Environ Geol (2007) 53:295–305
DOI 10.1007/s00254-007-0644-7

ORIGINAL ARTICLE

Quantitative identification and source apportionment of anthropogenic heavy metals in marine sediment of Hong Kong

Feng Zhou · Huaicheng Guo · Lei Liu

Received: 13 December 2006 / Accepted: 9 January 2007 / Published online: 6 February 2007

Abstract Based on ten heavy metals collected twice annually at 59 sites from 1998 to 2004, enrichment factors (EFs), principal component analysis (PCA) and multivariate linear regression of absolute principal component scores (MLR-APCS) were used in identification and source apportionment of the anthropogenic heavy metals in marine sediment. EFs with Fe as a normalizer and local background as reference values was properly tested and suitable in Hong Kong, and Zn, Ni, Pb, Cu, Cd, Hg and Cr mainly originated from anthropogenic sources, while Al, Mn and Fe were derived from rocks weathering. Rotated PCA and GIS mapping further identified two types of anthropogenic sources and their impacted regions: (1) electronic industrial pollution, riparian runoff and vehicle exhaust impacted the entire Victoria Harbour, inner Tolo Harbour, Eastern Buffer, inner Deep Bay and Cheung Chau; and (2) discharges from textile factories and paint, influenced Tsuen Wan Bay and Kwun Tong typhoon shelter and Rambler Channel. In addition, MLR-APCS was successfully introduced to quantitatively determine the source contributions with uncertainties almost less than 8%: the first anthropogenic sources were responsible for 50.0, 45.1, 86.6, 78.9 and 87.5% of the Zn, Pb, Cu, Cd and Hg, respectively, whereas 49.9% of the Ni and 58.4% of the Cr came from the second anthropogenic sources.

Keywords Heavy metals · Marine sediment · Source apportionment · Enrichment factors · Multivariate statistics · GIS

Introduction

Anthropogenic heavy metals are of particular concern worldwide because of their environmental persistence and the ecological risks they pose (González-Maćias 2006). Marine sediment has often been regarded as the final carriers of heavy metals (Sin et al. 2001). Therefore, correct discrimination of natural and anthropogenic sources, identification of source profiles, and the quantitative determination of source distributions for each element become more important for marine pollution control and prevention.

In the past three decades, enrichment factors (EFs) have been widely applied to distinguish the origins of heavy metals in air, water, sediments and soil environments (Zoller et al. 1974; Covelli and Fontolan 1997; Reimann and de Caritat 2000; Glasby et al. 2004; González-Maćias et al. 2006). Three assumptions of EFs are that the normalizer has a lower variability in concentration relative to the elements of interest, reference heavy metals have approximately similar spatial distributions, and crustal element ratios are relatively homogeneous in different mediums (Reimann and de Caritat 2005). However, many studies have ignored

F. Zhou · H. Guo (✉)
College of Environmental Sciences, Peking University, Room 208, Old Geosciences Building, Beijing 100871, People's Republic of China
e-mail: hcguo@pku.edu.cn

F. Zhou
e-mail: jardon.zhou@gmail.com

L. Liu
Department of Civil Engineering, Dalhousie University, Halifax NS B3J 1Z1, Canada

these intrinsic assumptions and have directly or indiscriminately normalized element concentrations to an average crust value, despite the fact that Blaser et al. (2000) suggested that local background values may provide a more meaningful basis for calculating EFs than average crust values, and Liu et al. (2003) demonstrated that multivariate analyses could be useful in determining the normalizer for anthropogenic heavy metals. Nevertheless, high EFs do not conclusively demonstrate human influence on heavy metals when the assumptions are not met (Reimann and de Caritat 2005).

Multivariate statistical models, as supplementary methods to costly field surveys, have been successfully used to identify anthropogenic pollution and source profiles in surface water, soils and sediments (Shine et al. 1995; Facchinelli et al. 2001; Simeonov et al. 2003; Astel et al. 2006; Han et al. 2006; Wang and Qin 2006; Zhou et al. 2006). For example, principal components analysis (PCA) allows the transformation and visualization of complicated datasets into meaningful variables without losing useful data (Peré-Trepat et al. 2006). However, multivariate analyses are sensitive to outliers and the non-normal distributions of geochemical datasets; thus, it is essential to examine the probability distributions of all variables and perform appropriate data transformations (Johnson and Wichern 2002). Yet few studies have considered these important factors (Peré-Trepat et al. 2006). In addition, receptor modelling of various potential pollution sources has been widely used in numerous source apportionment studies of air pollutants (Watson et al. 2002; Hopke 2003; Song et al. 2006). As eigenvector model, multivariate linear regression of absolute principal component scores (MLR-APCS) has been applied to the apportionment of water pollution sources (Simeonov et al. 2003; Pekey et al. 2004), demonstrating the reliability of applying receptor models to the aquatic environment.

In the present study, the integrated application of multivariate statistical models and GIS was introduced to distinguish anthropogenic heavy metals in marine sediment of entire Hong Kong based on 7-year (1998–2004) datasets. The four primary objectives of this study were to (1) compare the effects of different data transformations on the multivariate analyses, (2) discriminate between natural and anthropogenic heavy metals based on the correct application of EFs, (3) identify potential sources of pollution and their spatial impacts using GIS-based PCA with VARIMAX orthogonal rotation, and (4) quantitatively determine source distributions for each element derived from MLR-APCS.

Materials and methods

Study area and sampling methods

Hong Kong is located in southern China (22°90′–22°37′N, 113°52′–114°30′E), and has a land area of 1,104 and 1,651 km^2 of marine waters. Because of the dense human population, anthropogenic activities such as reclamation, fish culture, transportation and effluent disposal heavily impact the marine environment (Tanner et al. 2000). Since the mid-1950s, rapid growth of the manufacturing industries, population expansion and rapid urban development has resulted in substantial pollution, seriously contaminating Hong Kong's marine sediment. Although effective measures have used to control pollution levels and restore habitats, heavy metals in coastal areas and typhoon shelters still pose potential ecological risks to marine organisms and human health (Wong et al. 2000).

