Available online at www.sciencedirect.com
ScienceDirect
Journal of Geochemical Exploration 96 (2008) 117–131
JOURNAL OF GEOCHEMICAL EXPLORATION
www.elsevier.com/locate/jgeoexp

# Arsenic concentration in the soils of the Brenta Plain (Northern Italy): Mapping the probability of exceeding contamination thresholds

F. Ungaro [a,*], F. Ragazzi [b], R. Cappellin [b], P. Giandon [b]

[a] *CNR-IRPI, Istituto di Ricerca per la Protezione Idrogeologica, Unità Pedologia Applicata, Via Madonna del Piano 10, 50019 Sesto Fiorentino, Italy*
[b] *EPA Veneto Region, Soil Survey, Castelfranco Veneto, Italy*

Received 20 September 2006; accepted 5 March 2007
Available online 7 June 2007

**Abstract**

Traditional geochemical mapping is of little help when the uncertainty associated with the estimated values at unsampled locations is required to support decision making. In the Brenta Plain in Northeastern Italy (1350 km$^2$), soils are characterized by a pedo-geochemical background value for arsenic which is higher (36 mg kg$^{-1}$) than the regulatory threshold (20 mg kg$^{-1}$), and it can prove difficult to distinguish between geogenic enrichment and anthropogenic pollution. A simple indicator kriging with varying local means calibrated on the soil map 1:50,000 was used to infer the local conditional cumulative distribution function (ccdf) of As concentration in topsoil (0–40 cm) and subsoil (70–120 cm). The use of local uncertainty models based on the estimated ccdf allowed the assessment of the probabilities to exceed critical thresholds. At the probability levels corresponding to the observed marginal probabilities of values above the regulatory threshold, and using the local background values as reference thresholds, it was found that As concentrations exceed these values at 9 and 7% of the interpolated locations, respectively, for topsoil and subsoil. The computation of the top enrichment factor (TEF), calculated as the ratio between topsoil and subsoil estimated concentrations, and its combination with the probability map of exceeding the usual background value for topsoil, allowed the identification of areas affected by anthropogenic As enrichment, which characterize about 6% of the study area.


*Keywords:* Arsenic; Indicator kriging; Probability maps; Soil contamination; Top enrichment factor

## 1. Introduction

In the last decades, the increase of industrialization and the introduction of new technologies applied to activities traditionally characterized by a low environmental impact such as agriculture have enhanced the risk of soil contamination by pollutants. Among these, heavy metals, due to their persistence, accumulate in soils possibly reaching concentrations that are above regulatory contamination thresholds.

Recent legislative regulations (DM 471/99 for polluted sites reclamation, Ministero della Salute, 1999) introduced regulatory thresholds in Italy which, when exceeded, would require mandatory cleanup operations. A derogation from the law limits is allowed only if the natural background values for a given area are higher than those specified by the law. Within this

* Corresponding author. Tel.: +39 055 5226557; fax: +39 055 5226550.
*E-mail address:* fabrizio.ungaro@irpi.cnr.it (F. Ungaro).

0375-6742/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gexplo.2007.03.006

framework, the EPA of the Veneto Region started a surveying campaign in 2002, aimed at assessing heavy metals in the soils of the Veneto plain. The purpose was to evaluate background values for the different catchments in relationships with soil and terrain morphology. Owing to its environmental relevance, the first survey was carried out in the Venice lagoon watershed, whose soils were mainly formed over the Brenta river sediments. Previous studies (Giandon et al., 2000, 2004) pointed out that in the study area the levels of Sn and, to a greater extent, those of As, are above the regulatory limits. As is the 20th most abundant element in the Earth's crust ($4.01 \times 10^{16}$ kg, Matschullat, 2000) and is almost ubiquitous in soils. It is estimated that about 60% of arsenic present in the environment is of anthropogenic origin (Nriagu and Pacyna, 1988), with an annual addition to soil of about $98 \times 10^3$ kg As year$^{-1}$ (Bhumbla and Keefer, 1994). As occurs in soils at a global average concentration of 5–6 mg kg$^{-1}$ (Peterson et al., 1981), with a potential anthropogenic arsenic input in arable soils equal to 2.18 mg kg$^{-1}$ in the year 2000 (Han et al., 2003). The main factors affecting As concentration in soils are rock composition and human activities such as mining, smelting, combustion of fossil

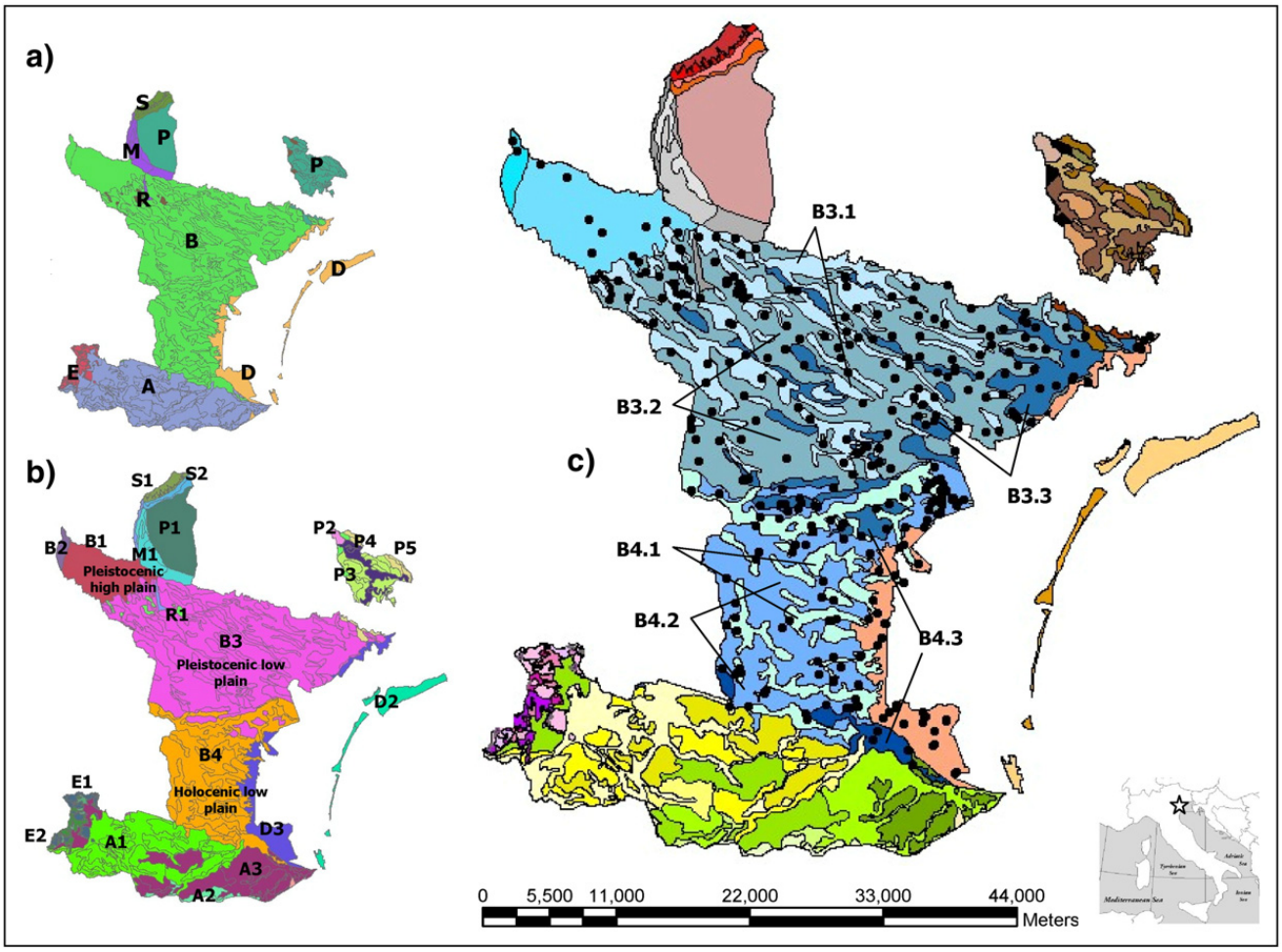


Fig. 1. Soil map of the Venice lagoon watershed: (1.a) landscape districts (B: Brenta alluvial plain; P: Piave alluvial plain; A: Adige alluvial plain; M: Musone alluvial plain; R: resurgences plain; D: lagoon and coastal plain; S: Asolo hills; E: Euganei hills); (1.b) landscape systems (B1: Brenta Pleistocenic high plain; B2. Brenta Holocenic high plain; B3: Brenta Pleistocenic low plain; B4: Brenta Holocenic low plain; P1: Piave Pleistocenic high plain; P2: Piave Holocenic high plain; P3: Piave Pleistocenic low plain; P4: Piave Holocenic low plain with partially decarbonated soils; P5: Piave Holocenic low plain with non-decarbonated soils; A1: Adige Holocenic low plain with partially decarbonated soils; A2: Adige Holocenic low plain with non-decarbonated soils; A3: Adige Holocenic low plain with hydromorphic soils and organic matter accumulation; M1: Musone Holocenic alluvial plain with decarbonated soils; M2: Musone Holocenic alluvial plain with non-decarbonated soils; R1: resurgences lowlands; D1: recent coastal plain with decarbonated soils and organic matter accumulation; D2: recent coastal plain with non-decarbonated soils; D3: reclaimed lagoons; S1: hill slopes on carbonatic substrata; S2: foothill fans with decarbonated soils; E1: hill slopes on silicatic substrata with acid soils; E2: hill slopes on carbonate substrata with decarbonated soils; E3: foothill fans); (1.c) Soil landscape units (B3.1: fluvial ridges; B3.2: undifferentiated alluvial plain; B3.3: depressions of the alluvial plain; B4.1: fluvial ridges; B4.2: undifferentiated alluvial plain; B4.3: depressions of the alluvial plain) and location of sampling points in the Brenta alluvial plain. Simple bold dots identify sampling points.

fuels, pesticides and herbicides applications. The parent material is the most important factor affecting As content. Other factors are soil texture and soil organic matter content (Chen et al., 2002). Literature data suggest an average value of 20 mg $kg^{-1}$ for Italian soils (range 1.8–60 mg $kg^{-1}$, Yan-Chu, 1994), but widespread hot spots occur in volcanic and mining areas (Costagliola et al., 2004; Donati et al., 2005).

