Nature Environment and Pollution Technology
An International Quarterly Scientific Journal

p-ISSN: 0972-6268 e-ISSN: 2395-3454	Vol. 16	No. 2	pp. 339-350	2017

Original Research Paper | Open Access

Spatial Distribution, Sources Identification and Risk of the Trace Metals in Surface Sediments of Chaohu Lake Using Multivariate Statistics and Geostatistics

Wei Wang*, Chenchang Du*†, Nengcheng Chen*, Zeqiang Chen* and Xian He**
*State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, 430079, China
**Dalian Academy of Reconnaissance and Mapping Co. Ltd., Dalian, 116061, China
†Corresponding author: Chenchang Du

Nat. Env. & Poll. Tech.
Website: www.neptjournal.com

Received: 15-06-2016
Accepted: 16-07-2016

Key Words:
Trace metals
Geostatistics
Semivariogram
Multivariate statistics
Sediment

ABSTRACT

This study evaluated the trace metal sources and their spatial distribution in surface sediments in Chaohu Lake, China, using multivariate statistical analysis and geostatistics. The results of the multivariate statistical analysis, including principal component analysis and cluster analysis, indicated that Cu, Pb and Zn were significantly influenced by anthropogenic sources, whereas, the remaining studied metals mainly originated from natural sources. Ordinary kriging was used to plot the spatial distribution of the trace metals and principal component scores, while indication kriging was applied to generate the probability of the trace metals exceeding their pollution threshold values. We found that natural origins and the grain size effect influenced the spatial distribution of Al, Cr, Fe, Mg, Ni, and V, while human sources combined with the former two factors affected the spatial distribution of Cu, Pb and Zn. The spatial distribution of each principal component scores represented the spatial impact of corresponding component on this lake. The possibility maps show that some areas had a high ecological risk caused by Cu, Pb and Zn. Multivariate statistical analysis and geostatistics are concluded to be effective tools to study trace metals.

INTRODUCTION

As the trace metals are non-biodegradable and can pose a large ecological risk, pollution caused by them has been an environmental issue of particular public concern (Wang et al. 2012). When entering aquatic systems, trace metals are adsorbed on particulate matter in water bodies and ultimately accumulate in sediments (Dassenakis et al. 2003). However, if the conditions of the sediments (e.g., perturbation) change, trace metals may be desorbed from them, threatening the food chain and ecosystems and ultimately endangering human health (Hu et al. 2014, Senesi et al. 1999). Therefore, it is important to survey trace metal concentrations and their spatial distribution in sediments in order to identify possible pollution control strategies.

When studying the spatial distribution of contaminants, it is impractical to survey every site in a study area, because of time and cost limitations (Yin et al. 2011). Hence, interpolation techniques are commonly employed to estimate the concentrations of contaminants at unsampled sites. Geostatistical methods are regarded as the optimal interpolation techniques owing to their unbiased estimation and accurate prediction of the interpolation errors. Therefore, such methods are usually used for the interpolation and generation of the spatial distribution of contaminants (Hani et al. 2011, Simasuwannarong et al. 2012).

Geostatistical methods have many variants such as ordinary kriging (OK) and indicator kriging (IK). OK is used to map the spatial distribution of contaminants, whereas IK, a nonparametric geostatistical method, can be employed to plot risk maps relative to the pollution threshold values. For instance, Shi et al. (2007) employed OK and log normal kriging to plot the spatial distribution of six contaminants in Changxing, China. Dash et al. (2010) applied OK to estimate groundwater depth and quality parameters in Delhi and used IK to generate hazardous maps of quality parameters in relation to the pollution threshold values. Arslan (2012) applied OK and IK to study the spatial pattern of ground salinity in Bafra Plain, Turkey.

Trace metals in sediments originate from natural origins and anthropogenic sources such as parent materials, industrial activities and agricultural practices (Nemr et al. 2006). Because of the interactive influences of natural and anthropogenic inputs, the sources of trace metals become more complicated and thus difficult to interpret (Liu et al. 2003).

Multivariate statistical analysis methods, including principal component analysis (PCA), cluster analysis (CA), are powerful tools for the identification of common patterns in dataset distributions, leading to dimensionality reductions for raw datasets and facilitating the interpretation of the results. Therefore, these methods are used to apply in studies of metals. For example, Lee et al. (2006) combined PCA with CA to discuss metal pollution in soils in Hong Kong, and found clearly different associations among metals in different study areas, while Gu et al. (2012) employed PCA to identify possible sources of metallic elements and detected two kinds of anthropogenic sources.

Because it overlooks the spatial relationships between sampling sites, multivariate statistical analysis may lose important information included in the spatial distribution. Meanwhile, only relying on geostatistical methods ignores the relationships between trace metals, making it hard to distinguish different origins. Therefore, many studies have combined the multivariate statistical analysis and geostatistics to study the trace metals (McGrath et al. 2004, Shrestha et al. 2007).

In recent decades, the rapid economic development, population growth and continuous urbanization in the Chaohu Lake watershed have caused environmental deterioration and pollution by trace metals, especially Cu, Pb, Zn. Pervious authors have studied the spatial distribution of heavy metals and their sources (Li et al. 2012, Yin et al. 2011, Zheng et al. 2010). Nevertheless, studies on the extent of the impact of different sources on this lake and the spatial distribution and probability maps of trace metals based on geostatistics are still needed.

The present study was conducted to understand the trace metal origins and their spatial distribution in surface sediments of Chaohu Lake using geostatistics, multivariate statistics. The aims of this study were to: 1) define the possible sources of trace metals and the spatial impacts of these sources on this lake; 2) map the spatial distribution of trace metals in surface sediments; 3) plot the risk maps of trace metals exceeding the pollution threshold values.

MATERIALS AND METHODS

Study area: Chaohu Lake, one of the five largest freshwater lakes in China, lies within 117°16'54' -117°51'46' E longitude and 31°25'28'-31°43'28' N latitude and is located in central Anhui Province. Chaohu Lake has an average depth of 2.7 m, with a water surface area of about 780 km^2 and a watershed area of 13,350 km^2. The watershed of this lake has an annual mean temperature of 16°C and an annual mean rainfall of 996 mm with a subtropical monsoon climate. The lake has 33 inflow rivers, of which the Dianbu River, Nanfei River, Pai River, Fengle River, Hangbu River, Baishishan River, and Zhao River are the major rivers, with only one outflow river (Yuxi River) feeding Yangtze River. Several cities (e.g., Hefei City, capital city of Anhui Province, and Chaohu City) surround this lake.