Bottom sediment samples in the fine-grained fraction (<63 μm) were collected by taking the top 10-cm layer of marine sediment using a grab sampler [Hong Kong Environmental Protection Department (HKEPD) 2005]. The marine sediment samples were analysed by HKEPD's laboratories and the government laboratory following standard methods (American Public Health Association 1995; American Society for Testing and Materials 2001). The selected datasets contain information on ten heavy metals monitored twice annually at 59 stations (Fig. 1) over 7 years (1998–2004). Table 1 listed the descriptive statistics for these ten heavy metals.

Enrichment factors

The formula to compute EFs can be generalized from Zoller et al. (1974) as follows:

$$\mathrm{EF} = \frac{(Y/X)_{\text{sample}}}{(Y/X)_{\text{reference}}}, \tag{1}$$

where Y is the element of interest and X is the reference element. Reference elements such as Al, Li, Fe and Mn have been used as normalizers in past studies (Covelli and Fontolan 1997; Tanner et al. 2000; Cobelo-García and Prego 2003; El Nemr 2006).

Prior to computing EFs, it was essential to test whether the EFs assumptions would be met by the Hong Kong datasets. First, variabilities in the distributions of different reference elements were checked using a robust nonparametric estimate of CV* that was not affected by outliers (Reimann and de Caritat 2005):

Fig. 1 Study area and its monitoring sites

$$\mathrm{CV}^* = \frac{\mathrm{Median}\{|x_i - a|\}}{a}, \tag{2}$$

where a is the median value of each element, $\mathrm{Median}\{|x_i - a|\}$ the median value of the absolute deviations of all x_i-values from the median value. Second, we used scatterplots to investigate the spatial patterns among the normalizers with 95% confidence intervals. For the last requirement, we assumed the conservation of crustal element ratios in Hong Kong, given the small size of the region (1,651 km^2). Once the EFs were determined to be suitable for the analysis, we selected the optimal normalizer and corresponding reference values (crust or local background) that were able to authentically represent the spatial distribution of contaminants.

Table 1 Descriptive statistics [mean, standard error (SE), standard deviation (SD), minimum (Min) and maximum (Max) values] of heavy metals in marine sediment (mg/kg *dry wt.*)

	Mean	SE	SD	Min	Max
Zn	147.73	3.44	98.75	17	790
Ni	24.72	0.68	19.49	5	220
Mn	523.99	6.64	190.83	47	1,600
Pb	53.56	1.03	29.74	9	260
Cu	118.68	12.12	348.36	1	4,000
Cd	0.33	0.02	0.57	0.1	5.3
Hg	0.19	0.01	0.36	0.05	8
Cr	48.93	2.04	58.77	5	560
Fe	28,617.82	227.62	6,541.81	3,400	79,000
Al	32,728.15	315.91	9,079.42	4,100	60,000

Multivariate analyses and data transformation

The PCA model was expressed as follows (Johnson and Wichern 2002):

$$Z_{ji} = \sum_{k=1}^{p} w_{jk} p_{ki}, \tag{3}$$

where i = 1,..., n samples; j = 1,..., m elements; k = 1,..., m sources; Z_{ji} is the standardized concentration; and w_{jk} and p_{ki} the factor loadings and factor scores, respectively. To improve the interpretation of pollution patterns, Kaiser's VARIMAX orthogonal rotation of PCs was performed to maximize the simplicity of the total loadings (Brūmelis et al. 2000; Peré-Trepat et al. 2006), and based on the rotated PCA, we identified the latent pollution sources and source profiles.

The MLR-APCS assumed a linear relationship between the total mass concentration and the contributions of each element (Thurston and Spengler 1985). Source contributions to the total concentration were estimated with multiple linear regression (MLR) using the de-normalized $(\mathrm{APCS})_{ki}$ values as follows:

Table 2 Probability distributions for the original, ln-transformed and Box–Cox transformed datasets

	Original data		Ln-transformed data[a]		Box–Cox transformed data		
	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis	Rounded λ
Zn	2.11	5.90	0.47	0.18	0.01	0.53	–0.27
Ni	5.38	36.84	1.04	4.71	0.04	2.83	–0.30
Mn	1.33	4.51	–0.84	3.61	0.09	1.80	0.39
Pb	2.23	7.61	0.48	0.47	0.02	0.57	–0.31
Cu	7.14	58.01	0.99	0.98	–0.07	–0.27	–0.32
Cd	5.32	33.89	1.26	1.09	–0.44	–1.51	–0.88
Hg	13.37	271.28	1.08	0.85	–0.22	–1.17	–0.50
Cr	5.22	31.30	1.70	5.22	–0.05	3.81	–0.43
Fe	0.07	5.15	–2.31	11.41	0.34	6.40	–
Al	–0.21	–0.12	–1.42	3.78	–0.05	–0.20	1.20

[a] $y = \ln(x)$

$$M_i = \delta_0 + \sum_{k=1}^{p} \zeta_k (\mathrm{APCS})_{ki} = \delta_0 + \sum_{k=1}^{p} \zeta_k \left(p_{ki} + \sum_{j}^{m} B_{kj} \frac{\overline{C}_j}{\sigma_j}\right), \tag{4}$$

where $\overline{C}_j$ and σ_j are the mean concentration and the standard deviation for element j, respectively, M_i the observed mass concentration in sample i, ζ_k the linear regression coefficients and δ_0 the contribution of unidentified sources.

Datasets in environmental science generally have far from normal or lognormal distributions (Zhang 2006); thus, the Box–Cox transformation, a power transformation, was used to correct the non-normality (Box and Cox 1962) as follows:

$$f_\lambda(b) = \begin{cases} \dfrac{b^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \ln(b), & \lambda = 0 \end{cases}, \tag{5}$$

where b is the original data point, f_λ (b) the transformed value and λ is the optimal coefficient.

