

# Mapping soil pollution by spatial analysis and fuzzy classification

Roberto Wagner Lourenço · Paulo M. Barbosa Landim · André Henrique Rosa · José Arnaldo F. Roveda · Antonio César Germano Martins · Leonardo Fernandes Fraceto

Received: 10 December 2008 / Accepted: 15 May 2009 / Published online: 6 June 2009  
© Springer-Verlag 2009

**Abstract** Spatial analysis and fuzzy classification techniques were used to estimate the spatial distributions of heavy metals in soil. The work was applied to soils in a coastal region that is characterized by intense urban occupation and large numbers of different industries. Concentrations of heavy metals were determined using geostatistical techniques and classes of risk were defined using fuzzy classification. The resulting prediction mappings identify the locations of high concentrations of Pb, Zn, Ni, and Cu in topsoils of the study area. The maps show that areas of high pollution of Ni and Cu are located at the northeast, where there is a predominance of industrial and agricultural activities; Pb and Zn also occur in high concentrations in the northeast, but the maps also show significant concentrations of Pb and Zn in other areas, mainly in the central and southeastern parts, where there are urban leisure activities and trade centers. Maps were also prepared showing levels of pollution risk. These maps show that (1) Cu presents a large pollution risk in the north–northwest, midwest, and southeast sectors, (2) Pb represents a moderate risk in most areas, (3) Zn generally exhibits low risk, and (4) Ni represents either low risk or no risk in the studied area. This study shows that combining

geostatistics with fuzzy theory can provide results that offer insight into risk assessment for environmental pollution.

**Keywords** Soil pollution · Geostatistics · Fuzzy classification · Risk analysis

## Introduction

Soil contamination poses an increasing threat to human health and environmental quality (Lourenço and Landim 2005). Although soil often acts as a filter, purifying and immobilizing many of the impurities deposited in it, its capacity is limited, so soil can be negatively affected by the cumulative effects of atmospheric pollutants, agrochemicals and fertilizers, industrial and domestic solid residues, and toxic and radioactive materials (Moreira-Nodermann 1987). Among soil and environmental pollutants, heavy metals have received considerable attention over the last few decades (Franssen et al. 1977). Metals, in particular lead, nickel, zinc, and copper, constitute a significant potential threat to human health. The environmental persistence of metals in concert with their intensive use by modern society has, over the years, created a concentration of metals in the biosphere. Thus, there is a great chance of exposure to toxic metals both in and out of the workplace, and several metals are known to be human carcinogens, including arsenic, chromium, and nickel. Further, many toxic effects of metals, including carcinogenicity, can be modified by concurrent exposure to other metals (Tang et al. 1999; Winneke et al. 2002; Stein et al. 2002; Yang et al. 2003; Rocha et al. 2003; Dias et al. 2006).

Since soil properties present a continuum in their spatial variations, it is difficult to categorize soil samples without

---

R. W. Lourenço (✉) · A. H. Rosa · J. A. F. Roveda · A. C. G. Martins · L. F. Fraceto  
Department of Environmental Engineering,  
São Paulo State University, Sorocaba Campus,  
3 de Março Avenue, n° 511, PO Box 18087-180,  
Sorocaba, SP, Brazil  
e-mail: robertow@sorocaba.unesp.br

P. M. B. Landim  
Department of Applied Geology,  
São Paulo State University,  
Rio Claro Campus, Rio Claro, SP, Brazil

introducing errors or over-simplifications. Therefore, class boundaries are usually chosen arbitrarily by imposing (1) an uncertainty about the accuracy of the critical threshold or range used to specify membership in a certain class and (2) an uncertainty about the quality of the input maps. Nevertheless, with crisp classifications, values close to class boundaries can fall into different classes, even though their uncertainties are the same within the range of the standard measurement or interpolation error; consequently, the resulting classifications can be erroneous. Moreover, in analyses that involve a succession of steps, the problems resulting from misclassification of soil units may be amplified (Dobermann and Oberthür 1997).

Prediction mappings of samples are often based on geostatistical methods, which calculate unbiased estimates at unsampled locations. This approach is increasingly used to characterize spatial variability of soil properties (Van Meirvenne and Goovaerts 2001; Webster and Oliver 2001; Lin et al. 2001; Romić and Romić 2003; McGraph et al. 2004). However, it remains difficult to define transition classes between polluted and unpolluted soil areas. As an alternative, fuzzy logic methods can be used to estimate the degree of membership in each class, thereby treating transition areas more realistically and eliminating imprecise and subjective concepts that are present in variables of the physical environment (Feng et al. 2006).

Fuzzy logic is an artificial intelligence method and, as its name indicates, it is based on soft, fuzzy values rather than crisp ones. Specifically, transitions between two numbers or sets of numbers are gradual, ranging from 1 to 0: unity indicates a total degree of membership while zero indicates no degree of membership (Zadeh 1965).

The first application of fuzzy sets and logic to environmental sciences was in land evaluation (Chang and Burrough 1987). Subsequently, the approach has been extended to many other applications. For example, at the Lacombe Experimental Farm in Alberta, fuzzy and Boolean sets were combined to generate maps of clay content in C horizon of the soil, interpolated by ordinary kriging. Both approaches were used to estimate soil pollution. In drainage net studies, a fuzzy approach was used as an alternative procedure for classifying abrupt transition data such as single pollution spots (Burrough 1992). Another reported application is the acquiring and representing of knowledge on soil–landscape relationships and applying that knowledge to digital soil mapping (Feng et al. 2006; Amini et al. 2005; Burrough and McDonnell 2004; Cattle et al. 2002).

The aim of this work is to present the spatial distributions of and pollution risks from heavy metals in soils of the southern coastal region of the State of São Paulo, Brazil. For that, data were interpolated using ordinary kriging and prediction mappings of polluted areas were

generated. A fuzzy approach was used to classify interpolated values into four levels of soil pollution risk: high risk, moderate risk, low risk, and unpolluted areas.

## Materials and methods

### Chemicals and reagents

All reagents used were of high-purity grade unless otherwise stated. Diluted acids and bases were prepared by diluting 30% hydrochloric acid (Suprapur, Merck AG, Darmstadt, Germany), 65% nitric acid (p.a. Merck AG; prepurified by sub-boiling distillation), and sodium hydroxide-monohydrate (Suprapur, Merck AG) with high-purity water (Millipore-Q system, Millipore GmbH, Eschborn, Germany). For metal determinations and their calibrations, synthetic standards were used (ICP multi-element standard solution IV, Merck AG).

