What about the practical implementation of geostatistics for contaminated sites and soils?

Hélène Demougeot-Renard[1], Michel Garcia[2] and Nicolas Jeannée[3]

Abstract For several years now, the application of geostatistics to soil contamination has been studied in detail and specific approaches have been developed. Geostatistics is now recognized as a helpful tool to support decision making, which provides the ability to map site contamination and to assess the volumes of soils requiring remediation, together with an estimation of the uncertainty. But what about their practical implementation? From our experience, real case studies raised three major types of difficulties. First, the quality of the available data may require some approximations, secondly the clients may raise questions requiring specific theoretical developments, and last but not least communication about the results may require the use of appropriate vocabulary and illustrations. Making use of our experience, solutions to overcome these difficulties are presented on the case of a former gasworks located in France.

1 Introduction

Geostatistics meets a growing interest for the remediation forecast of potentially contaminated sites, by providing adapted methods to perform both chemical and radiological pollution mapping, to estimate contaminated volumes, potentially integrating auxiliary information, and to set up adaptive sampling strategies.

However, the adequacy of the methodological framework is far from being the only key for success in operational contexts. Indeed, several issues are commonly met when applying geostatistics in such contexts: the quality of the available data may require some approximations, the clients may raise questions requiring specific theoretical developments and communication about the results may require the use of appropriate vocabulary and illustrations.

The case study presented in this paper is a case at hand. Several characterization campaigns have been performed on a former gasworks located in France. Forecasting the sale of this site, the owner asked for a geostatistical

---

[1] EODE, 7 chemin de Mont-Riant, CH-2000 Neuchâtel, Switzerland, helenedemougeotrenard@eode.ch

[2] KIDOVA, 155 avenue Roger Salengro, F-92370 Chaville, France, michel.garcia@kidova.com

[3] GEOVARIANCES, 49bis av. Franklin Roosevelt, F-77215 Avon, France, jeannee@geovariances.com

project aiming at both estimating the amount of soil volumes not compatible with remediation thresholds and locating these soil volumes.

The paper first presents the site and the geostatistical methodology which has been applied. Then, the key issues are discussed, underlining their practical consequences on the success of the project and how we attempted to solve them.

2 Material and Methods

Site presentation

The case study used to illustrate the practical problems encountered when applying geostatistics to contaminated sites is a former gasworks located in an urban environment. The industrial facilities cover an area of 16500 m^2. Soil is classically contaminated with polycyclic aromatic hydrocarbons (PAH), mono-aromatic hydrocarbons (e.g. benzene), cyanides and heavy metals. Reference thresholds were applied to PAH[4] and benzo(a)pyrene (BAP), recognized as carcinogenic, to delineate the areas requiring remediation. The owner asked for a geostatistical study with numerous purposes:

1. Assess and delineate the volume of contaminated soil using the investigation data, before remediation.
2. Compare this estimated volume to the “really” excavated volume (calculated according to the number of trucks of soil removed from the site), after remediation.
3. Assess and delineate the volume of contaminated soil remaining after remediation (residual contaminated soil).

Available investigation data (Figure 1) consisted mainly of 217 laboratory analysis of PAH and 221 laboratory analysis of BAP collected in boreholes and trenches in two stages. Location and number of the investigation data were guided by the industrial past of the former gasworks and expert judgment. Data collected during remediation consisted of 268 analysis of PAH and 279 analysis of BAP collected in the walls and the bottoms of large pre-defined excavations. Compounds were analyzed in a mobile laboratory during remediation, using quick procedures. Other information, such as lithology or organoleptic criteria, were also available and could be used to study their correlation with the two major variables PAH and BAP.

---

[4] Sum of 16 PAH compounds, as recommended by the US EPA (EPA610 standard) and the European community

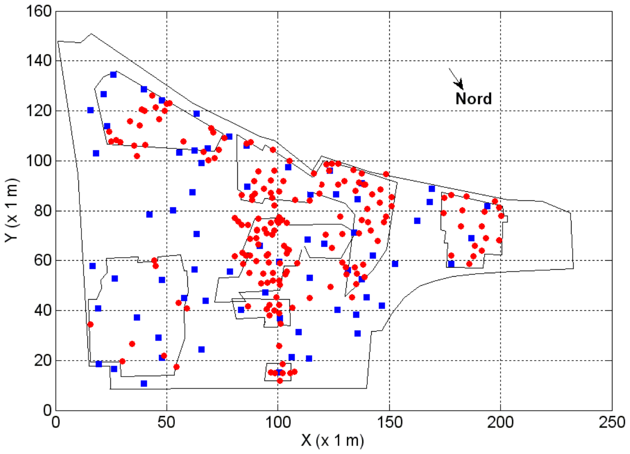


Figure 1 Horizontal distribution of the investigation data (blue squares) and remediation data (red points) in the study domain.

