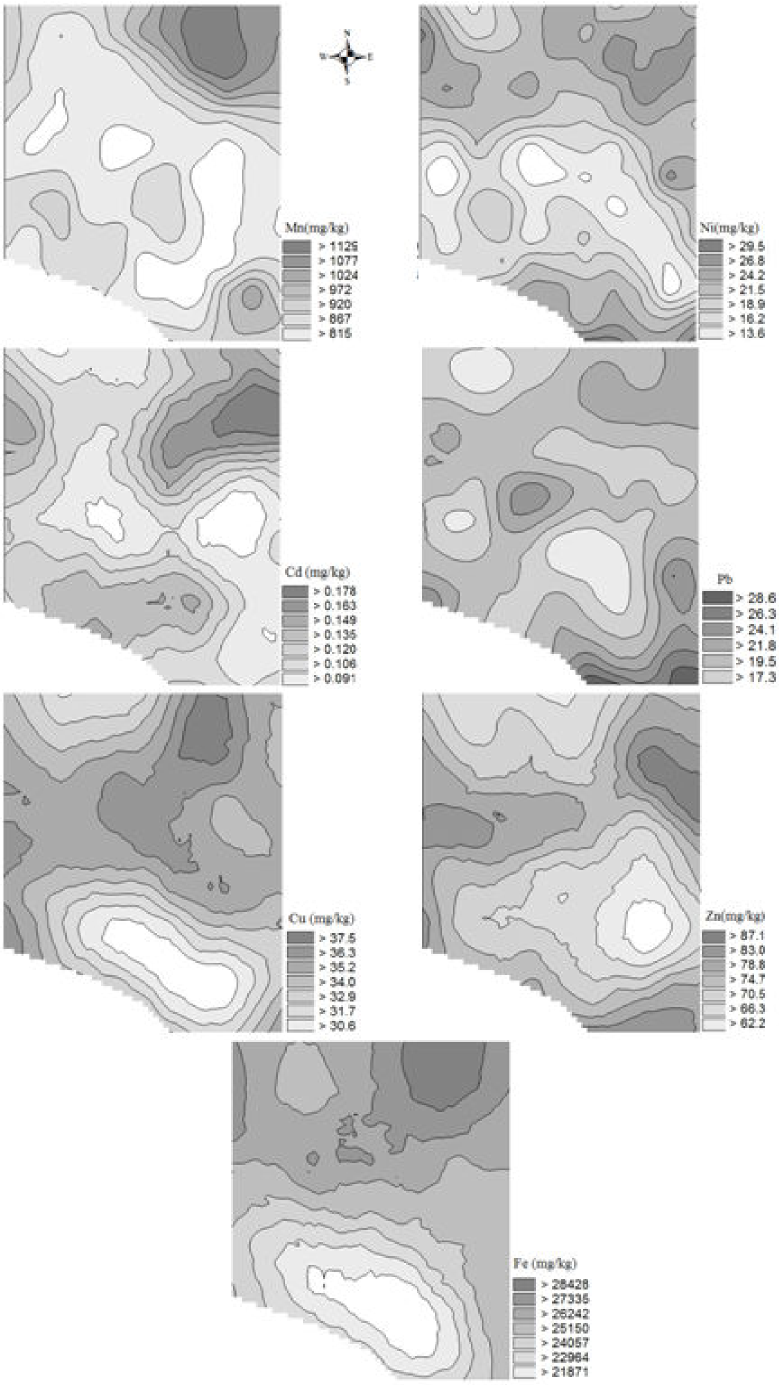

Figure 1. Zoning map of metals value manganese, nickel, lead, cadmium, copper, zinc and iron in the region

Conclusion

To investigate the spatial structure of heavy metals, empirical semivarogram of data was calculated and its curve was drown then the most appropriate model was chosen according to highest R^2 and lowest RMSE. This models for the metals zinc, iron, lead, copper and cadmium was fitted in the spherical form but for nickel and manganese was exponential. variography of semivariogram showed the average spatial structure for all heavy metals. As well impact range of different metal showed the different radius so that the maximum radius was obtained for iron (2638) and the minimum radius was for lead. In assessing of interpolation methods Kriging was more accurate than IDW and Cokriging that was compatible with the results of many researches (Xing Mei et al., 2006, Hooker and Nathanail, 2006, Xiaopeng and Lingqing, 2008). For example Xiaopeng and Lingqing (2008) used of ordinary Kriging for the spatial estimation of mercury pollution in the Baoji region, china and confirmed the high accuracy of this method. Among the evaluated method was shown that Kriging is preferred compared to other methods. but Kriging with IDW did'nt show differences for Cadmium and accuracy of IDW was assessed suitable for other metals and showed little difference with Kriging but Kriging method showed the less deviation compared to IDW for estimating. Cokriging due to lack of suitable correlation with soil factors did'nt have acceptable accuracy so using Cokriging depends on high correlation between evaluated and auxiliary variable. Totally Kriging method was selected as the best estimator to estimate the variables and study the spatial of heavy metal contamination in the region. Results of MBE, MAE and RMSE showed that this method had accurate enough for zoning. At the end after identifying the most appropriate method to estimate the amounts of metals in the region, the maps of each element were prepared.