Traditional geochemical mapping is of little help when the uncertainty associated to the estimated values at unsampled locations is required to support decision making. Although geostatistics have been successfully used in the past decades to model and display the spatial variability of the most different environmental variables, kriging variance as a measure of uncertainty for kriging estimates have been misused in most cases due to the limiting assumptions of normality of errors distribution and of homoscedasticity, i.e., that error variance is independent of data values (Van Meirvenne and Goovaerts, 2001). Such conditions are rarely met in practice, as environmental variables are characterized in most cases by skewed distributions.

Aim of this paper is to map the probability to exceed As regulatory thresholds and background values using a non-parametric geostatistical approach integrating soft information from the local soil map 1:50,000 in order to derive the local conditional cumulative distribution function (ccdf) of As concentration. Knowledge of the ccdf at any unsampled location allowed for uncertainty assessment; the adoption of a threshold probability level in combination with the top enrichment factor (TEF), calculated as the ratio between topsoil and subsoil estimated concentrations, provided a new criterion to identify anthropogenic As enrichment to soils with the specific objective to guide and improve remediation decision.

## 2. Material and methods

### 2.1. Study area

The study area (1350 $km^2$) is located in Northeast Italy (Fig. 1) and is the distal part of the Brenta river plain, which represents the central and more extended portion of the Venice lagoon watershed. The Brenta Plain is part of the widest Italian soil region (more than 50,000 $km^2$; European Soil Bureau, 1998), the Pianura Padano–Veneta, characterized by a temperate-suboceanic climate, with a minimum and a maximum average annual temperature of 8.4 and 17.4 °C respectively and a mean annual precipitation ranging from 725 to 1030 mm. The surveyed soils, mainly on quaternary alluvial deposits, are characterized by a high degree of heterogeneity, ranging from fine textured Inceptisosl with vertic characteristics (Aquertic Eutrudepts) to coarse textured Entisols (Typic Udipsamments), to soils characterized by organic horizons (Terric Sulfisaprists). The soil map of the Venice lagoon watershed 1:50,000 (ARPAV, 2004) based on a hierarchical soil landscape model, identifies 8 landscape districts (Fig. 1.a), distinguished in terms of main river catchments (with the Brenta, Piave and Adige being the most relevant), 23 landscape systems (Fig. 1.b), identified within each district based on the position within the landscape (for example high and low plains) and the age of the soil parent material (for example Holocenic and Pleistocenic deposits), and 56 soil landscape units, distinguished on the basis of soil and terrain morphology relationships, which are the soil mapping units (SMUs) displayed on the map (Fig. 1.c).

The main soil properties are summarized in Table 1. For topsoil and subsoil samples a wide range has been found for all parameters: sand 0–99%, silt 3–86%, clay

Table 1
Descriptive statistics for topsoil and subsoil soil properties

| | | | Organic carbon | $CaCO_3$ | Clay | Silt | Sand | CEC |
|---|---|---|---|---|---|---|---|---|
| | | pH | % | % | % | % | % | meq/100 g |
| Topsoil | Number of observations | 212 | 212 | 187 | 210 | 210 | 210 | 208 |
| | Mean | 7.93 | 1.17 | 10.06 | 21.25 | 45.12 | 33.62 | 16.35 |
| | SD | 0.49 | 0.96 | 9.59 | 10.63 | 13.06 | 17.98 | 7.77 |
| | Minimum | 5.58 | 0.13 | 0.00 | 1.10 | 2.40 | 0.60 | 0.00 |
| | Median | 8.10 | 0.95 | 6.00 | 18.61 | 47.20 | 31.69 | 14.95 |
| | Maximum | 8.81 | 9.07 | 43.00 | 59.38 | 79.20 | 96.50 | 43.63 |
| Subsoil | Number of observations | 212 | 210 | 188 | 207 | 207 | 207 | 207 |
| | Mean | 8.15 | 0.69 | 19.91 | 17.33 | 46.30 | 36.37 | 14.32 |
| | SD | 0.63 | 2.76 | 15.82 | 12.36 | 23.82 | 30.79 | 10.87 |
| | Minimum | 4.80 | 0.00 | 0.00 | 0.00 | 0.90 | 0.00 | 0.00 |
| | Median | 8.29 | 0.20 | 20.00 | 15.18 | 51.00 | 26.80 | 13.20 |
| | Maximum | 8.98 | 30.75 | 58.00 | 62.30 | 86.23 | 99.10 | 76.50 |

Table 2
Descriptive statistics for topsoil and subsoil heavy metals concentrations

| | | As | Cd | Co | Cr | Cu | Hg | Ni | Pb | Zn |
|---|---|---|---|---|---|---|---|---|---|---|
| Topsoil | Number of observations | 201 | 212 | 202 | 212 | 202 | 208 | 212 | 212 | 202 |
| | Mean (mg kg$^{-1}$) | 20.15 | 0.52 | 10.12 | 32.45 | 41.03 | 0.13 | 26.78 | 29.92 | 94.46 |
| | SD (mg kg$^{-1}$) | 7.97 | 0.27 | 2.86 | 16.32 | 24.75 | 0.22 | 23.66 | 12.54 | 24.93 |
| | Minimum (mg kg$^{-1}$) | 3.72 | 0.00 | 0.00 | 9.60 | 0.00 | 0.00 | 10.50 | 0.00 | 15.60 |
| | Median (mg kg$^{-1}$) | 18.80 | 0.49 | 9.94 | 29.95 | 33.50 | 0.06 | 23.80 | 27.90 | 93.20 |
| | Maximum (mg kg$^{-1}$) | 49.30 | 1.65 | 17.80 | 155.00 | 174.50 | 1.24 | 337.00 | 102.00 | 166.00 |
| Subsoil | Number of observations | 210 | 212 | 202 | 212 | 202 | 212 | 211 | 212 | 202 |
| | Mean (mg kg$^{-1}$) | 23.75 | 0.46 | 9.82 | 25.36 | 21.71 | 0.04 | 24.78 | 18.53 | 73.73 |
| | SD (mg kg$^{-1}$) | 20.39 | 0.38 | 3.54 | 13.85 | 11.95 | 0.14 | 11.92 | 11.70 | 30.03 |
| | Minimum (mg kg$^{-1}$) | 2.17 | 0.00 | 2.89 | 0.00 | 0.00 | 0.00 | 5.72 | 0.00 | 11.90 |
| | Median (mg kg$^{-1}$) | 20.25 | 0.43 | 9.33 | 23.80 | 20.50 | 0.00 | 23.30 | 17.35 | 71.40 |
| | Maximum (mg kg$^{-1}$) | 169.00 | 1.69 | 20.30 | 95.30 | 89.60 | 0.99 | 78.30 | 80.70 | 152.00 |

0–62%, organic carbon 0.06–30.8%, total carbonates 0–58%, pH 4.8–8.9, CEC 0.01–76.5 cmol(+) kg$^{-1}$.

For 212 representative profiles topsoil and subsoil were analyzed for heavy metals. Aqua regia soluble fraction was determined for As, Cd, Co, Cu Cr, Hg, Ni, Pb and Zn by ICP (Inductive Coupled Plasma; ISO/DIS 11466, 1995). Heavy metals contents were measured at two depth intervals: 0–40 cm and 70–120 cm. Sampling points (0.2 km$^2$) were selected following a "typological approach" (ISO/DIS 19258, 2004) based on parent material characteristics, soil type and land use. Sampling profiles were located only in agricultural land. Descriptive statistics for the analyzed metals are summarized in Table 2. Only in the case of As are mean values at both depths slightly above the legal threshold for contamination (20 mg kg$^{-1}$); in the case of subsoil samples the median value (20.3 mg kg$^{-1}$) is almost coincident with the regulatory threshold. Subsoil samples are assumed to be representative of the natural or pedo-geochemical background values, while topsoil samples have been use to infer the usual background values (ISO/DIS 19528, 2004). The frequency distribution histograms of As concentrations at the two sampling depths are shown in Fig. 2. The distributions are notably positively skewed at both depths but the upper tail is much more consistent for the deeper sampling interval (70–120 cm); it is interesting to note eventually that the mean subsoil concentration (23.7 mg kg$^{-1}$) is higher than the mean topsoil concentration (20.2 mg kg$^{-1}$), on the contrary of what occurs for the other measured metals.

Statistical correlation analysis and principal component analysis (PCA) were used to highlight the relationships between the concentration of As at the two depths and the concentrations of the other elements: the separation of topsoil and subsoil heavy metals concentrations in different components can be due to different prevailing origins (Borůvka et al., 2005) and can support the interpretation of the spatial distribution patterns.