Geostatistical modeling

The following approach was applied to assess and delineate the volumes of contaminated soils for the three steps mentioned above. It is now a well established approach [1, 2] composed of the main following tasks:

1. Use simple univariate and multivariate statistics and exploratory analysis of all available information to define the main variables (contaminant grades) and the conditions of their geostatistical modeling.
2. Characterize and model the spatial structure of the main variables by mean of conventional variographic tools.
3. Define a grid in the modeling domain, with blocks corresponding to remediation units.
4. Generate conditional stochastic simulations (or co-simulations) of the main variables in the modeling grid, taking into account the change of support between investigation data and remediation units.
5. Compute, for each grid block, the probability that at least the grade of one contaminant exceeds its reference threshold.
6. Classify the blocks as *safe* or *contaminated* according to this probability, as well as the risk of misclassification, for decision-making purposes.

The main differences between the volumes estimated in the three steps mentioned above depend on the dataset (only investigation data, or investigation and remediation data) and the modeling domain (with or without excavated sectors for the need of remediation).

3 Discussion

Data Quality issues and consequences for the geostatistical model

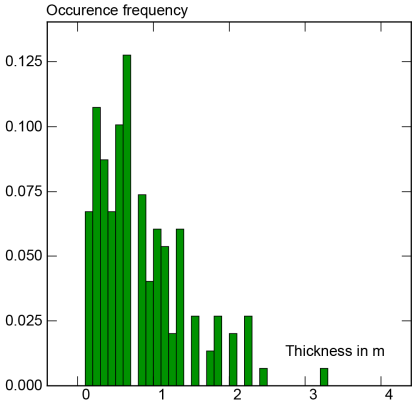
In most cases, geostatistical studies are asked at the end of the investigation stage, so that geostatisticians can usually not give their opinion on the sampling strategies. This situation is a disadvantage, increasing uncertainty on the final modeling. The former gasworks illustrates perfectly the main drawbacks encountered classically with data from polluted sites.

Support effect. Contaminant grades are usually measured on samples of various sizes, so that they do not constitute a statistically homogeneous population. The following rule was for example applied to the gasworks:

"Without any suspect organoleptic finding, a sample is collected in all different types of soils (anthropic materials or natural soil), as well as a combined sample, made of all encountered soils in the borehole.

In case of a suspect organoleptic finding, a sample is collected in all different types of soils, but also a specific sample, made of the suspected materials."

These sampling conditions mix preferential samples of very small size (decimetric) with averaged samples representative of a several meters height (Figure 2). Now it is well known that dispersivity of a given variable is usually decreasing while size of the sample is increasing (support effect). Data with various supports should thus not be mixed. But since no obvious mode could be found from the sample size distribution of the former gasworks, it was decided to keep all the data, being aware and informing of the uncertainty produced by this approximation, to provide nevertheless modeling results to the owner.



Number	149
Minimum	0.10 m
Average	0.83 m
Maximum	4.00 m
Standard deviation	0.64 m
Variation coefficient	0.77

Figure 2 Sample size of the investigation data (stage 1): histogram and elementary statistics.

Spatial structure of the PAH and BAP grades (after normal score transformation) was for example modeled approximately in the following way: the range of the variogram was defined using all the data, whatever their support, while the sill was fixed using only the data related to the smallest supports. The spatial continuity was indeed revealed only with all the dataset, while the subset composed a more homogeneous distribution whose variability could be interpreted as significant for the "point" data (thickness of 0.20 m).

Heterogeneous supports induce also approximations in simulation conditioning. The value of each conditioning data is usually assigned to the gravity center of each sample, whatever its size. This value may be then assigned wrongly to one block of the modeling grid, whose size may be smaller or larger from the data support. Other methods were also tested, such as simulations under linear constraints [3], but with no better results than this simplest one.