In the past three decades, Chaohu Lake has witnessed significant economic growth and urban development in its watershed. At the same time, this lake has experienced serious pollution because of the discharge of massive amounts of industrial and domestic sewage. For example, in 1985, the discharge was estimated to be about 1.5×10^8 tons, including 1.1×10^8 tons from Hefei City and 0.3×10^8 tons from Chaohu City (Bing et al. 2013). The pollution from Nanfei River, the major sewage discharge channel for Hefei City, is more serious than that from other rivers.

Sediment sampling and analysis: During April 2009 and April 2011, 61 sediment samples were obtained by a gravity corer from Chaohu Lake with the latitudes and longitudes positioned using the Global Positioning System (GPS). Considering the fact that the heavy metal pollution in the western lake is more serious than that in the eastern and central lake (Li et al. 2012, Liu et al. 2012), relatively more sampling sites were selected in the western lake. Sampling sites are mapped in Fig. 1. The top surface 2 cm of surface sediments was gathered, put into plastic bags, and placed into an icebox temporarily. After being transported to the laboratory, these sediments were immediately preserved below 4°C.

Sediments were freeze-dried and ground with an agate pestle and a mortar, sieved, and then were totally digested with HCl-HNO_3-HF-$HClO_4$. Metal concentrations were determined by Inductively Coupled Plasma-Atomic Emission Spectrometry (ICP-AES). The accuracy and precision of the experiment were ensured by using blanks, replicates, and standard reference sediments offered by the Chinese Academy of Geological Sciences. Particle size was measured by using a Malvern Mastersizer 2000.

Multivariate statistical analysis: Before multivariate statistical analysis, the normality of the variable was checked (Brumelis et al. 2000). If a variable is not normally distributed, a transformation, either a log transformation or a boxplot transformation, should be performed to reduce the skewness and kurtosis values of the variable.

PCA, one of the most popular multivariate statistical analysis methods, renders the original dataset dimensionless, extracting the most meaningful information from the original dataset with the minimum loss of useful information. PCA transforms the initial variables of the data into new, uncorrelated variables, known as principal components (PCs), which are linear combinations of the initial variables

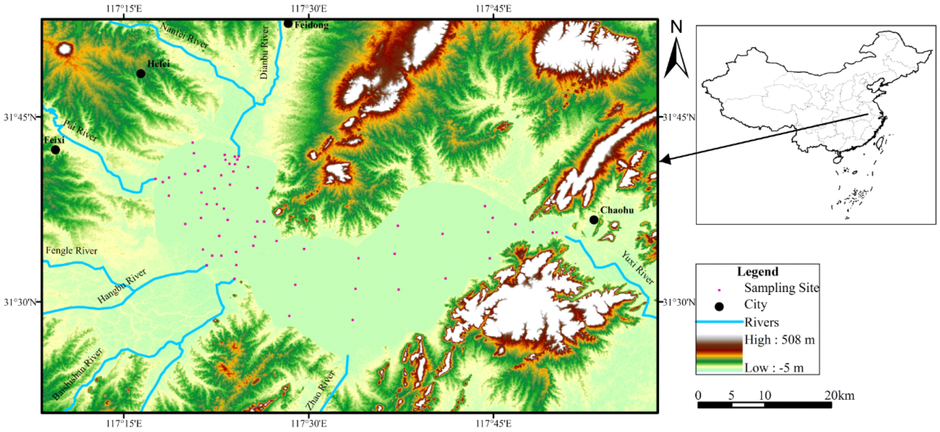

Fig. 1: Study area and distribution of sampling sites.

(Shrestha et al. 2007). In addition, Varimax rotation is usually employed to maximize the sum of the variance of the factor coefficients (Bhuiyan et al. 2011). Therefore, in the analysis of trace metals, some studies have employed PCA to identify and visualize the origins of trace metals (Bai et al. 2011, Zhou et al. 2007a).

CA, another multivariate statistical analysis method, is used to classify the cases or variables of a dataset into groups/clusters, based on the contribution from expected origins (Bhuiyan et al. 2011). The former mode that sorts the cases is named the Q-mode, while the latter mode that classifies the variables is termed the R-mode. The results generated by CA demonstrate that the degree of association is strong between members of the same group and weak between members of different groups. Before CA analysis, the initial data need to be standardized because different variables have different ranges. In this study, Pearson correlation was used to measure the similarity among groups integrated with the between-group linkages method as an agglomeration approach. Then, the results were visualized using a dendrogram which represented the hierarchical clustering procedure.

Geostatistical analysis: The analysis of experimental values was guided by the following four steps. First, the descriptive statistics of the data were generated, e.g., means. Second, the normality of the variables was examined with the Kolmogorov-Smirnov test. Third, the suitable theoretical semivariogram models for different trace metals were tried. Finally, OK and IK were applied to generate the spatial distribution of the trace metals and their probability maps in relation to their pollution threshold values.

Geostatistics is based on the regionalized variable theory that states that a regionalized variable has two properties, namely randomness and a spatial structure. The latter property implies spatial autocorrelation, meaning that the samples that exist far away in space are less like than those are close (McGrath et al. 2004). The most important aspect of geostatistics is the generation of a semivariogram, which represents the spatial variability of the regionalized variable and provides the input parameters for kriging interpolation (Zhang et al. 2004). The semivariogram can be calculated as follows (Goovaerts 1999, Shi et al. 2008):

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

Where, $y(h)$ expresses the semivariance value for the observed pairs at lag distance h, $N(h)$ states the number of observed pairs of sampling sites separated by distance h and $Z(x_i)$ is the value of the regionalized variable Z at site x_i.

Geostatistics has 11 different theoretical semivariogram models, such as the Hole Effect, and J-Bessel. Once the best-fit model for the semivariogram is determined, three semivariogram parameters, including the nugget (C_0), the sill (C_0+C), and the range (A_0), would be computed. C_0 represents the structural variance, and C_0+C the population variance. A_0 represents the spatial range across which the variable shows spatial dependence. Anisotropic spatial variability can be gained by computing the corresponding direction semivariogram. In this paper, however, only isotopic spatial variability was taken into consideration.

