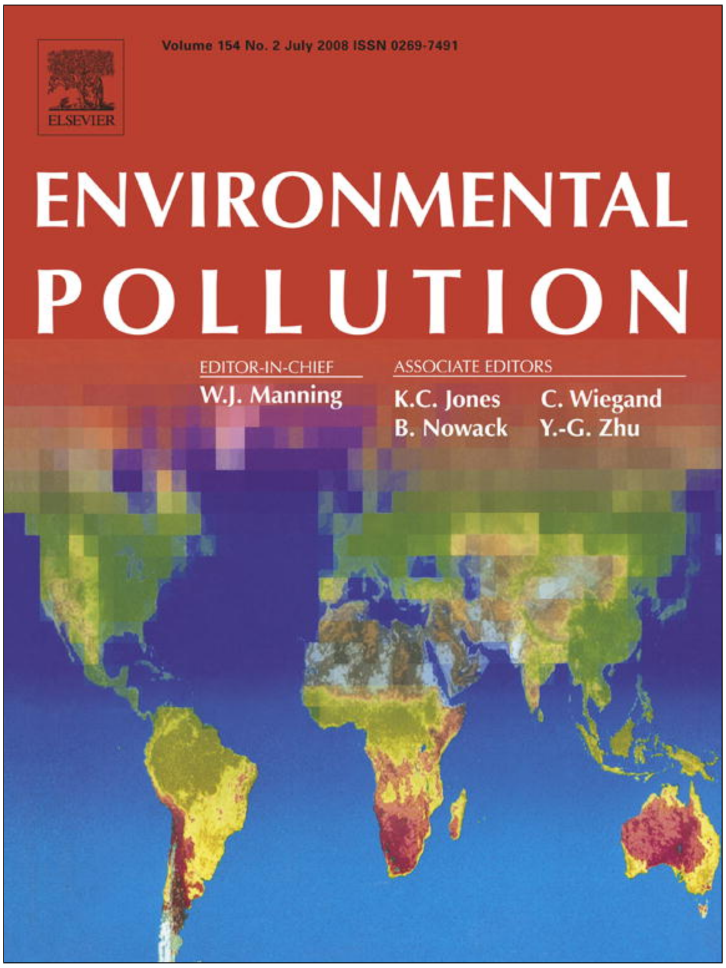
Volume 154 No. 2 July 2008 ISSN 0269-7491
ELSEVIER
ENVIRONMENTAL
POLLUTION
EDITOR-IN-CHIEF
W.J. Manning
ASSOCIATE EDITORS
K.C. Jones
C. Wiegand
B. Nowack
Y.-G. Zhu

Available online at www.sciencedirect.com
ScienceDirect
Environmental Pollution 154 (2008) 184–191
ENVIRONMENTAL POLLUTION
www.elsevier.com/locate/envpol

A multi-stage sampling strategy for the delineation of soil pollution in a contaminated brownfield

S. Verstraete*, M. Van Meirvenne

Ghent University, Department of Soil Management and Soil Care, Coupure 653, 9000 Gent, Belgium

Received 28 June 2007; received in revised form 26 September 2007; accepted 7 October 2007

A multi-stage sampling strategy based on geostatistics provides an efficient procedure to delineate a pollution in a contaminated brownfield.

Abstract

A multi-stage sampling strategy, based on sequential Gaussian simulation, was presented to optimize the step-wise selection of a small numbers of additional samples to delineate soil pollution. This strategy was applied to a Belgian brownfield of 5.2 ha polluted with lead (Pb). Starting from an initial number of 240 samples in stage 1, additional samples were added, 25 per stage, and the reduction of the uncertainty in the Pb delineation was monitored. Twenty stages were used. Already in stage 6 a local optimum was found based on the median conditional coefficient of variation. An independent validation confirmed that this index was to be preferred over the median conditional variance. So for the brownfield considered our procedure indicated that 365 selected samples would have been sufficient, representing a gain of 70.7% in sampling effort compared to current practice which resulted in a sampling effort of 1245 samples.

Keywords: Geostatistics; Lead pollution; Sequential Gaussian simulations

1. Introduction

One of the first steps of a soil sanitation project is the three-dimensional (3-D) delineation of polluted soil. However, mostly the available information about the pollution has to be extracted from a number of punctual observations, requiring an interpolation procedure. It has been shown repeatedly that geostatistical techniques are among the best available methods to perform this step (see Jones et al. (2003) for an overview). However, a recent survey in Belgium by De Boeck et al. (2005) revealed that 84% of the 32 Flemish soil remediation firms rarely, or never, use geostatistics in practical circumstances. They concluded that this is mainly because the regulatory organizations do not require a motivation of the used interpolation methodology. We expect this situation to be similar in many other countries. Therefore there still is a need for convincing examples showing the advantages of integrating spatial statistics into the practice of soil remediation. In this paper we focus on the application of geostatistics to optimize the soil sampling strategy.

A number of geostatistical interpolation techniques (generally termed kriging) are available. Due to the often very strong positive skewness of pollution data, the use of linear interpolation methods (like ordinary kriging) is not always optimal (Kerry and Oliver, 2007; Kishné et al., 2003; Saito and Goovaerts, 2000). Some of the more suitable methods are log-normal kriging, indicator kriging and sequential Gaussian simulation (SGS) (Goovaerts, 1997; Kishné et al., 2003; Saito and Goovaerts, 2000).

The localization of samples and their number involves a sampling strategy. Also, sampling aiming at quantifying and delineating soil pollution is rarely based on a single sampling event. Mostly a multi-stage sampling is used, calling for an optimization procedure. The use of geostatistics to optimize

* Corresponding author. Tel.: +32 92 646 042; fax: +32 92 646 247.
E-mail address: samuel.verstraete@gmail.com (S. Verstraete).

