Available online at www.sciencedirect.com
ScienceDirect
Journal of Geochemical Exploration 98 (2008) 15–21
JOURNAL OF GEOCHEMICAL EXPLORATION
www.elsevier.com/locate/jgeoexp

Geostatistical estimation of chemical contamination in stream sediments: The case study of Vale das Gatas mine (northern Portugal)

A.R. Salgueiro [a,*], P. Freire Ávila [b], H. Garcia Pereira [a], J.M. Santos Oliveira [b]

[a] CERENA, Natural Resources and Environment Centre, IST, Av. Rovisco Pais, 1049-001 Lisboa, Portugal
[b] INETI – National Institute of Engineer, Technology and Innovation, Laboratory, 4465-956 S. Mamede de Infesta, Portugal

Received 20 April 2006; accepted 27 October 2007
Available online 14 December 2007

Abstract

Based on an environmental geochemistry case study carried out in the neighbourhood of a W–Sn abandoned mine, the pollution in stream sediments was modelled through a Global Contamination Index. Such an index permits one to summarize the combination of deleterious elements in a single variable, obtained by the projection of samples onto the first axis of a PCASD (Principal Components Analysis of Standardized Data) applied to the entire $n \times p$ matrix containing the available concentrations of $p=16$ elements in the set of $n=220$ collected samples.

In order to provide a sound basis for a coherent planning of the remediation process which will be put in operation in the affected area, it is necessary to balance the costs of reclaiming with the probabilities of exceeding the upper limits accepted for concentrations of environmentally harmful elements in sediments. Given these limits, they are back-transformed in the index values, providing a practical threshold between 'clean' and 'contaminated' samples. On the other hand, the minimum dimension of the cell to be reclaimed is restrained by the selected remediation process to be applied in the affected area. Hence, to meet the constraints of such a remediation process, it is required to estimate the probabilities of exceeding the index threshold in technologically meaningful sub-areas. For this end, the Indicator Block Kriging technique was applied, producing a series of maps where sub-areas to be reclaimed can be spotted for different probability levels. These maps, on which the decision making remediation agency can rely for its cost-benefit analysis, take into account both the spatial structure of 'clean' *vs.* 'contaminated' samples and the constraints of the reclaiming process.

Keywords: Abandoned mines; Environmental remediation; Global Contamination Index; Indicator Block Kriging

1. Introduction and methodological overview

When assessing the impact of abandoned mines on the environment, an important issue is to map the concentration of deleterious elements in stream sediments affected by the spread of old tailings and wastes. Instead of mapping these elements individually, giving rise to unmanageable plots, the global contamination of a given area can be easily visualized by summarizing such elements in a global index, which may be provided by the PCASD (Principal Components Analysis of Standardized Data) of the entire data set, under the approach given in Barradas et al. (1992). In contrast with usual PCA, this approach (described in detail in Vairinho et al. 1990, p. 384), allows the n samples to be projected onto the same axes that are interpreted in terms of the p variables, linking R^p and R^n factorial spaces through transition relationships. The projection of samples onto the particular axis interpreted on the grounds of deleterious elements concentrations give the values of the required Global Contamination Index.

But, in addition to summarize the set of harmful elements into a single contamination index, it is also necessary, for remediation purposes, to locate in space significant areas where the index exceeds a certain threshold, derived from upper limits in the concentrations allowed for dangerous elements in sediments. Such areas, which must conform to technological constraints stemming from the selected remediation process (minimum dimension of a cell to be reclaimed and maximum contiguity between cells), can not be spotted from the sample map, which refers to index values in point supports.

* Corresponding author.
E-mail address: rita.salgueiro@ist.utl.pt (A.R. Salgueiro).

0375-6742/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gexplo.2007.10.005

For decision making in remediation planning, costs of reclaiming must be related to the probability of exceeding the maximum limits in dangerous elements. In order to approach this problem, providing a sound basis for cost-benefit reclaiming studies, an Indicator Block Kriging based methodology was developed, aimed at associating the probabilities of exceeding a given threshold with the correspondent sub-areas to be reclaimed, and producing a series of maps where these sub-areas are located in space. This methodology relies on the following assumptions:

(i) An indicator variable can be found to account for the probability of surpassing, in each sample point, a certain threshold derived from international regulations for upper limits in deleterious elements concentration.
(ii) The estimation of such a variable in technologically significant cells can provide a set of different sub-areas, formed by contiguous cells and associated with a range of estimated probabilities of exceeding the threshold.

The Indicator Kriging technique was used along the lines put forward by Journel (1983), and fully developed in environmental applications by Goovaerts (1997) and Mohammad et al. (1997).

