

## Determination of the geochemical baseline for the East River basin, China

LINCOLN FOK & MERVYN R. PEART

*Department of Geography, The University of Hong Kong, Pokfulam Road, Hong Kong*  
[mrpeart@hkucc.hku.hk](mailto:mrpeart@hkucc.hku.hk)

**Abstract** Baseline information serves as a point of reference against which future natural or anthropogenic changes to geochemistry may be evaluated. It may also serve to identify controls, such as rock type, upon geochemistry and has other applications. In 2007–2009, active stream sediment samples from the East River in China were collected and analysed in an attempt to develop a geochemical baseline. Sample density is around 1 per 200 km<sup>2</sup>, covering a total area of 27 600 km<sup>2</sup>. Situated in southern China, the East River basin contains three large reservoirs that are of great importance to flood control, power generation, shipping, and water supply in the region. The watershed is well vegetated and woodlands occupy ~75% of the basin area. The geology is heterogeneous: sedimentary units of various ages have been intruded by igneous rocks. In 2007, the population living within the watershed was some 12 million. The concentration ranges of the elements are consistent with average crustal values. The spatial variation of the geochemistry reveals that it is affected by both natural and anthropogenic factors.

**Key words** geochemical baseline; spatial distribution; East River, China

### INTRODUCTION

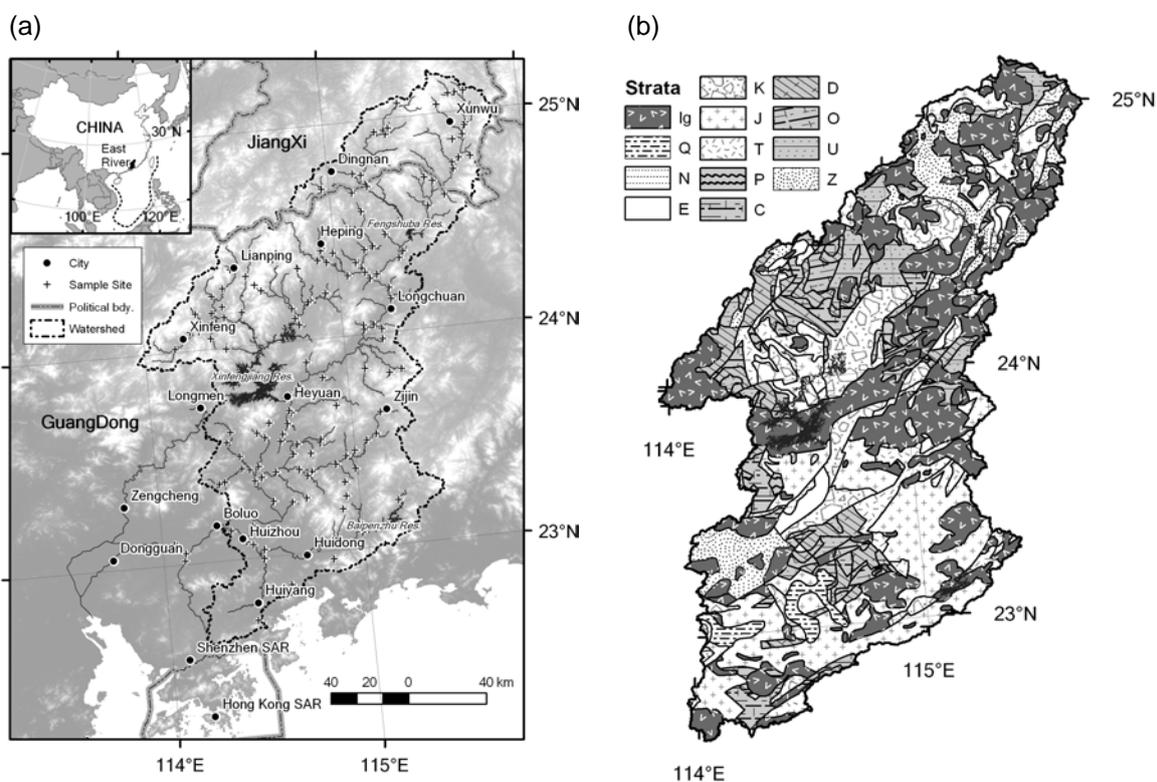
The geochemical baseline is a fundamental datum against which future natural or anthropogenic changes to the chemistry of the land surface may be evaluated (Darnley, 1995). Information on the spatial variation of surface geochemistry has a wide range of potential economic and environmental applications. The study of surface geochemistry originated from mineral (geochemical) exploration in the 1930s and its uses have been extended to address environmental issues (Plant & Hale, 1994). Recent geochemical surveys usually adopt a multi-purpose integrated approach and involve investigation of multi-media, multi-element data (e.g. G-BASE; British Geological Survey, 2009). In view of the numerous applications of geochemical baselines to a diversity of sample media, at various temporal and spatial scales, a consensus on a precise definition, as well as a coherent method for its assessment, has not been reached among geochemists (Reimann & Garrett, 2005). In the context of mineral prospecting, geochemical “background” refers to a reference level against which the “concentration” of a mineral deposit could be distinguished (Hawkes & Webb, 1962). Any deviation from this background value or range is termed an “anomaly” or “outlier”, and the precise level for their discrimination is called the “threshold”. To environmental geochemists, the ideal “background” should be the “pristine” level, solely of natural origin, against which anthropogenic contributions could be discriminated (e.g. Albanese *et al.*, 2006). Measurement of “natural background”, however, is fraught with difficulty, and it is even argued that “*natural background no longer exists on this planet*” (Reimann & Garrett, 2005). It may be more practical to determine the “ambient background” especially in a basin such as the East River. The latest definition of background given by the ISO is: “*Concentration of a substance characteristic of a soil type in an area or region arising from both natural sources and non-natural diffuse sources, such as atmospheric deposition*” (ISO 2005). This definition is synonymous with geochemical baseline proposed in the International Geological Correlation Programme (IGCP) project 259 (Darnley *et al.*, 1995), and is sometimes used interchangeably with “background”. Due to the variability of geochemistry and the implication of baseline as prescribing a single value, some researchers do not favour use of the term (e.g. Reimann & Garrett, 2005). The purpose of defining a “baseline” or “ambient background” is to take a snapshot of the geochemistry of an area, and this definition has been adopted in this study.