0269-7491/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envpol.2007.10.014

a sampling strategy is not new. Burgess et al. (1981) located new samples where the estimation variance obtained with ordinary kriging was the highest. But this kriging variance represents mainly the spatial configuration of the observations because it is not dependent on the local data values (i.e. it is homoscedastic, (Goovaerts, 1997)). Therefore it is less suitable as a local quality criterion of a prediction. Alternatively, the conditional variance ($\sigma_c^2(\mathbf{x}_0)$) obtained with indicator kriging or a sequential simulation method could be used, since this parameter is heteroscedastic. But since the conditional variance depends on the predicted value ($Z^*(\mathbf{x}_0)$) at location $\mathbf{x}_0$, the selection of additional sampling sites based on this parameter would principally be concentrated in areas with large values. Other approaches (D'Or, 2005; Demougeot-Renard et al., 2004; van Groeningen et al., 1999; Van Tooren and Mosselman, 1997; Juang et al., 2005) are based on the conditional probability of exceeding a critical threshold. Although this strategy results in an accurate localization of the iso-concentration line of the chosen threshold, it fails to identify any remaining pollution not identified during the first sampling steps. Goovaerts and Van Meirvenne (2001) proposed to use the conditional coefficient of variation $\{\mathrm{CV}_c(\mathbf{x}_0) = \sigma_c(\mathbf{x}_0)/Z^*(\mathbf{x}_0)\}$ as a criterion. These authors used indicator kriging but this procedure has some limitations, mainly related to the way the conditional cumulative distribution function (ccdf) is constructed from discrete indicator values. Therefore we applied a sequential simulation procedure (Goovaerts, 1997).

In Belgium, as in most neighboring countries, the identification of soil pollution proceeds along three legally regulated phases. In the first phase, called orientation phase, an exploratory sampling is conducted to investigate the nature of any potential pollution by analyzing a range of organic and inorganic variables. No delineation is required yet. If any remediation limit is exceeded, a second phase is initiated, called descriptive phase. In this phase, the extent of the pollution must be delineated, involving a specific sampling campaign. De Boeck et al. (2005) reported that 72% of the Belgian soil remediation projects involved three or more sampling stages during this phase. In the third phase, called executive phase, the pollution is cleaned. To reduce the risk of erroneous remediation, it is common to re-sample the area during this last phase, sometimes even exhaustively. Especially for large industrial sites, like brownfields, this procedure can result in quite extensive sampling efforts, often without an integrated processing or without an attempt to optimize this effort.

Taylor and Ramsey (2006) used documents by the EPA (United States Environmental Protection Agency), ISO (International Organization for Standardization), CL:AIRE (Contaminated land: Application in Real Environment organisation) and the Environment Agency of the UK to recommend sampling strategies for brownfields. They recognized that site investigation is a phased, multi-stage process. However, no further guidance on the selection of the extra sampling locations is provided. Only in the EPA guideline QA/G-5S (EPA, 2002) a strategy termed adaptive clustered sampling is described. This strategy prescribes additional sampling in areas where observations show elevated contaminant levels. Additional sampling is then done in neighboring grid nodes. The absence of geostatistics as a suitable tool to efficiently select additional sampling locations, confirms our proposition that convincing examples of a multi-phased sampling strategy based on geostatistics are still needed.

The aim of this study is to elaborate a practical method based on geostatistics to design a multi-stage sampling campaign with the objective to optimize the number of samples. Such a procedure is required especially for contaminated brownfields since these are often large sites with complex contaminated volumes which are sometimes sampled exhaustively.

2. Materials and methods

2.1. Variogram

A variogram is a measure for the spatial dissimilarity between data separated by a spatial vector **h**. The experimental variogram is calculated according to Goovaerts (1997).

Armstrong (1998) proposed to compose the 3-D variogram model by modeling the horizontal and vertical experimental variograms separately. This approach has also been used successfully by Van Meirvenne et al. (2003). Therefore, two directional variograms (assuming isotropy in the horizontal plane, else several directional horizontal variograms would be needed) are calculated using only pairs located in the horizontal plane or the vertical direction, respectively. After modeling these directional variograms, the models were combined into one anisotropic 3-D model. The variogram modeling was conducted by an automatic variogram fitting based on a least square error approach with weights proportional with the number of couples and inversely proportional with the magnitude of **h**, a procedure proposed by Cressie (1985).

2.2. Simulation

Sequential simulations are a generalization of the concept of Monte-Carlo simulations, with the integration of the spatial structure of the observations (via the variogram) (Dowd and Pardo-Iguzquiza, 1999). These simulations generate a number of equiprobable realizations, each of them reproducing the spatial variability of the data and more or less matching the original sample histogram and variogram (detailed descriptions of this procedure can be found in Goovaerts (1997) and Fagroud and Van Meirvenne (2002). At each location simulated values can be summarized into an estimated value and differences between the realizations provide a measure of the spatial uncertainty about the interpolation.

2.3. 3-D

Most delineation studies focus on 2-D (Demougeot-Renard et al., 2004; Saito and Goovaerts, 2003). Others, like the study by Garcia and Froidevaux (1997) have approximated the 3-D variability by a layered approach. But Van Meirvenne et al. (2003) has found this layered inferior compared to a full 3-D interpolation, especially to maintain the consistency between predicted values at different soil depths. Therefore we applied a full 3-D interpolation in the SGS.

2.4. Multi-stage sampling strategy

The multi-stage sampling strategy which we propose consists of six consecutive steps:

(a) The available data were normal score transformed (Deutsch and Journel, 1998). This resulted in values following a normal distribution with a mean of zero and a standard deviation of one.
(b) The 3-D variogram of these normal scores was calculated.