### Study area and sampling of soils

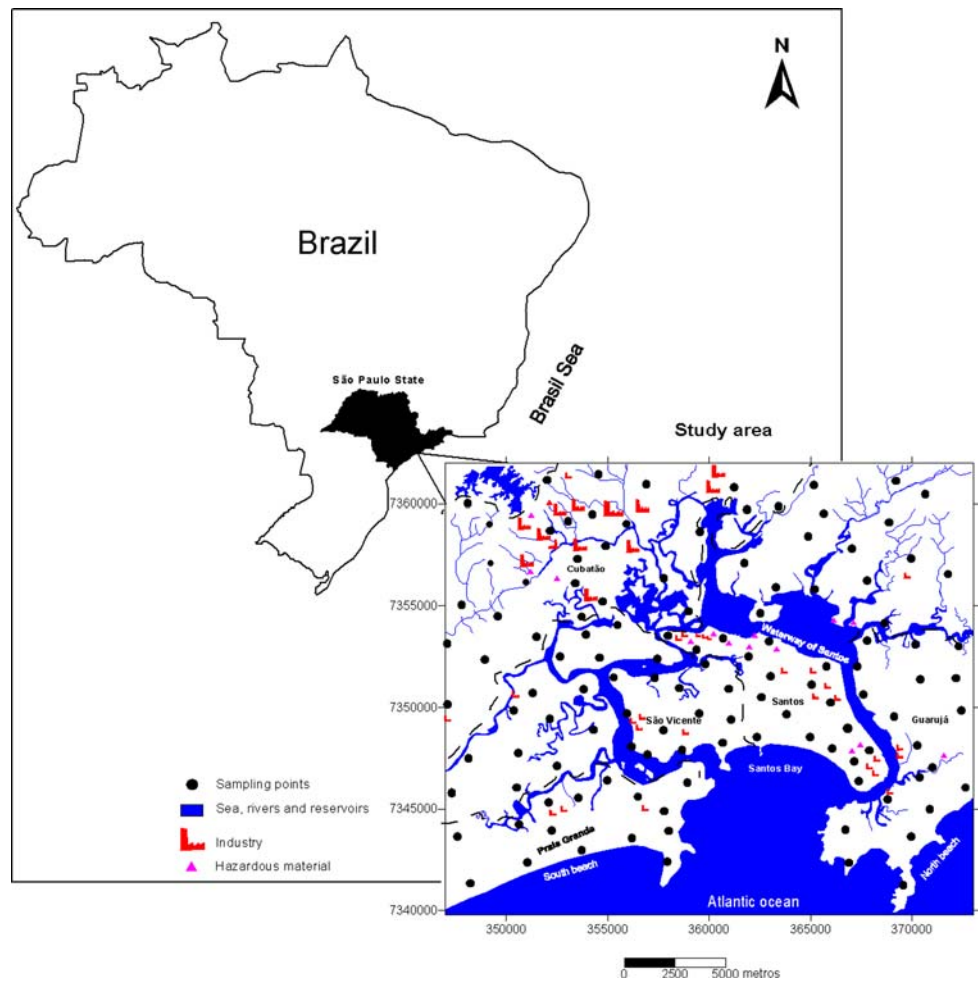
The research was conducted in the southern coastal region of the State of São Paulo, Brazil (Fig. 1). The study area covers about 572 km<sup>2</sup> and includes industrial, urban, and uncultivated lands (forest). Industrial activities are concentrated close to the rivers around Cubatão city in the northeast area; they include a large petroleum refinery and various chemical activities. Small factories are concentrated in the central area of the port region. Urban areas are concentrated mainly along the seacoast from southwest to southeast. Forest areas, with different types of vegetation, are concentrated in the northwest. There are several wastewater treatment plants in the region with high risks of pollution originating from sewage sludge and related compounds.

The parent rock materials are mainly recent terraces, recent alluvial deposits, and undifferentiated terraces (all of quaternary age) found in the southwest and southern part of the region. In the northwest there are Precambrian mountain rocks. The average annual rainfall is 1,200 mm and the average annual temperature in the region is 25.5°C.

### Sampling and preparation of soils

To generate prediction mappings of heavy metals in the whole region, a random unaligned sampling strategy was used to collect 123 soil samples from different locations in the study area. The sampling locations were separated by distances ranging from 95 to 650 m (Fig. 1). All soil samples were taken from a depth of 0–20 cm, as recommended by the soil reference book from the São Paulo State Basic Sanitation and Technology Company

**Fig. 1** The studied area of the present work, with the sampling locations



(CETESB 2001). Gravels, coarse organic matter, and plant–root residues were removed. Samples were thoroughly mixed and ground to pass a 2-mm sieve for analysis (Alloway 2001).

**Soil characterization**

Soil pH was measured using a pH meter with a glass electrode. The proportion of organic material in the samples was determined by calcination for 4 h at 650°C. The results showed about 5% (m/m) organic matter present in different samples. Cation exchange capacity (CEC) was measured for K, Ca, Mg, and H + Al (CETESB 2001). Available phosphorus (AP) and available potassium (AK) were extracted using HF–HNO<sub>3</sub>–HClO<sub>4</sub>–HCl and then measured by inductively coupled plasma mass spectrometry (ICP-MS).

To determine total metal concentrations, dried and ground soil samples were conventionally decomposed by aqua regia according to a standardized procedure (Alloway 2001). For this purpose, a weighed amount (approximately

2.00 g) of pre-dried soil sample was mixed with 21 mL of 30% HCl (Suprapur) and 7 mL of 65% HNO<sub>3</sub> (Suprapur) in a highly pre-purified quartz vessel (200 mL). The solution was heated first to 100°C and then to 120°C, where it was cautiously and largely evaporated. Subsequently, remainders of the sample were digested by an additional 20 mL of concentrated HNO<sub>3</sub> under reflux for 3 h. Finally, the digested samples were diluted with high-purity water to a final volume of 100 mL. Small soil remainders (approximately 5%) still undigested were removed by filtration. Metal determinations were usually carried out with 1/10 dilutions of the digestion solutions. The standard deviation of soil digestion by aqua regia (shown in Table 2 for a number of metals) was assessed from five replicates. Metal determinations were performed by ICP-OES. The chemical blanks, assessed by means of two aqua regia digestion runs without any sample, were subtracted from these results.

Multi-element determinations in acid extracts were preferably carried out by simultaneous ICP-OES (spectrometer: TJA IRIS AP, Thermo Jarrell Ash, Franklin, MA,

USA) according to routine recommendations of the manufacturer. To achieve low detection limits, relatively large sample volumes of up to 8 mL, each, were nebulized (Meinhard nebulizer) and long integration times of 30 s (three measurements) were used.