The assessment of the accumulation of elements in soils can be carried out in many ways, using indexes that compare soil concentrations to a reference value or element, or concentrations measured at different depths (Loska et al., 2002). In this study, in order to assess possible soil contamination due to anthropogenic As addiction, the top enrichment factor (TEF, Facchinelli et al., 2001) has been calculated as the ratio between topsoil and subsoil contents. Values above 2 are generally considered to point out relevant anthropogenic input to the soil.

### 2.2. Geostatistical modeling of local uncertainty

When modeling the uncertainty of an attribute $z$ at an unsampled location $\mathbf{u}$ (with $\mathbf{u}$ representing a vector of

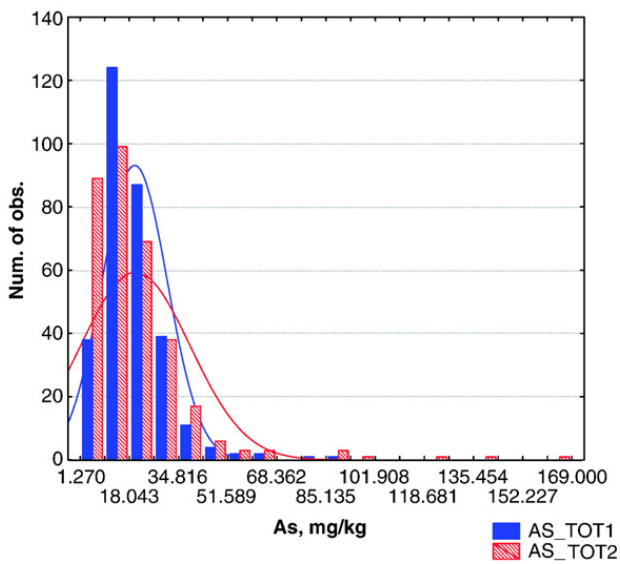


Fig. 2. Histograms of As concentration (mg kg$^{-1}$) at the two sampling depths (AS_TOT1: 0–40 cm; AS_TOT2: 70–120 cm).

coordinates), the available observations are considered as the realization of one set of $n$ spatially correlated random variables $Z(\mathbf{u})$. The conditional cumulative distribution function (ccdf) at any specific location $\mathbf{u}$ can be written as:

$$F(\mathbf{u} : z|(n)) = \text{Prob}\{Z(\mathbf{u}) \leq z|(n)\}$$

and it fully describes the uncertainty in $\mathbf{u}$ since it provides the probability that the unknown value is greater than any given threshold value $z$ made conditional to the information available ($n$) (Goovaerts, 1997).

The ccdf can be fully characterized adopting an analytical model described by a number of parameters or adopting a non-parametric approach which does not assume any particular shape for $F(\mathbf{u}: z|(n))$. In this second case the ccdf is modeled adopting a series of $K$ threshold values $z_k$ which discretizes the range of variation of $z$:

$$F(\mathbf{u} : z_k|(n)) = \text{Prob}\{Z(\mathbf{u}) \leq z_k|(n)\} \quad k = 1,\ldots,K$$

The above conditional probability can be viewed as the conditional expectation of an indicator random variable $I(\mathbf{u}; z_k)$, with $I(\mathbf{u}; z_k)=1$ if $Z(\mathbf{u}) \leq z_k$ and $I(\mathbf{u}; z_k)=0$ if $>z_k$, given the $n$ available observations. Thus it is possible to infer the ccdf by interpolation of indicator transforms of the data, once a set of $K$ thresholds has been defined using a multiple indicator kriging (Journel, 1983).

Preliminary to any further step in the spatial analysis is then the selection of a number of thresholds and their values, followed by coding of available data in indicator variables for each threshold. The number of thresholds should be between 5 and 15 (Goovaerts, 1997); it is common practice to adopt the nine deciles of the sample cumulative distribution plus critical values of practical relevance such as remediation thresholds. Each observation is then transformed into a vector of $K$ values, either 1 or 0 depending upon its relationship to a cutoff value $z_k$:

$$i(\mathbf{u}_\alpha; z_k) = \begin{cases} 1 & \text{if} z(\mathbf{u}_\alpha) \leq z_k \\ 0 & \text{otherwise} \end{cases} \quad k = 1,\ldots,K$$

Simple or ordinary kriging of the set of indicator-transformed values provides a resultant value between 0 and 1 which is, in effect, an estimate of the proportion of the values in the neighborhood which are smaller than the threshold value. Detailed descriptions of the indicator kriging and its formalisms are given by Journel (1983) and Goovaerts (1997).

Because of the impact of the different soil properties, As concentration shows a clear spatial trend that must be taken into account in the estimation procedure. For this reason, a simple indicator kriging with varying local means (Goovaerts and Journel, 1995; Goovaerts et al., 1997) was used. The local indicator means at each threshold were calculated for each soil mapping unit (SMU) of the 1:50,000 soil map, and the indicator residuals were calculated for the $K$ thresholds, as the difference between the indicator variables and the local indicator mean:

$$r(\mathbf{u}, z_k) = i(\mathbf{u}_\alpha; z_k) - y(\mathbf{u}_\alpha; z_k)$$

where $y(\mathbf{u}_\alpha; z_k)$ is the local mean of the indicator for the threshold $k$ at location $\mathbf{u}_\alpha$.

The experimental semivariograms of the indicator residuals at each threshold are then computed as:

$$\hat{\gamma} R(h; z_k) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} \left\{ r(\mathbf{u}_\alpha; z_k) - r(\mathbf{u}_\alpha + h; z_k) \right\}^2$$

The simple kriging estimator can be written as:

$$[F(\mathbf{u}; z_k)|(n+1)]^* = y(\mathbf{u}; z_k) + \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha(z_k) \{ i(\mathbf{u}_\alpha; z_k) - y(x_0; z_k) \}$$

where $y(\mathbf{u}; z_k)$ is the local indicator mean at the $k$ threshold at the unsampled location $\mathbf{u}$ to be estimated, and $\lambda_\alpha(z_k)$ are the kriging weights to be derived by solving a simple kriging system.

Once the ccdf has been derived at unsampled locations over a regular grid, there are several ways to incorporate the estimated uncertainty in the decision-making process. The unknown concentration $z(\mathbf{u})$ can be estimated by means of the ccdf using a least-square error criterion in order to obtain the so called E-type estimate (Deutsch and Journel, 1998):

$$z_E^*(\mathbf{x}_0) = \int_{-\infty}^{+\infty} z dF(\mathbf{x}_0; z|(n))$$

Furthermore, the definition of the complete ccdf allows for evaluating of the probability of exceeding a critical threshold $z_c$ at $\mathbf{u}$:

$$Prob\{Z(\mathbf{u}) > z_c|(n)\} = 1 - F(\mathbf{u}; z_c|(n))$$

and to assess the risk of misclassifying a location as safe or contaminated once a probability threshold has been identified.

All the geostatistical analyses presented in this paper were carried out with the geostatistical software Wingslib

1.3.1 (Statios, 2000), which works in conjunction with the GSLIB90 executables (Deutsch and Journel, 1998).

## 3. Results and discussion

### 3.1. Data analysis and As background values

Soil texture and related properties such as CEC exert a clear influence on As concentration in soils, and it appears clearly that the nature of soil parent material controls the distribution of As concentrations in soils to a relevant extent. Moderately strong and statistically significant correlations are detected between topsoil As concentration and clay content ($r=0.40$), sand content ($r=-0.39$), CEC ($r=0.32$), total carbonates ($r=-0.30$), and silt content ($r=0.23$). The correlations exhibited with pH, and organic matter are insignificant in the first case ($r=-0.01$) and very weak in the second ($r=0.13$). Subsoil As concentrations show similar correlations with soil properties, with the exception of total carbonates which exhibits a weaker correlation with As ($r=-0.18$). As an example, Fig. 3 shows the relationships between total As concentration in top and subsoil and clay content. The mean concentrations calculated for both topsoil and subsoil samples, grouping data in terms of soil textural classes defined according to the USDA particle size limits (USDA, 1993), showed that only sandy, loamy sandy, sandy loam and clay loam soil samples have a mean As concentration below the regulatory threshold of 20 mg kg$^{-1}$ (Fig. 4).

Figs. 5 and 6 show the As mean concentrations for land system units and soil landscape units respectively. Samples from the Brenta Pleistocenic low plain (landscape system unit B3) show a higher mean As concentration (26.9 mg kg$^{-1}$, Fig. 5) as compared to the mean concentration observed in the unit B4 (Brenta Holocenic low plain, 16.3 mg kg$^{-1}$) and D3 (drained lagoon basin,17.6 mg kg$^{-1}$). In terms of soil landscape units, it is interesting to note that within the landscape systems B3 and B4 of the Brenta Plain (Fig. 6), As mean concentrations are higher in the fine textured soils of the depressions of the plains (Aquic Eutrudepts, fine, mixed mesic, units B3.3 and B4.3) and lower in the coarser soils of the reliefs (Dystric Eutrudepts, coarse loamy, mixed mesic, units B3.1 and B4.1 ), with intermediate concentration characterizing the soils of the basal level of the plains (Oxyacquic Eutrudepts, fine silty, mixed mesic, units B3.2 and B4.2). The highest As mean concentrations are found in the depressions (30.2 mg kg$^{-1}$) and in the basal level (28.2 mg kg$^{-1}$) of the Brenta Pleistocenic low plain, while the lowest values characterizes the reliefs of the Brenta Holocenic plain (13.5 mg kg$^{-1}$) and the paleodunes (5.6 mg kg$^{-1}$).