(c) SGS was used to generate L realizations (typically L is in the order of 100–1000) of the normal scores in 3-D.
(d) These realizations were back-transformed to original units, using the inverse normal score transformation, and the realizations at every location were used to build the ccdf $[F(\mathbf{x}_0;\ z|(n))]^*$, where $|(n)$ represents the conditioning to the local information. Next, $Z^*(\mathbf{x}_0)$ (''E-type estimate'') and $\sigma_c^2(\mathbf{x}_0)$ were derived from this function (*20*) and $\mathrm{CV_c}(\mathbf{x}_0)$ was calculated. Finally, the conditional probability of exceeding a critical threshold z_c was obtained by $\mathrm{Prob}\{Z(\mathbf{x}_0) > z_c|(n)\} = 1 - F(\mathbf{x}_0;\ z_c|(n))$.
(e) Using $\mathrm{CV}_c(\mathbf{x}_0)$ and $\mathrm{Prob}\{Z(\mathbf{x}_0) > z_c|(n)\}$ as parameters of the selection criteria (see further), a number of locations were selected as additional sampling locations (in 3-D). This number should best be kept small to run the optimization procedure at a fine resolution.
(f) Once sampled and analyzed, these new data were added to the existing data set and the procedure was restarted at step (a), until the monitoring indices showed a local optimum (see further).

2.5. Study area and data sets

Our study area was a brownfield located in Belgium with an area of 5.2 ha. This site was investigated by a remediation firm in 2002 and cleaned in 2003. The site had been used for industrial activities during several decades and its soil was found to be contaminated a.o. by lead. Pb concentrations ranging between the analytical detection limit (1.19 mg kg^{-1}) and 11,500 mg kg^{-1}; in Flanders, the remediation threshold for lead is 500 mg kg^{-1} (VLAREBO, 1996). The pollution was caused by the erratically dumping of heavy metal containing residues of a cement producing industrial process.

During the remediation procedure, two sampling campaigns were carried out:

(1) A first campaign was conducted in the frame of the descriptive phase, resulting in 240 samples located between the soil surface and a depth of 20 m (from 17.5 m above sea level (a.s.l.) to 3 m below sea level); 60% of the samples were within the upper 3 m. The locations of these samples were selected by a remediation expert (Fig. 1a). On the basis of these data it was decided that this brownfield needed to be remediated and a first delineation was drawn (without any geostatistical processing).
(2) A second campaign took place in the frame of the executive phase. The remediation firm decided to sample exhaustively, i.e. at the center of every remediation unit (being a block of 10 by 10 by 1 m^3) down to 5 m below the soil surface (Fig. 1b). This resulted in 1005 soil samples. No use was made of the data of the first sampling campaign to optimize the localization of these samples.

During the first sampling campaign, soil samples were collected from a single borehole, resulting in point measurements at each sampling depth. For the second sampling campaign, a composite sample from four hollow-stem auger boreholes was collected at each depth and location.

The samples from both sampling campaigns were analyzed for the total concentration of lead according to the official national method. OVAM, the Public Waste Agency of Flanders (Belgium) requires that soil samples should be analyzed using a standard package of analytical determinations on the fine-earth fraction (<2 mm) of the soil samples. This standard package includes selected physico-chemical properties, a number of organic compounds and heavy metals, including Pb. For the total metal analysis, 0.5 g of air dry soil was subjected to microwave destruction with aqua regia destruction (6 ml 37% HCl, 2 ml 65% HNO_3 and 2 ml 40% HF) (OVAM, 1992; method CMA/2/II/A.3), a method in accordance with the ISO 11466 procedure (International Organization for Standardization, 1995). Analysis were performed with ICP-AES (inductively coupled plasma – atomic emission spectroscopy) (OVAM, 1992, method CMA/2/I/B.1). Readings below the detection limit were set to 1.19 mg Pb kg^{-1}. Measurements were accepted when individual repetitions were within a 10% boundary.

2.6. Applying the multi-stage sampling strategy to our test case

To test the multi-stage sampling strategy described above, three different data sets were used. The data set of the first sampling campaign was used to initiate the procedure (step 1). The data from the second sampling campaign were split randomly: one part was used to improve the delineation of the pollution (805 samples, referred to as ''extra data set''), the rest was kept for validation purposes (200 samples, referred to as ''validation data set'').

Starting from the 240 data of the first sampling campaign we added in each stage 25 data selected from the extra data set. Although this number was selected arbitrarily, we considered 25 samples to be a good balance between two opposing considerations: a sufficiently small resolution of the optimization procedure and a sufficiently large number to motivate a revisit of the site in practice. Each time SGS was run again, each time 500 realizations were generated and used to delineate the pollution in 3-D (steps 2–4).

The following combination of *selection criteria* was used (step 5). First, locations were selected satisfying the condition: $0.2 < \mathrm{Prob}\{Z(\mathbf{x}_0) > z_c|(n)\} < 0.8$ with $z_c = 500$ mg Pb kg^{-1}. Following the suggestion by Garcia and Froidevaux (1997), we considered that a location with $\mathrm{Prob}\{Z(\mathbf{x}_0) > z_c|(n)\} \geq 0.8$, or $(1 - \mathrm{Prob}\{Z(\mathbf{x}_0) > z_c|(n)\}) \geq 0.8$ could be classified with sufficient confidence as contaminated, or safe, respectively. Else it remained undefined, and thus additional information was required. The selected grid nodes were sorted using $\mathrm{CV}_c(\mathbf{x}_0)$ and the 25 locations with the largest value were taken as target locations for the next sampling stage. In our test case, these 25 added data were selected from the extra data set as the closest available observations to the selected localizations (the average deviation was 6.9 m).

The 3-D grid resolution used for interpolation was 2 by 2 by 0.5 m^3 and the grid dimensions were 160 (X) by 480 (Y) by 5 m^3 (from 17 down to 12 m a.s.l.), since during the descriptive phase the pollution was found to be concentrated in the top 5 m. This 3-D grid was somewhat reduced by the irregular boundary, yielding 117,814 locations to be simulated.