Geostatistics analysis

Geostatistics analysis methods are based upon the assumption that spatial variations in any continuous attribute are often too irregular to be modeled by a simple, smooth mathematical function. Instead the variation can be better described by a stochastic surface, known as a regionalized variable. Such variables apply to environmental properties such as soil types, variations in atmospheric pressure, elevation above sea level, and distributions of continuous demographic indicators. Geostatistical interpolation is known as kriging. The procedure is similar to that used in weighted moving average interpolation, except that the weights are derived from a geostatistical analysis of the data rather than from a general, and possibly inappropriate, model. The ‘true’ value  $\hat{z}(X_0)$  is given by

$$\hat{z}(X_0) = \sum_{i=1}^n \lambda_i \cdot z(x_i) \tag{1}$$

with  $\sum_{i=1}^n \lambda_i = 1$ . The weights  $\lambda_i$  are chosen so that (1) the estimate  $\hat{z}(X_0)$  is unbiased and (2) the estimation variance  $\sigma_e^2$  is less than that given by any other linear combination of the observed values.

The minimum variance of  $[\hat{z}(X_0) - z(X_0)]$ , the prediction error, or ‘kriging variance’ is given by

$$\hat{\sigma}_i^e = \sum_{i=1}^n \lambda_i \gamma(x_i, x_0) + \phi \tag{2}$$

and is obtained when

$$\sum_{i=1}^n \lambda_i \gamma(x_i, x_j) + \phi = \gamma(x_j, x_0) \quad \text{for all } j. \tag{3}$$

The quantity  $\gamma(x_i, x_j)$  is the semivariance of  $z$  between the sampling points  $x_i$  and  $x_j$ ;  $\gamma(x_i, x_0)$  is the semivariance between the sampling point  $x_i$  and the unvisited point  $x_0$ . Both quantities are obtained from the fitted variogram. The quantity  $\phi$  is a Lagrange multiplier required for the minimization. This method is known as ordinary kriging and is well described by many authors (Landim 2003; Olea 1999; Burrough et al. 1997; Goovaerts 1997; Isaaks and Srivastava 1989; Journel and Huijbregts 1978). In the present study, variogram models were used to analyze spatial patterns and ordinary kriging was used for mapping predictions of the concentrations of four heavy metals: Ni, Zn, Pb, and Cu.

Fuzzy classification

Fuzzy classification is a technique that operates on continuous classes with the purpose of reducing a complex system, represented by some sets of data, into explicitly defined classes. In using fuzzy set theory, observation data are grouped and individuals are assigned continuous class membership values instead of exactly defined (crisp) ones (McBratney et al. 1992; Burrough 1992). A membership value of 1 is assigned to individuals that exactly match strictly defined classes; individuals that do not match the strictly defined classes receive membership values depending on their degree of closeness to class centroids (or class means). Fuzzy classification based on the Semantic Import model (SI) is the common technique that has been used in various soil and environmental studies. In this approach, class limits are specified by experience or by conventionally imposed definitions before individuals are allocated on the basis of how close they match the requirements of the classes (Franssen et al. 1977; Feng et al. 2006; Burrough 1992; Burrough and Mcdonell 2004; Amini et al. 2005; McBratney et al. 1992; Odeh et al. 1992; Van Gaans and Burrough 1993; De Gruijter et al. 1997). In the present study, fuzzy classification was performed on the results of prediction mapping based on geostatistical methods. Heavy metal concentrations were grouped into four classes of pollution risk according to guide values (GV) and kriging standard deviations (KSD).

To obtain the assigned continuous class membership values, a set of linear models was used to describe the a priori membership function (MF). The basic form of such a function is given by

$$\varphi(x) = \begin{cases} \frac{x-a}{b-a} & \text{if } a \leq x \leq b \\ 1 & \text{if } b < x \leq c \\ \frac{x-d}{c-d} & \text{if } c < x \leq d \end{cases} \tag{4}$$

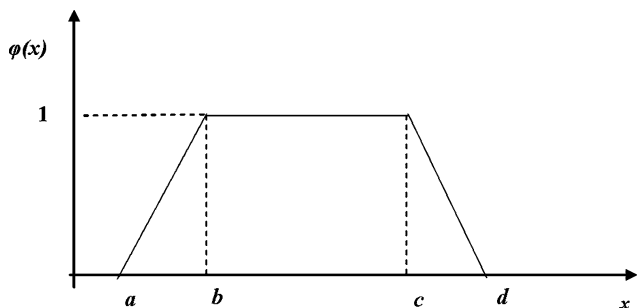
which has a general symmetrical shape. The parameters  $a$ ,  $b$ ,  $c$ , and  $d$  are such that  $[a, d]$  defines the support interval,  $[a, b]$  and  $[c, d]$  the boundaries, and  $[b, c]$  the core, as shown in Fig. 2.

Other variants of the MF are asymmetric types, which define the beginning and the ending member subsets of left and right curves defined by the following trapezoidal function:

$$\varphi(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq a \\ \frac{x-b}{a-b} & \text{if } a < x \leq b \end{cases} \tag{5}$$

and

$$\varphi(x) = \begin{cases} \frac{x-c}{d-c} & \text{if } c \leq x \leq d \\ 1 & \text{if } x > d \end{cases} \tag{6}$$



**Fig. 2** Form of a linear membership function (MF) where  $\phi(x)$  gives the membership value, with  $[a, d]$  being the support interval,  $[a, b]$  and  $[c, d]$  the boundaries, and  $[b, c]$  the core

To classify results from the ordinary kriging interpolation process for each element, four levels of soil pollution risk were constructed: high risk, moderate risk, low risk, and unpolluted areas. These were based on the legally defined limit levels for each element. The MF for each of the four  $F$  sets ( $MF^F$ ) was defined by values for  $a, b, c,$  and  $d$ . The idea underlying the choice of such values is as follows. Given a value  $v$  resulting from the interpolation process and having an intrinsic error KSD, its real value, according to statistical conditions, lies between  $v - \text{KSD}$  and  $v + \text{KSD}$  with a certain degree of confidence. If the interval  $[v - \text{KSD}, v + \text{KSD}]$  belongs to the core of  $MF^F$ , then  $v$  can be characterized by complete and full membership in the set  $F$ . However, if  $[v - \text{KSD}, v + \text{KSD}]$  is in the boundary area, then the membership of  $v$  is given by  $\phi(v)$ . Further, if  $[v - \text{KSD}, v + \text{KSD}]$  is outside the support interval, then the membership of  $v$  is set to 0. Consequently, if  $L_i$  and  $L_s$  define the inferior and superior limits of a class  $F$ , then it follows that

$$a = L_i - \text{KSD} \tag{7}$$

$$b = L_i + \text{KSD} \tag{8}$$

$$c = L_s - \text{KSD} \tag{9}$$

$$d = L_s + \text{KSD} \tag{10}$$

For each class,  $L_i$  and  $L_s$  are defined based on guide values (GV) that characterize them.

After the parameters for the four  $MF^F$  are defined, all points from the interpolated maps are given four

membership values, one for each defined class. To generate a classification map, the maximum fuzzy operator was used, and the classified values were linearly mapped to different ranges on an 8-bit grayscale image (Jiang and Eastman 2000).