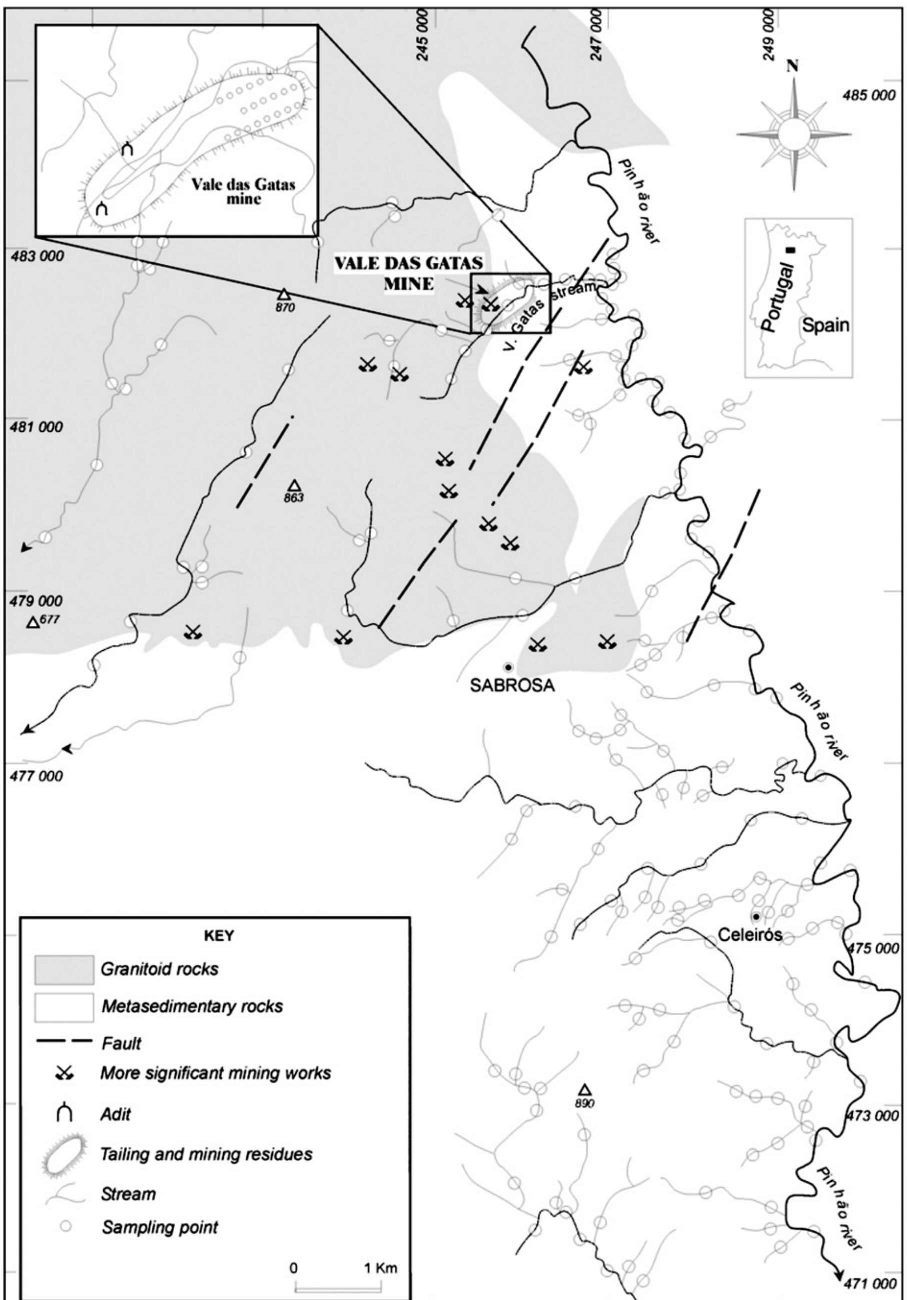


Fig. 1. General geological map of the Vale das Gatas area showing the sampling sites.

Table 1
Basic statistics of the concentrations in samples

Element (mg kg^{-1})	Minimum	Arithmetic mean	Maximum	Variation coefficient
Cu	11.00	149.72	1241.00	1.5640
Pb	11.00	627.96	9708.00	2.5327
Zn	32.00	159.52	1210.00	1.0955
Ag	0.30	16.28	222.20	2.4939
Ni	2.00	18.75	40.00	0.4724
Co	1.00	9.87	39.00	0.4673
Mn	70.00	419.83	1536.00	0.4507
Fe (%)	0.71	2.70	6.44	0.3140
As	2.00	931.92	18312.00	2.5355
Cd	0.20	1.54	15.40	1.6993
Bi	3.00	77.25	1314.00	2.6279
Cr	6.00	28.40	92.00	0.4156
W	2.00	192.27	2095.00	2.3454

The above outlined methodology is to be detailed in parallel with each step of the data modelling process applied in a case study referring to the Vale das Gatas mine, northern Portugal.

2. Case study presentation: geology and sampling

The Vale das Gatas mine is located in an important geotectonic segment of the Iberian Massif denoted Center-Iberian Zone, which is included in the Hercynian Chain (Lotze, 1945; Ribeiro and Pereira, 1982). The major lithological units outcropping in the study area comprise metasedimentary rocks of Pre-ordovician age (phyllites, psamites, greywackes, quartzites, limestones), which were intruded by Hercynian granitoids of diverse chemical and mineralogical composition (see Fig. 1). A number of late fracture systems (with predominant NNE–SSW directions) filled by quartz and aplito-pegmatites intersect all those lithologies (Sousa, 1982).

