

## The effects of soil pollution and its relation with morbidity in shoreline areas

R.W. Lourenço<sup>1</sup>, S. R. M. M. Roveda<sup>1</sup>, A. H. Rosa<sup>1</sup>, L. F. Fraceto<sup>1</sup>,  
A. Í. Ribeiro<sup>1</sup>, A. C. G. Martins<sup>1</sup>, M. R. Donalisio<sup>2</sup> & R. Cordeiro<sup>2</sup>

<sup>1</sup>*Department of Environmental Engineering, São Paulo State University,  
Sorocaba Campus, Brazil*

<sup>2</sup>*Department Social and Preventive Medicine, Campinas University,  
Brazil*

### Abstract

Soil and subsoil pollution is not only significant in terms of environmental loss, but also a matter of environmental and public health. Solid, liquid and gaseous residues are the major soil contamination agents. They originate from urban conglomerates and industrial areas in which it is impossible to emphasize the chemical, petrochemical and textile industry; thermoelectric, mining, and ironmaster activities. The contamination process can thus be defined as a compound addition to soil, from what qualitative and or quantitative manners can modify soil's natural characteristics and use, producing baneful and deteriorative effects on human health. Studies have shown that human exposition to high concentration of some heavy metals found on soil can cause serious health problems, such as pulmonary or kidney complications, liver and nervous system harm, allergy, and the chronic exposition that leads to death. The present study searches for the correlation among soil contamination, done through a geochemical baseline survey of an industrial contamination area on the shoreline of Sao Paulo state. The study will be conducted by spatial analysis using Geographical Information Systems for mapping and regression analysis. The used data are 123 soil samples of percentage concentration of heavy metals. They were sampled and spatially distributed by geostatistics methods. To verify if there is a relation between heavy metals soil pollution and morbidity an executed correlation and regression analysis will be done using the pollution registers as the independent variables and morbidity as dependable variables. It is expected, by the end of the study, to identify the areas relation between heavy



metals soil pollution and morbidity, moreover to be able to provide assistance in terms of new methodologies that could facilitate soil pollution control programs and public health planning.

*Keywords: heavy metals, soil pollution, spatial analysis, risk analysis, GIS.*

## 1 Introduction

Soil often acts as a filter, purifying and immobilizing a large part of the impurities deposited in it. However, because this capacity is limited, its quality can be negatively impacted by the cumulative effect of the deposition of atmospheric pollutants, agrochemicals and fertilizers, industrial and domestic solid residues, and toxic and radioactive materials [1]. Soil contamination is posing an increasing threat to human health and environmental quality [2]. Among soil and environmental pollutants, heavy metals have received considerable attention over the last few decades [3, 4]. Metals, in particular Cd, Pb and Hg constitute a significant potential threat to human health, both occupational and environmental. 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 outside the workplace. There are, in fact, several metals that are known to be human carcinogens, including arsenic, chromium and nickel. Many of the toxic effects of metals, including carcinogenicity, can be modified by concurrent exposure to other metals. [5–7].

Furthermore, an issue which dominates the political agenda in health is to know how to distribute wisely the increasingly scarce health care resources to meet priority areas, such as those that offer risk to human health due to environmental contamination. Some formal methods consider only the social preferences and economic evaluations of health programs and patients in need of care. This measure combines a simple numerical index of quantity of life (mortality) and quality of life (morbidity) of determined areas for the distribution of health resources.

The spatial analysis of certain environment pollutants and contaminants related to the occurrence of epidemiological morbidity events could be an important implement to assist in planning and management the resources that would meet such demand [8–11]

According to Assunção et al. [12] the analysis of the dispersion of a disease risk is mainly done through maps of incidence rates or some other measure of epidemiological risk. These maps are valuable tools in epidemiological studies and they can present three main objectives from its production: researches that relate environment and public health and also identify patterns of morbidity and mortality in the region of known sources of pollution; the exploratory ones, in order to suggest disease and unknown etiologic factors locations which may be formulated in terms of hypotheses to be further investigated; and finally those who aim to point out associations between potential sources of contamination and areas of high risk.