The OK method assumes a constant unknown mean and predicts the value of the unsampled site as a linear combination of the surrounding site values. The coefficients of this linear combination known as weights depend on 1) the distance between the corresponding surrounding site and the predicted site, and 2) the semivariogram of the variable (Dash et al. 2010). The related equations are given as follows:

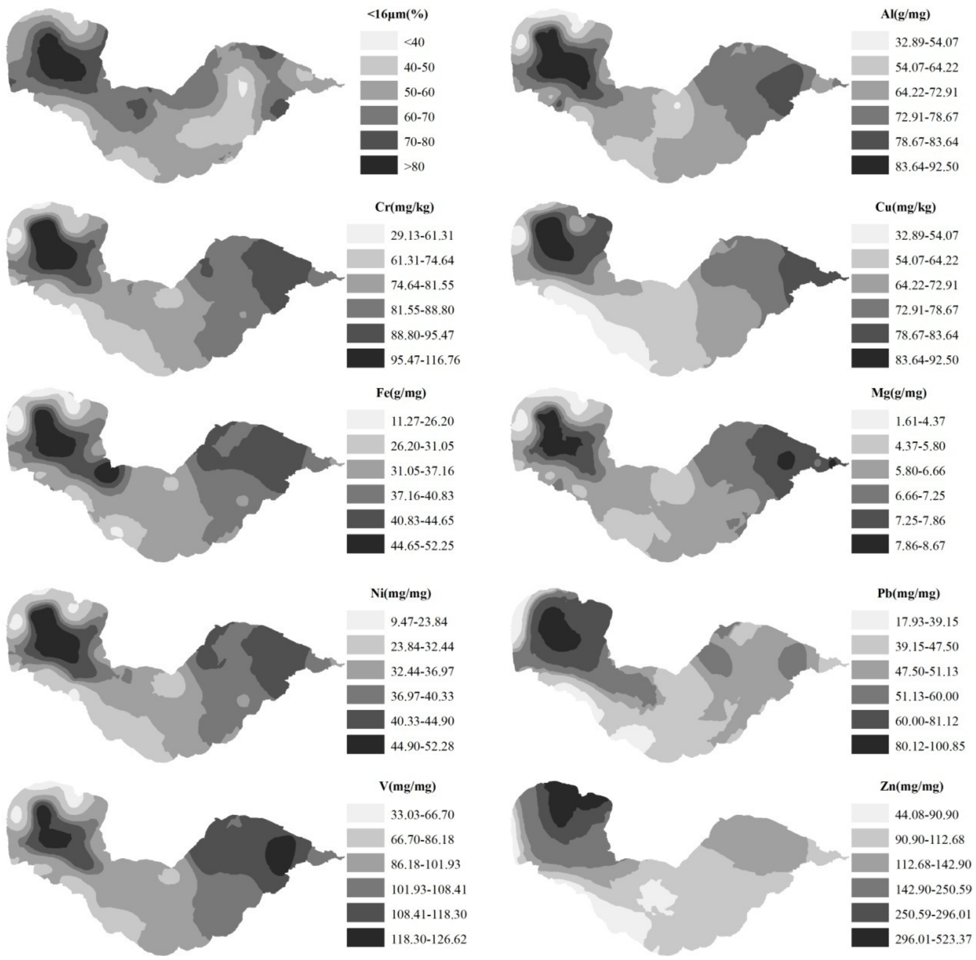

Fig. 2: Spatial distribution maps of the trace metals and fine-grained particles (OK).

$$Z(x_0)=\mu+\varepsilon(x_0) \quad ...(2)$$

$$Z(x_0)=\sum_{i=1}^{n}\lambda_i Z(x_i), \quad \sum_{i=1}^{n}\lambda_i=1 \quad ...(3)$$

Where, λ is an unknown constant, $Z(x_0)$ represents the predicted value of the regionalized variable at the predicted site x_0, $Z(x_i)$ expresses the measured value of the variable at site x_i, λ_i is the weight for $Z(x_i)$ and $\varepsilon(x_0)$ is the error in association with x_0.

The IK method does not demand that the raw dataset be normally distributed, and the binary function (0 or 1) is employed to guarantee this method is resistant to outliers (Hassan et al. 2011, Lee et al. 2007). Based on the pollution threshold value Z_p, IK transforms the regionalized variable into an indicator variable as follows (Lin et al. 2011):

$$I(x;Z_p)=\begin{cases}1, if\ Z(x)\geq Z_p \\ 0, otherwise\end{cases} \quad ...(4)$$

The expected value of $I(x;Z_p)|n)$, conditional on n surrounding data, can be expressed as:

$$E\{I(x;Z_p|n)\}=Prob\{Z(x)\geq Z_p|n\} \quad ...(5)$$

IK is based on an estimator, which can be written as:

$$Prob\{Z(x_0)\geq Z_p|n\}=\sum_{j=1}^{n}\lambda_j(Z_p)I(x_i;Z_p) \quad ...(6)$$

Where, $I(x_i;Z_p)n)$ represents the indicator value at the measured site x_i and $\lambda_j(Z_p)$ is the weight of $I(x_i;Z_p)n)$.

The cross-validation procedure is carried out as follows: Each time, a site is removed and its estimated value is calculated based on the values of the remaining sites, with the selected model and its parameters. Then the estimated values of all sites are combined with their measured values to generate the prediction errors and compute the correlation between the measured and estimated values. Afterwards, the predication errors and correlation are used to assess whether the selected model is the best-fit model for the current semivariogram. The best-fit model must meet the following requirement: The mean standardized error (SME) is closest to zero, and the root mean square error (RMSE) and average standard error (ASE) are minimized, and the root mean square standardized error (RMSSE) is closest to one.

Numerical sediment quality guidelines (SQGs) for freshwater and marine ecosystems have been developed to identify the relationships between contaminant concentrations and harmful biological effects (MacDonald et al. 2000). Herein, the consensus-based SQGs developed for freshwater ecosystems were employed to assess the harmful biological effects caused by the trace metals in Chaohu Lake (MacDonald et al. 2000). These guidelines have two threshold values, namely the threshold effect concentration (TEC) and the probable effect concentration (PEC). Below the TEC values, the harmful biological effects of metals are not expected to occur, whereas, above the PEC values, harmful biological effects are likely to be observed. In this study, the TEC values were selected as the pollution threshold values.

Data processing: Before conducting the multivariate statistical analysis, the normality of the original data and basic statistical parameters were calculated using the SPSS software. Multivariate statistical analyses, including PCA, CA and correlation analysis, were also performed in SPSS. Interpolation at unsampled sites as well as the spatial distribution maps and probability maps exceeding the pollution threshold values were depicted using ArcGIS 9.3 with Geostatistical Analyst Extensions.

RESULTS AND DISCUSSION

Descriptive statistics: The descriptive statistics of the trace metal concentrations and fine-grained particles are presented in Table 1. The percentage of fine-grained particles (<16 μm) varied from 12.28% to 97.04%. The concentration ranges of Al, Cr, Cu, Fe, Mg, Ni, Pb, V, and Zn were 32.59-92.50 g/kg, 29.13-116.76 mg/kg, 9.71-47.04 mg/kg, 11.27-52.25 g/mg, 1.61-8.67 g/mg, 9.47-52.68 mg/kg, 17.93-100.85 mg/kg, 33.03-126.62 mg/kg, and 44.08-523.37 mg/kg, respectively. In the analysis of the trace metals, it is necessary to compare the trace metal concentrations with their background values to determine the pollution levels of these metals. The trace metal concentrations of core sediments corresponding to the pre-industrial period were adopted as the background values (Liu et al. 2012). The

Table 1: Descriptive statistics of the trace metals and fine-grained particles.