Sampling strategy. Contaminant data are often located according to the documentary information about the industrial past of the site and the experience of the expert, in order to detect hot spots. Spatial distribution of the data is thus irregular, with lack of information in areas assumed to be *safe* and high data density in areas suspected as *contaminated*. This situation may alter the elementary statistics, such as an artificial increase of the average or a decrease of standard deviation, even if declustering techniques can be used to try to correct these effects. Data scarcity on the vertical direction is also an issue, more specifically at greater depths, when a 3D model of pollution is required. Moreover, sampling strategy at the investigation stage is often guided by *a priori* that eventually proved to be wrong. Investigation data from the former gasworks showed these three types of issues:

1. Despite data were distributed rather regularly in the horizontal plane, some areas were nevertheless left without any information (Figure 3).

2. Data were very few vertically (2 or 3 contaminant grades per borehole and lithology), so that vertical spatial structure was nearly impossible to model.
3. There were fewer samples in the anthropic materials than in the natural soil, and investigation data showed that contaminant grades exceeded the reference threshold only in the anthropic materials. These results proved to be wrong at the remediation stage.

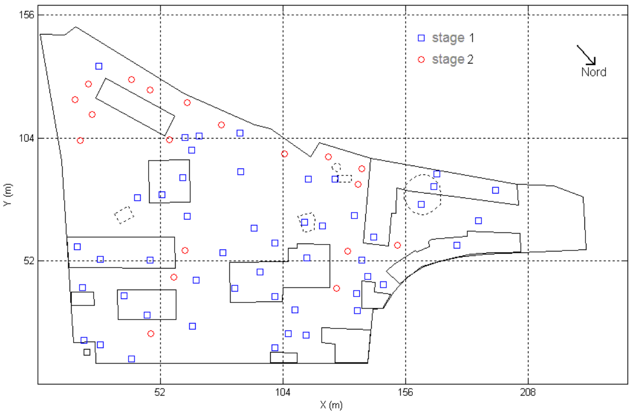


Figure 3 Horizontal distribution of the investigation data (stage 1 and 2) in the study domain.

These conditions induced findings of contaminated areas at the remediation stage that could not be expected from the investigation data, as well as underestimation of polluted volumes since the natural soil proved to be also contaminated.

Geometry of the modeling domain. The modeling domain is usually bounded by the topographic surface of the site and a lithologic boundary. Old building foundations remaining underground may also be excluded to calculate a realistic estimation of the total volume of soil in the domain. It appears that data provided to define this domain are often very imprecise, even if such information is crucial to get a significant estimation of *contaminated* volumes. In the case of the former gasworks, there were no data on the topographic surface, approximated location of old foundations and major gaps in the knowledge of the topographic surface after remediation. The boundaries of the modeling domain were then deduced from the lithologic data of the boreholes and a basic plan of a surveyor, using deterministic interpolation methods. This solution was not seen as totally satisfactory, since it increased uncertainty in the model without being able to quantify it, but was the best that could be done from the data.

From the geostatistical model to the remediation of contaminated soils: needs for specific theoretical developments

As already mentioned, the primary aim of the geostatistical study applied to the former gasworks was to estimate and delineate the volume of contaminated soils and to quantify the corresponding global and local uncertainties. Remediation plan was defined by the environmental consultants according to their experience of the domain. Site restoration was achieved by digging out the contaminated soils and off site treatment. In order to take into account spatial uncertainty, recommendations were made by the geostatisticians to start digging out soils where the probability of finding contaminated soils was very high (very likely contaminated soils) and to keep progressing until to reach not contaminated soils, control of contaminant grades being performed when necessary to confirm that they were below reference thresholds. In uncertain zones (Figure 4), away from the data used to build the geostatistical model, some boreholes were made to check the level of soil contamination. They allowed to identify a new source of contamination in one of these zones.