References

Abrahams PW, 2002. Soils: Their implications to human health. Sci. Total Environ. 291:1 – 32.

Atteia O and Dubois JP, 1994. Geostatistical analysis of soil contamination in the Swiss Jura. Env. Pollution. 86:315-327.

Baghaie AH, Khademi H, Mohammadi J, 2007. Geostatistical analysis of spatial variability of Lead and Nickel around two industrial factories in Isfahan province. J. Agric. Sci. Natur. Resour. 14(2).

Carlos C, Keith P, Martial B, Jerry M, 2003. Combining Soil C and N Spatial Variability and Modeling Approaches for Measuring and Monitoring Soil Carbon Sequestration. J. Environ. Manage. 33: 274–288.

Dayani M, Mohammadi J. Naderi M, 2009. Geostatistical Analysis of Pb, Zn and Cd concentration in soil of Sepahanshahr suburb (south of Esfahan). Journal of Water and Soil, 23, 4, 67-76.

Emmerich WE, Lound LJ, Page AL, Chang AC, 1982. Movement of heavy metals in sewage sludge-treated soils. J. Environ. Qual, 11, 174-178.

Fazeli M, 2010. Surveying and mapping of soil physical properties using Geostatistic. M. Sc. Thesis, Tehran University, Tehran.

Goovaerts P, 1998. Geostatistical tools for characterizing the spatial variability of microbiological and physico-chemical soil properties. Biol. Fertil. Soils, 27, 315-334.

Hajrasuliha Sh, Baniabassi N, Metthey J, Nielsen DR, 1980. Spatial variability of soil sampling for salinity studies insouthwest Iran. Irrig. Sci. 1: 197- 208.

Hooker PJ and Nathanail CP, 2006. Risk-based characterisation of lead in urban soils. Chem Geol. 226:340-35.

Isaaks EH, Srivastava RM, 1989. An introduction to applied geostatistics. New York: Oxford Univ. Press. 561 pp.

Jiachun S, Hazian W, Jianming X, Jinjun W, Xingmei L, Haiping Z, Shunlan J, 2006. Spatial distribution of heavy metals in soil: A case study of Changing, China. Environ Geol. 10:245-264.

Luo W, Lu Y, Wang G, Shi Y, Wang T, Giesy JP, 2008. Distribution and availability of arsenic in soils from the industrialized urban area of Beijing, China. Chemosphere 72:797 – 802.

Madrid L, Díaz-Barrientos E, Ruiz-Cortés E, Reinoso R, Biasioli M, Davidson CM, 2006. Variability in concentrations of potentially toxic elements in urban parks from six European cities. J. Environ. Monit. 8:1158 – 1165.

Mohammadi J, 2006. Pedometricsvol 2 (Spatial Statistics). Pelk publication. Tehran.

Poggio L, Vrščaj B, Schulin R, Hepperle E, AjmoneMarsan F, 2009. Metals pollution and human bioacces-sibility of topsoils in Grugliasco (Italy). Environ. Pollut. 157:680 – 689.

Vakernagel H, 2002. Multivariate Geostatistics. Springer press, 387pp.

Van Kamp I, Leidelmeijer K, Marsmana G, deHollander A, 2003. Urban environmental quality and human well-being. Towards a conceptual framework and demarcation of concepts. A literature study. LandscUrban Plan. 65:5 – 18.

Webster R, Oliver M, 2001. Geostatistics for environmental science. Wiley & Sons, chichester. 271 p.

Wei B, Yang L, 2009. A review of heavy metal contaminations in urban soils, urbanroad dusts and agricultural soils from China. Microchem J. doi:10.1016/j.microc.2009.09.014.

Wu S, Xia X, Lin C, Chen X, Zhou C, 2010. Levels of arsenic and heavy metals in the rural soils of Beijing and their changes over the last two decades (1985–2008). J. Hazard. Mater, 179:860–868.

Xiaopeng Y, Lingqing W, 2008. Spatial Analysis and Hazard Assessment of Mercury in Soil around the Coal-Fired Power Plant: A Case Study from the City of Baoji, China. Environ Geol. 53:1381–1388.

Xing Mei L, Jianjun WV, Jiangming XU, 2006. Characterizing the risk assessment of heavy metal and sampling uncertainty analysis in paddy fields by Geostatistics and GIS. Environ. Pollut. 41: 279-289.

Zhang C, 2006. Using multivariate analyses and GIS to identify pollutants and their spatial patterns in urban soils in Galway, Ireland. Environ. Pollut. 142:501 – 511.

Zhao Y, Wang Z, Sun W, Huang B, Shi X, Ji J, 2010. Spatial interrelations and multi-scale sources of soil heavy metal variability in a typical urban–rural transition area in Yangtze River Delta region of China. Geoderma 156: 216–227.