Element	Range	Mean	SD	Skewness	Kurtosis	CV(%)	BV	K–S
Al (g/kg)	32.89–92.50	70.12	14.20	-0.55	2.48	20.25	49.0	0.220
Cr (mg/kg)	29.13–116.76	80.64	18.07	-0.37	3.12	22.41	52.9	0.934
Cu (mg/kg)	9.71–47.04	28.42	8.27	-0.04	2.74	29.10	11.4	0.994
Fe (g/kg)	11.27–52.25	35.62	9.11	-0.51	2.59	25.59	22.4	0.565
Mg (g/kg)	1.61–8.67	6.22	1.69	-0.76	2.82	27.19	3.9	0.250
Ni (mg/kg)	9.47–52.68	35.31	9.97	-0.38	2.51	28.24	21.3	0.748
Pb (mg/kg)	17.93–100.85	56.40	20.33	0.56	2.65	36.04	21.3	0.300
V (mg/kg)	33.03–126.62	93.82	24.02	-0.67	2.37	25.60	nd	0.062
Zn (mg/kg)	44.08–523.37	187.04	108.80	0.78	2.92	58.17	41.2	0.056
<16μm(%)	12.28–97.04	60.70	-	-	-	-	-	-

SD: standard deviation; BV: background value; "-": no data.

Table 2: Best-fit semivariogram models of the trace metals and their parameters (OK).

Element	Model	Nugget (C_0)	Sill (C_0+C)	Range (m)	RNS (%)
Al	Pentaspherical	5.906	246.080	7417	2.40
Cr	Pentaspherical	77.380	351.280	7290	22.02
Cu	Rational Quadratic	22.608	65.514	12,210	34.51
Fe	Gaussian	14.549	87.128	5477	16.70
Mg	Circular	0.088	3.119	5590	2.82
Ni	Spherical	3.054	121.590	5943	2.51
Pb	J-Bessel	191.300	285.690	10,725	66.96
V	Gaussian	103.660	598.870	5954	17.31
Zn	Hole Effect	4287.800	9234.900	23,037	46.43
PC1	Spherical	0.237	7.944	5768	2.98
PC2	Rational Quadratic	1.874	4.251	12,643	44.08

Table 3: Cross-validation between measured and predicted values of the trace metals (OK).

Element	RMSE	ASE	MSE	RMSSE	Correlation
Al	9.923	12.320	0.0093	0.990	0.713[a]
Cr	12.750	17.600	0.0190	0.822	0.703[a]
Cu	5.243	7.468	-0.0023	0.776	0.772[a]
Fe	6.197	7.395	-0.0068	0.976	0.737[a]
Mg	1.125	1.349	0.0155	1.014	0.750[a]
Ni	6.386	8.702	0.0086	1.018	0.766[a]
Pb	12.610	18.290	0.0377	0.760	0.783[a]
V	15.500	18.880	-0.0096	0.976	0.771[a]
Zn	54.120	75.150	-0.0058	0.764	0.868[a]
PC1	1.717	2.260	0.0107	0.999	0.746[a]
PC2	1.326	2.014	0.0054	0.729	0.777[a]

[a] Significant at the 0.01 level

Table 4: The correlation matrix of the trace metals and fine-grained particles.

	Al	Cr	Cu	Fe	Mg	Ni	Pb	V	Zn
Cr	0.909[a]								
Cu	0.698[a]	0.879[a]							
Fe	0.950[a]	0.919[a]	0.718[a]						
Mg	0.949[a]	0.903[a]	0.675[a]	0.942[a]					
Ni	0.936[a]	0.981[a]	0.827[a]	0.953[a]	0.925[a]				
Pb	0.650[a]	0.796[a]	0.894[a]	0.637[a]	0.558[a]	0.751[a]			
V	0.939[a]	0.913[a]	0.671[a]	0.951[a]	0.962[a]	0.949[a]	0.531[a]		
Zn	0.142	0.368[a]	0.666[a]	0.152	0.049	0.286[b]	0.805[a]	0.005	
<16μm	0.738 [a]	0.790 [a]	0.649 [a]	0.772 [a]	0.696 [a]	0.802 [a]	0.673 [a]	0.731 [a]	0.316 [a]

[a]Significant at the 0.01 level (two-tailed test); [b]Significant at the 0.05 level (two-tailed test)

ratios of the mean concentrations of the nine metals to their background values were 1.43, 1.52, 2.49, 1.59, 1.59, 1.66, 2.65, and 4.54, respectively, with Cu, Pb, and Zn having higher ratios than the remaining metals.

The coefficient of variation (CV) of the trace metals in sediments had a wide range, varying from 20.25% to 58.17%, suggesting that they had light/moderate variation. Of the studied metals, Cu, Pb and Zn had high CV values, indicating that these metals may have a high probability of being polluted by human inputs (Wang et al. 2011). The remaining metals displayed low variation, implying they were likely to be from natural origins.

Multivariate statistical analysis and OK work best if the variable satisfies the normal assumption. Therefore, the Kolmogorov-Smirnov test for normality (K-S) was conducted to check the normality of the trace metals. The results demonstrated that the values of all metals were more than 0.05, indicating that they were normally distributed.

Ordinary kriging for trace metals: Because the trace metals in sediments belong to regionalized variables, geostatistics was employed to study their spatial structure and generate their spatial distribution (Chen et al. 2013). The semivariogram model attributes, including C_0, C_0+C and A_0, and best-fit models are given in Table 2. In addition, cross-validation results, including RMSE, ASE, MSE and RMSSE, are presented in Table 3. Meanwhile, the spatial distribution of the trace metals is presented in Fig. 2.

As Table 2 shows, different metals had different best-fit semivariogram models. The semivariogram models for the metals were suitable for a Pentaspherical model (Al, Cr), Rational Quadratic model (Cu), Gaussian model (Fe, V), Circular model (Mg), Spherical model (Ni), J-Bessel model (Pb), and Hole Effect model (Zn).

Nugget variance represents the experimental error and field variation within the minimum sampling spacing. The ratio of nugget/sill (RNS) can be employed to assess the level of spatial heterogeneity for metals in sediments, with an RNS less than 25% implying high spatial dependence, an RNS between 25% and 75% suggesting moderate spatial dependence, and an RNS more than 75% indicating weak spatial dependence (Liu et al. 2006). The spatial dependence of trace metals might be influenced by intrinsic factors (natural factors) and extrinsic factors (human factors). Generally, the high spatial dependence of trace metals can be ascribed to intrinsic factors and weak spatial dependence to extrinsic factors.