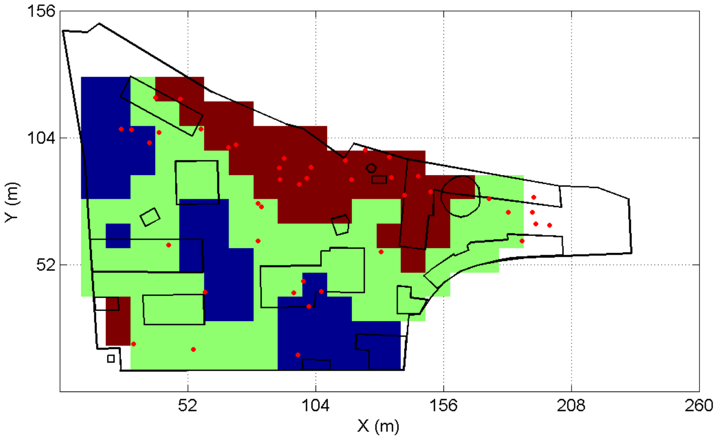


Figure 4 Remediation data (red points) superimposed to the geostatistical model built with investigation data. Soils are classified as very likely safe (blue), very likely contaminated (brown) and uncertain (green) in the model. Soils between 0.4 and 0.6 m depth.

The remediation of the site having been completed, with new data available from the analyses of soil samples collected during the remediation process, we were asked to:

1. demonstrate that the geostatistical model was providing good predictions,
2. assess the volume and location of residual contaminated soils, which was expected to be very small.

Such requests are all but surprising. They can be seen as natural quests to quality-control of the good achievement of the remediation: be sure that the contaminated soils have correctly been identified (validity of the geostatistical model), and that the residual contamination, if any, is insignificant (no forgotten contaminated soils).

Addressing these two questions seems to be easy. On the one hand, the simulated distribution of contaminant grades are to be compared with measured grades. On the other hand, the geostatistical model only needs to be calculated again by integrating the new remediation data.

Actually, they raise a number of issues that may become very consequential when the success of a geostatistical study is precisely considered as depending on good predictions and successful remediation of (almost) all contaminated soils. Should one of these easy to understand objectives appear as not satisfactory fulfilled, it is the whole geostatistical approach which is depreciated, the preference being then given to simpler empirical methods. The latter are generally unsatisfactory, whether they ignore spatial uncertainty or are unable to account for auxiliary information. Nevertheless, they are easier and faster to use than geostatistical methods, and tend to be preferred when the overall benefits of geostatistics cannot definitely be understood.

These aspects are discussed in the next sections.

Confronting the geostatistical model to remediation data. To fully demonstrate its validity, both the volume and the location of contaminated soils predicted by the model should be compared with the volume of remediated soils and with the contaminant grades measured in soil samples collected during the remediation process. If all the remediation data are within the predicted confidence intervals, the geostatistical model can be deemed to be good. If not, reasonable and understandable explanations must be found to justify the gaps. The error may be due to inappropriate assumptions or choices made in the geostatistical model, based on the data available before remediation. The remediation data themselves, however, may be the concern, making ineffective or partially effective the comparisons with the geostatistical model.

A common issue results from the location of remediation data. When confronting a geostatistical model with new contaminant grade data, all data do not have the same meaning or importance whether they lie in zones where the uncertainty is low or high. The higher the local spatial uncertainty, the higher should be the risk of error. It is one thing to say it to a geostatistician, it is another one to say it to persons non familiar with spatial uncertainty and who already tend to be doubtful about innovative approaches like for instance geostatistics!

As seen from Figure 5, the new remediation data cannot simply be compared with the geostatistical model, one by one. The relative location of each data, within the range of simulated values, is not much helpful to validate the model.