The mineralisation, which occurs in quartz veins that fill up some of those fractures, is mainly composed of wolframite (hubnerite, but also ferberite), cassiterite and schellite (Gaspar and Bowles, 1985), which were the main components of the exploited ore, during the period of the mine's life. significant amounts of sulphides (arsenopyrite, pyrite, chalcopyrite, pyrrothite, sphar-elite, galena, stannite) and Bi–Pb–Ag sulphosalts as matildite, pavonite and neyite are present as accessory minerals.

The mine ceased its economical activity in 1986, and no environmental remediation action has been taken since then. As expected, chemical oxidation and leaching of sulphides from the old tailings produced a severe acid mine drainage (AMD). This had an important contamination effect on stream sediments located in areas where land use may be jeopardized by such an environmental impact (Freire Ávila et al., 2001), being the results presented in this paper the environmental-geochemical data model to be used as a basis for launching the reclaiming project.

The data collected in this case study was obtained by two sampling campaigns. The first refers to stream sediments (199 samples), and the second, and more specific, is composed of 21 samples aimed at tailings characterization (see Fig. 1). The samples were dried at 40 °C and sieved at 80 mesh. After this, they were crushed, homogenized and sieved, retaining the <200 mesh fraction for chemical analysis, as described in Freire Ávila et al. (2005). The analytical methodology was supported on the multielement chemical analysis of stream sediment and tailing samples. Accredited procedures (based in Quality Systems) of Inductively Coupled Plasma-Atomic Emission Spectrometry (ICP-ES) were adopted for the total analysis of minor and trace elements: Cu, Pb, Zn, Ag, Ni, Co, Mn, Fe, As, Cd, Bi, V, P, Cr, Ti and W.

The geochemical data obtained for the Vale das Gatas area was extensively discussed by Freire Ávila (2003). From this study it was concluded that As, Cu, Pb and Zn appear as the most harmful elements from an environmental point of view.

Table 1 summarizes the statistical characterization of the data set on which the approach used in this paper relies (Freire Ávila et al., 2005). The whole data set of 220 samples was used jointly, given that the relationship structure between elements is the same in tailings and sediment samples, as deduced from Fig. 2, where a separate PCASD was performed for each subset. It is worth noting that the similarity between the two sources concentration structure for most elements is likely to stem from the intense leaching process that occurred in the area, producing resiliently contaminated sediments far away from the mine site.

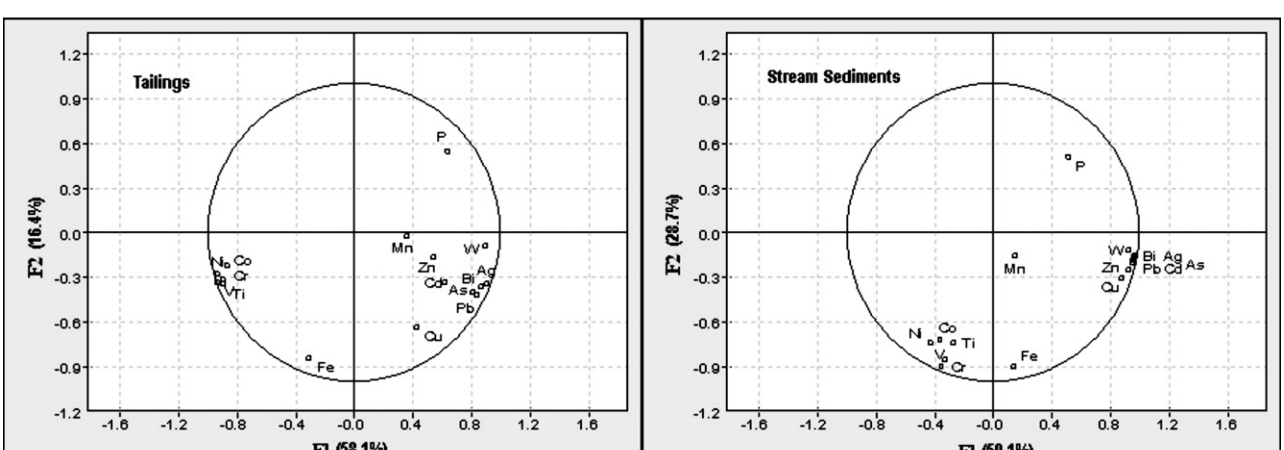


Fig. 2. Separate PCASD for tailings and sediment samples.

3. Global Contamination Index construction and discussion of its estimation method

Stream sediments have been widely considered as good environmental indicators. The interpretation of this type of data by multivariate data analysis is commonly accepted, hence it constitutes a successful tool to quantify anthropogenic effects (Anazawa et al., 2004; de Carlo et al. 2005; Ali et al. 2006; Wu et al. 2007; among others).