The RNSs for nine metals ranged from 2.40% to 66.96%. The RNSs of Cu, Pb and Zn, between 25% and 75%, showed moderate spatial dependence, which suggests that human factors may weaken their dependence. The RNSs for Al, Cr, Fe, Mg, Ni, and V, less than 25%, showed moderate dependence, implying they may result from natural factors. Although the RNS can provide useful information on the spatial dependence of a spatial variable, it is less sensitive to the variable than the range value (Chen et al. 2008). The range values for trace metals varied from 5,477 to 23,037 m. As Table 3 shows, the correlation between the estimated values and measured values of the trace meals was significant, implying that the prediction was accurate.

The spatial distribution of the trace metals, with the exception of Zn, exhibited similar spatial features, with high concentrations in the western lake, medium concentrations in the eastern lake and low concentrations in the central lake. In addition, the highest concentrations occurred in the central part of the western lake. Previous research has demonstrated that the concentrations of trace metals in sediments are strongly impacted by grain size, termed the grain size effect, because as sediment becomes finer, its surface specific surface area tends to rise, resulting in an increase in the concentrations of trace metals (Liu et al. 2014, Zhang et al. 2009a). In this study, these metals were highly correlated with fine-grained particles (Table 4), and the spatial distribution of these trace metals was similar to that of fine-grained particles (Fig. 2), implying that the grain size effect influences these metals. Therefore, it can be inferred that the spatial distribution of these metals is affected by the grain size effect. Considering that Al, Cr, Fe, Mg, Ni and V had low CV values close to their background values, these metals are likely to be from natural origins as well as influenced by the grain size effect. By contrast, Cu and Pb had high CV values and were apparently higher than their background values, indicating that they are affected by human inputs besides natural origins and the grain size effect.

The spatial distribution of Zn was different from that of other metals, with high concentrations in the western lake and low concentrations in the eastern lake and central lakes. The concentrations in the western lake decreased drastically from northwest to southeast, with the highest concentrations occurring at the mouth of Nanfei River. Meanwhile, the spatial distribution of Zn did not follow that of the fine-grained particles, indicating the existence of other sources influencing Zn. In addition, Zn had high a CV value that was distinctly higher than its background value, implying that human inputs influenced it in addition to natural ori-

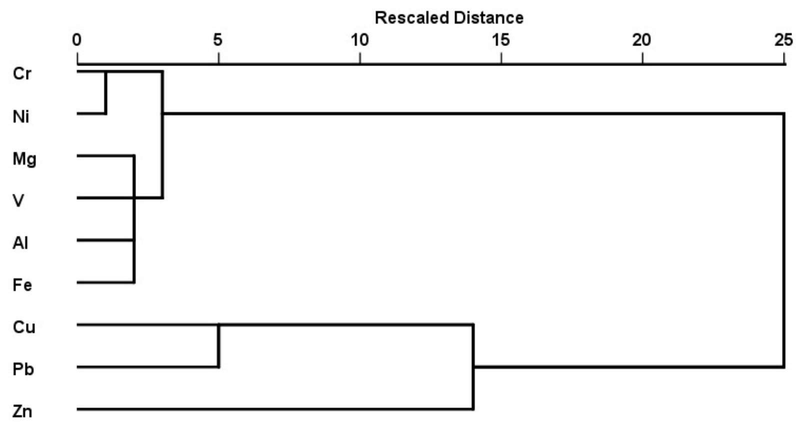

Fig. 3: Dendrogram of the cluster analysis of the trace metals.

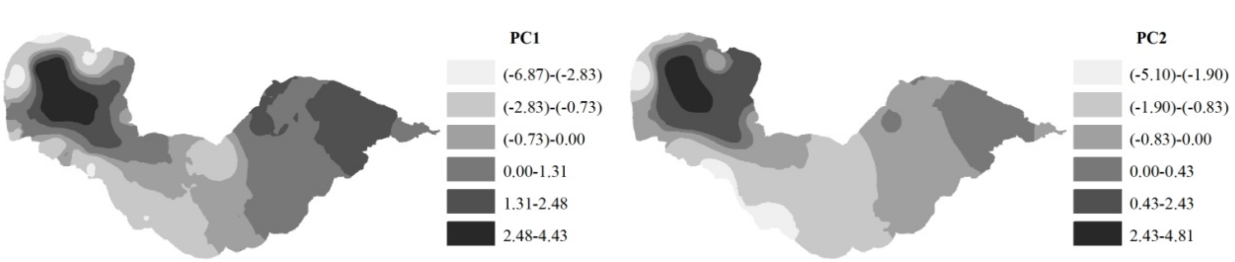

Fig. 4: Spatial distribution maps of PC1 and PC2 (OK).

Table 5: Total variance explained and rotated component matrix of the trace metals.

Element	PC1	PC2
Al	0.949	0.210
Cr	0.879	0.451
Cu	0.604	0.746
Fe	0.952	0.216
Mg	0.973	0.118
Ni	0.921	0.367
Pb	0.472	0.861
V	0.988	0.085
Zn	-0.078	0.986
Eigen value	5.944	2.720
% of Variance	66.04	30.22
Cumulative %	66.04	96.26

gins and the grain size effect.

PCA analysis: PCA was conducted to identify the possible origins of the trace metals and explore the hidden relationships among the studied metals. Two PCs with eigen values greater than 1 were extracted. These two PCs represented all variables well because they explained 96.26% of the total variance. The PCA results are presented in Table 5. Before the interpretation of these results, the Kaiser-Meyer-Olkin test and Bartlett's test were conducted to check the applicability of PCA (Zhou et al. 2007b). The results were 0.88 and 1,209 (df=36, $p<0.05$), respectively, suggesting that PCA can be used to reduce dimensionality.

The first PC (PC1) accounted for 66.04% of the total variance and had strong loading for Al, Cr, Cu, Fe, Mg, Ni and V. Among these metals, Al and Fe generally originated from parent rock. As Table 4 shows, these elements were strongly correlated with each other, indicating that they may share common sources. Moreover, these metals had low CV values and their mean concentrations were close to the background values, with the exception of Cu. Therefore, it is reasonable to deduce that PC1 may result from natural sources.

The second PC (PC2) explained 30.22% of the total variance and it was contributed by Cu, Pb and Zn. Pb and Zn showed low values in PC1, indicating that other sources that influence the concentrations of Pb and Zn exist. Cu, Pb, and Zn had high CV values and high concentrations compared with their background values. Some studies have associated high concentrations of Cu, Pb and Zn in sediments with anthropogenic inputs (Zhang et al. 2009b). In addition, previous research has also pointed out that the pollution of Cu, Pb and Zn in Chaohu Lake is related to the industrial and domestic wastewater (Yin et al. 2011). Thus, it can be concluded that PC2 is controlled by human inputs.

Cu exhibited a merged relationship with PC1 and PC2. In addition, Cu was significantly correlated with the remaining members of PC1, while, it was also strongly related to Pb and Zn in PC2. All these findings confirmed that this metal was dominated by a merged source.