These results show that local trends in As concentration are clearly detectable as related to soil texture and landscape features. These thus provide valuable information to integrate into the procedure to assess As spatial distribution and to map the probability to exceed regulatory thresholds.

The correlation matrix of metals concentrations in topsoil and subsoil is shown in Table 3. Focusing on As,

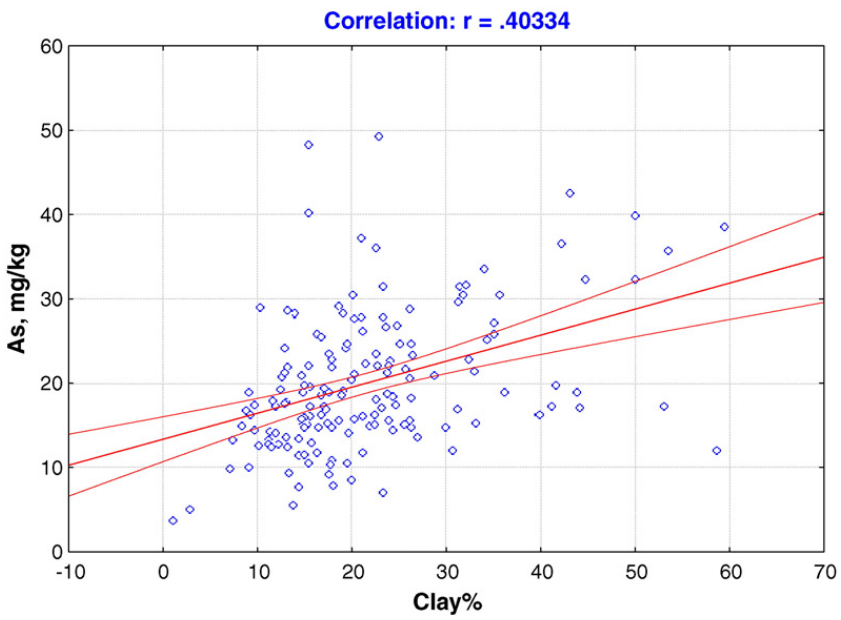


Fig. 3. Clay content (%) vs. As concentration (mg kg$^{-1}$).

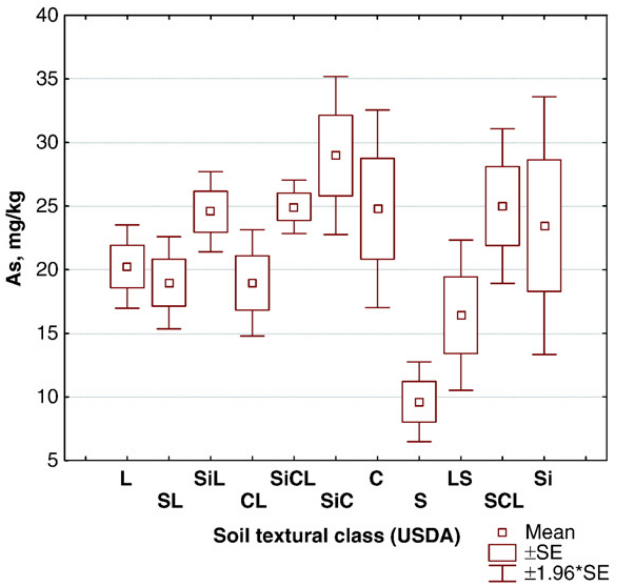


Fig. 4. As concentrations (mg kg$^{-1}$) for the different soil textural class (USDA, 1993). C: clay; CL: clay loam; L: loam; LS: loamy sand; S: sand; SL: sandy loam; SCL: sandy clay loam; Si: silt; SiL: silt loam; SiCL: silty clay loam.

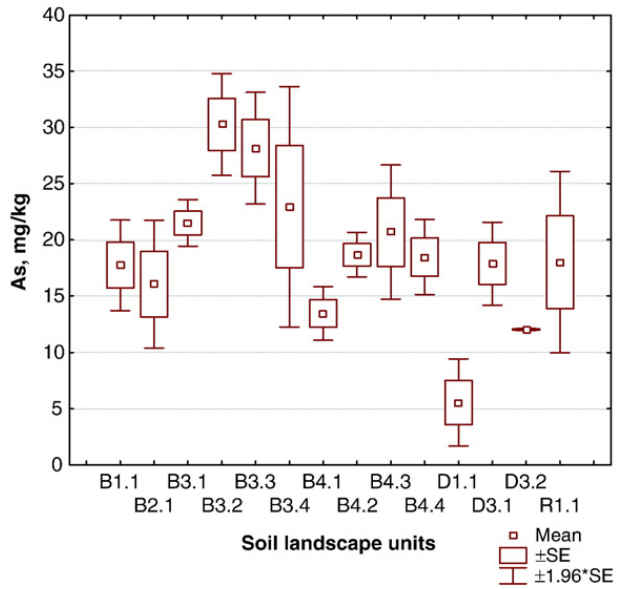


Fig. 6. As concentrations (mg kg$^{-1}$) for the different soil landscape units.

in the case of topsoil samples significant correlations ($p<0.05$) are observed with Pb ($r=0.57$) and Zn ($r=0.33$) and secondarily with Co ($r=0.30$) and Cd ($r=0.22$). In the case of subsoil samples, As concentration is significantly correlated only with Cd concentration ($r=0.60$). The relationships between the metals are quite complex, differentiated at the two sampling depths and difficult to explain on an individual basis, but they would suggest a different origin for topsoil and subsoil

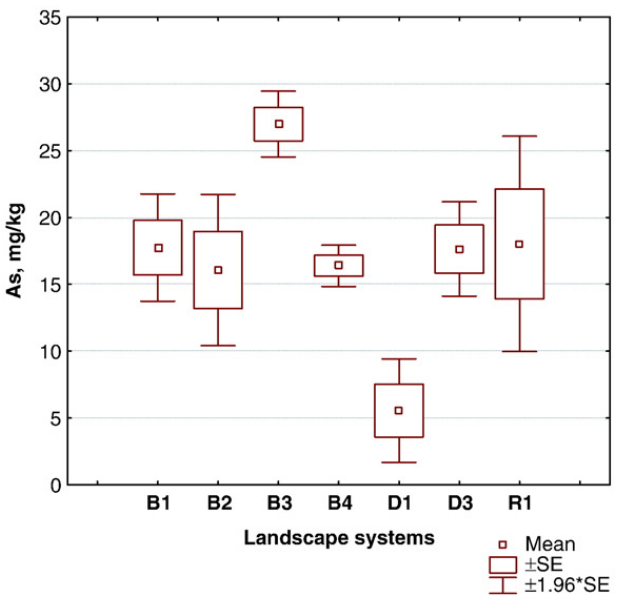


Fig. 5. As concentrations (mg kg$^{-1}$) for the different landscape systems.

As. This is confirmed by the PCA results shown in Table 4. The analysis was applied to the whole data set, standardized to zero mean and unit variance, treating the concentration in topsoil and subsoil as separate variables. The components were rotated using a Varimax normalized rotation (Kaiser, 1958) which maximizes the variances of the squared normalized factor loadings across variables for each factor. In order to account for more than 70% of the observed variance, the first five principal components with eigenvalues greater than one were selected. The largest loadings for the first component, which accounts for more than 37% of observed variance, was observed for subsoil Co, Cr, Cu, Ni and Zn. The second component showed the largest loadings for topsoil Cu, Pb and Zn followed by topsoil As and Cd. The third component was again mostly represented by topsoil concentrations with the largest loadings for Ni and Cr. The fourth component was the only one to include topsoil and subsoil concentrations, but for a single element, namely Hg. The fifth and last component had large loadings for subsoil As and Cd followed by topsoil As. The separation of topsoil and subsoil heavy metals concentrations observed for all the elements with the exception of Hg indicates that the prevailing origins of the element in the topsoil should be different from those of the subsoil. In the case of As, though, topsoil concentration is on one side clearly related to topsoil Pb and Zn, as shown by the loadings of the second component, but on the other side, it is to some extent related to subsoil As concentration, since its loading ranks third in those calculated for the fifth

Table 3
Correlation matrix of total concentrations in topsoil and subsoil

| | | As | Cd | Co | Cr | Cu | Hg | Ni | Pb | Zn |
|---|---|---|---|---|---|---|---|---|---|---|
| Topsoil | As | 1.000 | | | | | | | | |
| | Cd | **0.217** | 1.000 | | | | | | | |
| | Co | **0.297** | **0.170** | 1.000 | | | | | | |
| | Cr | −0.010 | **−0.159** | **0.418** | 1.000 | | | | | |
| | Cu | −0.037 | **0.307** | 0.120 | −0.099 | 1.000 | | | | |
| | Hg | 0.056 | **0.361** | 0.054 | −0.089 | 0.090 | 1.000 | | | |
| | Ni | −0.037 | **−0.159** | **0.650** | **0.740** | 0.012 | −0.106 | 1.000 | | |
| | Pb | **0.573** | **0.339** | **0.361** | **0.171** | **0.535** | **0.213** | 0.012 | 1.000 | |
| | Zn | **0.328** | **0.575** | **0.470** | 0.118 | **0.354** | **0.400** | 0.100 | **0.535** | 1.000 |
| Subsoil | As | 1.000 | | | | | | | | |
| | Cd | **0.596** | 1.000 | | | | | | | |
| | Co | 0.244 | **0.557** | 1.000 | | | | | | |
| | Cr | −0.065 | 0.246 | **0.687** | 1.000 | | | | | |
| | Cu | 0.141 | **0.415** | **0.633** | **0.808** | 1.000 | | | | |
| | Hg | −0.183 | **0.465** | 0.192 | 0.178 | 0.194 | 1.000 | | | |
| | Ni | 0.117 | 0.199 | **0.678** | **0.922** | **0.805** | −0.082 | 1.000 | | |
| | Pb | 0.225 | **0.561** | **0.859** | **0.614** | **0.729** | 0.352 | **0.531** | 1.000 | |
| | Zn | 0.093 | **0.513** | **0.844** | **0.813** | **0.777** | **0.455** | **0.693** | **0.894** | 1.000 |

Correlations significant at $p<0.05$ are in bold.

component, which nevertheless represent only 5.5% of the total variance.