### Results

Guide values were used to quantify concentrations of heavy metals in soil (CETESB 2001). Table 1 shows levels of specific chemical elements, defining the maximum, alert, and critical values allowed for heavy metals in soil. Those levels constitute the universe of “soil pollution” in this study. For example,  $13 \text{ mg kg}^{-1}$  is the maximum acceptable concentration for Ni, values between 13 and  $30 \text{ mg kg}^{-1}$  are classified as alert concentrations, and values greater than  $30 \text{ mg kg}^{-1}$  are critical concentrations of Ni for agricultural, urban, and industrial areas.

In this study, any data that were more than three standard deviations from the average were considered exceptional (Liu et al. 2006; Shi et al. 2007). Those data were replaced with the maximum or minimum values within  $[A - 3s, A + 3s]$  in the raw data set. Here  $A$  denotes the average value for each heavy metal and  $s$  is the standard deviation (SD). This was necessary to implement linear geostatistics (McGraph et al. 2004; Landim 2003; Gringarten and Deutsch 2001; Zhang and Selinus 1998; Zhang et al. 1995; Isaaks and Srivastava 1989).

The representative statistical summary of the available data sets for soil attributes is in Table 2. Note that the kurtosis (1.79–3.50) and skewness (1.47–5.82) for Ni, Zn, Pb, and Cu were high. This suggests the presence of outliers and non-normal distributions, which can impair the variogram structure and kriging results. Therefore, raw data sets were logarithmically transformed before performing geostatistical analysis, resulting in reduced values for the kurtosis (−0.45 to 0.17) and skewness (0.08–1.04) of Ni, Zn, Pb, and Cu. The transformed data sets passed the log-normal test.

The coefficients of variation (CV) for Cu, Zn, and Ni were 65.22, 55.33, and 50.32%, respectively. These were

**Table 1** Guideline values for levels of heavy metals in soil ( $\text{mg kg}^{-1}$ )

São Paulo State Basic Sanitation and Technology Company (CETESB 2001)

Soil attributes	Maximum acceptable value	Alert	Critical value		
			Agricultural	Domestic	Industrial
Ni	13	30	50	200	300
Zn	60	300	500	1,000	1,500
Pb	17	100	200	350	1,200
Cu	35	60	100	500	700

**Table 2** Basic statistical values obtained from the data

Soil attributes	Interval	Mean	SD	<i>K</i>	<i>S</i>	CV (%)	Guide value <sup>a</sup>	Alert <sup>b</sup>	>Alert <sup>c</sup>
Ni (mg/kg)	0.71–69	15.18	8.81	3.50	2.22	50.32	64	54	5
Log Ni	1.05–1.66	1.18	0.14	0.17	1.03	–	–	–	–
Zn (mg/kg)	19–587	150.1	113.1	1.79	1.47	55.33	19	89	15
Log Zn	1.27–2.68	2.06	0.30	–0.45	0.08				
Pb (mg/kg)	10–935	93.17	102.6	3.22	5.24	18.05	04	92	27
Log Pb	1.01–1.66	1.83	0.30	0.25	1.04				
Cu (mg/kg)	6–684	48.87	87.39	3.26	5.82	65.22	63	19	41
Log Cu	0.77–2.49	1.46	0.35	0.35	0.77	–	–	–	–
pH	3.91–4.62	4.22	0.30	–1.22	0.22	7	–	–	–
OM (%)	0.32–16.13	4.99	6.78	1.73	1.39	–	–	–	–
CEC (meq/100 cm <sup>3</sup> )	4.60–19.85	11.74	6.92	–2.90	0.08	–	–	–	–

*SD* standard deviation, *K* kurtosis, *S* skewness, *CV* coefficient of variation

<sup>a</sup> Unpolluted

<sup>b</sup> Polluted

<sup>c</sup> Very polluted

**Table 3** Semivariogram models of heavy metals and their parameters

Soil attributes	Model	$C_0$	$C + C_0$	$C_0 / (C + C_0)$	Range (m)	$R^2$
Ni	Exponential	0.22	0.0155	0.354	900	0.30
Zn	Exponential	0.32	0.90	0.355	600	0.85
Pb	Exponential	0.345	0.96	0.359	900	0.72
Cu	Exponential	0.30	0.91	0.329	640	0.81

$C_0$  nugget variance,  $C$  structural variance,  $(C + C_0)$  sill variance

higher than those for Pb, suggesting that Cu, Zn, and Ni had greater variations among the soil samples; such high variations could possibly be influenced by extrinsic factors, such as human activities. Values for soil pH, organic matter (OM), and CEC are also shown in Table 2.

Table 3 shows the semivariogram parameters for the soil samples that were logarithmically transformed. The range values of semivariograms for Ni and Pb were similar and around 900 m; these were much greater than those for Zn and Cu (around 600 m). All of the Nug/Sill ratios for the four metals were less than 0.36 and  $R^2$  was between 0.30 and 0.85.

The experimental semivariograms for the heavy metals in soil are compared with the fitted models in Fig. 3. The results showed that soil containing Ni, Zn, Pb, and Cu were best fit with the exponential model.

The ordinary kriging technique was used here to predict attribute values at unsampled points in the study area. Figure 4 presents the spatial patterns generated from the semivariograms for the four heavy metals in the studied area. The spatial distribution maps showed similar geographical trends, especially for Ni (Fig. 4a), Pb (Fig. 4c),

and Cu (Fig. 4d), with higher concentrations in the north-east and decreasing presence toward the southwest. However, Zn (Fig. 4b) showed an inverse trend with higher concentrations in the southwest and decreasing presence toward the northeast.