Results and discussion

Data pretreatment

To improve the effectiveness of the multivariate analyses, the following data pretreatments were performed: (1) missing data were estimated using average values, (2) values below the limits were replaced by the limits of detection (Fharnham et al. 2002), and (3) the normality of each variable's distribution was checked by analysing kurtosis and skewness (Johnson and Wichern 2002). Most variables were heavily skewed, with kurtosis values much greater than zero ($p < 0.05$), as shown in Table 2. Ln- and Box–Cox transformations of these variables improved normality except for Fe and Al, and Fe, respectively. Compared with the ln-transformation, the Box–Cox transformation significantly reduced the skewness of the data, which was beneficial to the multivariate analysis. Meanwhile, the negative effects of outliers in the heavily skewed raw data were reduced by the Box–Cox transformation (Zhang 2006). Finally (4) the Box–Cox transformed datasets were standardized for PCA to minimize the effects of differences in measurement units and variance and to render the data dimensionless (Johnson and Wichern 2002).

Differentiation between natural and anthropogenic elements

According to a robust nonparametric estimate, the CV* for the reference elements Mn, Fe, Al and typical anthropogenic pollutants Pb, Cu and Hg were 22, 14, 18.1, 27, 57 and 50%, respectively, indicating that the variability of reference elements, especially Fe, was significantly less than that of anthropogenic pollutants. Thus, the EFs results were expected to be predominantly influenced by the distribution of polluted elements rather than reference elements. In the linear regression plots between reference elements (Mn, Fe, Al; Fig. 2), most data points fell between the upper and lower 95% prediction limits; the strength of the positive linear relationship in terms of the correlation coefficient (R^2) partially confirmed the spatial distribution pattern similarities of all the typical normalizers. Moreover, compared with the mapping of regional distributions in Reimann and de Caritat (2005), scatter plots of linear regressions with 95% confidence intervals would accurately and quantitatively reflect spatial relationships because a

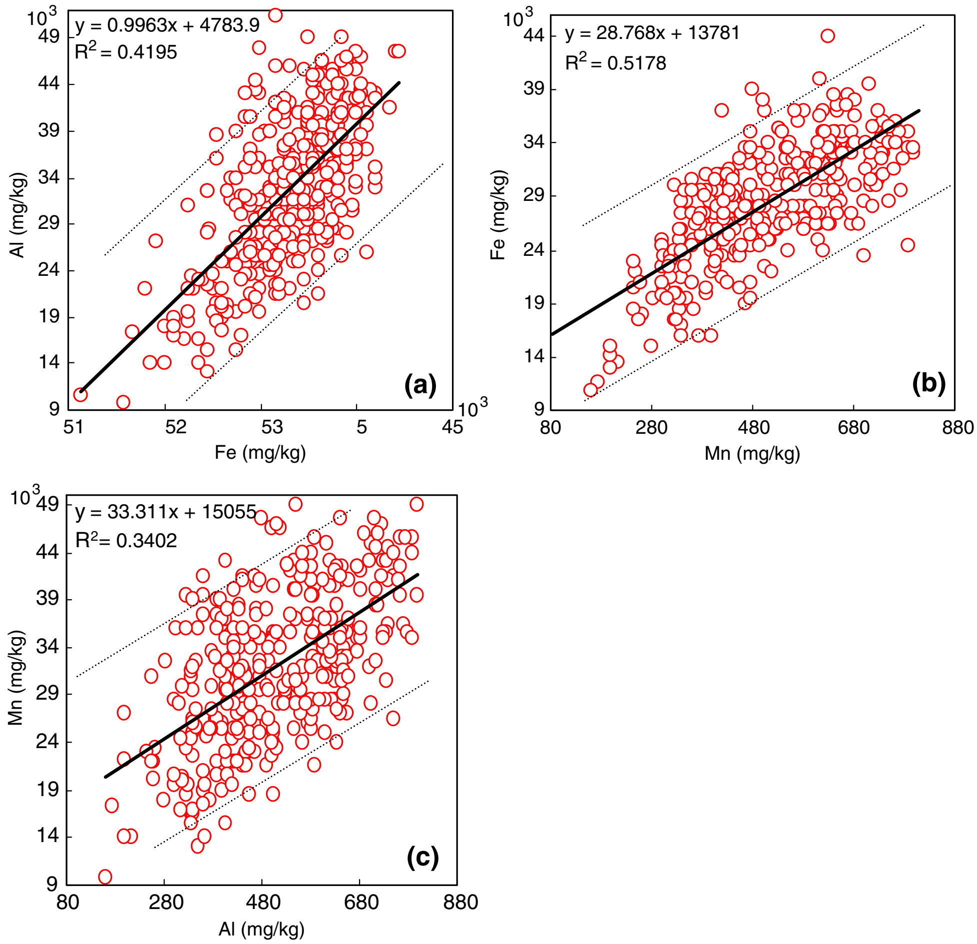

Fig. 2 Scatter plots of the linear regression models between different annual means (N = 413): Al–Fe (**a**), Mn–Fe (**b**) and Mn–Al (**c**). *Note:* dotted lines indicate the upper 95% prediction interval based on the linear regression model; the outliers were deleted when doing linear regression

few outliers that can skew analyses are inevitable in huge datasets.

The above investigations indicated that EFs were effective in this study if based on the congenial normalizer and associated reference values. To select the best normalizer, the inverse distance weighted (IDW) technique was employed to create continuous surfaces to accurately identify the overall spatial patterns. As one of typical pollutants, Pb was chosen to compare the effect of using different normalizers (Fe, Mn and Al). According to Fig. 3a, high concentrations of Pb were located around Victoria Harbour, inner Tolo Harbour, Cheung Chau (SS7) and inner Deep Bay. When computing EFs using the Pb/Mn, Pb/Fe and Pb/Al ratios, the spatial patterns were approximately the same as with Pb, and most high EFs-values also occurred throughout Victoria Harbour, inner Tolo Harbour, Cheung Chau and inner Deep Bay. However, some spurious values of Pb/Mn and Pb/Al ratios were also generated: higher values occurred around southern Lantau Island and outer Deep Bay (Fig. 3d), where Pb concentrations were lower, and lower values were found along outer Tolo Harbour (Fig. 3b). Fe seemed to be the best normalizer (Fig. 3c). To determine the corresponding reference mediums (local background or crust values), we compared the dissimilarity between the Pb/Fe ratio in average crust (Fig. 3c) and in local background samples (Fig. 4a); their spatial distributions and ratio values never differed significantly. However, the Pb/Mn ratio (Fig. 4b) for the local background was much lower than that for the average crust, further suggesting the advantage of using Fe as the normalizer for the Hong Kong sediments. The EFs of all other heavy metals relative to Fe in average crust and in local background indicated that for most heavy metals, except Pb and Al, the values were underestimated using average crust as a reference medium (Fig. 5). EFs based on local background were more conservative for differentiating between natural and anthropogenic elements in Hong Kong (Fig. 5). Figure 5 showed that Al, Mn and Fe originated primarily from natural sources, whereas Zn, Ni, Pb, Cu, Cd, Hg and Cr were due to anthropogenic impacts.