Cluster analysis: R-mode CA was carried out to determine the origins of the trace metals in sediments. The CA results are presented in Fig. 3. Overall, they are consistent with those derived from the PCA. The nine metals were merged into two clusters. Cluster 1 was made up of Al, Cr, Fe, Mg, Ni and V. This cluster was in agreement with PC1 in the PCA analysis. Cluster 2 comprised of Cu, Pb, and Zn. These three metals were interpreted as anthropogenic metals. Of them, Zn had the furthest linkage distance.

Based on CA, PCA, and the spatial distribution of the trace metals, the origins of the trace metals could be defined. We found that Al, Cr, Fe, Mg, Ni and V mainly originated from natural origins, while, Cu, Pb, and Zn were from human inputs in addition to natural inputs.

Although the spatial distribution of Cu and Pb was similar to that of Al, Cr, Fe, Mg, Ni and V, the origins of Cu and Pb were not the same as those of these metals. By only relying on the spatial distribution of trace metals, it is hard to identify their origins. However, it is practicable to combine PCA, CA, and geostatistics to explore the possible origins of trace metals.

Spatial impacts for PC1 and PC2: Component scores state the cumulative contribution for each metal loaded on the corresponding PC. To understand the spatial variability of impact for PC1 and PC2, the maps of the two PC scores were generated using the OK method. The semivariogram attributes are shown in Table 2 and the cross-validation results are listed in Table 3. Meanwhile, the spatial distribution of the PC1 and PC2 scores are mapped in Fig. 4.

PC1 was fitted with a Spherical model, whereas PC2 was suitable for a Hole Effect model. The RNS for PC1, less than 25%, showed strong dependence owing to intrinsic factors, while the RNS for PC2, in the range of 25%-75%, exhibited moderate dependence, suggesting that extrinsic factors might influence its spatial dependence.

The PC1 scores map represents the extent of the impact of natural origins on Chaohu Lake. The PC1 scores varied from -6.87 to 4.43, and 58.08% of areas had positive scores, indicating the importance of natural origins there. PC1 showed a similar spatial distribution as its members such as Al, Cr. As discussed earlier, the spatial distribution of PC1 might be influenced by the grain size effect because of the impact of this effect on its members.

The PC2 scores map (Fig. 4) demonstrates the extent of the effect of human sources on this lake. The PC2 scores

Table 6: Best-fit semivariogram models of the trace metals and their parameters (IK).

Element	Models	Nugget (C_0)	Sill (C_0+C)	Range (m)	Nugget ratio
Cu	Hole Effect	0.107	0.114	10,014	94.47
Pb	Gaussian	0.047	0.143	13,913	32.87
Zn	J-Bessel	0.075	0.296	34,264	25.34

Table 7: Cross-validation between measured and predicted values of the trace metals (IK).

Element	RMSE	ASE	MSE	RMSSE	Correlation
Cu	0.332	0.417	0.0384	0.832	0.643[a]
Pb	0.335	0.264	0.0045	1.238	0.377[a]
Zn	0.241	0.303	-0.0383	0.806	0.756[a]

[a] Significant at the 0.01 level

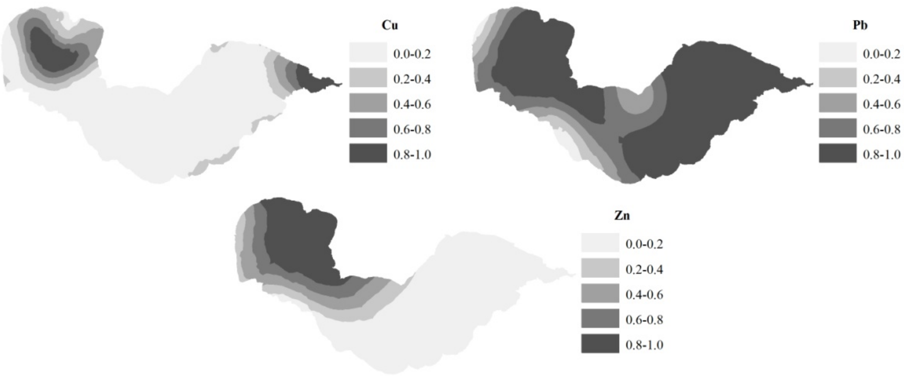

Fig. 5: Probability maps for the trace metals (IK).

ranged from -5.10 to 4.81, and 31.13% of areas had factor scores above zero, implying that human activities significantly influenced these areas. The PC2 scores in the western lake decreased in the northwest-southeast direction. The highest PC2 scores occurring in the northwest part of the western lake might be attributed to the proximity of the mouth of Nanfei River. As discussed previously, Nanfei River receives a large amount of industrial and municipal sewage from Heifei City. Thus, in these areas, the highest PC2 scores can be expected. The PC2 scores in the eastern lake were markedly lower than those in the western lake, but higher than those in the central lake. The PC2 scores in the eastern lake ranged from -0.83 to 0.43, suggesting a moderate pollution effect from Chaohu City.

In this study, multivariate statistical analysis and geostatistical methods were combined to explore the origins of trace metals in sediments and understand their spatial dependence. In some studies, geostatistical methods have only been employed to interpolate values at unsampled points, ignoring the spatial characteristics (Guo et al. 2012, Yuan et al. 2014). Indeed, when semivariogram attributes have been used to understand the spatial dependence of the trace metals, they can also be employed for the identification of the origins of the trace metals as in multivariate statistical analysis (Chen et al. 2008). Therefore, studies of the origins of trace metals should analyse their dependence.

IK for trace metals in relation to the pollution threshold values: Considering that Cu, Pb and Zn are influenced by human sources and might have harmful ecological effects, probability maps relative to the pollution threshold values were generated based on IK. The pollution threshold values for Cu, Pb and Zn were 31.6, 35.8 and 121.0 mg/kg, respectively. The attributes of the semivariograms for these metals, cross-validation results and estimations of the areas potentially affected by each metal are summarized in Table 6-8, respectively, while the probability maps for Cu, Pb and Zn exceeding their pollution threshold values are plotted in Fig. 5.

The Hole Effect model was the best-fit model for the semivariogram for Cu; the Exponential model was for Pb; and the J-Bessel model was for Zn. The RNSs for Cu, Pb, and Zn were between 25% and 75%, indicating that each metal

Table 8: Probability ranges of areas exceeding the threshold values of the trace metals (IK).

Probability range	Area(%)		
	Cu	Pb	Zn
0.0-0.2	71.7	2.0	58.7
0.2-0.4	10.0	3.7	8.0
0.4-0.6	7.1	8.9	8.5
0.6-0.8	5.0	18.4	7.3
0.8-1.0	6.2	67.0	17.5

shows moderate spatial dependence. The range values for these metals ranged from 10,014 to 34,264 m. The cross-validation demonstrated the accuracy of the estimation, with a significant correlation between the estimated values and measured values of the trace metals.