In order to illustrate the above description, a study in an area traditionally known for being one of the most polluted regions in Brazil will be presented, using soil data and information on general morbidity. Geoprocessing tools will be used using a combination of geostatistics and spatial linear regression. The main objective was to determine whether a correlation between the distribution of contaminants in the soil with the spatial distribution of infant morbidity in children less than one year old affected by diseases of the respiratory and intestinal tract exists.

## 2 Study area

The research was conducted in the Cubatão city, southern coastal region of the State of São Paulo, Brazil. The study area is strong industrial activities in the northeast area with a big petroleum refinery and various chemical activities. There are several wastewater treatment plants in the region with high risk of pollution associated to sewage sludge, and compounds (Figure 1).

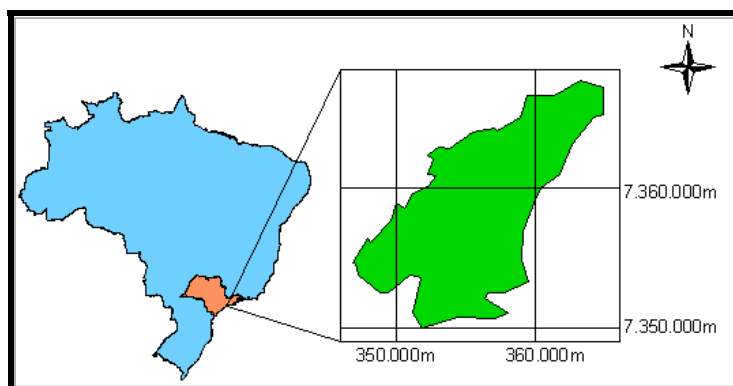


Figure 1: Study area.

In the mid-80s the city was worldwide famous for numerous health and environment problems. It has been stigmatized by the paradigm of “Valley of Death”. By this time, the population lived with 75 different types of pollutants emitted by different sources, including water, soil and air pollution responsible for high local morbidity rates.

Morbidity is a term used to designate the number of cases of a disease or the sum of health problems that affect a group of individuals in a given time within a specific location. An easier way to determine the morbidity of a disease is through the simple count of individuals affected by a given disease. It allows understanding the magnitude of the problem and its short term trend in a population. In this work, morbidity was determined as the health damages to the movement of hospitalization and outpatient care of the study area in the year 2007.

## 2.1 Soil sampling

Soil samples were collected from different locations of the study area at distances between sampling locations ranging from 95 to 650m (Figure 2) and it has been determined the total metal concentrations Cd, Pb and Hg in accordance to the standardized procedure [10].

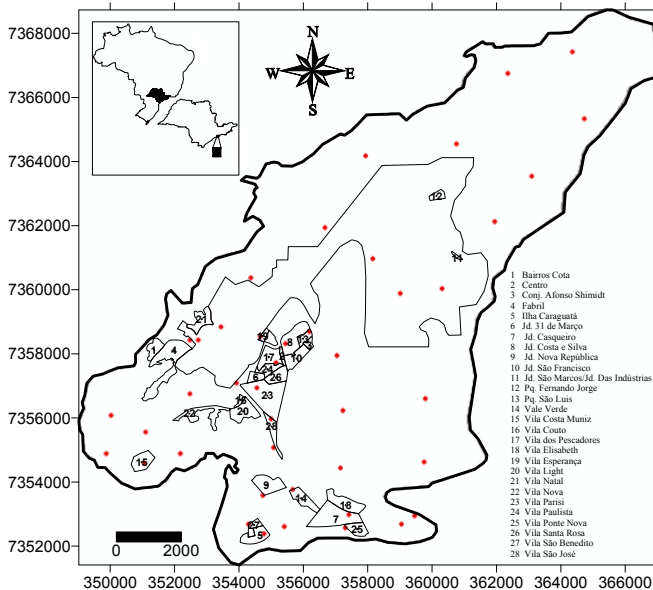


Figure 2: Samples area (read point is soils samples).