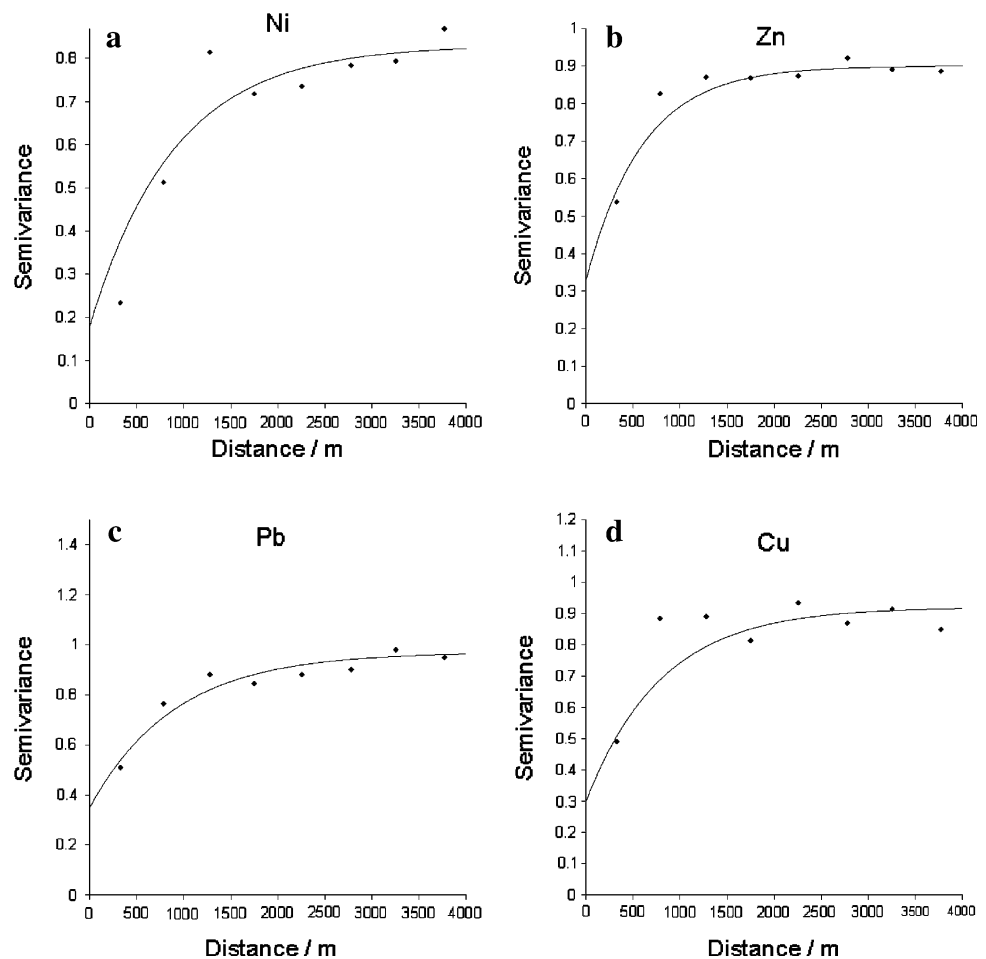
To assess pollution risks, the fuzzy classification approach described in “Materials and methods” was used to compare the guide values in Table 1 with estimates of soil pollutant content obtained by ordinary kriging. The value for the KSD of Ni was 1.09 mg kg<sup>-1</sup>, that for Zn was 16.73 mg kg<sup>-1</sup>, that for Pb was 19.86 mg kg<sup>-1</sup>, and Cu had 20.11 mg kg<sup>-1</sup>. The resulting MFs for these heavy metals are shown in Fig. 5.

After each pixel from the spatial distribution maps was fuzzy classified, 8-bit grayscale images were generated in such a way that unpolluted values were mapped to the interval [0, 65], low risk to [66, 127], moderate risk to [128, 191], and high risk to [192, 255]. For example, if a certain pixel has membership values of 0.2 for the unpolluted class, 0.9 for the low-risk class, 0 for the moderate-risk class, and 0 for the high-risk class, then the maximum fuzzy operator selects 0.9 (the low-risk class), and in the output image the value 121 is assigned to that pixel (Fig. 6).

## Discussion

Spatial variability can be theoretically estimated by the ratio Nug/Sill in geostatistical studies, and the result may be used as a criterion for measuring the spatial dependence of regional variables. The results in Table 2 show spatial dependence between 0.32 and 0.36. These values

**Fig. 3** Experimental semivariograms for heavy metals in soil compared with fitted models: **a** Ni, **b** Zn, **c** Pb, and **d** Cu



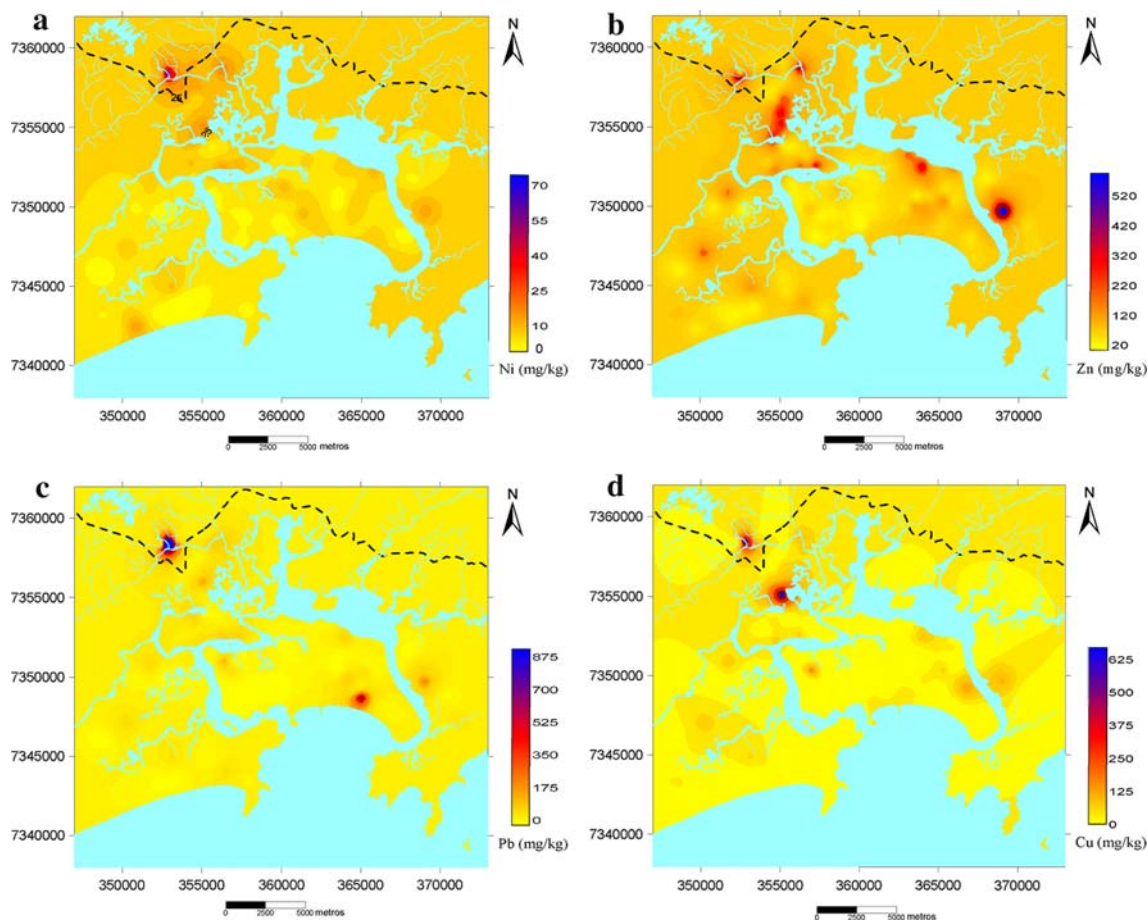
characterized moderate spatial dependence between concentrations of heavy metals in soil. This is a feature present in areas under anthropic and industrial influence (Chang et al. 1998; Chien et al. 1997; Cambardella et al. 1994).