To evaluate the multi-stage sampling strategy, two *monitoring indices* were used (step 6):

(1) The median conditional variance $\mathrm{me}(\sigma_c^2)$ of the 117,814 conditional variances $\sigma_c^2(\mathbf{x}_0)$, being an absolute measure of the overall uncertainty.
(2) The average conditional coefficient of variation $(\overline{CV_c})$ of the 117,814 conditional coefficients of variations $\mathrm{CV}_c(\mathbf{x}_0)$, reflecting the relative overall uncertainty.

Both indices should be as small as possible. By plotting their value versus the number of additional samples, local optima (i.e. local minimum values) could be identified.

To validate the procedure we used, and to check which monitoring index performed best, we calculated the root mean square estimation error (RMSEE):

$$\mathrm{RMSEE} = \sqrt{\frac{1}{n_v}\sum_{i=1}^{n_v}\left[Z^*(\mathbf{x}_i) - Z(\mathbf{x}_i)\right]^2}$$

where n_v is the number of validation points (i.e. 200), $Z(\mathbf{x}_i)$ are the true Pb concentrations of the validation data set and $Z^*(\mathbf{x}_i)$ is the estimated Pb concentration at location $\mathbf{x}_i$. Obviously, this error should be as small as possible.

3. Results and discussion

3.1. Data distributions

The cumulative frequency distributions of the lead data of both sampling campaigns are given in Fig. 2 and some descriptive statistics can be found in Table 1. A 3-D cell declustering analysis (Goovaerts, 1997) did not change the estimated mean value significantly, so no declustering was performed. Generally, both distributions were very similar. But the second sampling campaign contained a few very large values resulting in somewhat larger values for all statistics. This result indicated that the first sampling campaign missed some of the

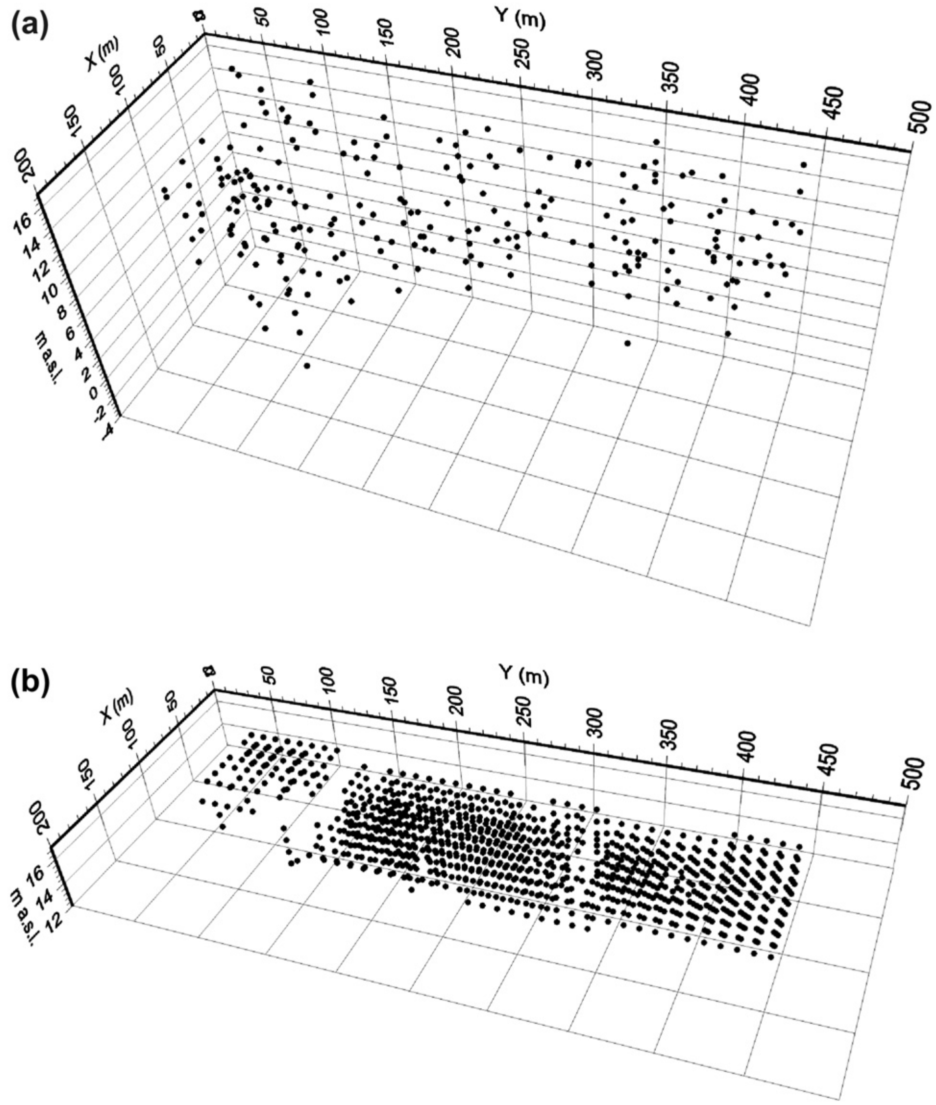

Fig. 1. The sampling locations of the first (a) and second (b) sampling campaign.

