Nature Environment and Pollution Technology
An International Quarterly Scientific Journal

ISSN: 0972-6268	Vol. 13	No. 1	pp. 73-77	2014

Original Research Paper

Source Identification of Heavy Metals in River Sediments by Using Factor Analysis in Combination with K-means Cluster Analysis

Linhua Sun* and Herong Gui
School of Earth Science and Engineering, Suzhou University, Suzhou-234000, China (*Corresponding Author)

Nat. Env. & Poll. Tech.
Website: www.neptjournal.com
Received: 14-6-2013
Accepted: 24-7-2013

Key Words:
Heavy metals
Sediments
Factor analysis
K-means cluster analysis

ABSTRACT

Source identification of heavy metals in river sediments and spacial polluted sample separation are important for either river system protection or remediation. In this study, concentrations of five heavy metals (Fe, As, Cr, Cu, Pb) in the sediments from Bianhe River, northern Anhui Province, China have been measured and analysed by factor and cluster analysis for tracing their sources. The results suggest that there are three kinds of sources for these metals: natural, anthropogenic one and two. Fe is mainly contributed by natural source, 14 points are mainly polluted by As and Cr and four points are polluted by Cu and Pb. In comparison with the location of sampling, the anthropogenic As and Cr are mainly supplied by urban activities, whereas the anthropogenic Cu and Pb are mainly related to traffic, to a lesser extent, point pollution. The study demonstrated that a combination use of factor and K-means cluster analysis can provide reliable information for identifying the source of heavy metals and the location with specific pollution.

INTRODUCTION

Rivers play important roles for the development of human society because they are main source of water supply during the long history. However, with the development of modern industry and agriculture, large amount of uncontrolled metal inputs from either point or nonpoint sources have been contributed to the river systems, and cause a considerable number of the world's rivers severely contaminated (Theofanis et al. 2001, Arribere et al. 2002, Akcay et al. 2003, Susana et al. 2005). Therefore, heavy metal pollution of rivers had attracted an increased attention with the studies focused on heavy metals in river water, sediments and living creatures (Begum et al. 2009, Peng et al. 2009, Alinnor & Obiji 2010).

In aquatic environment, heavy metal is usually distributed as water-soluble species, colloids, suspended forms and sedimentary phases. Among these phases, more than 90% of the anthropogenic metals are bound to particulate matters and deposited on the bed, synchronously with the debris from the weathered mother rock and soil in the catchment (Gomez-Parra et al. 2000, Amin et al. 2009). And therefore, sediments at the bottom are important in the river systems because they can act both as sinks and secondary sources of trace metals (Botsou et al. 2011).

Identification of the source of anthropogenic and natural heavy metals in the sediments is important for either river system protection or remediation. Therefore, some statistical approaches (e.g. factor and principal component analysis) and other methods (e.g. normalizing the metals concentrations by the inert metals) have been used for solving this issue (Liaghati et al. 2003, Chen et al. 2004). Additionally, separation of natural and anthropogenic affected samples is also an important work; take an instance, if we want to establish the natural background, the anthropogenic affected samples should be removed first (Reimann et al. 2005, Apitz et al. 2009) because natural distribution of these elements is always overlain by anthropogenic inputs. Moreover, before remediation, we do not only want to know the degree of the pollution but also where or which point has been polluted and by what. However, higher concentration of metals does not mean more contribution from anthropogenic activities because of the inhomogeneous nature of mother rocks.

In this study, concentrations of five heavy metals (Fe, As, Cr, Cu, Pb) in the sediments from Bianhe River, northern Anhui Province, China have been analysed, and a combination use of factor and K-means cluster analysis has been applied for identification of the sources of heavy metals and separating the natural and anthropogenic affected samples.

MATERIALS AND METHODS

Bianhe River is an artificial river, which was built between 1966 and 1968. The length of the river is 127 km and flowing through Suzhou, Lingbi, Sixian and Sihong in northern Anhui Province, China (Fig. 1). It is not only an important water supplier for agricultural and industrial use, but also an important transport channel in the area. However, with the development of the economic society along with the river, large amount of wastewater from agricultural, industrial and domestic activities has been discharged into the river making it polluted.