Principal Component Analysis of Standardized Data (PCASD) was applied to the analytical results organized in a matrix of 220 samples by 16 variables: Cu, Pb, Zn, Ag, Ni, Co, Mn, Fe, As, Cd, Bi, V, P, Cr, Ti and W. The results obtained revealed the existence of two particular element associations (see Fig. 3). The first axis is interpreted on the grounds of the elements that are characteristic of the mineral paragenesis of Vale das Gatas mine (Pb, Ag, Bi, W, Cd, As, Zn e Cu). Apart from the above identified harmful elements (As, Cu, Pb and Zn), this group of strongly inter-correlated variables contains also W from wolframite, Bi/Ag from sulphosalts, and Cd from spharelite. It is worth noting that this analysis, referring to the whole data set, gives rise to the same results shown in Fig. 2 only for the sub-set of stream sediment samples, suggesting that tailings and mine waste material is intermingled with sediments, as a consequence of the severe leaching process. The second axis is explained by the association Cr, Fe, V, Ni, Co and Ti, suggesting a signature related with the local geology (Freire Ávila, 2003). The plot of samples represented in Fig. 4 shows a dense cloud along axis 2 which is related to variations in the geology, and a set of points, scattered along the positive semi-axis 1, corresponding to contaminated sample sites.

In order to establish a Global Contamination Index – IND – for the study area, the coordinate of each sample in the first axis (Fig. 4) was taken as a synthetic measurement of the effect produced by the association of the above given group of elements.

This coordinate was confined between 0 and 1 by:

$$IND = \frac{(\text{Coord}-\text{Min})}{(\text{Max}-\text{Min})} \text{ where,}$$

IND is the required index and Coord is the projection of the sample onto the first axis, contained between Max and Min.

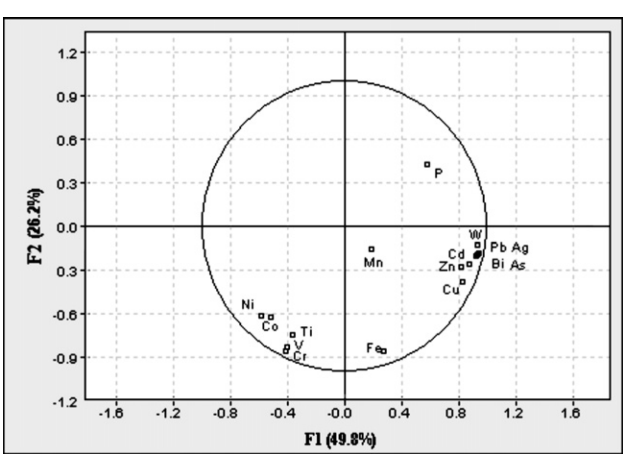


Fig. 3. Graphic representation of variables in the factorial plan defined by axes 1 and 2.

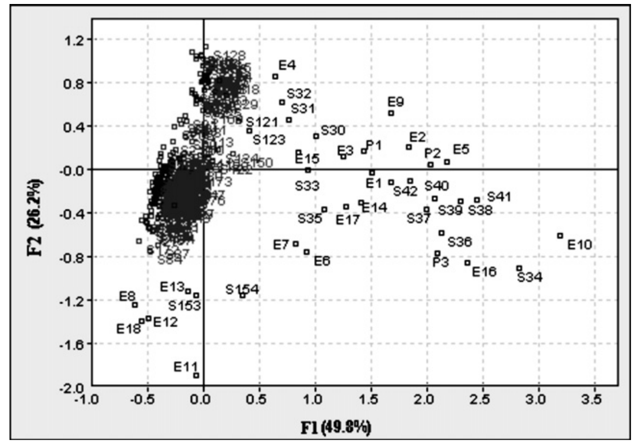


Fig. 4. Graphic representation of samples in the factorial plan defined by axes 1 and 2.

The histogram of the proposed Global Contamination Index is given in Fig. 5, showing a clear two-populations distribution: the low classes correspond to the geological background and the high values to contaminated samples.

The spatial distribution of the variable index in sample sites is given in Fig. 6, representing the level of contamination of each sample by a visual plot.

The combined analysis of Figs. 5 and 6 shows that high values sub-population samples tend to cluster spatially in particular segments of the entire area of the study, which correspond obviously to contaminated sites.