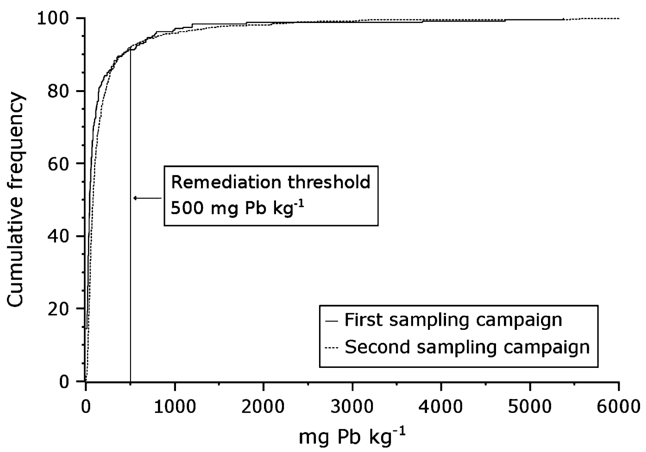

Fig. 2. Cumulative frequency distribution of the first and the second sampling campaign.

locations with high lead concentrations which were included in the exhaustive second sampling campaign. So the descriptive phase resulted in an underestimation of the magnitude of pollution of this site. In both distributions a large proportion of the data values was smaller than the threshold of 500 mg Pb kg^{-1} (91.3% of the first and 91.8% of the second campaign). So most of the soil did not need cleaning.

Table 1
Some descriptive statistics of both sampling campaigns

	First sampling campaign	Second sampling campaign
Number of data	240	1005
Mean (mg Pb kg^{-1})	187.9	224.1
Standard deviation (mg Pb kg^{-1})	558.6	609.5
Median (mg Pb kg^{-1})	43.8	80.0
Maximum (mg Pb kg^{-1})	5378	11,499
Coefficient of skewness	6.9	9.9

3.2. First stage of the multi-stage sampling strategy

The first step of the implemented multi-stage sampling strategy was based on the 240 data of the first sampling campaign. These data were transformed into normal scores and the transformation table was kept for back transformation. Both the vertical and horizontal (horizontally omni directional) experimental variograms of the normal score values were calculated and a theoretical spherical model was fit to them. The vertical variogram was modeled with a nugget variance of 0.44, a sill of 1.07 and a range of 5.8 m. The horizontal variogram was fit using a double spherical model with a nugget variance of 0.50, a first structural variance of 0.25 with a short range of 30 m, and a second structural variance of 0.19 and a range of 120 m. By combining these two variogram models a 3-D anisotropic model was obtained (Fig. 3) providing variogram values for any distance and angle β, where β = 0° (or 180°) for the horizontal direction and β = 90° for the vertical direction.

SGS was used to create 500 realizations of the 3-D grid using only these 240 data. The search neighborhood was defined as an ellipsoid with a horizontal radius of 40 m and a vertical axis of 4 m. Every node was simulated using at least 4 and maximum 20 Pb measurements and a maximum of 16 previously simulated grid nodes. The simulated values of each of the 500 realizations were back-transformed into the original units using an inverse normal score transformation. Fig. 4a shows the estimated values obtained from these realizations. Only a small volume of soil, distributed over a few isolated small volumes, was found to exceed the threshold of 500 mg Pb kg^{-1}. The two monitoring indices and the RMSEE were calculated and used as a starting value (sampling stage 1) of the multi-stage sampling strategy.

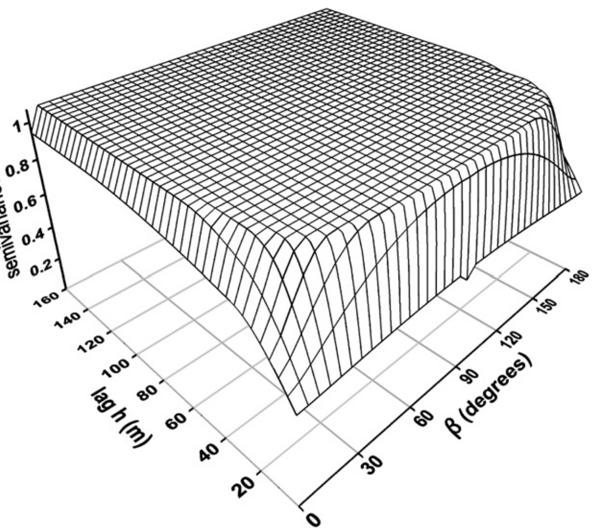

Fig. 3. 3-D variogram model of the normal scores of the data of the first sampling campaign (β represents the dip angle: 0° and 180° are the horizontal orientations, 90° is the vertical).

3.3. Subsequent stages of the multi-stage sampling strategy

To initiate the second stage, all locations of the estimated grid of the first stage were selected based on the condition $0.2 < \mathrm{Prob}\{Z(\mathbf{x}_0) > z_c|(n)\} < 0.8$. Of these, the 25 locations with the largest $\mathrm{CV}_c(\mathbf{x}_0)$ were retained. For each of these locations the nearest observation of the extra data set was taken and added to the data set of the first sampling campaign, yielding 265 data. The variogram was recalculated and remodeled, SGS was run again and the two monitoring indices and the RMSEE were recalculated (yielding their values for sampling stage 2). This sequence was repeated 19 times, resulting in 715 data in the 20th stage.

3.4. Monitoring indices

The evolution of $\mathrm{me}(\sigma_c^2)$ and $\overline{CV_c}$ over the different stages is shown in Figs. 5 and 6, respectively.

Up to the 3rd stage both indices did not improve: the $\mathrm{me}(\sigma_c^2)$ even increased, while the $\overline{CV_c}$ remained almost constant. This was due to the underestimation of the magnitude of the Pb pollution during the first, descriptive sampling phase, as observed earlier. Yet, this first sampling stage identified some of the polluted volumes, guiding the additional samples to locations with a large uncertainty. Accidentally some of these locations contained large Pb concentrations.