For Cu, 11.2% area of the lake fell within the high ecological risk category (i.e., probability between 0.6 and 1.0). These areas were mainly situated in the central part of the western lake and the eastern part of the eastern lake because of anthropogenic origins of Cu in these areas. This finding is consistent with the spatial distribution of Cu, which is considered to be a severe contaminant in aquatic ecosystems. Therefore, more attention should be paid to the excessive accumulation of Cu in these area. The remaining areas, nearly 85% of the lake, had a low probability (Table 8), indicating a low ecological risk there. For Pb, the areas with a high ecological risk were nearly 86% of the lake. The mean concentration of Pb was clearly higher than its pollution threshold. Therefore, a number of areas with a high risk can be expected. Pb is also a severe contaminant that could cause many diseases such as colic and anemia. Therefore, Pb in this lake requires further monitoring to control pollution. For Zn, the areas with a high ecological risk were about 25% of the lake. These areas were distributed in the northwest part of the western lake. The remaining areas with low risk, can be considered to be safe (possibility was less than 0.4).

CONCLUSIONS

Sixty-one surface sediments gathered from Chaohu Lake were analysed for trace metal concentrations by ICP-AES. The concentrations of all studied metals were normally distributed. Cu, Pb and Zn had high CV values and were markedly higher than their background values. Cu, Pb and Zn had moderate spatial dependence, implying that human inputs weaken their spatial dependence, whereas the remaining metals had high spatial dependence, implying the importance of natural origins. Al, Cr, Fe, Mg, Ni and V had similar spatial characteristics, with high concentrations in the western lake, medium concentrations in the eastern lake, and low concentrations in the central lake. These may originate from natural origins and are influenced by the grain size effect. Cu, Pb and Zn were thus impacted by human inputs in addition to natural origins and the grain size effect. Two component factors were extracted by PCA, which represented 96% of the total variance. PC1 represented natural origins, whereas PC2 represented human inputs. The CA results verified those of PCA, with cluster 1 containing Al, Cr, Fe, Mg, Ni and Cr and cluster 2 including Cu, Pb and Zn. PC1 and PC2 exhibited a similar spatial distribution to that of their own members. For Cu, 11.2% of the lake, in the central part of the western lake and the eastern part of eastern lake, had a high ecological risk. For Pb, 85% the lake (i.e., in nearly entire lake), had a high risk. For Zn, 25% of the lake, mainly in the eastern lake, had a high ecological risk. These results may provide useful information for pollution control and policy making about this lake.

ACKNOWLEDGEMENT

This study was financially supported by Major Science and Technology Program for Water Pollution Control and Treatment (2013ZX0750300104), The National High Technology Research and Development Program of China (2013AA01A608), China Torch Program (2012GH722002), the National Nature Science Foundation of China program (41301441) and China Postdoctoral Science Foundation funded project (2014M562050 and 2015T80829).

REFERENCES

Arslan, H. 2012. Spatial and temporal mapping of groundwater salinity using ordinary kriging and indicator kriging: The case of Bafra Plain, Turkey. Agricultural Water Management, 113: 57-63.

Bai, J.H., Cui, B.S., Chen, B., Zhang, K.J., Deng, W., Gao, H.F. and Xiao, R. 2011. Spatial distribution and ecological risk assessment of heavy metals in surface sediments from a typical plateau lake wetland, China. Ecological Modelling, 222(2): 301-306.

Bhuiyan, M.A.H., Suruvi, N.I., Dampare, S.B., Islam, M.A., Quraishi, S.B., Ganyaglo, S. and Suzuki, S. 2011. Investigation of the possible sources of heavy metal contamination in lagoon and canal water in the tannery industrial area in Dhaka, Bangladesh. Environmental Monitoring and Assessment, 175(1-4): 633-649.

Bing, H.J., Wu, Y.H, Liu, E.F. and Yang, X.D. 2013. Assessment of heavy metal enrichment and its human impact in lacustrine sediments from four lakes in the mid-low reaches of the Yangtze River, China. Journal of Environmental Sciences-China, 25(7): 1300-1309.

Brumelis, G., Lapina, L., Nikodemus, O. and Tabors, G. 2000. Use of an artificial model of monitoring data to aid interpretation of principal component analysis. Environmental Modelling & Software, 15(8): 755-763.

Chen, T., Liu, X.M., Zhu, M.Z., Zhao, K.L., Wu, J.J., Xu, J.M. and Huang, P.M. 2008. Identification of trace element sources and associated risk assessment in vegetable soils of the urban-rural transitional area of Hangzhou, China. Environmental Pollution, 151(1): 67-78.

Chen, X.D., Lu, X.W., Li, L. Y. and Yang, G. 2013. Spatial distribution and contamination assessment of heavy metals in urban topsoil from inside the Xi'an second ringroad, NW China. Environmental Earth Sciences, 68(7): 1979-1988.

Dash, J.P., Sarangi, A. and Singh, D.K. 2010. Spatial variability of groundwater depth and quality parameters in the national capital territory of Delhi. Environmental Management, 45(3): 640-650.

Dassenakis, M., Andrianos, H., Depiazi, G., Konstantas, A., Karabela, M., Sakellari, A. and Scoullos M. 2003. The use of various methods for the study of metal pollution in marine sediments, the case of Euvoikos Gulf, Greece. Applied Geochemistry, 18(6): 781-794.

Goovaerts, P. 1999. Geostatistics in soil science: state-of-the-art and perspectives. Geoderma, 89(1-2): 1-45.

Gu, Y.G., Wang, Z.H., Lu, S.H., Jiang, S.J., Mu, D.H. and Shu, Y.H. 2012. Multivariate statistical and GIS-based approach to identify source of anthropogenic impacts on metallic elements in sediments from the mid Guangdong coasts, China. Environmental Pollution, 163: 248-255.

Guo, G.H., Wu, F.C., Xie, F.Z. and Zhang, R.Q. 2012. Spatial distribution and pollution assessment of heavy metals in urban soils from southwest China. Journal of Environmental Sciences-China, 24(3): 410-418.

Hani, A. and Pazira, E. 2011. Heavy metals assessment and identification of their sources in agricultural soils of Southern Tehran, Iran. Environmental Monitoring and Assessment, 176(1-4): 677-691.

Hassan, M.M. and Atkins, P.J. 2011. Application of geostatistics with indicator kriging for analyzing spatial variability of groundwater arsenic concentrations in Southwest Bangladesh. Journal Of Environmental Science and Health Part a-Toxic/Hazardous Substances & Environmental Engineering, 46(11): 1185-1196.