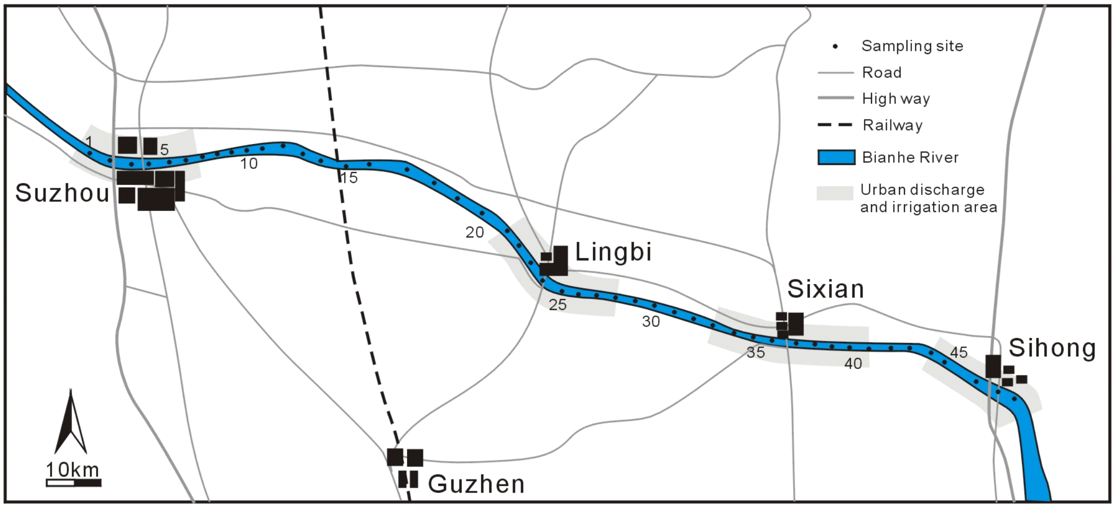

Fig. 1: Location of the study area and the distribution of the sampling sites.

A total of 45 surface sediments (0 to 10 cm) were collected from the river by using a home-made sediment sampler in May, 2011. The samples were first naturally air-dried for one day and debris of animals and plants was removed by hand. The samples were then sieved to obtain a grain size fraction smaller than 63 μm by hand with nylon sieve after parching in a drying oven with forced convection for 24 h at 80°C.

About 1 g samples were dissolved with 30 mL HNO_3 and HF mixture (1:2) in an open container for 10 h at room temperature and then heated at 50°C for 2 d. The samples were then cooled at room temperature, and then filtered through 0.45 μm paper filter and diluted to 50 mL with HNO_3 (3%). The concentrations of Fe, As, Cr, Cu and Pb were determined by inductively coupled plasma-mass spectrometry (ICP-MS) (Agilent 7700). Standard reference samples (GBW08301, river sediment) were analysed simultaneously for calibration.

Statistical analysis of the data includes factor and K-means cluster analysis, which were performed by Mystat (version 12). The former was applied for identifying the correlation between metals and the probability of sources of them in combination with the use of density plots. K-means cluster analysis was used for separating the natural and anthropogenic affected samples and, absolute distance between clusters was chosen as the method.

RESULTS AND DISCUSSION

Descriptive statistics: The analytical results of these five kinds of heavy metals are synthesized in Table 1. As can be seen from the table, the concentrations of Fe, As, Cr, Cu and Pb are 1.44-5.17%, 7.51-72.5, 77.9-291, 23.9-121 and 7.48-113 mg/kg, respectively. In comparison with the soil environmental background values of China (CEPA 1990), the enrichment factors for average concentrations of Fe, As, Cr, Cu and Pb are 1.10, 2.22, 2.87, 2.44 and 1.27, respectively, and indicating that Fe and Pb are light-moderate pollution, whereas As, Cr and Cu are moderate pollution (Hakanson 1980). Moreover, Fe has the most insignificant spatial variation because it has the smallest coefficient of variation (0.22) relative to other heavy metals with CVs equal to 0.52 (As), 0.26 (Cr), 0.31 (Cu) and 0.52 (Pb), respectively.