## 2.2 Soil variable mapping

Mapping of sample soil was based on geostatistical methods. Geostatistics analysis methods are based upon the assumption that the spatial variation of any continuous attribute is 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, or the distribution of continuous demographic indicators. Geostatistical interpolation is known as *kriging* [11, 13–16]. 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  $z(X_0)$  is given by:

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

with  $\sum_{i=1}^n \lambda_i = 1$ . The weights  $\lambda_i$  are chosen so that the estimate  $\hat{z}(X_0)$  is unbiased, and that the estimation variance  $\sigma_e^2$  is the least, compared to 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:

$$\sigma_e^2 = \sum_{i=1}^n \lambda_i \gamma(x_i, x_0) + \phi \quad (2)$$

and is obtained when

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

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 minimalization. This method is known as *ordinary kriging* and it is very well described by many authors [17–20].

### 2.3 Morbidity variable mapping

The data of the movement of hospital and outpatient care of the study area provided by the Health System Agency Brazilian Government Health System (SUS-SUS) in the year 2007 have been analyzed. These data were filtered to obtain only the data of hospital admissions according to the 10<sup>th</sup> revision of International Classification of Diseases (ICD-10), for hospitalizations related to some kind of disease that can be caused by direct or indirect contact with a high degree of degradation areas, such as in areas with contaminated soils [21].

The map of the spatial distribution contains all the cases for the year 2007 was made the using the interpolation method of the inverse of the distance with power squared values [17]. This procedure was carried out in order to obtain a surface continue distribution morbidity that could be compared in a space with other maps of soil pollutants

In order to analyze the relationship between the morbidity spatial distribution maps and of soil maps the procedure of linear regression analysis space were used.



### 3 Results and discussion

#### 3.1 Soil mapping

The ordinary kriging technique was used to predict the soil maps (Figure 3a-c). The maps showed similar geographical trends, especially for Cu (a) and Hg (b), with higher concentration in the west area and decreasing presence towards northeast. Meanwhile despite Pb (c) showed similar spatial trend, the intensity is higher in the west area and also the southern area is emerging as an important local pollution.

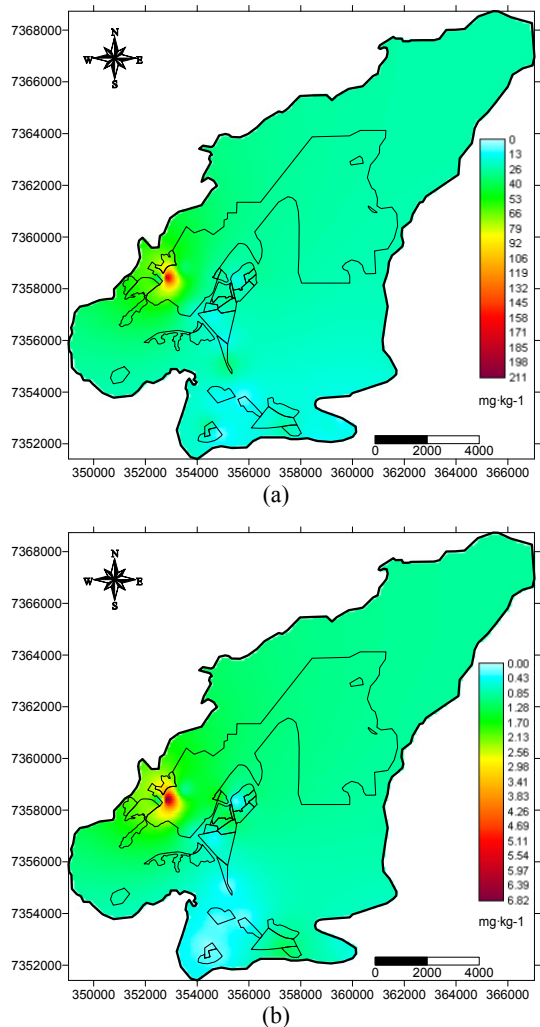


Figure 3: Prediction mapping of soil (a) Cu, (b) Hg and (c) Pb generated by ordinary Kriging.