As to contributions of anthropic processes in the studied area, we analyzed correlations between heavy metal variables and soil properties (OM, CEC, AP, AK, and pH) using a significance level of 5%. Significant positive correlations were found between Ni and Pb ( $r = 0.625$ ), between Ni and Cu ( $r = 0.682$ ), and between Cu and Pb ( $r = 0.608$ ). The correlations for these three heavy metals are related to the geographical distributions in the study area, indicating a trend of spatial concentrations increasing from southeast to northwest (Fig. 4a, c, d). However, correlations of Zn with the other heavy metals were weaker, with values below  $r = 0.35$ . In Fig. 4b the geographical distribution of Zn shows an inverse trend, with values increasing from northwest to southeast.

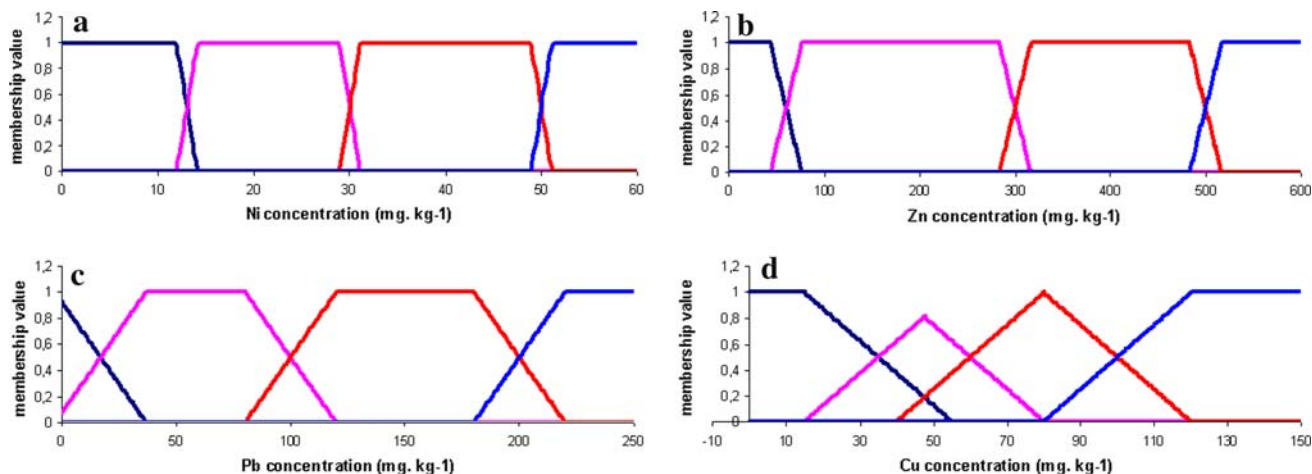
Among heavy metals and some soil properties, correlation measures were all positive. For instance, values for pH were between 0.192 and 0.554. For Cu and AP the value of  $r = 0.196$  was obtained and with Pb a value of

$r = 0.252$ . The heavy metal correlations in soil and soil properties suggest that there are locations in which anthropic factors cause significant changes in environmental characteristics (Liu et al. 2006; Carnelo et al. 1997; Yang et al. 1995).

The mean and maximum contents of Ni, Zn, Pb, and Cu in Table 2 were considered high compared to the corresponding guide values suggested by maximum levels of concentration (São Paulo State Basic Sanitation 2001). Thus, these four elements exhibit a risk for environmental pollution and pose a possible threat to human health. Of the 123 samples, the following numbers of samples had concentrations exceeding the maximum levels set by the guide values: 55 for Ni, 89 for Zn, 92 for Pb, and 19 for Cu. These results suggest that some cultivated lands in the study area are polluted. In addition, the following numbers of samples had concentrations that exceeded the minimum critical values: 5 for Ni, 15 for Zn, 27 for Pb, and 41 for Cu. These results suggest that some industrial and urbane areas are polluted. These conclusions match results from risk assessment. In this study, the tested soils were taken mainly from plains and valleys in the studied area. The parent



**Fig. 4** Spatial distribution maps generated by ordinary kriging for heavy metals in soil: **a** Ni, **b** Zn, **c** Pb, and **d** Cu



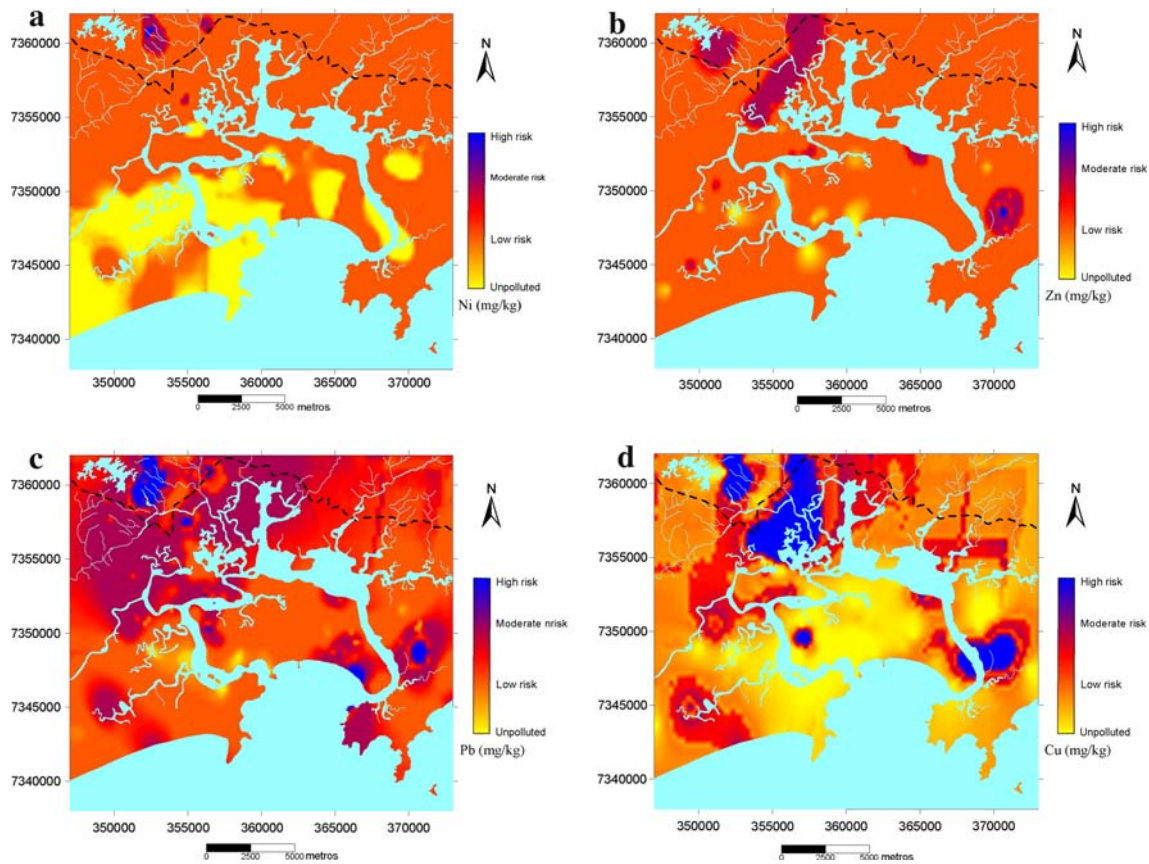
**Fig. 5** Membership functions: **a** Ni, **b** Zn, **c** Pb, and **d** Cu