Additionally, Fe obtains the lowest skewness value

Table 1: Descriptive statistics of heavy metal concentrations in sediments of Bianhe River.

	Min	Max	Mean	SD	EF	CV	SK	AD test	p-value
As (mg/kg)	7.51	72.5	24.9	12.8	2.22	0.52	1.49	1.54	<0.01
Cr (mg/kg)	77.9	291	175	45.8	2.87	0.26	0.68	0.85	0.03
Cu (mg/kg)	23.9	121	55.1	17.1	2.44	0.31	1.69	1.42	<0.01
Pb (mg/kg)	7.48	113	33.0	17.3	1.27	0.52	2.16	1.37	<0.01
Fe (%)	1.44	5.17	3.23	0.72	1.10	0.22	0.35	0.47	>0.15

Note: SD-standard deviation; EF-Enrichment factor (concentration/background value); CV-coefficient of variation; AD test- Anderson-Darling normal distribution test.

Table 2: Results of factor analysis (after varimax rotation).

Factor	1	2
As	-0.033	0.876
Cr	-0.020	0.891
Cu	0.829	-0.122
Pb	0.888	-0.109
Fe	0.737	0.447
Explanation of variance	40.1%	36.2%
Eigenvalue	2.049	1.760

Table 3: Results of K-means cluster analysis.

	N	As	SD	Cr	SD	Cu	SD	Pb	SD
Cluster 1	31	19.4	8.12	151	25.2	53.7	11.3	31.0	9.90
Cluster 2	14	37.7	14.2	233	32.3	47.2	10.7	26.1	12.4
Cluster 3	4	22.6	3.94	164	25.9	93.3	25.1	72.9	27.1

Table 4: Descriptive statistics of heavy metal concentrations in sediments without significant pollution.

	N	Min	Max	Mean	SD	AD	p-value
As	35	7.51	40.3	19.8	7.78	0.61	0.11
Cr	35	77.9	202	152	25.3	0.39	>0.15
Cu	45	10.6	77.7	47.6	14.2	0.27	>0.15
Pb	45	7.48	56.7	29.5	10.9	0.32	>0.15

(0.345) relative to other heavy metals and, in combination with the results obtained by Anderson-Darling normal distribution test, indicating that only Fe has normal distribution because its p-value is higher than 0.05 (95% confidence level), whereas other heavy metals cannot pass the normal distribution test because they have p-values lower than 0.05. Such results suggest that these heavy metals except for Fe are originated from multi sources, including either natural or anthropogenic, whereas the source of Fe is considered to be natural or geogenic dominant (Reimann et al. 2005, Nakic et al. 2007, Apitz et al. 2009).

Source identification: As can be seen from the density plots (Fig. 2), three peaks have been identified for As: the first one with peak value near 20 mg/kg, and is considered to be representative of natural source, whereas the second and third ones with As concentrations between 40 and 80 mg/kg are considered to be originated from anthropogenic activities. Similarly, either natural or anthropogenic sources are identified for Cr, Cu and Pb (Fig. 2).

Additionally, two factors with eigen value higher than one have been obtained by factor analysis, and the total explanation of variance is 76.2% (Table 2). The first one is 40.1% and is predominantly participated by Fe, Cu and Pb, whereas the second one is dominated by As and Cr with an explanation of 35.2%. The results suggest that at least two sources are responsible for the heavy metals in these samples.

Just because Fe has not been dramatically affected by human activities, the first factor (including Fe, Cu and Pb) is probably a natural or geogenic factor, and factor two (including As and Cr) is an anthropogenic factor. However, as mentioned in the descriptive statistics, this conclusion is probably wrong because these metals except for Fe are considered to be affected by anthropogenic activities.

Additionally, as can be seen from Table 2, Fe has higher loading in factor one (0.737) than in factor two (0.447), this probably reflects that natural process is responsible for either Cu-Pb or As-Cr, but their contribution degrees are different. Most probably, natural contributions are higher for Cu and Pb than for As and Cr in the sediment samples. Therefore, it can be concluded that at least two anthropogenic and one natural sources are responsible, one is mainly contributed of Cu and Pb, another one is main contributor of As and Cr, whereas the natural source are responsible for the concentrations of all metals.