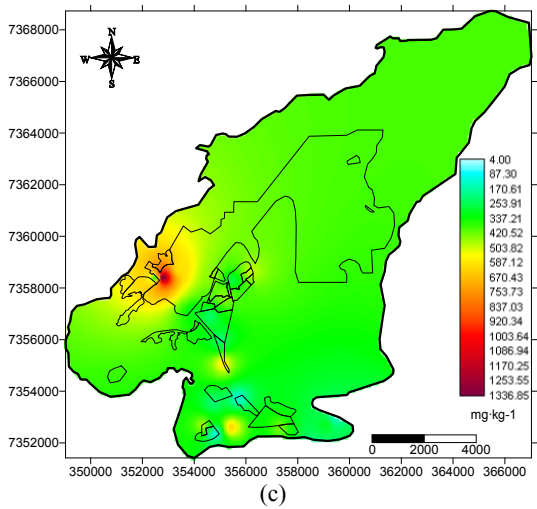


Figure 3: Continued.

3.1.1.1 Morbidity mapping

Figure 4 shows the map of the surface remains of morbidity data for the study area for the studied period using the interpolation method of the inverse of the distance with power squared values.

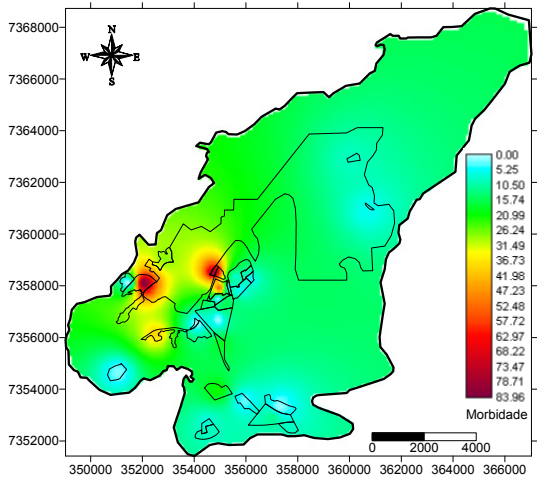


Figure 4: Surface remains of morbidity data for the study area for the studied period.



3.2 Measures of spatial relationship between morbidity and soil pollution

In order to analyze the relationship between the morbidity spatial distribution maps and of soil maps the procedure of linear regression analysis space was used. Figures 5 to 7 show the regression graphs and the statistical parameters of spatial relationship.

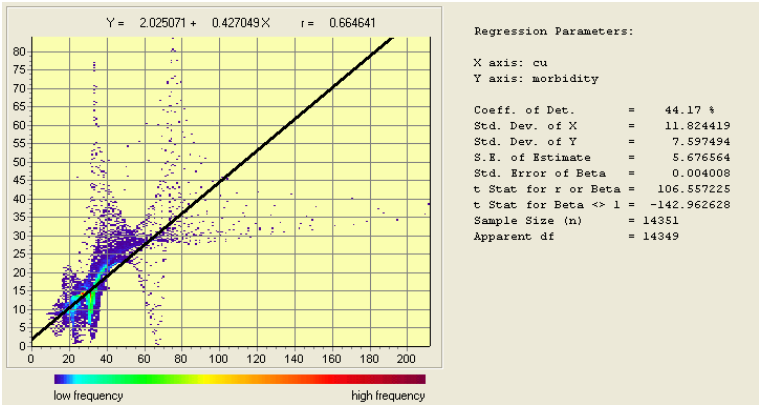


Figure 5: Graph of spatial linear regression between morbidity and Cu.

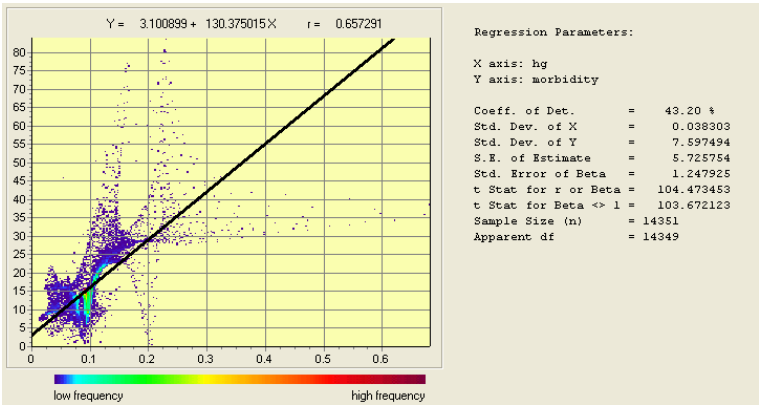


Figure 6: Graph of spatial linear regression between morbidity and Hg.