From stage 4 onwards, both indices decreased, but with a different pattern. The $\mathrm{me}(\sigma_c^2)$ decreased continuously towards stage 11 where a first local minimum was obtained. Then this index increased slightly, up to stage 13, after which it decreased again continuously until stage 20. The $\overline{CV_c}$ dropped abruptly in stage 5, resulted in a first local minimum. For the other stages its behavior was similar to the $\mathrm{me}(\sigma_c^2)$. Based on the $\mathrm{me}(\sigma_c^2)$ one would decide to stop the multi-stage sampling in stage 12, i.e. after having observed the first local minimum at stage 11. But when the $\overline{CV_c}$ would be used, sampling would be stopped in stage 6, i.e. after the local minimum of stage 5. The latter decision would have resulted in 125 additional samples to be taken in 5 consecutive stages of 25 samples each, yielding 365 samples in total.

Fig. 4b,c show the estimated grid obtained in the 5th and 11th stage, respectively. It will be clear that in stage 1 (Fig. 4a) the volume of soil with Pb concentrations exceeding 500 mg Pb kg^{-1} was underestimated and very patchy. In stage 5 (Fig. 4b, based on 340 samples) some new patches with high Pb concentrations were identified and some continuity formed between the patches identified in stage 1. Stage 11 (Fig. 4c, with 490 samples) confirmed the patterns identified in stage 5. This suggests that the decision based on the $\overline{CV_c}$ to stop the sampling at stage 6 was justified.

3.5. Validation

Using the 200 data of the validation data set, the RMSEE was calculated for each stage (Fig. 7). As could be expected,

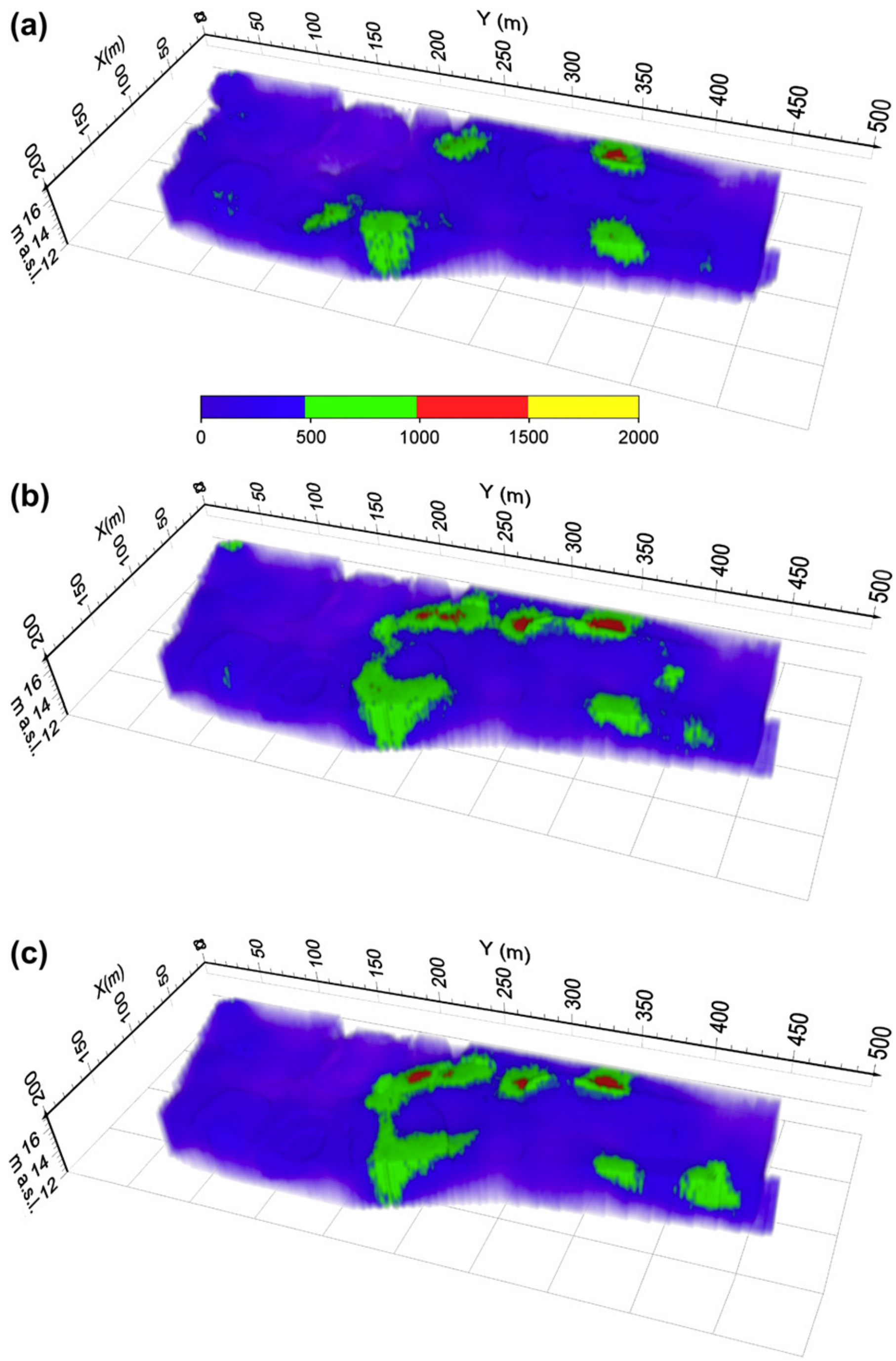

Fig. 4. Estimated Pb content (mg kg^{-1}) over the 3-D grid of the first stage (a), the 5th stage (b) and the 11th stage (c) (the transparency of the color of very low values was increased to improve the visibility of subsurface colors).

the RMSEE was largest for stage 1. Adding extra data reduced the RMSEE strongly up to stage 5 where a first (small) minimum was obtained. Then the RMSEE decreases further until stage 11 where it reached a second minimum. In the subsequent stages this index fluctuated around this minimum, at stage 15 it was smaller, at stage 18 it was larger.