Given the specific statistical and spatial characteristics of the index variable, there is no point to estimate it by any linear kriging method, as proposed, in a rather a different setting, by Reed et al. (2001) for anthropogenic contamination and by Wackernagel et al. (1988) for soil data. In fact, even though the above constructed index may be viewed as a Regionalized Variable summarizing the concentrations of set of meaningful elements (avoiding by this token any type of co-kriging), it does not meet stationary and homogeneous requirements, as evidence provided by the geochemical-environmental study supporting Figs. 5 and 6 fully demonstrates. Also, any type of

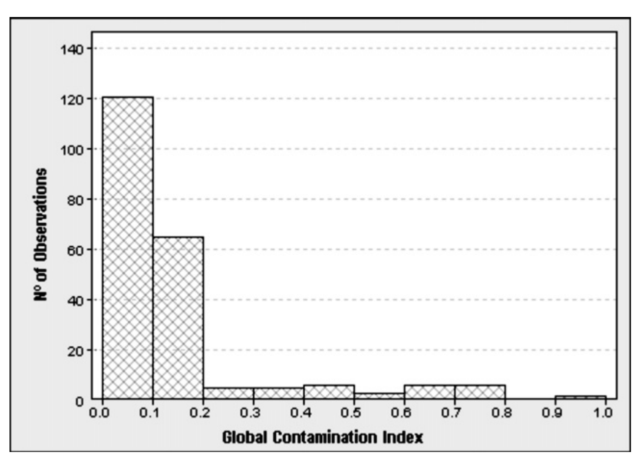


Fig. 5. Global Contamination Index Histogram.

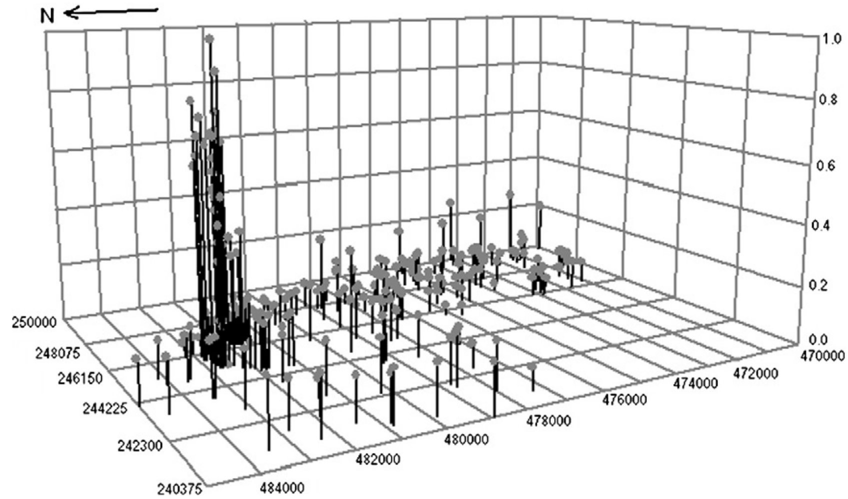


Fig. 6. Perspective diagram allowing the visualisation of the spatial distribution of the level of contamination in sample points.

conditional simulation, as advocated by Bailly et al. (2006) for a similar context, does not pursue the objectives of this case study, since such a method does not identify the location of high values, accounting instead for global variability. Moreover, the universal kriging and external drift models are also inappropriate for approaching the problem, since the patchy spatial distribution of the variable does not exhibit any discernible trend (the values of the index do not follow the hydrological network, as it could be expected, being controlled by a blend of factors that gives rise to the spiky spatial distribution of Fig. 6).

Hence, the above mentioned approaches were discarded, in favor of the Indicator Block Kriging technique, a non-linear estimation method that, not only suits the peculiar conditions of the data setting, but also matches the objective of the reclaiming process. In fact, Indicator Block Kriging allows one to split the index into a binary variable that accounts for the separation in two populations by a given threshold, and also to estimate, within technologically meaningful units, the probability of exceeding such a threshold.

4. Indicator Block Kriging estimation and production of probability maps

The first step of the application of Indicator Kriging to a synthetic variable like the Global Contamination Index is to derive an operational threshold to split the global population into two sub-sets: the 'contaminated' and the 'clean' samples. Since the index is an 'artificial' variable, such a threshold must be derived from the concentrations of the elements that contribute to its construction. The set of harmful elements identified in section 3 that take part in the index are As, Cu, Pb and Zn. For those elements there are international norms that define upper limits in their concentration (mg kg^{-1}), below which the sediment is considered as 'clean' : As – 33, Cu – 150, Pb – 130, Zn – 460, *cf.* WDNR (2003).

Hence, to get a threshold for the variable IND, a minimum risk scenario was adopted: the sediment is considered 'clean' if none of the above given limits is exceeded. The particular form of PCA used here allows to back-transform variable coordinates into sample projections by the transition relationships linking factorial spaces R^p and R^n (Vairinho et al., 1990, p. 384), providing the values of the index corresponding to particular values of the original variables. Therefore, choosing the minimum value of IND corresponding to the prescribed limits, the practical threshold of 0.10 is found for the Global Contamination Index.

At this stage, an Indicator variable can be established for the Global Contamination Index: it takes the value 1 if IND $\geq$ 0.1 and 0, otherwise. The representation of this indicator variable in the sample points is given in Fig. 7, showing that contaminated points are clustered around the mine and fairly scattered in the remaining area.