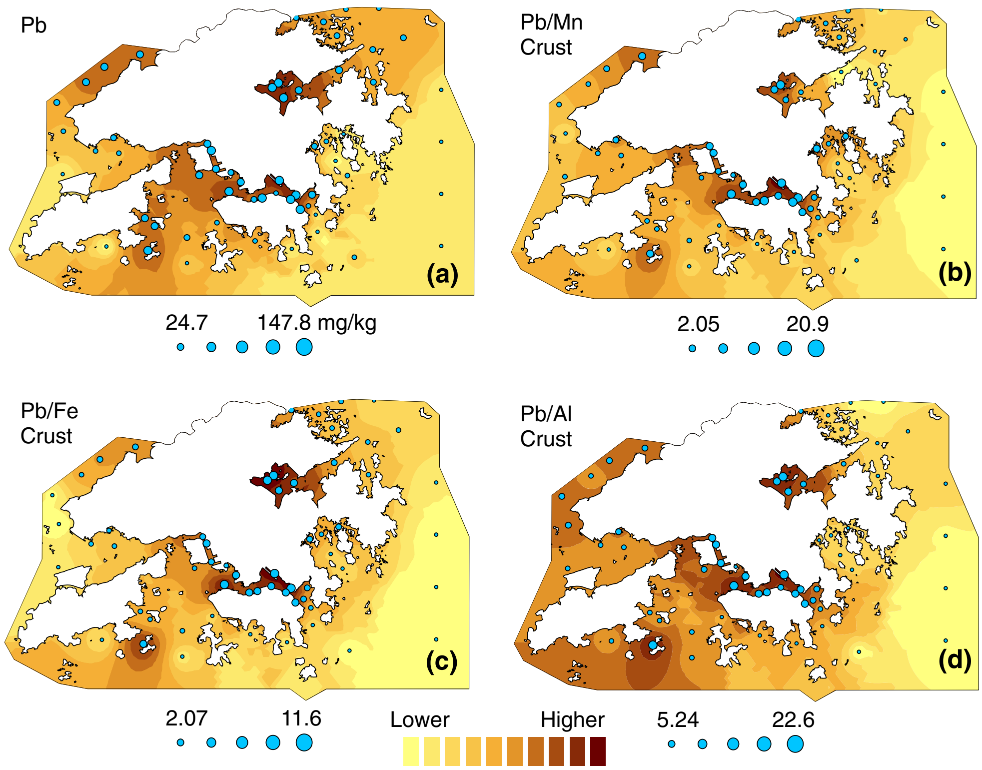

Fig. 3 Regional distributions of Pb concentrations (**a**), the EFs calculated for Pb relative to the Pb/Mn (**b**), Pb/Fe (**c**) and Pb/Al (**d**) ratio in average crust. *Note:* the legend for (**a**) and (**b**) are not same, divided with smart quantile of value

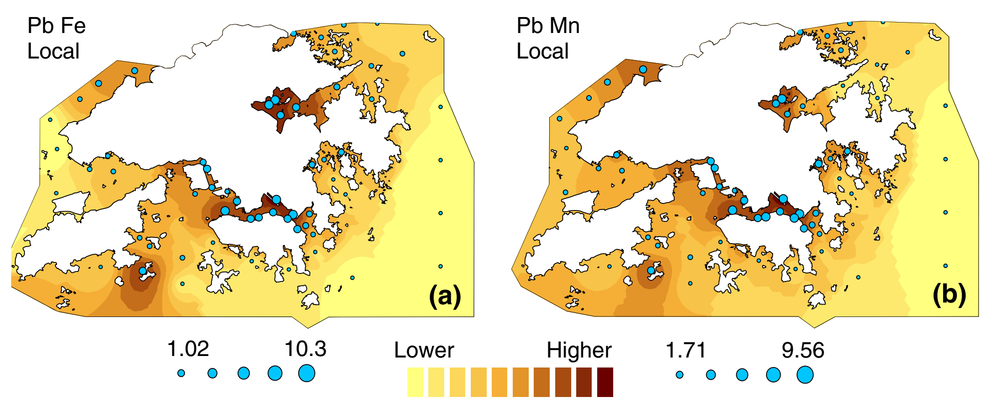

Fig. 4 Regional distributions of the EFs calculated for Pb relative to the Pb/Fe ratio in average crust (**a**), and Pb/Mn ratio in local background (**b**). *Note:* the legend for (**a**) and (**b**) are same, divided with equal interval of value

Potential source identification

Before performing PCA, Kaiser–Meyer–Olkin (KMO) and Bartlett's sphericity tests were used to examine the validity of PCA (Zhou et al. 2006). KMO and Bartlett's results were 0.84 and 8,426 ($df = 45$, $p < 0.05$), respectively, indicating that PCA would be effective in reducing dimensionality.