Separation of baseline from anomalies (outliers) is central to the approach. In general, baseline delineation methods can be classified into geochemical and statistical (Matschullat *et al.*, 2000). It is the latter which has been adopted for this study, and is described in the methods section.

## STUDY AREA AND METHODS

The East River, locally known as the Dongjiang, drains an area of 35 340 km<sup>2</sup>; it is one of the three main tributaries of the Pearl River (Fig. 1(a)). In recent decades, a series of water conservancy projects have been built within the watershed. The three major reservoirs in the basin are of great importance to flood control, power generation, irrigation, shipping, and water supply. Apart from the watershed itself, the East River also supplies potable water to the adjacent area of the Huangpu region in Guangzhou, as well as to Shenzhen and Hong Kong. The watershed is well vegetated and woodlands occupy 75% of the basin area. The geology of the East River basin is heterogeneous (Fig. 1(b)). Sedimentary units of various ages have been intruded by igneous rocks (Fan & Du, 1999). The average annual runoff at Boluo station is 32.4 billion m<sup>3</sup>, with a distinctive seasonal variation due to the monsoonal rainfall regime. The average sediment concentration of the East River is 130 mg L<sup>-1</sup>, which is the lowest among the major tributaries of the Pearl River system. More than 90% of the annual sediment load is transported between April and September. In 2007, there were 12.3 million registered inhabitants living within the East River drainage basin.