Separation of naturally and anthropogenically affected samples: Because there are three sources responsible for the heavy metal concentrations in this study, including natural, anthropogenic one and two, the group number is therefore set to be three during K-means analysis and the results are listed in Table 3 and the profile plots are shown in Fig. 4.

As can be seen from the table, there are 31, 14 and 4 samples are classified into the first, second and third clusters, respectively. The highest mean concentrations of As and Cr are observed in cluster two, this factor can be assigned as As and Cr pollution samples (including case 23-27, 35-39 and 45-48). The highest mean concentrations of Cu and Pb are observed in cluster three, therefore this factor can be assigned as Cu and Pb pollution samples (including case 2, 4, 20 and 28). Rest of the samples are classified into cluster one, they are characterized by medium heavy metal concentrations relative to cluster two and three, and therefore, it is assigned as natural cluster.

A similar conclusion can also be archived from Fig. 3. Metal concentrations in cluster one have a small range relative to each other, As and Cr in cluster two have much higher concentrations, whereas Cu and Pb in cluster three have highest concentration. It can also be obtained from Fig. 3 that the variation of Fe concentrations in cluster one is similar to each other, whereas Fe concentrations in cluster two and three are different with As-Cr and Cu-Pb, respectively. However, there are some overlapping parts of Fe variation relative to the variation of As-Cr and Cu-Pb, indicating that although the samples in these two clusters have been affected by human activities, parts of them are contributed by natural processes.