Fig. 8 shows the evolution of the difference in RMSEE (Δ RMSEE) calculated as:

$$\Delta\,\mathrm{RMSEE}(i) = \mathrm{RMSEE}(i) - (i - 1)$$

with i the sampling stage. It can be observed that the biggest difference in RMSEE occurred in the first stages, with a first positive Δ RMSEE between stages 5 and 6.

These validation results confirm that the conclusions based on $\overline{CV_c}$ to stop at stage 6 was justified, and that $\overline{CV_c}$ is a better monitoring index than me(σ^2).

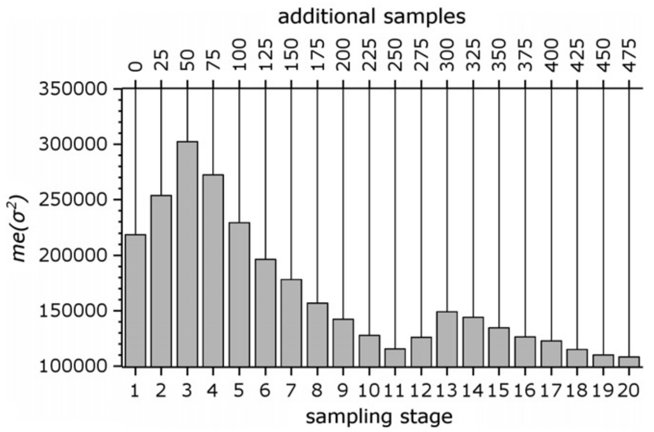

Fig. 5. The median conditional variance in respect to the sampling stages and the number of extra samples.

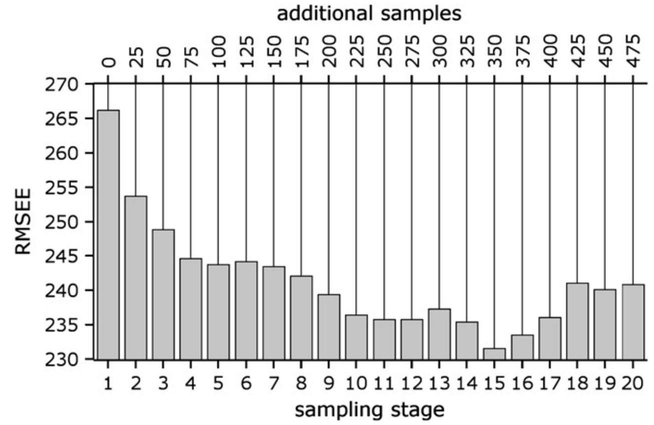

Fig. 7. The RMSEE in respect to the sampling stages and the number of extra samples.

4. Conclusions

If the procedure we presented would have been applied to the brownfield of our test case, 365 soil samples (240 samples taken in the first stage followed by 5 stages of 25 samples each) would have been sufficient to delineate the Pb pollution with an accuracy which could be improved only marginally by taking many more samples. According to current practice, 1245 samples were taken, so we realized a reduction by 70.7%. This illustrates the gain which could be realized by including a geostatistically based optimization procedure in to the sampling strategy.

To implement our multi-stage sampling strategy in a new situation, we recommend the following procedure:

(1) Start the multi-stage sampling strategy with an initial sampling. Depending on the circumstances one of the following strategies could be applied. In the case where historical information is available, the first sampling campaign should be best based on a judgmental sampling strategy as described by Taylor and Ramsey (2006). When no historical information is available, a preliminary survey with sensors could be followed by a judgmental sampling strategy. In those cases where a sensor approach is not possible a non-judgemental sampling strategy should be chosen (Taylor and Ramsey, 2006). This initial sampling should contain no less than 100 (2-D study) to 200 (3-D study) samples, to stabilize sufficiently the variogram.

(2) Subsequently, the results of this initial sampling can be used to design the multi-stage sampling strategy, i.e. the number of additional samples in every stage and the number of stages. The number of additional samples in every stage can vary, but typically should be kept as limited (we used 25). In that way, a fine-grained monitoring of the number of samples and stages is assured.

Even when the optimal sampling scheme can not be achieved due to practical or budget restrictions, the above mentioned strategy will allow attaining as much certainty as is possible with a limited sampling effort.

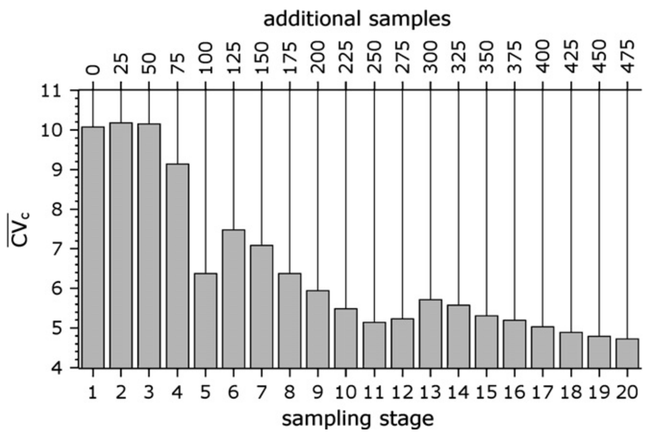

Fig. 6. The average coefficient of variation ($\overline{CV_c}$) in respect to the sampling stages and the number of extra samples.