The samples used in this study were intended for an on-going sediment tracing (fingerprinting) project. The data from 137 sampling sites have been identified for inclusion in this paper. Sample sites were distributed throughout the drainage network and those located in small drainage basins were selected to take account of the spatial variation of geology (Fig. 1(a)). Bed



**Fig. 1** (a) Location of the East River basin and the sampling sites. (b) The geology of the East River basin. Ig: Igneous, Q: Quaternary, N: Neogene, E: Paleogene, K: Cretaceous, J: Jurassic, T: Triassic, P: Permian, C: Carboniferous, D: Devonian, O: Ordovician, U: Cambrian, Z: Sinian.

sediment was chosen as the sampling media because depositional environments are relatively unaffected by temporal variations and are coherent with stream sediment in terms of their geochemical composition (Ridgway *et al.*, 2003). Sampling methods are based on the US Geological Survey *Open-file Report 94-458* (Shelton & Capel, 1994) with minor modifications to address local conditions. The sampling principle is to obtain fine-grained, recently deposited, surficial sediments in natural depositional zones during low-flow conditions, within a 100-m river reach. The top 10 cm of river deposits were taken from five or more depositional features within each of the sites. The samples were composited and homogenized on site, in order to minimize local geochemical variability. The <63- $\mu\text{m}$  fraction was obtained by wet sieving through a 63- $\mu\text{m}$  nylon mesh in the laboratory, and subsequently air dried. Multi-elemental geochemical analysis by ICP, ICP/MS, and INAA was carried out by the Activation Labs in Canada.

Statistical methods for baseline determination, generally regarded as being more objective, were adopted. Traditionally, the mean  $\pm 2$  standard deviations (sdev) is used to screen out the top and bottom 2.5% of the data. The use of the mean  $\pm 2$ sdev, in most cases, is ill-defined, due to its dependency on data normality. This issue is valid for the current study as the Kolmogorov-Smirnov test revealed that most analytes belong neither to a normal or a lognormal distribution, although a closer resemblance to the latter was observed. Reimann *et al.* (2005) suggest that more robust measures outperform the traditional mean  $\pm 2$ sdev method. Two robust methods of threshold determination were used. The first one is the median  $\pm 2$  median absolute deviation (*MAD*), which is defined as the median of the absolute deviations from the data's median:  $MAD = \text{median}_i(|\text{median}(x) - x_i|)$ . The second method involves using the upper and lower "whisker" of the boxplot (Tukey, 1977). The boxplot is a graphical illustration of data distribution. The centre box of a boxplot is bound by the upper (*UH*, 75% percentile) and lower hinge (*LH*, 25% percentile), and so approximates the interquartile range (also called the hinge width, *HW*). The upper and lower whiskers are defined as  $UH + 1.5HW$  and  $LH - 1.5HW$ , respectively. Since a lognormal distribution better describes the data set, the boxplot limits are defined using log values.

The spatial distribution of geochemical data is perhaps of the same order of importance as compared to its variability in concentration. Using the same boxplot thresholds in baseline determination for symbol class definition, the symbolic map allows spatial identification of anomalies.

## RESULTS AND DISCUSSION

The median  $\pm 2MAD$  (M+) method always detected a larger percentage of positive extreme values compared to the boxplot method (UW) (Table 1). This is because the median  $\pm 2MAD$  method is more sensitive to the symmetry of the data. However, the median  $\pm 2MAD$  method also is likely to overestimate the number of outliers from a distribution with less than 40% outliers. Reimann *et al.* (2005) suggest that the boxplot method performs best with up to 25% outliers. Since the other method (median  $\pm 2MAD$ ) only detected up to 15% outliers, the boxplot method should, therefore, be the best method for baseline concentration assessment for this study.

The geochemistry of the current data set generally is consistent with average upper crustal geochemical abundances (Wedepohl, 1995). Baseline levels of Y and Zn, along with most rare earth elements (REEs), exhibit enrichment when compared with crustal values, although the deviations are small. Most major elements are comparable with the range given for the East River basin in the *Environmental Geochemistry Atlas of China* (AGC; Li & Wu, 1999), and with the average values of suspended matter in world rivers (Thomas & Meybeck, 1996). The relatively low values of calcium, as indicated from both the current data set as well as the AGC range, are due to the scarcity of calcareous rocks in the area. Zinc values for the data set are just above those found in China (AGC), and are comparable to the range for Hong Kong urban soils (Lee *et al.*, 2006: 23–930 ppm). The local baseline is, however, below the zinc guideline value given by the Interim Sediment Quality Values for marine sediment developed for Hong Kong (Chapman *et al.*,

**Table 1** Geochemical baseline median and range determined by median  $\pm 2MAD$  (M-, M+) and lower, upper whisker (LW, UW) for the <63- $\mu\text{m}$  samples (n = 137), along with their respective percentage of detected outliers (M%out, B%out).