The omnidirectional variogram of the above defined indicator variable is presented in Fig. 8, depicting the global spatial structure of contaminated *vs.* 'clean' points. Given that

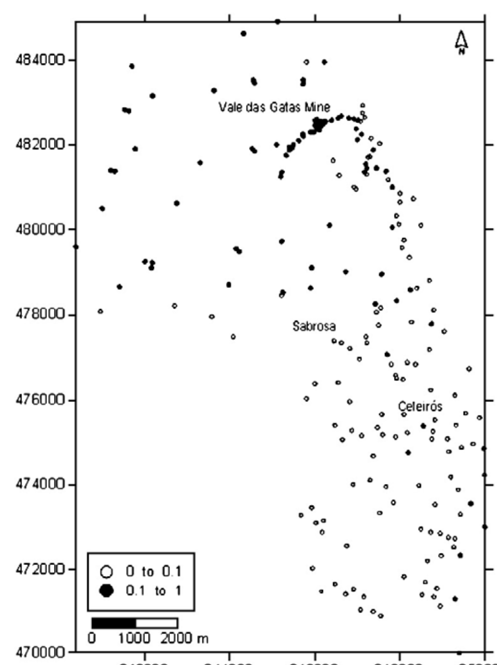
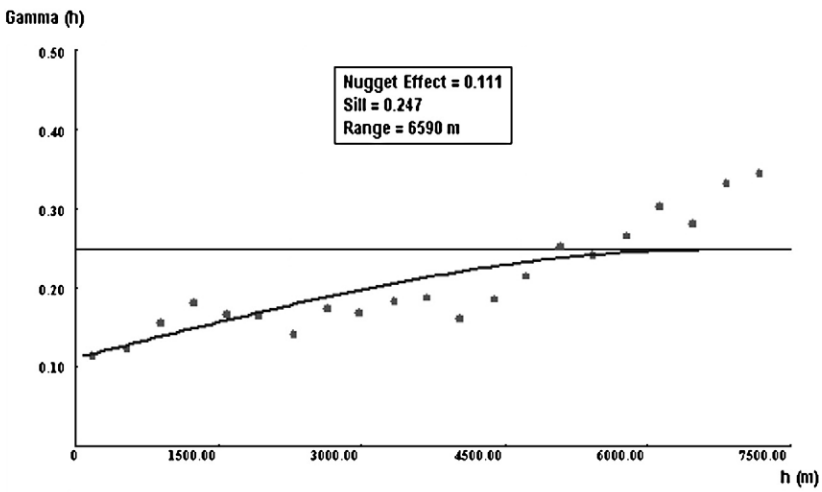


Fig. 7. Representation of the indicator variable in sample points.



Class Dist.	h	Gamma(h)	N (Points)
Class 1	170.732	0.1135	502
Class 2	529.0491	0.1225	494
Class 3	886.9951	0.1551	777
Class 4	1231.7112	0.1809	865
Class 5	1578.2585	0.1668	1025
Class 6	1913.5052	0.1634	878
Class 7	2278.4927	0.1414	824
Class 8	2633.3918	0.1735	948
Class 9	2970.6174	0.1683	918
Class 10	3336.2612	0.1825	893
Class 11	3672.6111	0.1871	922
Class 12	4022.2603	0.1605	885
Class 13	4386.5039	0.1854	1003
Class 14	4716.4121	0.2146	806
Class 15	5079.7939	0.2518	993
Class 16	5420.4507	0.2404	890
Class 17	5772.0518	0.2649	853
Class 18	6116.8833	0.3027	854
Class 19	6467.6113	0.2805	656
Class 20	6820.7163	0.3314	694
Class 21	7172.4238	0.3441	744

Fig. 8. Experimental omnidirectional variogram of the indicator variable and fitted model accounting for the number of pairs of points used for each class.