The PCA with VARIMAX rotation determined three latent pollution sources shaping the pollution patterns of heavy metals in marine sediment and explaining 85.35% of the total variance (Table 3). VF1 (53.28% of the total variance) had strong positive loadings on Zn, Pb, Cu, Cd and Hg and moderately positive loadings on Cr, which represented anthropogenic sources according to the EF results; VF2 (24.27% of the total variance) had strong positive loadings on Mn, Fe and Al, which distinctly represented natural sources, while VF3 (7.80% of the total variance) had strong positive loadings on Ni and Cr, likely representing other human impacts. In addition, Table 3 indicated that analysis of non-transformed datasets also produced three pollution sources, but with dissimilar source profiles and less of the total variance explained, strongly indicating that the outliers in datasets greatly influenced the multivariate analysis results (Peré-Trepat et al. 2006). We then made use of factor scores and GIS techniques to display the spatial

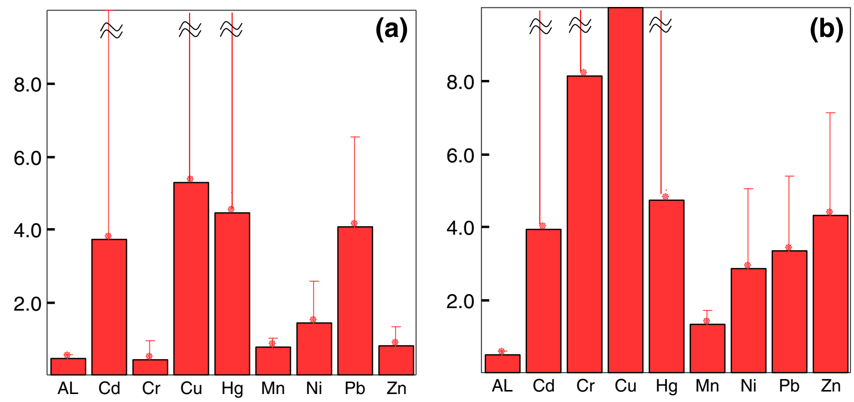

Fig. 5 Average and standard deviation of EFs of heavy metals relative to Fe in average crust (**a**), and in local background (**b**). *Note:* Error bars show +1/–1 SD; values were normalized using Fe and local background levels in Hong Kong (Lau et al. 1993) and earth crust average as references (Wedepohl 1995). ''Approximately equal'' symbol indicates that the value is greater than the maximum scale

Table 3 Loadings of ten variables on VARIMAX rotated factors of Original and different transformed datasets

Parameters	Box–Cox transformed			Ln-transformed			Original		
	VF1	VF2	VF3	VF1	VF2	VF3	VF1	VF2	VF3
Zn	**0.90**[a]	–0.23	0.25	**0.92**	0.22	0.19	0.61	0.14	**0.68**
Ni	0.25	–0.31	**0.85**	0.33	0.33	**0.84**	**0.92**	0.15	0.07
Mn	0.24	**0.86**	–0.15	–0.23	**0.88**	0.09	–0.10	**0.81**	–0.22
Pb	**0.86**	–0.36	0.12	**0.89**	0.32	0.02	0.40	0.18	**0.78**
Cu	**0.91**	0.15	0.29	**0.90**	–0.14	0.35	**0.88**	–0.08	0.31
Cd	**0.90**	0.03	0.03	**0.89**	–0.07	0.23	**0.87**	–0.06	0.40
Hg	**0.81**	0.26	0.31	**0.83**	–0.20	0.28	0.16	–0.17	**0.76**
Cr	0.61	–0.15	0.73	0.65	0.13	**0.71**	**0.95**	0.01	0.24
Fe	0.02	**0.69**	–0.55	–0.02	**0.79**	0.40	0.10	**0.85**	–0.03
Al	–0.35	**0.78**	–0.11	0.35	**0.80**	–0.01	0.04	**0.79**	0.31
Eigenvalue[b]	4.46	2.22	1.86	4.62	2.38	1.67	3.85	2.11	2.11
Cumulative (%) variance	53.28	77.54	85.35	54.96	78.90	86.74	49.96	70.94	80.77

[a] The loadings whose absolute value is more than 0.65 of the total variance was bold

[b] Eigenvalues after VARIMAX rotation

impacts of pollution sources on each monitoring site. Although a few errors occurred in spatial interpolation using IDW, Fig. 6 presented most of the local patterns. The areas polluted by the first pollution sources (VF1) were the entire Victoria Harbour, inner Tolo Harbour, the Eastern Buffer, inner Deep Bay and Cheung Chau (Fig. 6a). According to Owen et al. (2000), Tam and Wong (2000), Tanner et al. (2000) and HKEPD (2005), Victoria Harbour, Tolo Harbour and the Eastern Buffer have been affected by numerous industrial influents generated by the rapidly growing electronic industries, including electroplating, dyeing and printed circuit board manufacturing. In addition, inner Deep Bay receives discharge from the Shenzhen River and Yuen Long River and is also influenced by riparian runoff and vehicle exhaust (HKEPD 2005; Qu and Kelderman 2001). Cheung Chau is an old harbour shelter in Hong Kong, affected by boat antifouling paints (El Nemr et al. 2006). Thus, the main human impacts in these areas were industrial pollution, riparian runoff and vehicle exhaust. The regions impacted by natural sources (VF2) were generally distributed throughout the entire area, especially around Tolo Harbour and the entire Mirs Bay (Fig. 6b). Most monitoring sites in this area were relatively far from dense populations and pollution sources, and were mainly influenced by the weathering of parent rock (Cobelo-García and Prego 2003), further corroborating the EF results. The regions influenced by VF3 were mainly located in Tsuen Wan Bay (VS10), the Kwun Tong typhoon shelter (VS14) and Rambler Channel (VS17 and VS9), sites near the old industrial centres (Fig. 6c; HKEPD 2005); thus, the principal human impacts on these sites were effluent discharge from textile factories and paint (Owen et al. 2000).