Analyte	Median	M-	M+	M%out	LW	UW	B%out	AGC-l	AGC-h	Crust	Susp.
SiO <sub>2</sub>	550000	440000	690000	2.9	430000	710000	1.5			620000	590000
Al <sub>2</sub> O <sub>3</sub>	200000	130000	310000	0	140000	290000	1.5	100000	190000	150000	170000
Fe <sub>2</sub> O <sub>3</sub>	69000	42000	110000	4.4	36000	110000	4.4	32000	39000	63000	74000
MnO	1800	500	6300	2.2	520	7500	2.2	160	720	1000	1300
MgO	7600	3700	16000	9.5	2800	17000	3.6	2800	5800	37000	19000
CaO	3600	1400	9000	10	1200	11000	9.5	560	3400	55000	33000
Na <sub>2</sub> O	2300	790	6700	2.2	900	7700	3.6	140	4300	32000	9300
K <sub>2</sub> O	28000	18000	43000	4.4	17000	44000	2.9	4500	30000	24000	25000
TiO <sub>2</sub>	8800	6000	13000	12	5300	14000	8	4900	6700	6800	10000
P <sub>2</sub> O <sub>5</sub>	1900	940	3800	12	700	4500	5.1	390	930	1800	3300
S	290	110	760	13	80	1100	4.4			950	
Be	4	1.7	9.4	4.4	3	10	28			3.1	
Ta	2.7	1.1	6.6	8	0.8	8	1.5			1.5	
Th	27	9.5	79	8.8	7.4	110	2.2			10	
U	8.3	2.5	27	2.9	2.5	37	1.5			2.5	
Y	65	20	210	4.4	28	280	4.4			21	
Zn	140	67	280	14	67	340	12	49	67	52	110
La	90	37	220	8	46	280	5.8			32	
Ce	150	76	300	5.1	85	360	3.6			66	
Nd	70	29	170	5.8	25	220	5.1			26	
Sm	12	4.1	34	4.4	5.2	48	3.6			4.7	
Eu	2.2	1.1	4.3	7.3	1.1	5.4	6.6			0.95	
Tb	1.6	0.55	4.7	7.3	0.8	5.9	8			0.5	
Yb	5.8	2.1	16	5.1	3	23	3.6			1.5	
Lu	0.86	0.33	2.3	7.3	0.43	2.8	4.4			0.27	