materials were recent terraces, recent alluvial deposits, and undifferentiated terraces (all of quaternary age), with characteristics of being easily enriched by heavy metals.

In recent years, rapid industrialization in the study area has led to more township enterprises, such as cement

production, a chemical-petroleum industry, and factories for fertilizers, plaster, paper, and steel. Most of these industries are located in the northwest-southwest section (Fig. 1). These industries would inevitably present potential risks from heavy metal pollution. Based on the similar





**Fig. 6** Risk pollution maps of soil generated by Fuzzy classification: **a** Ni, **b** Zn, **c** Pb, and **d** Cu

geographical distribution between Pb and Cu risk and the township enterprises, especially cement and petroleum-chemical plants, it could be extrapolated that pollution discharge from industrial production might partially contribute to Ni and Zn pollution. These findings should be taken into account in decision making related to adjustments in the industrial infrastructure.

Using (1) Fuzzy classification to stratify geochemical data into four classes, from unpolluted to highly polluted, (2) legally prescribed guide values, and (3) KSD (Fig. 5), we have produced maps that clearly indicate the pollution risk levels of chemical elements (Fig. 6). The maps show that the presence of Cu and Pb are associated with older urban areas and areas of dense traffic. The concentrations of Ni are randomly distributed and are low throughout the entire sampled area, except in the southeastern section. Concentrations of Zn are intermediate in urban areas and decrease toward the edges of those areas. The presence of these four elements is obviously associated with anthropogenic activity. This is supported by the fact that the highest concentrations of the four elements occur at the port and at oil refineries, steelworks, and fertilizer plants.

This work demonstrates that fuzzy classification methods can be successfully combined with geostatistical

methods to analyze data. Since the results represent the most likely membership of a certain sampled point into one of four risk classes, we have a certain degree of confidence that a sampled point has been classified correctly. Nevertheless, the samples were point data and some form of interpolation was necessary to produce continuous spatial maps. Although data interpolation always casts some doubt on the results, the general trends in the original data are preserved. Consequently, this type of mapping can be used to help optimize public health policies for urban areas.

**References**

Alloway BJ (2001) Heavy metals in soils. Wiley, New York  
 Amini M, Afyuni M, Fathianpour N, Khademi H, Flüher H (2005) Continuous soil pollution mapping using fuzzy logic and spatial interpolation. *Geoderma* 124:223–233. doi:10.1016/j.geoderma.2004.05.009  
 Burrough PA (1992) Development of intelligent geographical information systems. *Int J Geogr Inf Sci* 6:1–12. doi:10.1080/02693799208901891  
 Burrough PA, Mcdonell R (2004) Principles of geographical information systems. Oxford University Press, Oxford  
 Burrough PA, Van Gaans P, Hootsmans RJ (1997) Continuous classification in soil survey: spatial correlation, confusion and