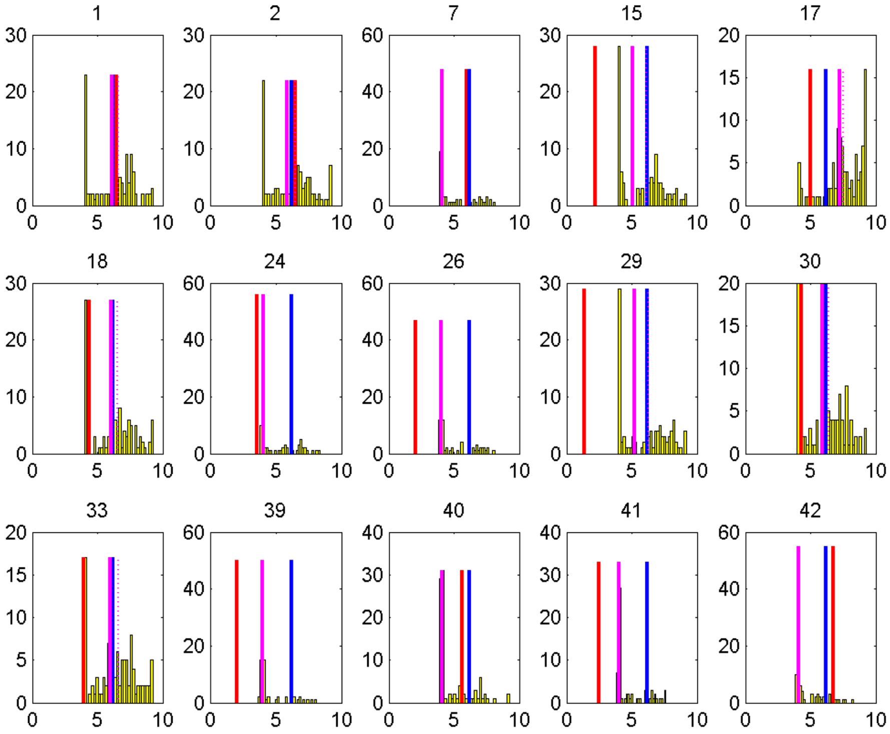


Figure 5 One by one comparison of the remediation data with the corresponding local distributions of 100 simulated contaminant grade values (point support). Red line = contamination data, pink line = median of simulated values, pink dot line = mean of simulated values, reference threshold = blue line.

Proportions of new data that are correctly predicted can be calculated by taking into account confidence intervals. Advanced notions like the one of precision and accuracy proposed by [4, 5] can be used, which also need to be explained to non-experts. Figure 6 shows the type of scatterplots that can be produced.

It can be also noticed that the new remediation data used to validate a geostatistical model are not consistently chosen for this purpose, but result from remediation goals that may not be compatible. From a statistical standpoint, the locations of remediation data are generally biased because they are preferentially located where non contaminated soils are expected (in the walls and the bottoms of predefined excavations). Small or intermediate contaminant grades are mostly measured, giving little chance to the new data to be higher than simulated grades. Such data may wrongly be interpreted as a tendency of the model towards overestimation (Figure 7). Whereas the investigation data are preferentially located in contaminated areas (high contaminant grades), the remediation data are more frequently found in *safe* soils (small and intermediate grades). Beside the overestimation trend issue previously mentioned, a non-stationarity issue may be faced: the geostatistical model relies on data mostly representative of contaminated soils, the remediation data used to validate the model mostly represent non- or poorly contaminated soils.

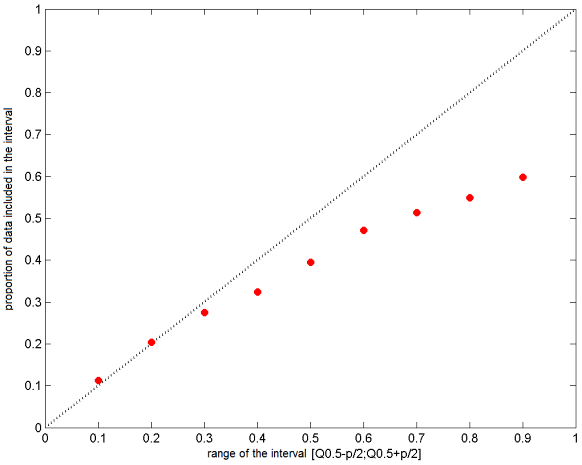


Figure 6 Precision and accuracy scatterplot of the proportion of data within the confidence interval vs. the confidence interval.