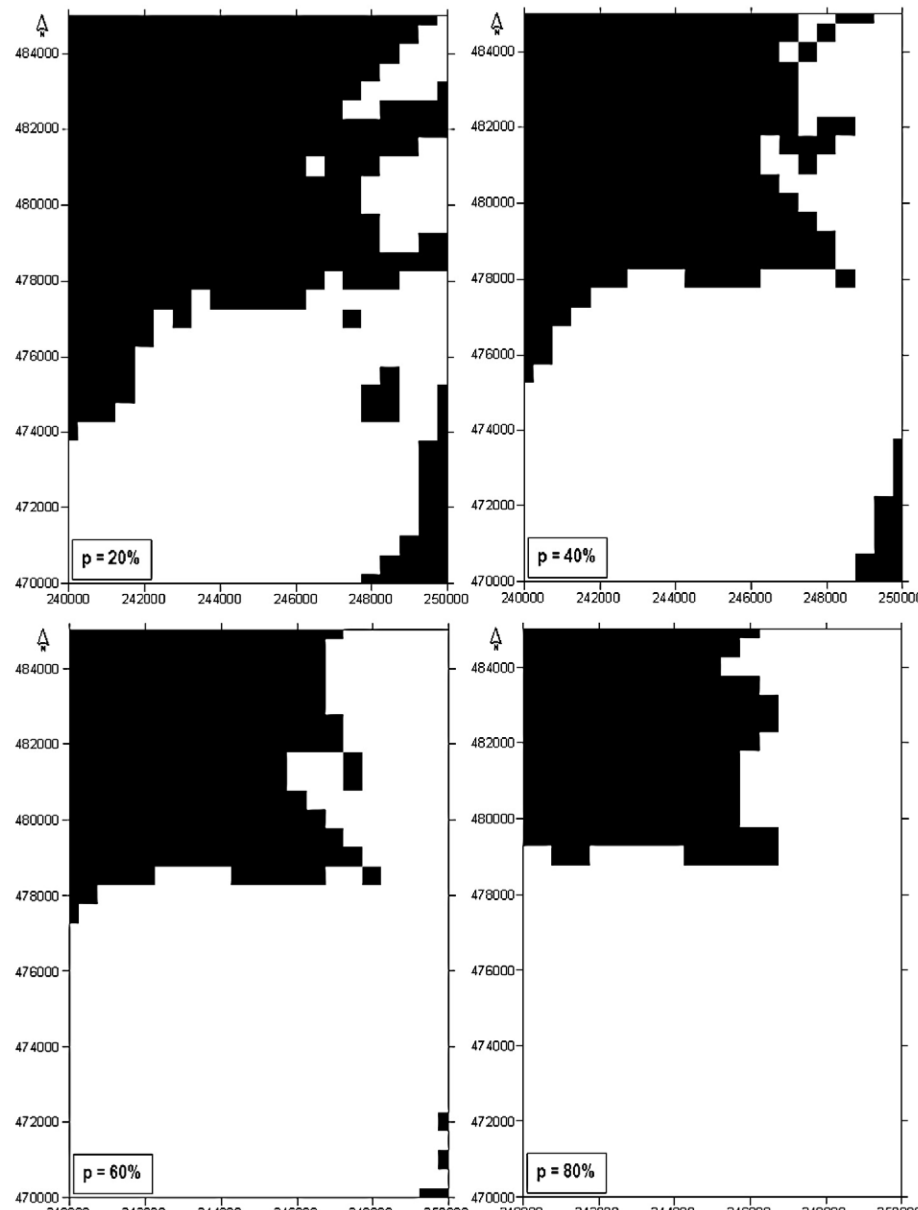


Fig. 9. Estimation map showing in black the contaminated cells for 20%, 40%, 60% and 80% of probability of exceeding the threshold.

the indicator variable is not a continuous one, but reflects only the effect of an 'external' cut-off applied to an 'artificial' variable, it is not surprising that no significant hydrological driving force is controlling its two-phase spatial structure. However, for the sake of acquiring a reasonable confidence in the omnidirectional model (whose parameters are also given in Fig. 8, together with the number of pairs averaged for each lag), it was decided to test the global behavior of the model through a cross-validation process, as described by Webster and Oliver (2001), p. 189–190. This process consists of calculating the diagnostic statistics Mean Error (ME) and the Mean Squared Deviation Ratio (MSDR), which should take the values of 0 and 1, respectively, for the 'perfect' fit. The experimental values obtained for the model of Fig. 8 were 0.0051 and 1.11, respectively. This statistical validation, jointly with the phenomenological interpretation of the spatial structure of contaminated sites, guarantees in practical terms the consistency of the omnidirectional variogram model to be used in the Indicator Kriging technique.

Once obtained a reliable variogram model, the kriging plan was established: the block to be estimated is a square of 500×500 m, corresponding to the minimum size of a cell to be 'cleaned' by the technological reclaiming process; given the cell size and the variogram range, a search radius of 7.000 m was selected and a minimum of 4 samples was established for the number of points to be weighted in the index estimation of a given cell, approximated by a regular mesh of 25 points.

The estimated maps for different levels of probabilities (0.2, 0.4, 0.6, 0.8) of exceeding the above derived threshold are given in Fig. 9.

5. Conclusions

This case study illustrates a methodology combining multivariate data analysis with non-linear geostatistics to calculate a Global Contamination Index and to estimate the probabilities that such an index exceeds a given threshold, within technologically meaningful units for a given remediation method.

The estimated probability maps provide a sound basis to perform cost-benefit analyses focused on the specific reclaiming process to be used in the contaminated area. Apart from the total area to be reclaimed, which obviously shrinks when increasing the level of accepted probability of surpassing the minimum risk scenario put forward for the threshold definition, it is also necessary to take into account the contiguity of the cells corresponding to each sub-area for calculating the corresponding remediation costs. Such an important remediation planning parameter can also be deduced from the maps provided by the Indicator Block Kriging method.

Acknowledgments

The initial version of this paper was noticeably improved by the effort put by referees into its critical revision. The present version was made possible by adopting a new angle, stemming from the re-assessment of the problem under a fresh approach, as suggested by the referees, to whom the authors are deeply thankful.