The regions affected by the various pollution sources varied annually (Fig. 7a). However, some monitoring sites were seriously polluted in all years; ES3, ES5, TS2, TS3, TS7, VS6, VS12–VS14 and VS17 were influenced by the first type of pollution source, MS4–MS6, MS17, TS2, TS4 and TS5 by the second type, and

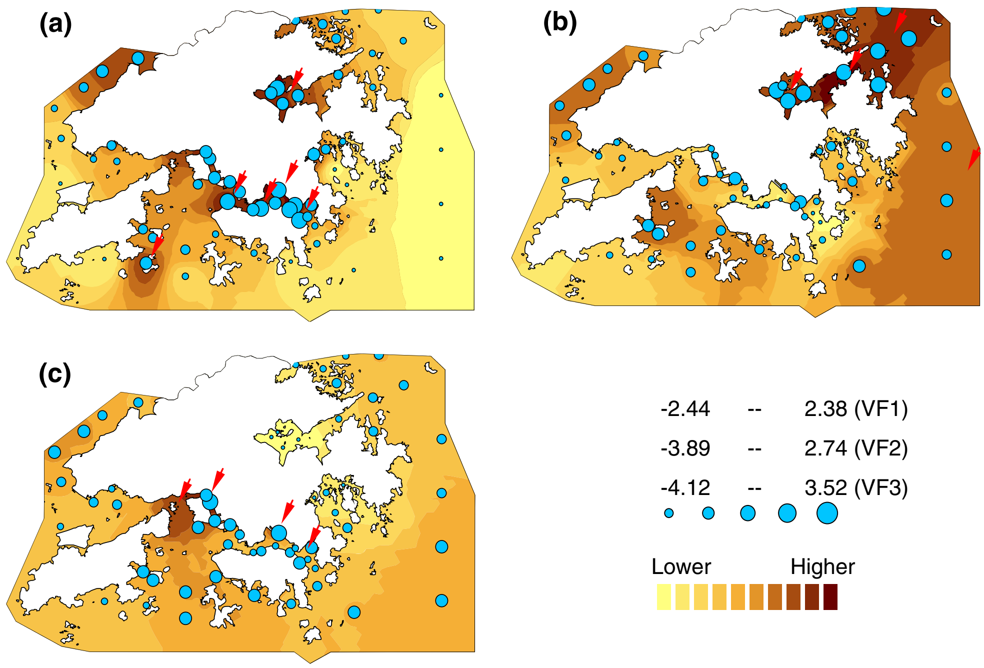

Fig. 6 Spatial interpolations of factor scores of VF1 (**a**), VF2 (**b**) and VF3 (**c**) using IDW. *Note:* the red arrow represents the "hot spot"

TS10, VS9, VS14 and VS17 by the third type. We also compared PCA results based on Box–Cox transformed datasets with those based on ln-transformed or original datasets. Although the differences in impacted regions between Box–Cox transformed (Fig. 7a) and ln-transformed datasets (Fig. 7b) were not initially distinct, the factor scores based on the ln-transformed datasets were more influenced by contaminated samples (e.g. VS13, VS14), which underestimated or neglected the degree of contamination at some obviously contaminated sites (e.g. TS2, TS3, TS4). This influence became more significant when datasets were not transformed before standardization (Fig. 7c): samples with larger factor scores obscured other differences among the remaining samples because they appeared close to the *y*-axis, which complicated further interpretation. These results also demonstrated that the Box–Cox transformation was effective at normalizing the data, thus decreasing the effects of outliers and facilitating the accurate identification of similar contamination patterns (Peré-Trepat et al. 2006; Zhang 2006).

Source contributions for each element

After determining the possible sources and source profiles by rotated PCA, source contributions of anthropogenic heavy metals were then computed using MLR-APCS, a proven effective approach for supplying quantitative information regarding the contributions of each source type (Pekey et al. 2004; Song et al. 2006). In the present study, the ratio between mean estimated and observed values of each element was used to test the accuracy of MLR-APCS. The results (Table 4) showed that all estimated values, except that of Cd, were close to measured values and varied between 0.92 (Hg) and 1.01 (Zn). Most uncertainties in prediction were less than 8%. However, the higher uncertainty of Cd (32%) was derived from the models themselves and the Cd dataset quality (e.g. model assumptions and missing data). Although Cd and Hg had moderate correlation coefficients (R^2 = 0.76 and 0.76), the others often had higher values of R^2, from 0.88 to 0.93, suggesting acceptable results. Also, these R^2 values were very close to the original PCA communality values.

Industrial pollution, riparian runoff and vehicle exhaust were the main sources of most elements: 50.0, 45.1, 86.6, 78.9 and 87.5% of Zn, Pb, Cu, Cd and Hg, respectively; pollution from textile factories and paint were the primary contributors to Ni (49.9%) and Cr (58.4%). Unidentified sources also contributed three elements to marine sediment, Zn, Ni and Pb at 33.0, 37.0, 45.1%, respectively, representing another major source of pollution along with those identified above (Simeonova et al. 2003); therefore, field surveys are essential to further identify the sources of this pollution.

Conclusion

Box–Cox transformation was more effective at improving normality characteristics and decreasing the negative affects of outliers on multivariate analyses. MLR-APCS was proved as a feasible method in source