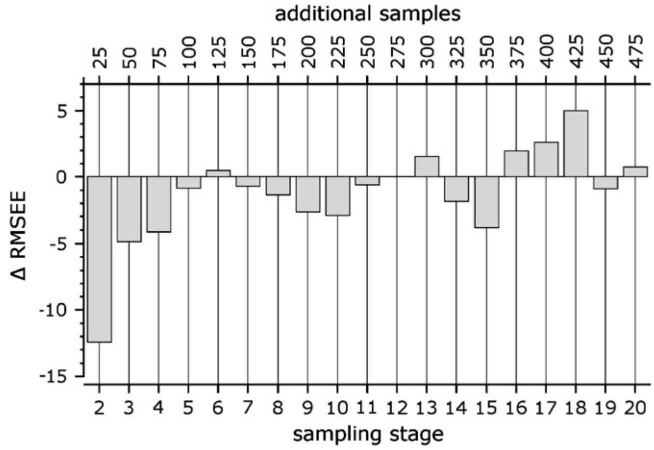

Fig. 8. Δ RMSEE in respect to the sampling stages and the number of extra samples.

Acknowledgements

The authors thank Geosan N.V. and the Jan De Nul Group for providing the data for this study and "*Het Bijzonder Onderzoeksfonds*" (BOF-fund) of the Ghent University for the financial support of this investigation.

References

Armstrong, M., 1998. Basic Linear Geostatistics. Springer, Berlin Heidelberg New York.

Besluit van de Vlaamse Regering dd. 5 maart 1996 houdende Vlaams Reglement betreffende de bodemsanering (VLAREBO, In Dutch).

Burgess, T.M., Webster, R., McBratney, A.B., 1981. Optimal interpolation and isarithmic mapping of soil properties. IV Sampling strategy. Journal of Soil Science 32, 643–659.

Cressie, N., 1985. Fitting variogram models by weighted least squares. Mathematical Geology 17 (5), 563–586.

De Boeck, K., De Lat, B., Van Camp, N., Van Ransbeek, N., 2005. Afbakeningsstrategieën voor bodemsanering in Vlaanderen: praktijk. wetgeving en theorie (In Dutch).

Demougeot-Renard, H., de Fouquet, C., Renard, P., 2004. Forecasting the number of soil samples required to reduce remediation cost uncertainty. Journal of Environmental Quality 33, 1694–1702.

Deutsch, C., Journel, A., 1998. GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press.

D'Or, D., 2005. Towards a real-time multi-phase sampling strategy optimization. In: Renard, P., et al. (Eds.), geoENV V – Geostatistics for Environmental Applications. Springer-Verlag, Berlin, Heidelberg, pp. 355–366.

Dowd, P., Pardo-Iguzquiza, E., 1999. The incorporation of model uncertainty in risk assessment using geostatistical simulation. In: Atkinson, P.M., Riding, A.E. & Tate, N.J. (Eds.), Proceedings of the Geostats-UK '99 Conference. University of Southampton, pp. 29–43.

EPA, 2002. Guidance on Choosing a Sampling Design for Environmental Data Collection for use in Developing a Quality Assurance Project Plan, EPA QA/G-5S.

Fagroud, M., Van Meirvenne, M., 2002. Accounting for soil spatial autocorrelation in the design of experimental trials. Soil Science Society of America Journal 66, 1134–1142.

Garcia, M., Froidevaux, R., 1997. Application of geostatistics to 3-D modeling of contaminated sites: a case study. In: Soares, A., et al. (Eds.), geoENV I – Geostatistics for Environmental Applications. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 309–325.

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

Goovaerts, P., Van Meirvenne, M., 2001. Delineation of hazardous areas and additional sampling strategy in presence of a location-specific threshold. In: Soares, A., et al. (Eds.), geoENV III – Geostatistics for Environmental Applications. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 125–136.

International Organization for Standardization, 1995. ISO 11466. Soil Quality: Extraction of Trace Elements Soluble in Aqua Regia. ISO, Geneva.

Jones, N.L., Davis, R.J., Sabbah, W., 2003. A comparison of three-dimensional interpolation techniques for plume characterization. Ground Water 41, 411–419.

Juang, K.-W., Lee, D.-Y., Teng, Y.-L., 2005. Adaptive sampling based on the cumulative distribution function of order statistics to delineate heavy-metal contaminated soils using kriging. Environmental Pollution 138, 268–277.

Kerry, R., Oliver, M.A., 2007. Determining the effect of skewed data on the variogram I. Underlying asymmetry. Computers and Geosciences 33, 1212–1232.

Kishné, A.S., Bringmark, E., Bringmark, L., Alriksson, A., 2003. Comparison of ordinary and lognormal kriging on skewed data of total cadmium in forest soils of Sweden. Environmental Monitoring and Assessment 86, 243–263.

OVAM, 1992. Compendium voor Monsterneming en Analyse ter uitvoering van het Afvalstoffendecreet en het bodemsaneringsdecreet, Openbare Afvalstoffenmaatschappij voor het Vlaamse Gewest, Mechelen (in Dutch).

Saito, H., Goovaerts, P., 2000. Geostatistical interpolation of positively skewed and censored data in a dioxin-contaminated site. Environmental Science & Technology 34, 4228–4235.

Saito, H., Goovaerts, P., 2003. Selective Remediation of contaminated sites using a two-level multiphase strategy and geostatistics. Environmental Science & Technology 37, 1912–1918.

Taylor, P.D., Ramsey, M.H., 2006. Sampling strategies for contaminated brownfield sites. Soil Use and Management 21, 440–449.

van Groeningen, J.W., Siderius, W., Stein, A., 1999. Constrained optimisation of soil sampling for minimisation of the kriging variance. Geoderma 87, 239–259.

Van Meirvenne, M., Maes, K., Hofman, G., 2003. Three-dimensional variability of soil nitrate-nitrogen in an agricultural field. Geoderma 102, 75–100.

Van Tooren, C.F., Mosselman, M., 1997. A framework for optimization of soil sampling strategy and soil remediation scenario decisions using moving window kriging. In: Soares, A., et al. (Eds.), geoENV I – Geostatistics for Environmental Applications. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 259–270.