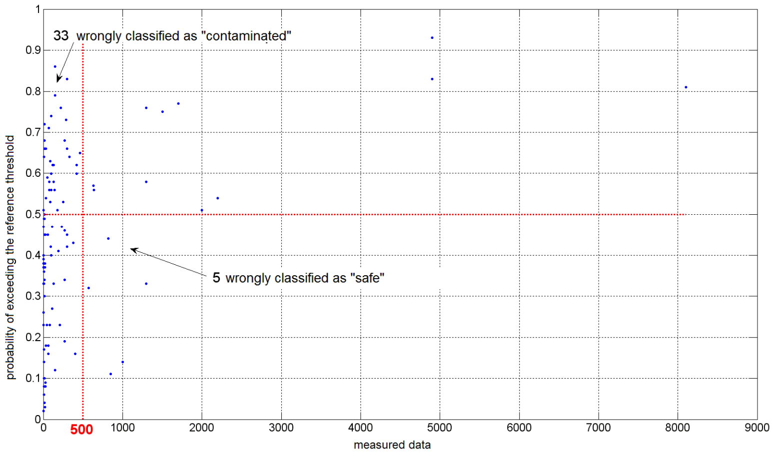


Figure 7 Scatterplot of the probability that the contaminant grade be greater than the reference threshold vs. the measured grade at the same location.

Other issues can affect the comparison of a geostatistical model with remediation data.

- Support effect: the remediation data are often related to composite soil samples representing a larger support than the “point” support of the geostatistical model based on small sample sizes. Whether the representative volume of the composite sample is not clearly defined, the resolution of the geostatistical model is not high enough to perform a reliable upscaling, or the level of uncertainty is varying within the composite sample volume, the comparison may become complex and imprecise.
- Short scale variability of contaminant grades: depending on the behavior of contaminants in the environment, the short scale variability of contaminant grades may be more or less important. Data a few tens of centimeters away can show very contrasted grades, as it can be seen in the former gasworks where PAH show a very high density and a very low mobility (Figure 8).
- Approximate estimation of remediated soil volumes: during the remediation process, the way the volume of contaminated soils is estimated may be questionable. Volumes are estimated by counting the number and the weight of trucks sent to the different treatment centers and waste deposits A few and non-systematic samples are collected for lab or on site analyses, or indirect measurements or organoleptic criteria are used to check rapidly and at a minimum cost the soil quality. Such data may not be reliable.

All these issues can affect the validation of a geostatistical model using remediation data. The model may be relatively good, based on statistical considerations, but part of the remediation data may be irrelevant to prove it. The gaps between model and data can be quantified, and we can argue for the gaps between the model and the remediation data. Making the arguments understandable and convincing to non-specialists is another challenge.

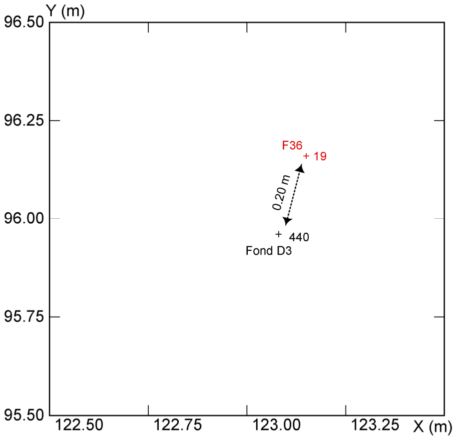


Figure 8 Example of contrasted investigation (red) and remediation (black) PAH grade data, corresponding to soil samples 20 cm apart.

Estimating and validating the potential residual soil contamination. The estimation of the volume of residual contaminated soils after remediation is something that is supposed to be easily done. It just consists in building (simulating) again the geostatistical model by taking into account all former (investigation) and new (remediation) data. Knowing the location of remediated soils, it can be obtained from it the volume and location of potential residual contaminated soils.

In practice, however, this simple exercise may become difficult for reasons already mentioned in the previous section.

1. Difference of support sizes between investigation and remediation data.
2. Non-stationarity of the geostatistical model away from the contaminated zone.

Nevertheless, an updated geostatistical model of the residual soil contamination can be generated, from which it can be derived statistics about the volume of residual contaminated soils and probability maps of their most likely locations.

What may be more a concern is the value that is given to such a residual soil contamination model. Especially, the appreciation of what is a small volume of residual contaminated soils is something very subjective and a potential source of disagreement between the owner of the site, who paid to have the soil contamination remediated, the engineer who conducted the remediation and was in charge of making the site “fully safe”, and the geostatistician who carried out the geostatistical studies. Our experience is that great cautions must be taken to update the geostatistical model, taking into account non-stationarity aspects as much as possible, and to present residual contamination results. All good results previously obtained can just be forgotten, and the geostatistical approach be fully depreciated, if the residual contamination is finally deemed to be too excessive.