Hu, Y.N., Wang, D.X., Wei, L.J. and Song, B. 2014. Heavy metal contamination of urban topsoils in a typical region of Loess Plateau, China. Journal of Soils and Sediments, 14(5): 928-935.

Lee, C. S., Li, X.D., Shi, W.Z., Cheung, S.C. and Thornton I. 2006. Metal contamination in urban, suburban, and country park soils of Hong Kong: A study based on GIS and multivariate statistics. Science of the Total Environment, 356(1-3): 45-61.

Lee, J.J., Jang, C.S., Wang, S.W. and Liu, C.W. 2007. Evaluation of potential health risk of arsenic-affected groundwater using indicator kriging and dose response model. Science of the Total Environment, 384(1-3): 151-162.

Li, G.L., Liu, G.J., Zhou, C.C., Chou, C.L., Zheng, L.G. and Wang, J.Z. 2012. Spatial distribution and multiple sources of heavy metals in the water of Chaohu Lake, Anhui, China. Environmental Monitoring and Assessment, 184(5): 2763-2773.

Lin, Y.P., Cheng, B.Y., Chu, H.J., Chang ,T.K. and Yu, H.L. 2011. Assessing how heavy metal pollution and human activity are related by using logistic regression and kriging methods. Geoderma, 163(3-4): 275-282.

Liu, E.F. and Shen, J. 2014. A comparative study of metal pollution and potential eco-risk in the sediment of Chaohu Lake (China) based on total concentration and chemical speciation. Environmental Science and Pollution Research, 21(12): 7285-7295.

Liu, E.F., Shen, J., Yang, X.D., and Zhang, E.L. 2012. Spatial distribution and human contamination quantification of trace metals and phosphorus in the sediments of Chaohu Lake, a eutrophic shallow lake, China. Environmental Monitoring and Assessment, 184(4): 2105-2118.

Liu, W. X., Li, X. D., Shen, Z. G., Wang, D. C., Wai, O. W. H. and Li, Y. S. 2003. Multivariate statistical study of heavy metal enrichment in sediments of the Pearl River Estuary. Environmental Pollution, 121(3): 377-388.

Liu, X.M., Wu, J.J. and Xu, J.M. 2006. Characterizing the risk assessment of heavy metals and sampling uncertainty analysis in paddy field by geostatistics and GIS. Environmental Pollution, 141(2): 257-264.

MacDonald, D.D., Ingersoll, C.G. and Berger, T. A. 2000. Development and evaluation of consensus-based sediment quality guidelines for freshwater ecosystems. Archives of Environmental Contamination and Toxicology, 39(1): 20-31.

McGrath, D., Zhang, C.S. and Carton, O.T. 2004. Geostatistical analyses and hazard assessment on soil lead in Silvermines area, Ireland. Environmental Pollution, 127(2): 239-248.

Nemr, A.E., Khaled, A. and Sikaily, A.E. 2006. Distribution and statistical analysis of leachable and total heavy metals in the sediments of the Suez Gulf. Environmental Monitoring and Assessment, 118(1): 89-112.

Senesi, G.S., Baldassarre, G., Senesi, N. and Radina, B. 1999. Trace element inputs into soils by anthropogenic activities and implications for human health. Chemosphere, 39(2): 343-377.

Shi, G.T., Chen, Z.L., Xu, S.Y., Zhang, J., Wang, L., Bi, C.J. and Teng, J.Y. 2008. Potentially toxic metal contamination of urban soils and roadside dust in Shanghai, China. Environmental Pollution, 156(2): 251-260.

Shi, J.C., Wang, H.Z., Xu, J.M., Wu, J.J., Liu, X.M., Zhu, H.P. and Yu, C.L. 2007. Spatial distribution of heavy metals in soils: a case study of Changxing, China. Environmental Geology, 52(1): 1-10.

Shrestha, S. and Kazama, F. 2007. Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. Environmental Modelling & Software, 22(4): 464-475.

Simasuwannarong, B., Satapanajaru, T., Khuntong, S. and Pengthamkeerati, P. 2012. Spatial distribution and risk assessment of As, Cd, Cu, Pb, and Zn in topsoil at Rayong Province, Thailand. Water Air and Soil Pollution, 223(5): 1931-1943.

Wang, C., Liu, S.L., Zhao, Q.H., Deng, L. and Dong, S.K. 2012. Spatial variation and contamination assessment of heavy metals in sediments in the Manwan Reservoir, Lancang River. Ecotoxicology and Environmental Safety, 82: 32-39.

Wang, H.Y. and Lu, S.G. 2011. Spatial distribution, source identification and affecting factors of heavy metals contamination in urban-suburban soils of Lishui city, China. Environmental Earth Sciences, 64(7): 1921-1929.

Yin, H.B., Deng, J.C., Shao, S.G., Gao, F., Gao, J.F. and Fan, C.X. 2011. Distribution characteristics and toxicity assessment of heavy metals in the sediments of Lake Chaohu, China. Environmental Monitoring and Assessment, 179(1-4): 431-442.

Yuan, G.L., Sun, T.H., Han, P., Li, J. and Lang, X.X. 2014. Source identification and ecological risk assessment of heavy metals in topsoil using environmental geochemical mapping: Typical urban renewal area in Beijing, China. Journal of Geochemical Exploration, 136: 40-47.

Zhang, C.S. and McGrath, D. 2004. Geostatistical and GIS analyses on soil organic carbon concentrations in grassland of southeastern Ireland from two different periods. Geoderma, 119(3-4): 261-275.

Zhang, W.G., Feng, H., Chang, J.N., Qu, J.G., Xie, H.X. and Yu, L.Z. 2009a. Heavy metal contamination in surface sediments of Yangtze River intertidal zone: An assessment from different indexes. Environmental Pollution, 157(5): 1533-1543.

Zhang, X.Y., Lin, F.F., Wong, M. T.F., Feng, X.L. and Wang, K. 2009b. Identification of soil heavy metal sources from anthropogenic activities and pollution assessment of Fuyang County, China. Environmental Monitoring and Assessment, 154(1-4): 439-449.

Zheng, L.G., Liu, G.J., Kang, Y. and Yang, R.K. 2010. Some poten-

tial hazardous trace elements contamination and their ecological risk in sediments of western Chaohu Lake, China. Environmental Monitoring and Assessment, 166(1-4): 379-386.
Zhou, F., Guo, H.C. and Hao, Z.J. 2007a. Spatial distribution of heavy metals in Hong Kong's marine sediments and their human impacts: A GIS-based chemometric approach. Marine Pollution Bulletin, 54(9): 1372-1384.
Zhou, F., Guo, H.C. and Liu, L. 2007b. Quantitative identification and source apportionment of anthropogenic heavy metals in marine sediment of Hong Kong. Environmental Geology, 53(2): 295-305.