To check of the suitability of these separations, Anderson-

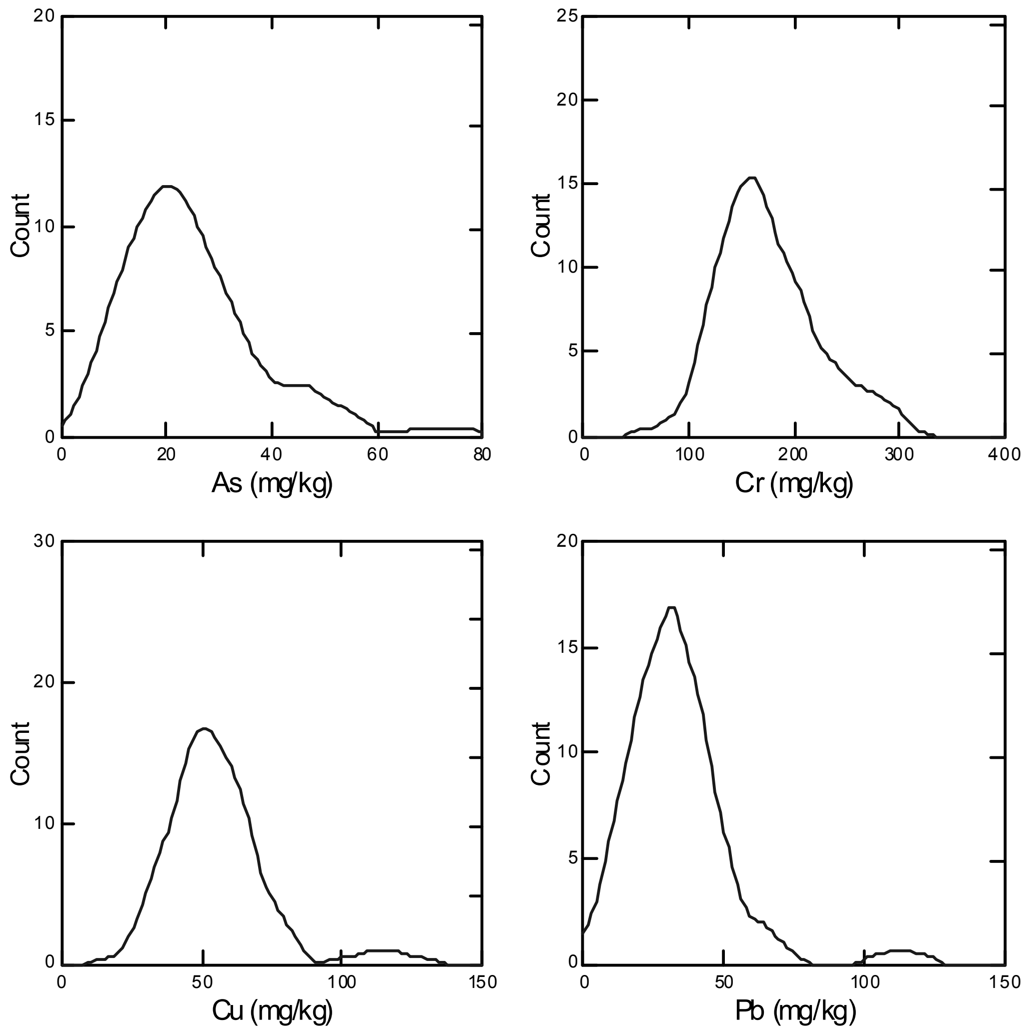

Fig. 2: Density plots of heavy metal concentrations.

Darling normal distribution test is performed for rest of the data without specific anthropogenic pollutions: samples with As-Cr pollution (14 samples) have been removed and then the As and Cr concentrations of the remaining samples (35 samples) are tested for normal distribution, whereas samples with Cu-Pb pollution (4 samples) have been removed and then the rest samples have been tested (45 samples). As can be seen from Table 4, As, Cr, Cu and Pb can all pass the normal distribution test because their p-values are all higher than 0.05 (95% confidence level).

Further discussions: In combination with the sample distribution (Fig. 1), it can be concluded that As and Cr pollutions are related to the distribution of cities, e.g. samples with As and Cr pollution are located near Lingbi, Sixian and Sihong. Therefore, wastewater discharges related to domestic and industrial activities in urban areas are considered to be their main sources. However, points with Cu and Pb pollution are limited, two of them are located near the road across the river, and they are considered to be related to traffic. Moreover, other two of them are located in normal place without any special characteristic, therefore, they are considered to be originated from point pollution.

CONCLUSIONS

Based on factor and K-means cluster analysis of concentrations of five kinds of heavy metals (Fe, As, Cr, Cu, Pb) in the sediments from Bianhe River, northern Anhui Province, China, a series of conclusions have been made:

1. There are three major sources contributing for the concentrations of these five heavy metals: Fe is mainly supplied by natural processes, As and Cr are originated from either natural processes and urban activities, whereas Cu

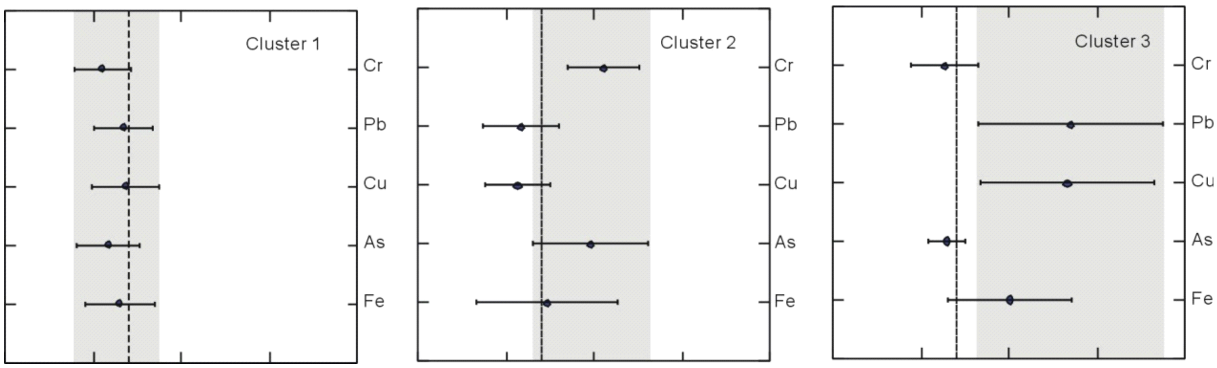

Fig. 3: Profile plots of clusters.

and Pb are contributed from natural processes and point pollution.

2. Single use of factor analysis gives only probable sources of heavy metals, and if the number of polluted samples is limited, it may give misunderstanding about their sources (e.g. Fe, Cu and Pb in this study). However, the combination use of factor and K-means cluster analysis can provide not only the probable source of heavy metals, but also can be used for tracing the location with specific pollution.