Communication

Communication issues between the geostatistician and the other parties might occur at several stages of a geostatistical project. Most of these issues are due to the lack of geostatistical knowledge/background of the interlocutors. As a consequence, the latter might not understand some geostatistical assumptions, requirements, methods or results. This might not be problematic if the project results are pleasant to the contractor, but arguing for the complexity of the approach is quite easy in case the results differ from the contractor’s expectations.

Therefore, pedagogy is a key point at each stage of the project:

- Before the project, the geostatistician needs to be involved in the preparation of the sampling campaigns: sampling organization, homogeneity of sampling protocols, justification of these needs (support effect, representativeness of collected data); also,

expectations about the geostatistical project and its results should be clearly defined;
- During the project, methodological choices of importance should be explained as clearly as possible and most technical details should probably be put between brackets to avoid losing the interlocutor; in some cases, results might be simplified or expressed in meaningful ways: for instance, it is probably easier to classify soils based on a probability map to exceed a target threshold, instead than involving the risk of misclassification which requires to define tricky thresholds; being able to quantify the uncertainty about a soil pollution usually lead to explicitly show that whatever the decision is (except if the whole site is excavated), there is a residual pollution; accepting this idea might be difficult for the contractor or the administration.
- At the end of the project, ensuring that the geostatistical results can be used in an operational context is crucial to validate the added value of this approach.

Also, it is usually difficult for a geostatistician to face the absence of spatial structure, which can be due to inappropriate sampling but also to a random contamination (mixed backfill). Identifying this kind of situation as soon as possible is important to avoid irrelevant results (inappropriate variogram choice, overestimation of spatial continuity).

4 Conclusion

Nowadays, the operational application of geostatistics for the characterization of soil pollutions is becoming frequent. Although successful in some cases, this application is facing several issues which are discussed in this paper, based on a real example. Solutions applied to overcome some of the difficulties are presented. On the whole, it seems that the following efforts should be provided by the different stakeholders to improve the results of modeling:

- The owner and the field expert should ensure data efficiency by involving the geostatistician in the preliminary discussions about the sampling campaigns and the aims of the modeling.
- The geostatistician should inform the owner as soon as possible that the situation is not favorable to build a satisfactory geostatistical model, due for example to a lack of spatial structure of the phenomenon, data quality issues, etc.
- Knowledge of the field expert and information collected in the field should always be confronted carefully to the results of modeling, from an objective point a view of both the field expert and the geostatistician. The discrepancy between field information and modeling should be discussed in detail to find explanations.

- The geostatistician should learn how to explain simply the results of sophisticated theoretical developments, putting into brackets some technical details, while the owner and the expert should learn the basic principles of geostatistics to be able to differentiate the different sources of uncertainties of a model (data scarcity, data heterogeneity, underlying hypothesis of modeling, etc.).
- At the end of the project, the geostatistician should be able to provide directly usable results for decision making in an operational context.

In most cases, we think that these efforts should guarantee that the use of geostatistics for contaminated sites turns into a success story.

Bibliography

[1] Chilès J.P., Demougeot-Renard H., Fouquet De C., D'Or D., Garcia M., Le Loc'h G., Jeannée N., and Peraudin JJ. (2005). Géostatistique appliquée aux sites et sols pollués. Manuel méthodologique GeoSiPol et exemples d'application. http://www.geosipol.org

[2] D'Or D., Demougeot-Renard H., and Garcia M. (2009). An integrated geostatistical approach for contaminated site and soil characterization. *Mathematical Geoscience, 41*(3), pp. 307-322.

[3] Gómez-Hernández J. F. (2005). Exact conditioning to linear constraints in kriging and simulation. In O. Leuangthong & C.V. Deutsch (eds), *Proceedings of Geostatistics Banff 2004.* Banff: Springer.

[4] Deutsch C.V. (1997). Direct assessment of local accuracy and precision. In E.Y. Baafi & N.A. Schofield (eds), *Proceedings of the Fifth International Geostatistics Congress*, pp. 115-125, Wollongong, Sept. 1996: Kluwer.

[5] Shmaryan L. E., Deutsch C. V. and Journel A. G. (1998). A new Markov model for cokriging and simulation. *SCRF Report, 2*(11).