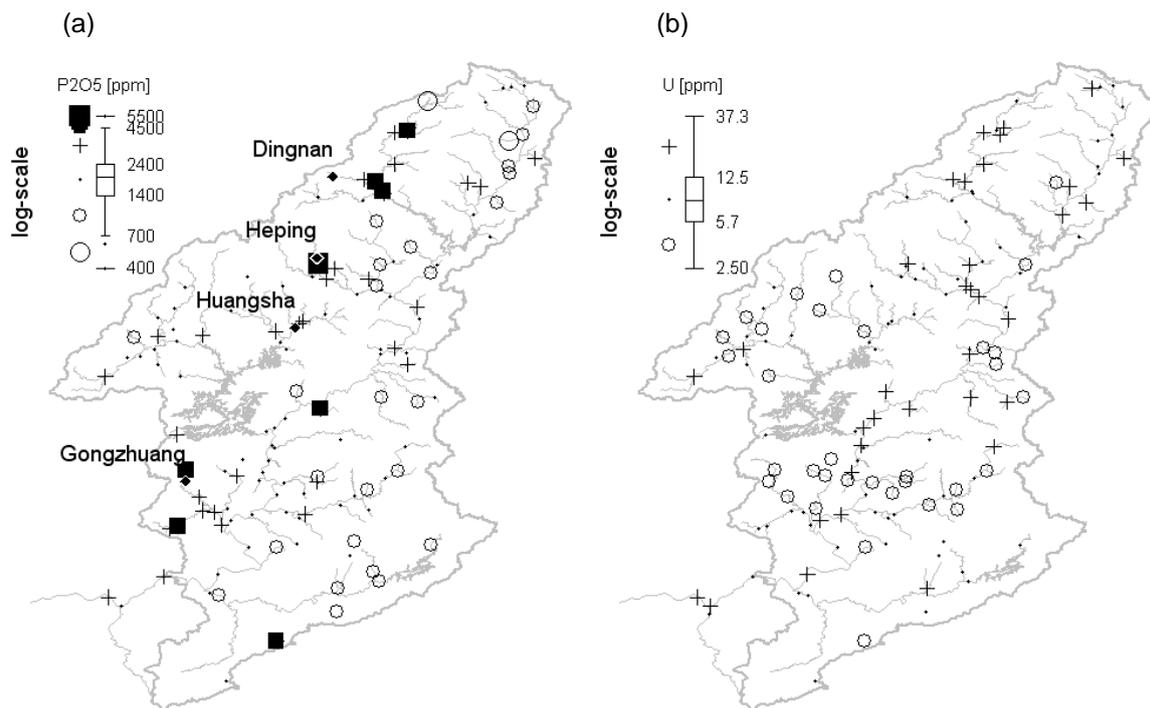
Concentration range from the *Atlas of the Ecological Environmental Geochemistry of China* (AGC-l/h); Mean concentration of the upper crust (Crust) and river suspended matter (Susp.) are also included for comparison. Units in ppm.

1999), and the Severe Effect Level (SEL) for freshwater sediments from the *Ontario Screening Level Guidelines* (Ontario Ministry of the Environment, 1993), which are 410 and 820 ppm, respectively.

Although the current sampling programme is not specifically designed for geochemical mapping, which would require a grid-based site selection across the entire study area, two interesting patterns can be observed. Firstly, the spatial patterns for zinc and sulfur are similar to those of phosphorus, with the distribution of the latter being shown in Fig. 2(a). The "hotspots" for these analytes are located on the smaller rivers near some of the major urban settlements, which may indicate anthropogenic origins (Fig. 2(a)). The localized effect indicates that sediment mixing in the channel system may have diluted the signature of these elements. In addition, using samples obtained along the main channel of the East River system, a significant correlation is found between these elements and the basin area (0.70, 0.59 and 0.93 for P<sub>2</sub>O<sub>5</sub>, S and Zn respectively). This may suggest that human activity, in the form of elemental inputs, leads to downstream accumulations, and may play a significant role. Secondly, the distribution of Si, Al, Be, Ta, Th, U, Y, and the REEs, and to a lesser extent Na, are enriched in the central and northeastern part of the basin. These areas correspond to an igneous lithology. The U-Th grouping, in particular, is associated with I- and A-type granites (Sewell, 1999). The geochemical variation of uranium, plotted in Fig. 2(b), is typical among these elements.

## CONCLUSION

Data on geochemical baselines and their spatial distribution in the East River basin are presented. The concentration ranges are consistent with average crustal values. The spatial variation of the



**Fig. 2** Symbolic map representation for: (a) phosphorus and (b) uranium.

elements also reveal that the geochemistry of the area is affected by both natural and anthropogenic factors. In particular, discrimination of the geochemical signatures for igneous and sedimentary units is promising, and may aid in the ongoing spatial source determination of sediment in this large basin with a heterogeneous lithology. It also shows that geochemical baselines may not be characterized by a “single” descriptive value such as the median. Account needs to be taken of its variability.

**Acknowledgement** This work was supported by Grant no. 200611159130 from the Seed Funding Programme for Basic Research Project of the Hong Kong General Research Fund.