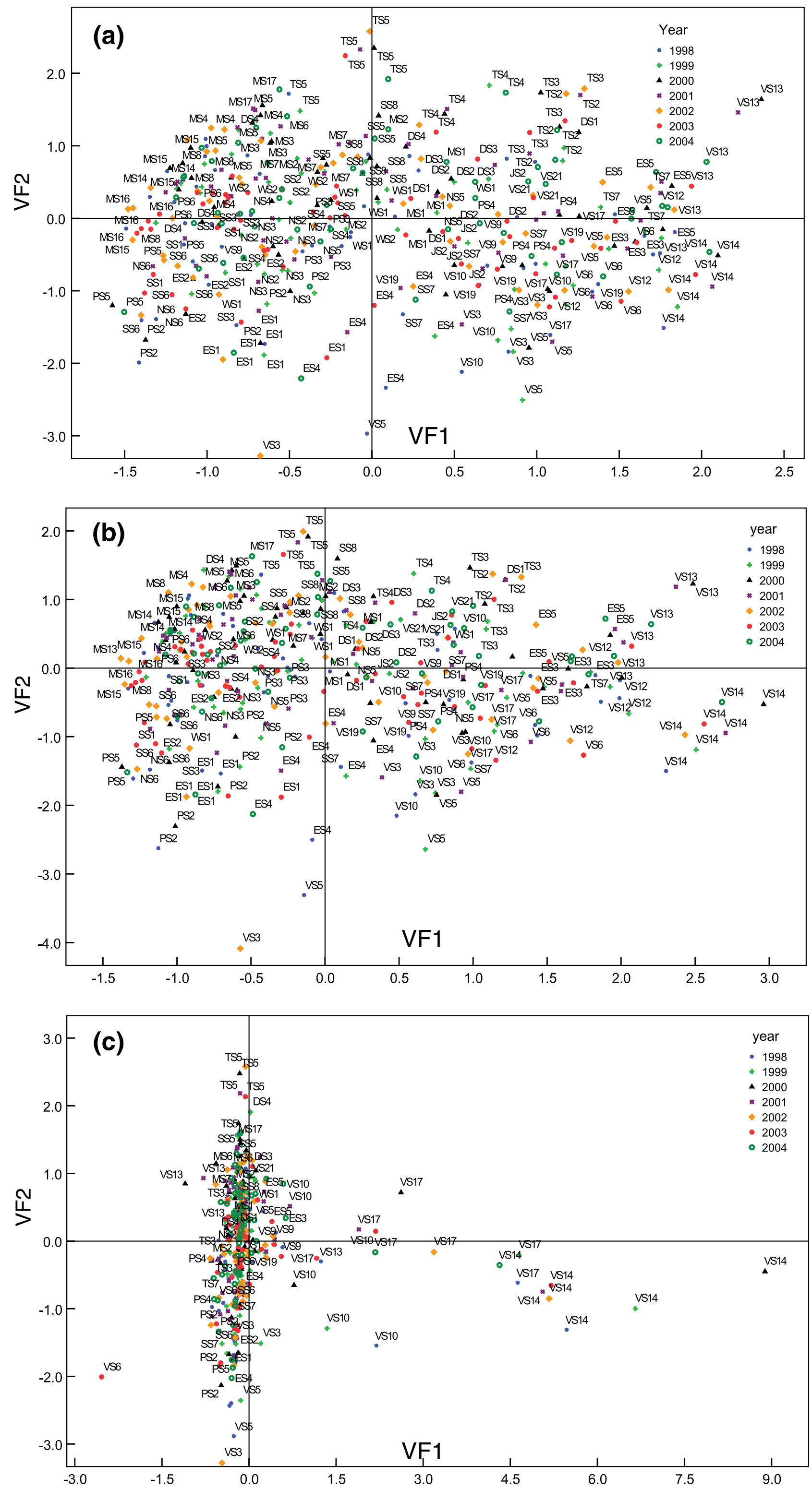

Fig. 7 VF1 verses VF2 factor scores for Box–Cox transformed (**a**), Ln-transformed (**b**) and original (**c**) datasets in different years

apportionment of heavy metals in marine sediment; moreover, continuous surface maps generated by GIS were also useful for accurately understanding spatial distribution patterns of contaminants.

After verifying the intrinsic assumptions, EFs were confirmed to be suitable for differentiating the natural or anthropogenic sources of heavy metals in marine sediment of Hong Kong, and the optimal normalizer

Table 4 Source contribution (in %) for anthropogenic heavy metals using MLR-APCS

	VF1	VF2	VF3	US	E/O[a]	R^2	Communality[b]
Zn	50.0	0.7	16.3	33.0	1.01	0.93	0.93
Ni	12.3	0.7	49.9	37.0	1.00	0.88	0.88
Pb	45.1	1.0	7.5	46.4	1.00	0.88	0.88
Cu	86.6	0.9	12.5	0.0	0.93	0.89	0.93
Cd	78.9	0.0	21.1	0.0	1.32	0.76	0.81
Hg	87.5	1.9	10.5	0.0	0.92	0.76	0.82
Cr	41.1	0.5	58.4	0.0	0.99	0.93	0.93

VF1 industrial pollution, riparian runoff and vehicle exhaust, *VF2* parent rocks weathering, *VF3* discharge from textile factories and paint, *US* unidentified sources

[a] Mean ratio of average estimated to observed values of element

[b] Derived from rotated PCA

and corresponding reference values were Fe and local background values, respectively. The EFs results showed that Al, Mn and Fe originated from natural sources, whereas Zn, Ni, Pb, Cu, Cd, Hg and Cr were due to anthropogenic sources. The first anthropogenic pollution sources were industrial pollution, riparian runoff and vehicle exhaust; these sources primarily impacted the areas around all of Victoria Harbour, inner Tolo Harbour, the Eastern Buffer, inner Deep Bay and Cheung Chau, and their contributions to Zn, Pb, Cu, Cd and Hg were 50.0, 45.1, 86.6, 78.9 and 87.5%, respectively. The second human pollution sources were discharge from textile factories and paint, which seriously influenced Tsuen Wan Bay, the Kwun Tong typhoon shelter and Rambler Channel; their contributions to Ni and Cr were 49.9 and 58.4%, respectively. Further studies should focus on regional field surveys to identify additional pollution sources and crustal element ratios in different mediums.

Acknowledgments We would like to thank Prof. Shu Tao and Mr. Yu Yang for their valuable suggestions. This paper was supported by the China Scholarship programs (2006100766) for Feng Zhou and the National Basic Research Program (No.2005CB724205). The opinions in this paper are those of the authors and do not reflect the views or policies of HKEPD.