The regression trend line shows the stronger the linear relationship to the data at soil pollution with morbidity. The correlation coefficient (“r”) next to the equation tells us the same numerically. This is the case with our data since as morbidity increases, then soil pollution also increases. In this study the correlation coefficients ranged around 0.6 indicating a strong relationship between soil pollution and morbidity.





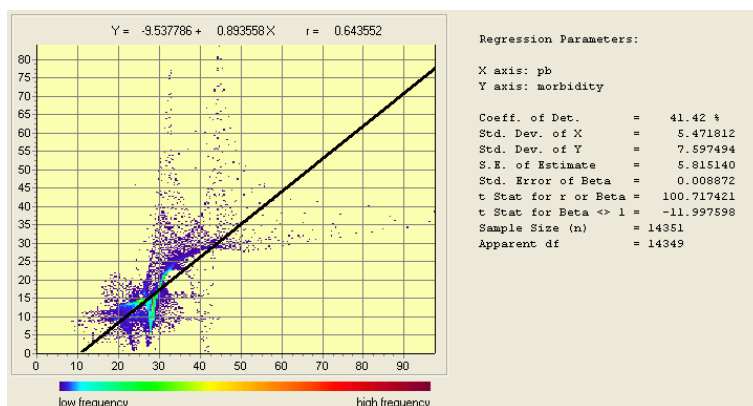


Figure 7: Graph of spatial linear regression between morbidity and Pb.

Now we will examine if the morbidity is a variable that varies function of concentration of the soil pollution. There are several explanatory (or independent) variables which help to predict the variable of concern, the dependent variable. In multiple regression, a linear relationship is assumed between the dependent variable and the independent variables. This study was used the independent variables the soil pollution and dependent variable the morbidity for multiple regression linear. A regression equation is showed down:

$$\text{morbidity} = 0.1326 + 0.4777 \cdot \text{cu} - 0.0024 \cdot \text{hg} + 0.0027 \cdot \text{pb} \quad (4)$$

The coefficients the regression equation indicate the effects of each of the independent variables on the dependent variable. For example, if the Cu value emission for an area increases by 100 units because of the industry, then the morbidity percentage increases by 47,77% (i.e., -100 multiplied 0,4777).

The multiple correlation coefficient between the independent variables (ie, Cu, Hg and Pb) and the dependent variable (morbidity) was  $R = 0.91$  and the extent of variability in the dependent variable explained by all of the independent variables was  $R^2 = 84\%$ , i.e., 84% of the variance in the morbidity is explained by independent variables soil pollutions.

The significance of the coefficient is expressed in the form of a  $t$ -statistic. In this study, the  $t$ -statistic has to exceed the following critical values in order for the independent variable to be significant. The 99% confidence level with  $\infty$  degrees of freedom the value is 2.57, and Cu coefficient has a  $t$ -statistic of 8.71, the Pb  $t$ -statistic is 4.75 and the Hg  $t$ -statistic is 3.66 indicating that the all variables are highly significant (99%). The  $t$ -statistic is the most common test used in estimating the relative success of the model and for adding and deleting independent variables from a regression model.

## 4 Conclusions

The study has showed that areas with high concentrations of pollutants in soil may be associated with disease events, especially those related to the intestinal tract and allergy. In general, the study area has serious problems related to soil use and occupation, as well as socio-environmental issues, since the high number of hospital service causes considerable resources demand to care for the population.

The methodology presented is interesting when it requires a tool that aims to demonstrate the areas of emergency action need in the public sector, and the environmental point of view and from the standpoint of human health.

Overall, understanding the processes of soil and public health issues proves very evident in this work, thus opening many other options to deal with problems and the tools of spatial analysis through Geographic Information Systems can contribute significantly in understanding the problems and are an important tool management and urban planning.