## REFERENCES

- Albanese, S., Vivo, B. D., Lima, A. & Cicchella, D. (2006) Geochemical background and baseline values of toxic elements in stream sediments of Campania region (Italy). *J. Geochemical Exploration* **93**(1), 21–34.
- British Geological Survey (2009) *Geochemical Baseline Survey of the Environment (G-BASE)*. From [www.bgs.ac.uk/gbase/](http://www.bgs.ac.uk/gbase/).
- Chapman, P. M., Allard, P. J. & Vigers, G. A. (1999) Development of sediment quality values for Hong Kong Special Administrative Region: A possible model for other jurisdictions. *Marine Pollution Bull.* **38**(3), 161–169.
- Darnley, A. G. (1995) A global geochemical reference network: the foundation for geochemical baselines. *J. Geochemical Exploration* **60**(1), 1–5.
- Darnley, A. G., Björklund, A., Bølviken, B., Gustavsson, N., Koval, P. V., Plant, J. A., Steenfelt, A., Tauchid, M., Xie, X.-j., Garrett, R. G. & Hall, G. E. M. (1995). *A Global Geochemical Database for Environmental and Resource Management*. Final report of IGCP Project 259. UNESCO Publishing, Paris, France.
- Fan, Y. & Du, X.-r. (eds) (1999) *The National Physical Atlas of China*. Cartographic Publishing House, Beijing, China.
- Hawkes, H. E. & Webb, J. S. (1962) *Geochemistry in Mineral Exploration*. Harper, New York, USA.
- ISO (2005) *Soil Quality – Vocabulary*. BS ISO 11074:2005. BSI, London, UK.
- Lee, C. S.-l., Li, X., Shi, W., Cheung, S. C.-n. & Thornton, I. (2006) Metal contamination in urban, suburban, and country park soils of Hong Kong: A study based on GIS and multivariate statistics. *Sci. Total Environ.* **356**(1-3), 45–61.
- Li, J.-X. & Wu, G.-j. (eds) (1999) *Atlas of the Ecological Environmental Geochemistry of China*. Geological Publishing House, Beijing, China.
- Matschullat, J., Ottenstein, R. & Reimann, C. (2000) Geochemical background – can we calculate it? *Environ. Geology* **39**(9), 990–1000.

- Ontario Ministry of the Environment (1993) *Guidelines for the Protection and Management of Aquatic Sediment Quality in Ontario*. Queen's Printer for Ontario, Ontario, Canada.
- Plant, J. & Hale, A. M. (1994). Introduction: The foundations of modern drainage geochemistry. In: *Handbook of Exploration Geochemistry 6* (by M. Hale & J. A. Plant), 3–9. Elsevier, Amsterdam, The Netherlands.
- Reimann, C., Filzmoser, P. & Garrett, R. G. (2005) Background and threshold: critical comparison of methods of determination. *Sci. Total Environ.* **346**, 1–16.
- Reimann, C. & Garrett, R. G. (2005) Geochemical background – concept and reality. *Sci. Total Environ.* **350**, 12–27.
- Ridgway, J., Breward, N., Langston, W. J., Lister, R., Rees J. G. & Rowlett, S. M. (2003) Distinguishing between natural and anthropogenic sources of metals entering the Irish Sea. *Appl. Geochemistry* **18**, 283–309.
- Sewell, R. J. (1999) *Geochemical Atlas of Hong Kong*. Geotechnical Engineering Office, Hong Kong.
- Shelton, L. R. & Capel, P. D. (1994) Guidelines for Collecting and Processing Samples of Stream Bed Sediment for Analysis of Trace Elements and Organic Contaminants for the National Water-Quality Assessment Program (*USGS Open-File Report 94-458*). USGS, Sacramento, USA.
- Thomas, R. & Meybeck, M. (1996) The use of particulate material. In: *Water Quality Assessments* (by D. Chapman), 127–174. E & FN Spon, London, UK.
- Tukey, J. W. (1977) *Exploratory Data Analysis*. Addison-Wesley, Reading, UK.
- Wedepohl, K. H. (1995) The composition of the continental crust. *Geochim. Cosmochim. Acta* **59**(7), 1217–1232.