ACKNOWLEDGEMENT

This work was financially supported by National Natural Science Foundation of China (41173106), Natural Science Foundation of Anhui Province (1308085QE77), and the Program for Innovative Research Team in Suzhou University (2013kytd01).

REFERENCES

Alinnor, I.J. and Obiji, I.A. 2010. Assessment of trace metal composition in fish samples from Nworie River. Pakistan Journal of Nutrition, 9(1): 81-85.

Amin, B., Ismail, A., Arshad, A., Yap, C.K. and Kamarudin, M.S. 2009. Anthropogenic impacts on heavy metal concentrations in the coastal sediments of Dumai, Indonesia. Environmental Monitoring and Assessment, 148: 291-305.

Akcay, H., Oguz, A. and Karapire, C. 2003. Study of heavy metal pollution and speciation in Buyak Menderes and Gediz river sediments. Water Research, 37: 813-822.

Apitz, S.E., Degetto, S. and Cantaluppi, C. 2009. The use of statistical methods to separate natural background and anthropogenic concentrations of trace elements in radio-chronologically selected surface sediments of the Venice Lagoon. Marine Pollution Bulletin, 58: 402-414.

Arribere, M.A., Ribeiro, G.S., Sanchez, R.S., Gil, M.I., Roman, R.G., Daurade, L.E., Fajon, V., Horvat, M., Alcalde, R. and Kestelman, A.J. 2002. Heavy metals in the vicinity of a chlor-alkali factory in the upper Negro river ecosystem, Northern Patagonia, Argentina. Science of the Total Environment, 301: 187-203.

Begum, A., Ramaian, M., Khan, I. and Veena, K. 2009. Heavy metal pollution and chemical profile of Cauvery River water. Journal of Chemistry, 6(1): 47-52.

Botsou, F., Karageorgis, A.P., Dassenakis, E. and Scoullos, M. 2011. Assessment of heavy metal contamination and mineral magnetic characterization of the Asopos River sediments. Marine Pollution Bulletin, 62: 547-563.

CEPA (Chinese Environmental Protection Administration) 1990. Elemental background values of soils in China. Beijing: Environmental Science Press.

Chen, Z.Y., Saito, Y., Kanai, Y., Wei, T.Y., Li, L.Q. and Yao, H.S. 2004. Low concentration of heavy metals in the Yangtze estuarine sediments, China: A diluting setting. Estuarine, Coastal and Shelf Science, 60: 91-100.

Gomez-Parra, A., Forja, J.M., Delvalls, T.A., Saenz, I. and Riba, I. 2000. Early contamination by heavy metals of the Guadalquiver estuary after the Aznalcollar mining spill (SW Spain). Marine Pollution Bulletin, 40: 1115-1123.

Hakanson, L. 1980. An ecological risk index for aquatic pollution control: A sediment to logical approach. Water Research, 14(8): 975-1001.

Liaghati, T., Preda, M. and Cox, M. 2003. Heavy metal distribution and controlling factors within coastal plain sediments, Bells Creek Catchment, southeast Queensland, Australia. Environment International, 29: 935-948.

Nakic, Z., Posavec, K. and Bacani, A. 2007. A visual basic spreadsheet macro for geochemical background analysis. Ground Water, 45(5): 642-647.

Peng, J.F., Song, Y.H., Yuan, P., Cui, X.Y. and Qiu, G.L. 2009. The remediation of heavy metals contaminated sediments. Journal of Hazardous Materials, 161: 633-640.

Reimann, C., Filzmoser, P. and Garrett, R.G. 2005. Background and threshold: Critical comparison of methods of determination. Science of the Total Environment, 346: 1-16.

Susana, O.R., Daniel, D.L.R., Lazaro, L., David, W.G., Katia, D.A., Jorge, B. and Francisco, M. 2005. Assessment of heavy metal levels in Almendares River sediments-Havana City, Cuba. Water Research, 39: 3945-3953.

Theofanis, Z.U., Astrid, S., Lidia, G. and Calmano, W.G. 2001. Contaminants in sediments: Remobilisation and demobilization. Science of the Total Environment, 266: 195-202.