References

Ali, K., Cheng, Q., Li, W., Chen, Y., 2006. Multi-element association analysis of stream sediment geochemistry data for predicting gold deposits in south-central Yunnan Province, China. Geochemistry: Exploration, Environment, Analysis 6 (4), 341–348.

Anazawa, K., Kaida, Y., Shinomura, Y., Tomiyasu, T., Sakamoto, H., 2004. Heavy-metal distribution in river waters and sediments around a "Firefly Village", Shikoku, Japan: application of multivariate analysis. Analytical Sciences 20, 79–84.

Bailly, J.-S., Monestiez, P., Lagacherie, P., 2006. Modelling spatial variability along drainage networks with geostatistics. Mathematical Geology 38, 515–539.

Barradas, J., Fonseca, E., Silva, E., Pereira, H.G., 1992. . Identification and mapping of pollution indexes using a multivariate statistical methodology. Applied Geochemistry 7, 563–572.

de Carlo, E.H., Tomlinson, M.S., Anthony, S.S., 2005. Trace elements in streambed sediments of small subtropical streams on O'ahu, Hawai'i: results from the USGS NAWQA program. Applied Geochemistry 20 (12), 2157–2188.

Freire Ávila, P., 2003. Dispersão de elementos vestigiais na envolvente da mina abandonada do Vale das Gatas (Sabrosa, Norte de Portugal): implicações de ordem ambiental. PhD thesis. Aveiro University, Portugal.

Freire Ávila, P., Santos Oliveira, J.M., Ferreira da Silva, E., Cardoso Fonseca, E., 2001. Metais na Mina do Vale das Gatas: aspectos sobre a sua dinâmica de dispersão e de transferência em várias interfaces. Proc. VI Congresso de Geoquímica dos Países de Língua Portuguesa, pp. 417–421.

Freire Ávila, P., Santos Oliveira, J.M., Ferreira da Silva, E., Cardoso Fonseca, E., 2005. Geochemical signatures and mechanisms of trace elements dispersion in the area of the Vale das Gatas mine (Northern Portugal). Journal of Geochemical Exploration, 85, 17–29.

Gaspar, O., Bowles, J.W., 1985. Nota preliminar sobre a paragénese dos sulfossais de Bi, Pb e Ag do jazigo de tungsténio de Vale das Gatas. Estudos, Notas e Trabalhos do ServicĐo de Fomento Mineiro 27, 49–54.

Journel, 1983. Non-parametric estimation of spatial distributions. Mathematical Geology 15 (3), 445–468.

Goovaerts, P., 1997. Geostatistics for Natural resources Evaluation. New York, Oxford University Press.

Lotze, F., 1945. Observation respect a la division de los Variscides de la Meseta Ibérica. Publicaciones Extranjeras Sobre Geologia de Espana 5, 149–166.

Mohammadi, J., Van Meirvenne, M., Goovaerts, P., 1997. Mapping cadmium concentration and the risk of exceeding a local sanitation threshold using indicator geostatistics. In: Soares, A., Gómez-Hernandez, J., Froidevaux, R. (Eds.), Proc. geoENV I – Geostatistics for Environmental Applications. Kluwer Academic Publishers, pp. 327–337.

Reed, J., Chappell, A., French, J.R., Olivier, M.A., 2001. Geostatistical analysis of PCB-contaminated sediment in a commercial dock, Swansea, UK. In: Monestiez, P., Allard, D., Froidevaux, R. (Eds.), Proc. geoENV III, Geostatistics for Environmental Applications. Kluwer Academic Publishers, pp. 487–497.

Ribeiro, A., Pereira, E., 1982. Controlos paleogeográficos, petrológicos e estruturais na génese dos jazigos portugueses de estanho e volfrâmio. Geonovas 1 (3), 23–31.

Sousa, B., 1982. Litoestratigrafia e estrutura do Complexo Xisto Grauváquico ante-Ordovícico Grupo do Douro (NE Portugal). PhD thesis. Coimbra University, Portugal.

Vairinho, M., Fonseca, E., Pereira, H.G., 1990. Discrimination of gossans using Principal Components Analysis of standardized data. Journal of Geochemical Exploration, 38, 375–394.

Wackernagel, H., Webster, R., Olivier, M.A., 1988. A geostatistical method for segmenting multivariate sequences of soil data. In: Bock, H.H. (Ed.), Classification and Related Methods of Data Analysis. Elsevier Science Publishers B.V., North Holland.

WDNR, 2003. Consensus-Based Sediment Quality Guidelines: Recommendations for Use & Application (WT-732 2003).Contaminated Sediment Standing Team.

Webster, R., Oliver, M., 2001. Geostatistics for Environmental Scientists. England, Wiley.

Wu, Y., Hou, X., Cheng, X., Yao, S., Xia, W., Wang, S., 2007. Combining geochemical and statistical methods to distinguish anthropogenic source of metals in lacustrine sediment: a case study in Dongjiu Lake, Taihu Lake catchment, China. Environmental Geology 52, 1467–1474.