With the exception of Co and Ni, topsoil and subsoil mean concentrations are significantly different at $p<0.05$ (non-parametric Kolmogorov–Smirnov two samples test), but only in the case of As subsoil concentrations are, on the average, higher than topsoil concentrations, suggesting a prevailing geogenic origin for this element. The correlation between As concentration at the two depths is significant at $p<0.05$ ($r=0.39$), indicating that to some extent topsoil concentration is affected by that of the soil parent material. All elements

Table 4
Factor loadings (FL), estimated communalities and eigenvalues

| | FL1 | FL2 | FL3 | FL4 | FL5 | Communality |
|---|---|---|---|---|---|---|
| *Topsoil* | | | | | | |
| As | 0.122 | 0.581 | 0.142 | −0.125 | 0.529 | 0.405 |
| Cd | −0.040 | 0.577 | −0.004 | 0.252 | 0.375 | 0.374 |
| Co | 0.326 | 0.418 | 0.244 | −0.040 | 0.012 | 0.780 |
| Cr | 0.034 | 0.049 | **0.935** | 0.015 | 0.137 | 0.559 |
| Cu | 0.208 | **0.809** | −0.168 | 0.042 | −0.116 | 0.804 |
| Hg | 0.042 | 0.292 | −0.155 | **0.731** | 0.033 | 0.649 |
| Ni | 0.204 | 0.137 | **0.915** | −0.056 | −0.105 | 0.781 |
| Pb | 0.012 | **0.809** | 0.122 | 0.087 | 0.181 | 0.525 |
| Zn | 0.212 | **0.777** | 0.255 | 0.295 | 0.220 | 0.350 |
| *Subsoil* | | | | | | |
| As | 0.394 | 0.275 | −0.085 | −0.007 | **0.715** | 0.583 |
| Cd | 0.371 | 0.021 | 0.079 | 0.125 | **0.793** | 0.834 |
| Co | **0.830** | 0.189 | 0.039 | 0.038 | 0.273 | 0.480 |
| Cr | **0.755** | −0.091 | 0.094 | 0.067 | 0.010 | 0.743 |
| Cu | **0.783** | 0.335 | −0.008 | 0.093 | 0.205 | 0.619 |
| Hg | 0.200 | −0.012 | 0.093 | **0.849** | 0.056 | 0.690 |
| Ni | **0.788** | 0.022 | 0.188 | 0.030 | 0.117 | 0.904 |
| Pb | 0.652 | 0.202 | 0.006 | 0.124 | 0.493 | 0.609 |
| Zn | **0.777** | 0.267 | 0.036 | 0.260 | 0.401 | 0.443 |
| Eigenvalues | 6.68 | 2.22 | 1.91 | 1.27 | 1.00 | 11.13 |
| Cumulative %variance | 37.11 | 49.47 | 60.10 | 67.15 | 72.69 | |

Loadings >0.70 are in bold.

show higher average topsoil concentrations, with TEF mean values always >1, ranging from 1.1 for Co to 6.7 for Cd, with Cu, Hg and Pb with mean TEF>2. This suggests local surface addiction of Cd, Cu, Hg and Pb of anthropogenic origin. In the case of As the TEF has a mean value over the entire study area of 1.22, ranging from a minimum of 0.19 to a maximum of 8.90, with a standard deviation equal to 0.91 and a consistent upper tail with 13% of values above 2 and a skewness of 3.94, suggesting localized topsoil addictions of As of anthropogenic origin. In this case statistically significant differences among the soil mapping units were not detected, and locations characterized by TEF>2 appear randomly scattered in the area.

Usual and natural background values for As concentrations were determined following the ISO international standards (ISO/DIS 19528, ISO, 2004). Descriptive statistics identified as outliers values >40 mg kg$^{-1}$ in the case of topsoil samples, and values >51.3 mg kg$^{-1}$ in the case of subsoil samples; after log transformation of data values, the outliers were checked once more, and if they were true outliers they were not retained for the calculation of the 90th percentile which was assumed to represent the upper boundary of the background values. The upper limit of the pedo-geochemical background value defined for subsoil samples following the ISO international standards is higher (36 mg kg$^{-1}$) than the regulatory threshold (20 mg kg$^{-1}$), and the same occurs for the upper limit of the usual background value defined for topsoil samples (30 mg kg$^{-1}$).

### 3.2. Spatial analysis and exceedence of As regulatory thresholds

As the first step in the spatial analyses, moving windows statistics (MWS) were computed for As concentrations at the two sampling depths, adopting windows of 3000×3000 m in size and 50% overlapping. MWS revealed a strong proportional effect due to anomalies in local variability, highlighting a linear direct relationship between mean values and variances at both depths with $R^2$ of 0.40 and 0.36, respectively for topsoil and subsoil. This has a great impact on prediction uncertainty since in this case kriging variance is not a reliable measure of estimates reliability and cannot be used to calculate the probability of exceedence regulatory thresholds in order to delineate hazardous or potentially contaminated areas. In this case a more appropriate approach is provided by indicator kriging (Journel, 1983). For both topsoil and subsoil samples, the nine deciles of the sample cumulative distributions have been adopted as cutoffs (Table 5), along with the two regulatory thresholds of 20 and 50 mg kg$^{-1}$. The transformed variables were used to calculate the indicator means for each soil landscape unit at each cutoff at the two sampling depths: these represent the trend surfaces at the nodes of the estimation grid. Eventually the difference between the indicator transforms and the indicator means, provided the indicator residuals for calculating the semivariograms. The parameters of the eleven fitted indicator residuals of the As concentration at the two sampling depths are shown in Table 6. The experimental semivariograms and their fitted models for As subsoil concentrations are shown in Fig. 7. The omnidirectional variogram models were used to infer the ccdfs in correspondence of 500×500 m blocks for a total of 5323 estimation blocks with a simple kriging with varying local means, the latter being provided by the mean indicators values for each mapping unit. Simple block kriging was implemented using a search radius of 7600 m equal to the average largest range of the fitted semivariograms, with the minimum number of neighboring observations to estimate a block set to four and the maximum set to 32 neighbors within the search radius. Each block was discretized in 16 points and an affine correction was used as volume correction for the change of support, adopting a variance reduction factor equal to 0.75 (Deutsch and Journel, 1998). Due to the non-linearity of the indicator transform, the block ccdf cannot be derived as a linear average of the ccdfs at the discretization points. For this reason, the composite probability at each of the selected thresholds $z_k$ is calculated as a linear combination of point indicator data using the simple kriging weights (Goovaerts, 1997). The means of the As concentration ccdfs at both depths are shown in Fig. 8a and b. The topsoil map (Fig. 8a) shows higher As concentrations clearly localized in the depressions of

Table 5
Cutoffs for As concentration indicator transforms

| zk Topsoil | As topsoil (mgkg$^{-1}$) | zk Subsoil | As subsoil (mgkg$^{-1}$) |
|---|---|---|---|
| Decile 1 | 3.9 | Decile 1 | 7.5 |
| Decile 2 | 11.6 | Decile 2 | 10.5 |
| Decile 3 | 14.1 | Decile 3 | 13.2 |
| Decile 4 | 15.7 | Decile 4 | 16.4 |
| Decile 5 | 17.2 | Decile 5=DM 471/99 | 20.3 |
| Decile 6 | 18.9 | Decile 6 | 23.5 |
| DM 471/99 | 20.0 | | |
| Decile 7 | 22.1 | Decile 7 | 26.8 |
| Decile 8 | 27.0 | Decile 8 | 31.7 |
| Decile 9 | 31.6 | Decile 9 | 37.4 |
| DM 471/99 | 50.0 | DM 471/99 | 50.0 |

the Brenta Pleistocenic low plain with a few hotspots in the southeastern portion of the plain characterized by concentrations >30 mg kg$^{-1}$. At this depth, a clear distinction between the Pleistocenic and the Holocenic plains is not as evident as in the case of subsoil As concentration (Fig. 8b). In this case it appears that the basal level of the Pleistocenic plain is characterized by As concentrations between 30 and 40 mg kg$^{-1}$, while in the Holocenic low plain, with the exception of two evident hotspots in the southernmost part, values are between 15 and 25 mg kg$^{-1}$. The probabilities of exceeding the natural background values (36 mg kg$^{-1}$) for topsoil and subsoil As concentration are shown in Fig. 9a and b respectively. The mean probability to exceed the natural background value is of 0.07 in the case of topsoil and of 0.15 in the case of subsoil. These figures would rise up to 0.33 and 0.67 respectively when considering the regulatory threshold of 20 mg kg$^{-1}$ suggesting the existence of an unrealistic heavy As contamination, diffuse in most of the area. In this case in order to avoid the application of unrealistically low mandatory guideline levels, it would be advisable, as suggested also by other authors (Kabata-Pendias and Pendias, 1992), to adopt the site-specific background values defined for the soils developed on the Brenta alluvial deposits.