- boundaries. *Geoderma* 77:115–136. doi:[10.1016/S0016-7061\(97\)00018-9](https://doi.org/10.1016/S0016-7061(97)00018-9)
- Cambardella CA, Moorman TB, Novak JM, Parkin TB, Turco RF, Konopka AE (1994) Field-scale variability of soil properties in central Iowa soils. *Soil Sci Soc Am J* 58:1501–1511
- Carnelo LGL, de Miguez SR, Marba'n L (1997) Heavy metals input with phosphate fertilizers used in Argentina. *Sci Total Environ* 204:245–250. doi:[10.1016/S0048-9697\(97\)00187-3](https://doi.org/10.1016/S0048-9697(97)00187-3)
- Cattle AJ, McBratney AB, Minasny B (2002) Kriging methods evaluation for assessing the spatial distribution of urban soil lead contamination. *J Environ Qual* 319:1576–1588
- CETESB (2001) São Paulo State Basic Sanitation and Technology Company. Relatório de estabelecimento de valores orientadores para solos e águas subterrâneas no Estado de São Paulo. CETESB, São Paulo, p 70
- Chang L, Burrough PA (1987) Fuzzy reasoning: a new quantitative aid for land evaluation. *Soil Surv Land Eval* 7:69–80
- Chang YH, Scrimshaw MD, Emmerson RHC, Lester JN (1998) Geostatistical analysis of sampling uncertainty at the Tollesbury managed retreat site in Blackwater Estuary, Essex, UK: kriging and cokriging approach to minimize sampling density. *Sci Total Environ* 221:43–57. doi:[10.1016/S0048-9697\(98\)00262-9](https://doi.org/10.1016/S0048-9697(98)00262-9)
- Chien YL, Lee DY, Guo HY, Houg KH (1997) Geostatistical analysis of soil properties of mid-west Taiwan soils. *Soil Sci* 162:291–297. doi:[10.1097/00010694-199704000-00007](https://doi.org/10.1097/00010694-199704000-00007)
- De Gruijter JJ, Walvoer DJJ, Van Gaans PFM (1997) Continuous soil maps: a fuzzy set approach to bridge the gap between aggregation levels of process and distribution models. *Geoderma* 77:169–195. doi:[10.1016/S0016-7061\(97\)00021-9](https://doi.org/10.1016/S0016-7061(97)00021-9)
- Dias NL, do Carmo DR, Rosa AH (2006) Selective sorption of mercury (II) from aqueous solution with an organically modified clay and its electroanalytical application. *Sep Sci Technol* 41:733–746. doi:[10.1080/01496390500526896](https://doi.org/10.1080/01496390500526896)
- Dobermann A, Oberthür T (1997) Fuzzy mapping of soil fertility a case study on irrigated riceland in the Philippines. *Geoderma* 77:317–339. doi:[10.1016/S0016-7061\(97\)00028-1](https://doi.org/10.1016/S0016-7061(97)00028-1)
- Feng Qi, A-Xing Zhu, Harrower M, Burt JE (2006) Fuzzy soil mapping based on prototype category theory. *Geoderma* 136:774–787. doi:[10.1016/j.geoderma.2006.06.001](https://doi.org/10.1016/j.geoderma.2006.06.001)
- Franssen H, Eijnsbergen AC, Stein A (1977) Use of spatial prediction techniques and fuzzy classification for mapping soil pollutants. *Geoderma* 77:243–262. doi:[10.1016/S0016-7061\(97\)00024-4](https://doi.org/10.1016/S0016-7061(97)00024-4)
- Goovaerts P (1997) *Geostatistics for natural resources evaluation*. Oxford University Press, Oxford
- Gringarten E, Deutsch CV (2001) Teacher's aide: variogram interpretation and modeling. *Math Geol* 33:507–534. doi:[10.1023/A:1011093014141](https://doi.org/10.1023/A:1011093014141)
- Isaaks E, Srivastava RM (1989) *An introduction to applied geostatistics*. Oxford University Press, Oxford
- Jiang H, Eastman JR (2000) Application of fuzzy measures in multi-criteria evaluation in GIS. *Int J Geogr Inf Sci* 14(2):173–184. doi:[10.1080/136588100240903](https://doi.org/10.1080/136588100240903)
- Journel AG, Huijbregts C (1978) *Mining geostatistics*. Academic Press, London
- Landim PMB (2003) *Statistical analysis of geological data*. Ed. UNESP, São Paulo
- Lin YP, Chang TK, Teng TP (2001) Characterization of soil lead by comparing sequential Gaussian simulation, simulated annealing simulation and kriging methods. *Environ Geol* 41:189–199. doi:[10.1007/s002540100382](https://doi.org/10.1007/s002540100382)
- Liu XM, Wu JJ, Xu JM (2006) Characterizing the risk assessment of heavy metals and sampling uncertainty analysis in paddy field by geostatistics and GIS. *Environ Pollut* 141:257–264. doi:[10.1016/j.envpol.2005.08.048](https://doi.org/10.1016/j.envpol.2005.08.048)
- Lourenço RW, Landim PMB (2005) Risk mapping of public health through geostatistics methods. *Rep Public Health* 21(1):150–160
- McBratney AB, De Gruijter JJ, Brus DJ (1992) Spatial prediction and mapping of continuous soil classes. *Geoderma* 54:39–64. doi:[10.1016/0016-7061\(92\)90097-Q](https://doi.org/10.1016/0016-7061(92)90097-Q)
- McGraph D, Zhang CS, Carton O (2004) Geostatistical analyses and hazard assessment on soil lead in Silvermines, area Ireland. *Environ Pollut* 127:239–248. doi:[10.1016/j.envpol.2003.07.002](https://doi.org/10.1016/j.envpol.2003.07.002)
- Moreira-Nodermann LM (1987) *Geochemistry and environment*. *Geoquímica Brasileira* 1:89–107
- Odeh IOA, McBratney AB, Chittleborough DJ (1992) Fuzzy-c-means and kriging from mapping soil as a continuous system. *Soil Sci Soc Am J* 56:1848–1854
- Olea RA (1999) *Optimum mapping techniques using regionalized variable theory*. Kansas Geological Survey, Kansas
- Rocha JC, Sargentini E, Zara LF, Rosa AH, dos Santos A, Burba P (2003) Reduction of mercury (II) by tropical river humic substances (Rio Negro). Part II. Influence of structural features (molecular size, aromaticity, phenolic groups, organically bound sulfur). *Talanta* 61:699–707. doi:[10.1016/S0039-9140\(03\)00351-5](https://doi.org/10.1016/S0039-9140(03)00351-5)
- Romic M, Romic D (2003) Heavy metals distribution in agricultural topsoils in urban area. *Environ Geol* 43:795–805
- Shi J, Wang H, Xu J, Wu J, Liu X, Zhu H, Yu C (2007) Spatial distribution of heavy metals in soils: a case study of Changxing, China. *Environ Geol* 52:1–10. doi:[10.1007/s00254-006-0443-6](https://doi.org/10.1007/s00254-006-0443-6)
- Stein J, Schettler T, Wallinga D, Vallenti M (2002) In harm's way: toxic threats to child development. *J Dev Behav Pediatr* 1(Suppl 1):S13–S22
- Tang HW, Huel G, Compagna D, Hellier G, Boissinot C, Blot P (1999) Neurodevelopment evaluation of 9-month old infants exposed to low levels of Pb in vitro: involvement of monoamine neurotransmitters. *J Appl Toxicol* 19:167–172. doi:[10.1002/\(SICI\)1099-1263\(199905/06\)19:3<167::AID-JAT560>3.0.CO;2-8](https://doi.org/10.1002/(SICI)1099-1263(199905/06)19:3<167::AID-JAT560>3.0.CO;2-8)
- Van Gaans PFM, Burrough PA (1993) The use of fuzzy logic and continuous classification in GIS applications. In: Harts JJ, Ottens HFL, Scholten HJ (eds) *Proceedings of UGIS'93*, Utrecht, Amsterdam
- Van Meirvenne M, Goovaerts P (2001) Evaluating the probability of exceeding a site-specific soil cadmium contamination threshold. *Geoderma* 102:63–88. doi:[10.1016/S0016-7061\(00\)00105-1](https://doi.org/10.1016/S0016-7061(00)00105-1)
- Webster R, Oliver MA (2001) *Geostatistics for environmental scientists*. Wiley, Chichester, pp 89–96
- Winneke G, Walkowiak J, Lillenthal H (2002) PCB induced neurodevelopmental toxicity in human infants and its potential mediation by endocrine dysfunction. *Toxicology* 181–182:161–165. doi:[10.1016/S0300-483X\(02\)00274-3](https://doi.org/10.1016/S0300-483X(02)00274-3)
- Yang HD, Nong SW, Cai SM (1995) Environmental geochemistry of sediments in Lake Dong Hu, Wuhan. *Chin Sci Abstr Ser B* 14:23
- Yang JH, Derr-Yellin EC, Kodavanti PR (2003) Alterations in brain protein kinase c isoforms following developmental exposure to a polychlorinated biphenyl mixture. *Mol Brain Res* 111:123–135. doi:[10.1016/S0169-328X\(02\)00697-6](https://doi.org/10.1016/S0169-328X(02)00697-6)
- Zadeh LA (1965) Fuzzy sets. *Inf Control* 8(3):338–353. doi:[10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)
- Zhang CS, Selinus O (1998) Statistics and GIS in environmental geochemistry—some problems and solutions. *J Geochem Explor* 64:339–354. doi:[10.1016/S0375-6742\(98\)00048-X](https://doi.org/10.1016/S0375-6742(98)00048-X)
- Zhang CS, Zhang S, Zhang LC, Wang LJ (1995) Background contents of heavy metals in sediments of the Changjiang River system and their calculation methods. *J Environ Sci* 7:422–429