## Acknowledgements

The authors thank the State of São Paulo Research Foundation (FAPESP) and Development Foundation of São Paulo State University (Fundunesp) by the support given.

## References

- [1] Moreira-Nodermann, L.M. Geochemistry and environment. *Brasilienses Geochemistry*, v.1,n.1, 89-107, 1987.
- [2] Lourenço R.W. & Landim, P.M.B. *Risk mapping of public health through geostatistics methods*. Reports in Public Health (CSP), Vol.21, no.1, p.150-160, 2005.
- [3] Alloway, B.J. *Heavy Metals In Soils*. New York: John Wiley, 339p, 2001.
- [4] Franssen, H., Eijnsbergen, A.C. & Stein, A. *Use of spatial prediction techniques and fuzzy classification for mapping soil pollutants*. *Geoderma* 77: 243-262, 1997.
- [5] Tang, H.W., Huel, G., Compagna, D., Hellier, G., Boissinot, C. & Blot, P. *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, 1999.
- [6] Winneke, G., Walkowiak, J., & Lilienthal, H. PCB induced neurodevelopmental toxicity in human infants and its potential mediation by endocrine dysfunction. *Toxicology* 181-182, 161-165, 2002.
- [7] Stein, J., Schettler, T., Wallinga, D. & Vallenti, M. *In harm's way: toxic threats to child development*. *J. Dev. Behav. Pediatr. (Suppl. 1)*, S13-S22, 2002.



- [8] Zhang, C.S. & Selinus, O. *Statistics and GIS in environmental geochemistry—some problems and solutions*. Journal Geochemistry Explorer 64:339–354, 1998.
- [9] Amini, M., Afyuni, M., Fathianpour, N. Khademi, H. & Flühler, H. *Continuous Soil Pollution mapping Using Fuzzy Logic and Spatial Interpolation*. Geoderma 124: 223–233, 2005.
- [10] Lourenço, R.W., Rosa, A.H., Roveda, J.A.F., Martins, A.C.G. & Fraceto, L.F.. Mapping soil pollution by spatial analysis and fuzzy classification. *Environmental Earth Sciences*. DOI: 10.1007/s12665-009-0190-6, 2010.
- [11] Mcgraph, D., Zhang, C.S. & Carton, O. *Geostatistical analyses and hazard assessment on soil lead in Silvermines, area Ireland*. Environ Pollut 127:239–248, 2004.
- [12] Assunção, R.M., Barreto, S.M., Guerra, H.L. & Sakurai, E. Mapas de taxas epidemiológicas: uma abordagem Bayesiana. Cad. Saúde Pública, vol.14, no.4, p.713–723, 1998.
- [13] Van Meirvenne, M. & Goovaerts, P. *Evaluating the probability of exceeding a site-specific soil cadmium contamination threshold*. Geoderma 102, pp. 63–88, 2001.
- [14] Webster, R. & Oliver, M.A. *Geostatistics for environmental scientists*. Wiley, Chichester, pp 89–96, 2001.
- [15] Lin, Y.P., Chang, T.K. & Teng, TP. *Characterization of soil lead by comparing sequential Gaussian simulation, simulated annealing simulation and kriging methods*. Environ Geol 41:189–199, 2001.
- [16] Romic, M. & Romic, D. *Heavy metals distribution in agricultural topsoils in urban area*. Environmental Geology, 43:795–805, 2003.
- [17] Burrough, P.A. & Mcdonell, R. *Principles of Geographical Information Systems*. Oxford, Oxford University Press, 2004.
- [18] Goovaerts P. *Geostatistics For Natural Resources Evaluation*. Oxford University Press, 1997.
- [19] Isaaks, E. & Srivastava, R.M. *An Introduction to Applied Geostatistics*. Oxford University Press, Oxford, 1989.
- [20] Journel, A.G. & Huijbregts, C. *Mining Geostatistics*. Academic Press, 1978.
- [21] Kakkar, P. & Jaffery, F.N. *Biological markers for metal toxicity*. Environmental Toxicology and Pharmacology 19: 335–349, 2005.