In order to identify the hazardous or potentially contaminated areas and prioritize possible reclamation interventions, the uncertainty associated to the probability map must be taken into account. A reference threshold could be represented by the marginal probability of exceeding the regulatory threshold, equal to 0.65 and 0.50 respectively for topsoil and subsoil. Given these figures, in the case of subsoil As concentration, 29.3% of the area shows values above the natural background, but only 7% of locations has a probability >0.5 to exceed the natural background. This figure drops to 0.5% in the case of As topsoil concentrations, but in this case it would be safer to consider the usual background value (30 mg kg$^{-1}$): at a probability level of 0.5 about 19% of the area is then characterized by As topsoil estimated concentrations above the usual background value.

In order to identify areas of anthropogenic As accumulation in soils, TEF has been calculated based on the E-type estimates at the two depths, and its distribution has been combined with that of the probability above 0.5 to exceed the usual background value. The locations characterized by a TEF>1 and a probability level >0.5 to exceed the usual background area shown in Fig. 10: they represent about 6% of the study area (312 km$^2$) and should be the area to address

Table 6
Parameters of the eleven fitted residual indicator variograms of the As concentration

| zk (mgkg$^{-1}$) | Nugget | Model 1 | Range1 | Sill1 | Model 2 | Range2 | Sill2 | Total sill | Nugget/Sill |
|---|---|---|---|---|---|---|---|---|---|
| *Topsoil (<40 cm)* | | | | | | | | | |
| 3.9 | 0.38 | Sph | 2800 | 0.36 | Sph | 7600 | 0.28 | 0.64 | 0.59 |
| 11.6 | 0.41 | Sph | 2800 | 0.50 | Sph | 8200 | 0.10 | 0.60 | 0.68 |
| 14.1 | 0.56 | Sph | 3200 | 0.45 | | | | 0.45 | 1.24 |
| 15.7 | 0.70 | Sph | 5400 | 0.31 | | | | 0.31 | 2.26 |
| 17.2 | 0.62 | Sph | 4400 | 0.39 | | | | 0.39 | 1.59 |
| 18.9 | 0.53 | Sph | 1800 | 0.35 | Sph | 10400 | 0.14 | 0.49 | 1.08 |
| 20.0 | 0.68 | Sph | 8000 | 0.34 | | | | 0.34 | 2.00 |
| 22.1 | 0.72 | Sph | 7600 | 0.30 | | | | 0.30 | 2.40 |
| 27.0 | 0.66 | Sph | 4400 | 0.35 | | | | 0.35 | 1.89 |
| 31.6 | 0.52 | Sph | 3800 | 0.48 | | | | 0.48 | 1.08 |
| 50.0 | 0.42 | Sph | 2800 | 0.63 | | | | 0.63 | 0.67 |
| *Subsoil (>70 cm)* | | | | | | | | | |
| 7.5 | 0.27 | Sph | 3000 | 0.49 | Sph | 8000 | 0.25 | 0.74 | 0.36 |
| 10.5 | 0.29 | Sph | 2800 | 0.53 | Sph | 6800 | 0.19 | 0.72 | 0.40 |
| 13.2 | 0.34 | Sph | 2400 | 0.38 | Sph | 7000 | 0.30 | 0.68 | 0.50 |
| 16.4 | 0.19 | Sph | 2600 | 0.56 | Sph | 7200 | 0.26 | 0.82 | 0.23 |
| 20.3 | 0.24 | Sph | 2800 | 0.36 | Sph | 7200 | 0.40 | 0.76 | 0.32 |
| 23.5 | 0.40 | Sph | 2000 | 0.39 | Sph | 6600 | 0.21 | 0.60 | 0.67 |
| 26.8 | 0.57 | Sph | 2000 | 0.16 | Sph | 6400 | 0.27 | 0.43 | 1.33 |
| 31.7 | 0.40 | Sph | 5600 | 0.39 | Sph | 9600 | 0.21 | 0.60 | 0.67 |
| 37.4 | 0.14 | Sph | 3000 | 0.53 | Sph | 7000 | 0.33 | 0.86 | 0.16 |
| 50.0 | 0.56 | Sph | 2800 | 0.19 | Sph | 7400 | 0.26 | 0.45 | 1.24 |

Sph: spherical model.

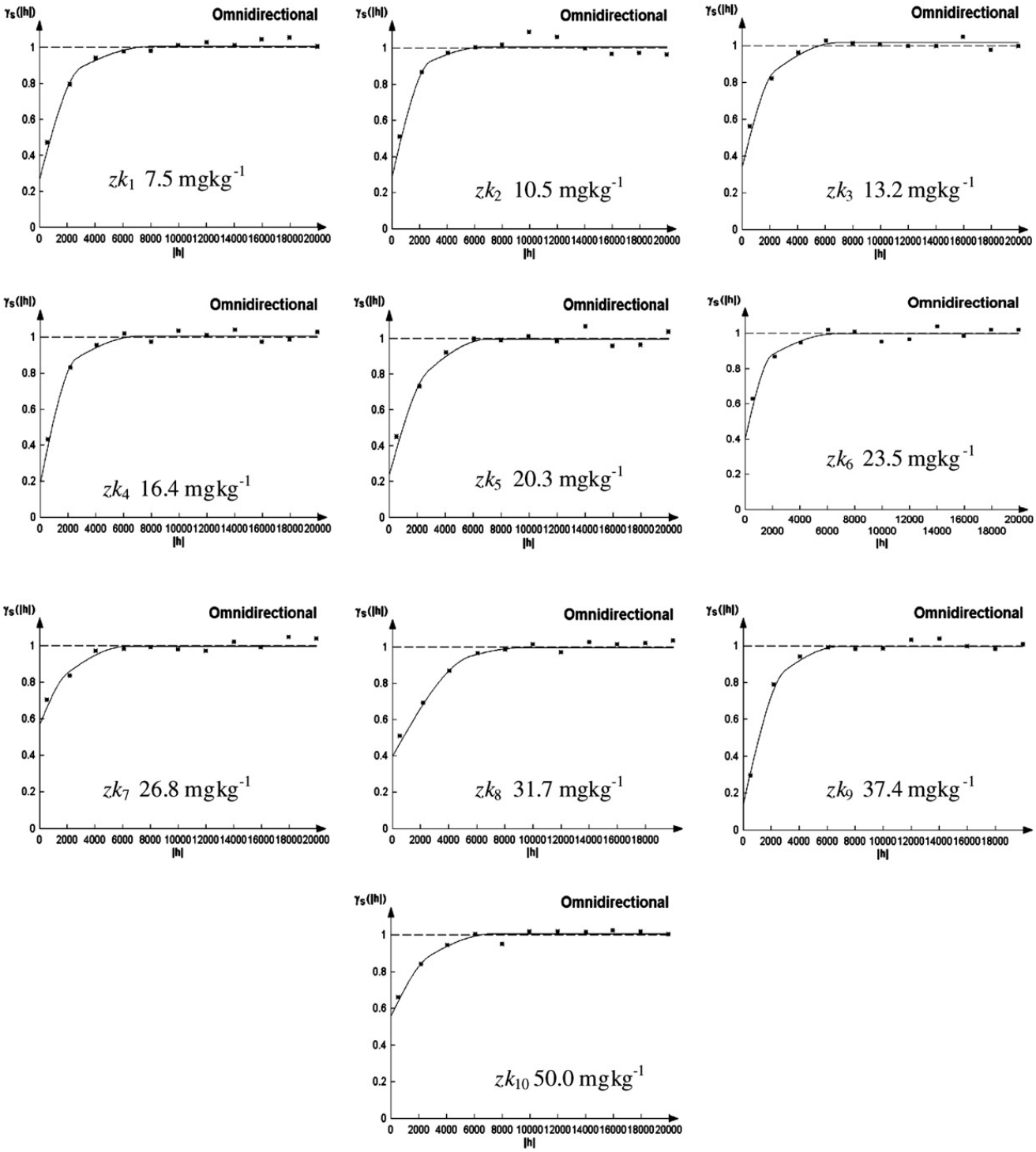


Fig. 7. Experimental and fitted indicator residual semivariograms for subsoil As concentration.

first, in case of cleanup operations planning. For example in Fig. 10b the circle indicates the localization of two distinct sites contaminated with materials from mining activities, and sludge from a local chemical plant for the production of sulfuric acid and sodium sulfate. In one of these areas pyrite ashes, which are a well-known source of As, from the production of sulfuric acid piled up over the years to form mounds between 7 and 16 m above the field level (Provincia di Venezia, 2006).