References

American Public Health Association (1995) Standard methods for the examination of water and wastewater, 19th edn. Washington, DC

American Society for Testing and Materials (2001) American society of testing and materials standards. New York

Astel A, Biziuk M, Przyjazny A, Namiesnik J (2006) Chemometrics in monitoring spatial and temporal variation in drinking water quality. Water Res 8:1706–1716

Blaser P, Zimmermann S, Luster J, Shotyk W (2000) Critical examination of trace element enrichments and depletions in soils: As, Cr, Cu, Ni, Pb, and Zn in Swiss forest soils. Sci Total Environ 249:257–280

Box GEP, Cox DR (1964) An analysis of transformations. J R Stat Soc A Sta Ser B 26:211–252

Brūmelis G, Lapina L, Nikodemus O, Tabors G (2000) Use of an artificial model of monitoring data to aid interpretation of principal component analysis. Environ Modell Softw 15:755–763

Cobelo-García A, Prego R (2003) Heavy metal sedimentary record in a Galician Ria (NW Spain): background values and recent contamination. Mar Pollut Bull 46:1253–1262

Covelli S, Fontolan G (1997) Application of a normalization procedure in determining regional geochemical baselines. Environ Geol 30(1–2):34–45

El Nemr A, Khaled A, El Sikaily A (2006) Distribution and statistical analysis of leachable and total heavy metals in the sediments of the Suez Gulf. Environ Monit Assess 118:89–112

Fharnham IM, Singh AK, Stetzenbach KJ, Lohannesson KH (2002) Treatment of nondetects in multivariate analysis of groundwater geochemistry data. Chemometr Intell Lab 60:265–281

Facchinelli A, Sacchi E, Mallen L (2001) Multivariate statistical and GIS-based approach to identify heavy metal sources in soils. Environ Pollut 114:313–324

González-Maćias C, Schifter I, Lluch-Cota DB, Mendez-Rodriguez L, Hernandez-Vazquez S, (2006) Distribution, enrichment and accumulation of heavy metals in coastal sediments of Salina Cruz Bay, Mexico. Environ Monit Assess 118:211–230

Glasby GP, Szefer P, Geldon J, Warzocha J (2004) Heavy-metal pollution of sediments from Szczecin Lagoon and the Gdansk Basin, Poland. Sci Total Environ 330:249–269

Han YM, Du PX, Cao JJ, Posmentier ES (2006) Multivariate analysis of heavy metal contamination in urban dusts of Xi'an, Central China. Sci Total Environ 355:176–186

Hong Kong Environmental Protection Department (2005) Marine water quality in Hong Kong, Hong Kong Government Printer, Hong Kong

Hopke PK (2003) Recent developments in receptor modeling. J Chemometr 17:255–265

Johnson RA, Wichern DW (2002) Applied multivariate statistical analysis, 5th edn. Prentice-Hall, NJ

Lau MM, Rootham RC, Bradley GC (1993) A strategy for the management of contaminated dredged sediment in Hong Kong. J Environ Manage 38:99–114

Liu WX, Li XD, Shen ZG, Wang DC, Wai OWH, Li YS (2003) Multivariate statistical study of heavy metal enrichment in sediments of the Pearl River Estuary. Environ Pollut 121:377–388

Owen RB, Sandhu N (2000) Heavy metal accumulation and anthropogenic impacts on Tolo Harbour, Hong Kong. Mar Pollut Bull 40:174–180

Pekey H, Karakaş D, Bakoğlu M (2004) Source apportionment of trace metals in surface waters of a polluted stream using multivariate statistical analyses. Mar Pollut Bull 49:809–818

Peré-Trepat E, Olivella L, Ginebreda A, Caixach J, Tauler R (2006) Chemometrics modelling of organic contaminants in fish and sediment river samples. Sci Total Environ 371:223–237

Reimann C, de Caritat P (2000) Intrinsic flaws of element enrichment factors (EFs) in environmental geochemistry. Environ Sci Technol 34:5084–5091

Reimann C, de Caritat P (2005) Distinguishing between natural and anthropogenic sources for elements in the environment:

regional geochemical surveys versus enrichment factors. Sci Total Environ 337:91–107

Qu W, Kelderman P (2001) Heavy metal contents in the Delft canal sediments and suspended solids of the river Rhine: multivariate analysis for source tracing. Chemosphere 45:919–25

Shine JP, Ika RV, Ford TE (1995) Multivariate statistical examination of spatial and temporal patterns of heavy-metal contamination in New-Bedford Harbor marine-sediments. Environ Sci Technol 29:1781–1788

Simeonov V, Stratis JA, Samara C, Zachariadis G, Voutsa D, Anthemidis A, Sofoniou M, Kouimtzis T (2003) Assessment of the surface water quality in Northern Greece. Water Res 37:4119–4124

Sin SN, Chua H, Lo W, Ng LM (2001) Assessment of heavy metal cautions in sediments of Shing Mun River, Hong Kong. Environ Int 26:297–301

Song Y, Xie SD, Zhang YH, Zeng LM, Salmon LG, Zheng M (2006) Source apportionment of PM25 in Beijing using principal component analysis/absolute principal component scores and UNMIX. Sci Total Environ DOI 10.1016/j.scitotenv. 2006.08.041

Tam NFY, Wong YS (2000) Spatial variation of heavy metals in surface sediments of Hong Kong mangrove swamps. Environ Pollut 110:195–205

Tanner PA, Leong LS, Pan SM (2000) Contamination of heavy metals in marine sediment cores from Victoria Harbour, Hong Kong. Mar Pollut Bull 40:769–779

Thurston GD, Spengler JD (1985) A quantitative assessment of source contributions to inhalable particulate matter pollution in Metropolitan Boston. Atmos Environ 19:9–25

Wang XS, Qin Y (2006) Use of multivariate statistical analysis to determine the relationship between the magnetic properties of urban topsoil and its metal, S, and Br content. Environ Geol 51:509–516

Watson JG, Zhu T, Chow JC, Engelbrecht J, Fujita EM, Wilson WE (2002) Receptor modeling application framework for particle source apportionment. Chemosphere 49:1093–1136

Wedepohl KH (1995) The composition of the continental crust. Geochim Cosmochim Acta 59:1217–1232

Wong CKC, Yeung HY, Cheung RYH, Yung KKL, Wong MH (2000) Ecotoxicological assessment of persistent organic and heavy metal contamination in Hong Kong coastal sediment. Arch Environ Contam Toxiol 38:486–493

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

Zhou F, Liu Y, Guo HC (2006) Application of multivariate statistical methods to the water quality assessment of the watercourses in the Northwestern New Territories, Hong Kong. Environ Monit Assess DOI 10.1007/s10661-006-9497-x

Zoller WH, Gladney ES, Duce RA (1974) Atmospheric concentrations and sources of trace metals at the South Pole. Science 183:199–201