## 4. Conclusions

Decisions concerning the delineation of As contaminated areas and their possible remediation should take

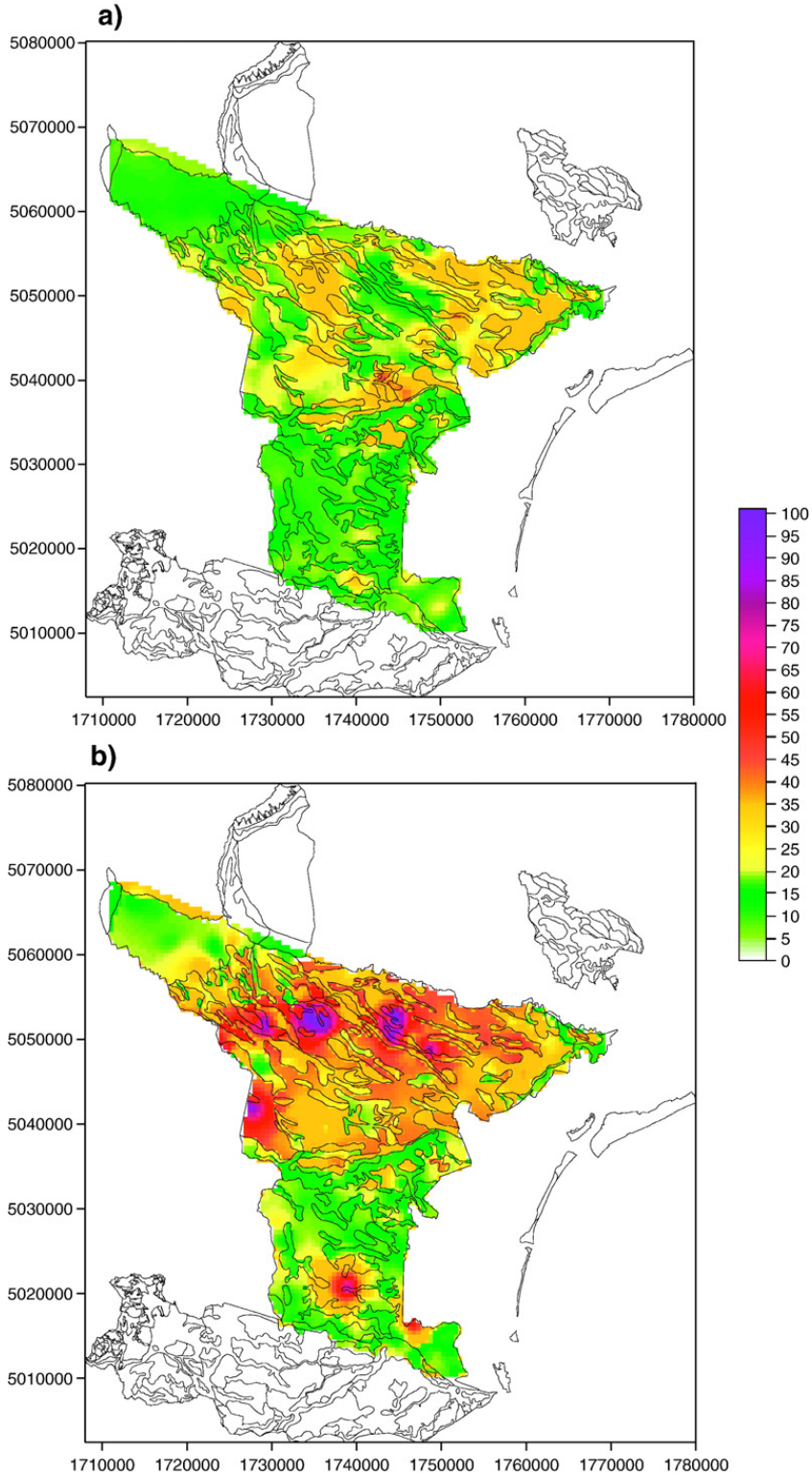


Fig. 8. Map of As E-type estimates (mg kg$^{-1}$); (a) topsoil (0–40 cm); (b) subsoil (70–120 cm).

into account the uncertainty about potential exceedence of relevant thresholds at unsampled locations.

The indicator approach provided estimates of the risk of contamination by kriging the indicator transforms of the measured concentrations, supporting the available data by incorporating soft information derived from the soil map 1:50,000. The calibration of the local soil map allowed the explicit use of the denser soft information derived from the local soil map, once the relevance of soil properties in determining As concentration was assessed.

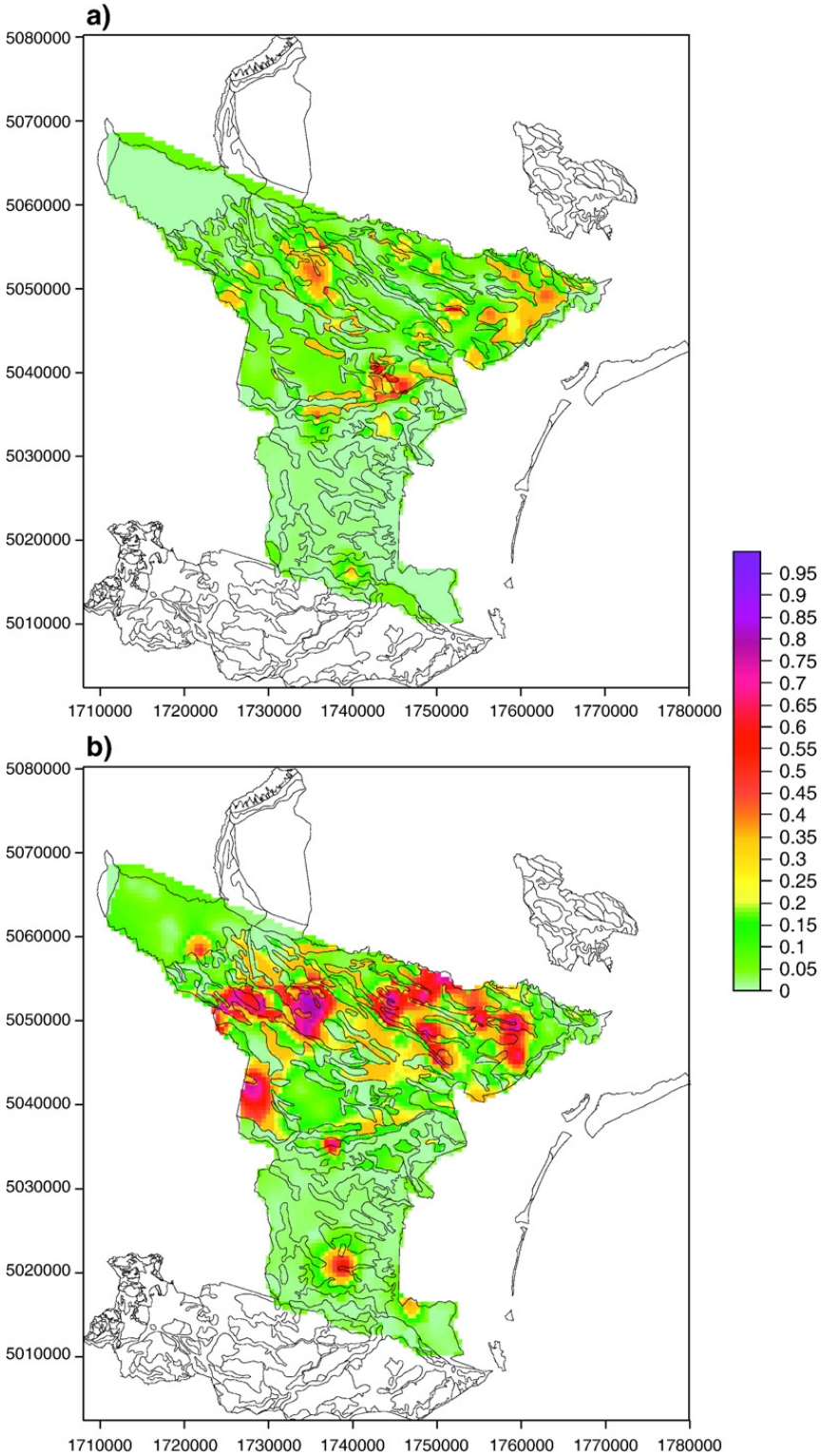


Fig. 9. Probability maps to exceed the natural background value (36 mg $kg^{-1}$); (a) topsoil (0–40 cm); (b) subsoil (70–120 cm).

Results indicate that in the case of As its high concentration in soils is mainly due to lithogenic enrichment, with a spatial distribution which is mainly related to the nature and age of the soil parent material. Nevertheless, topsoil As contamination due to anthropogenic addiction is present in localized areas. A probability threshold of 0.5 to exceed the usual background value combined with a TEF > 1 indicated that about 6% of the area is likely to be interested by As accumulation of anthropogenic origin. Nevertheless, in evaluating these

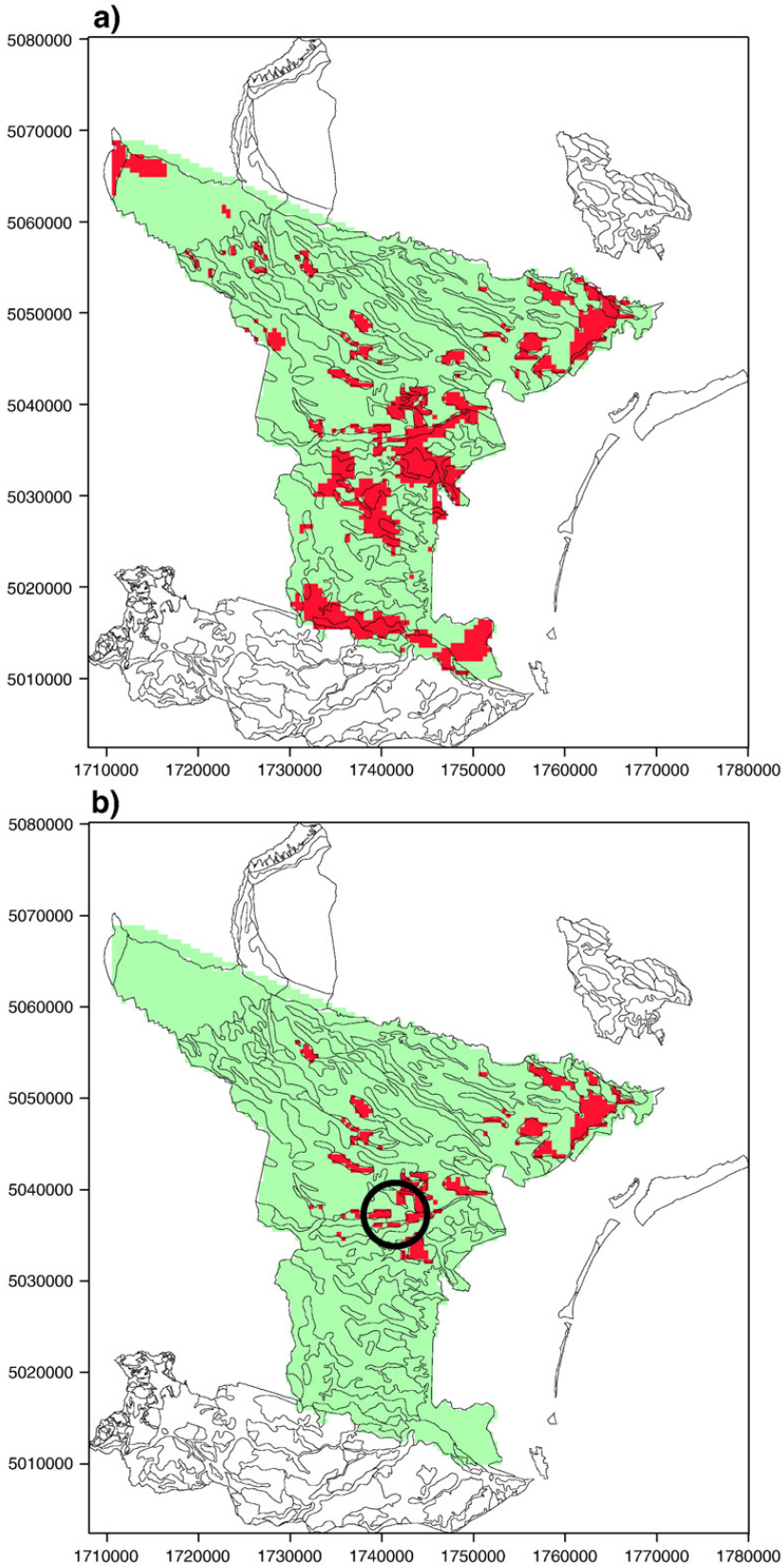


Fig. 10. Top enrichment factor (TEF) map (a) and map of the locations with TEF > 1 and with probability above 0.5 of As concentration > 30 mg $kg^{-1}$ (b).

results, it must be considered that the choice of a probability threshold is always subjective and depends upon political and social decisions as to the acceptance of a given level of toxicity. Furthermore, results show that the use of a single value based on regulatory thresholds for assessing As contamination in soils cannot be adequate and that thresholds based on local variability should be used for properly assessing As contamination.

## Acknowledgements

This research was funded by the Environmental Protection Agency of the Veneto Region (ARPAV), within the framework of the research project "Modelling soil properties affecting groundwater vulnerability to nitrate pollution and heavy metals contamination in the in the Venice Lagoon watershed". The authors whish to thank two anonymous reviewers for their precious suggestions. Finally the authors thank V.L. Palchetti Beard for reviewing the English text.

## References

ARPAV, 2004. Carta dei suoli del bacino scolante in laguna di Venezia. Agenzia regionale per la protezione dell'ambiente del Veneto, Osservatorio regionale suolo. 399 pp.

Bhumbla, D.K., Keefer, R.F., 1994. Arsenic mobilization and bioavailability in soils. In: Nriagu, J.O. (Ed.), Arsenic in the Environment: Part I. Cycling and Characterization. John Wiley and Sons, Inc., pp. 51–82.

Borůvka, L., Vecek, O., Jehlička, J., 2005. Principal component analysis as a tool to indicate the origin of potentially toxic elements in soils. Geoderma 128, 289–300.

Chen, M., Ma, L.Q., Harris, W.G., 2002. Arsenic concentration in Florida surface soils: influence of soil type and properties. Soil Science Society of America Journal 66, 632–640.

Costagliola, P., Benvenuti, M., Benvenuti, M.G., Innocenti, A., Mascaro, L., Paolieri, M., Rossato, L., Tanelli, G., 2004. Arsenic distribution in the Quaternary sediments of the median valley of the Pecora stream (Grosseto, Italy). In: Donati, A., R., Rossi, C., Brebbia, C. (Eds.), Brownfield Sites: II. Assessment, Rehabilitation and Development, Proceedings of the 2nd International Conference on Brownfield Sites: Assessment, by Rehabilitation, and Development, Siena, Italy, pp. 169–177.

Deutsch, C.V., Journel, A.G., 1998. GSLIB, Geostatistical Software Library and User's Guide, 2nd Ed. Oxford University press, New York. 369 pp.

Donati, A., Pulselli, F.M., Riccobono, F., Dallai, L., Francovich, R., Tiezzi, E., 2005. Origin of arsenic pollution in Southwest Tuscany: comparison of fluvial sediments. Annali di Chimica 95 (3–4), 161–166.

European Soil Bureau, 1998. Georeferenced Soil Database of Europe. Manual of procedures. Version 1.0. EUR 18092 EN. European Communities.

Facchinelli, A., Sacchi, E., Mallen, L., 2001. Multivariate statistical and GIS-based approach to identify heavy metal sources in soils. Environmental Pollution 114, 313–324.

Giandon, P., Vinci, I., Fantinato, L., 2000. Heavy metal concentration in the soils of the basin draining the Venice Lagoon. Bollettino della Società Italiana di Scienza del Suolo 49 (1–2), 359–366.

Giandon, P., Cappellin, R., Ragazzi, F., Vinci, I., 2004. Confronto tra il livello naturale e il livello antropico dei metalli pesanti nei suoli della pianura veneta in relazione al materiale di partenza. Bollettino della Società Italiana di Scienza del Suolo 53 (1–2), 540–544.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York. 483 pp.

Goovaerts, P., Journel, A.G., 1995. Integrating soil map information in modeling the spatial variation of continuous soil properties. European Journal of Soil Science 46, 397–414.

Goovaerts, P., Webster, R., Dubois, J.P., 1997. Assessing the risk of soil contamination in the Swiss Jura using indicator geostatistics. Environment and Ecological Statistics 4, 31–48.

Han, F.X., Su, Y., Monts, D.L., Plodinec, M.J., Banin, A., Triplett, G.E., 2003. Assessment of global industrial-age anthropogenic arsenic contamination. Naturwissenschaften 90 (9), 395–401.

ISO, 1995. Soil quality. Extraction of Trace Elements Soluble in Aqua Regia. Dir. ISO/DIS 11466. 6 pp.

ISO, 2004. Soil Quality—Guidance on the Determination of Background Values. Dir. ISO/DIS 19258. 23 pp.

Kabata-Pendias, A., Pendias, H. (Eds.), 1992. Trace elements in soils and plants, 2nd ed. CRC Press, Boca Raton, FL.

Kaiser, H.F., 1958. The Varimax criterion for analytic rotation in factor analysis. Psychometrika 23, 187–200.

Loska, K., Wiechula, D., Barska, B., Cebula, E., Chjnecka, A., 2002. Assessment of arsenic enrichment of cultivated soils in southern Poland. Polish Journal of Environmental Studies 12 (2), 187–192.

Journel, A.G., 1983. Nonparametric estimation of spatial distributions. Mathematical Geology 15 (3), 445–468.

Matschullat, J., 2000. Arsenic in the geosphere—a review. The Science of Total Environment 249, 297–312.

Ministero della Salute, 1999. Decreto ministeriale 25/10/1999, N. 471 Regolamento recante criteri, procedure e modalità per la messa in sicurezza, la bonifica e il ripristino ambientale dei siti inquinati, ai sensi dell'articolo 17 del decreto legislativo 5 febbraio 1997, n. 22, e successive modificazioni e integrazioni. Gazz. Uff. Suppl. Ordin. n° 293 del 15/12/1999, 51 pp.

Nriagu, J.O., Pacyna, J.M., 1988. Quantitative assessment of worldwide contamination of air, water and soils by trace metals. Nature 333, 134–139.

Peterson, P.J., Benson, L.M., Zieve, R., 1981. Metalloids: Section 3. Arsenic. In: Lepp, N.W. (Ed.), Effect of Heavy Metal Pollution in Plants, vol. 1. Applied Science Publ., London, pp. 299–322.

Provincia di Venezia, 2006. Siti inquinati—Interventi di bonifica delle aree inquinate. Retrieved July 14, 2006 from the Province of Venice Environmental Policy web site: http://politicheambientali.provincia.venezia.it/rifiuti/siti/siti.html.

Statios, 2000. WinGslib version 1.3. Statios Software and Services, San Francisco, CA.

USDA, 1993. Soil Survey Manual. United States Department of Agriculture, Handbook, vol. 18. Government Printer, Washington, DC. 437 pp.

Van Meirvenne, M., Goovaerts, P., 2001. Evaluating the probability of exceeding site-specific soil cadmium contamination threshold. Geoderma 102, 75–100.

Yan-Chu, H., 1994. Arsenic distribution in soils. In: Nriagu, J.O. (Ed.), Arsenic in the Environment: Part I. Cycling and Characterization. John Wiley and Sons, Inc., pp. 17–49.