COMPARATIVE ASSESSMENT OF GEOSTATISTICAL TOOLS: A SUDBURY CASE STUDY

José F. Saavedra, Graeme Spiers, Paul G. Dunn

Abstract: Soil sampling, more than any other medium, reflects the total historical metal accumulation from point sources. In Sudbury, the mining and processing of mineral deposits for over a century has resulted in concerns about the potentially high levels of heavy metals within the soil environment. Sampling, analysis and interpretation of information describing metal levels within the footprint of the Sudbury smelter region is a pre-requisite for future developments in ecological and human risk assessment projects. In this paper a set of tools for incorporating spatial and temporal coordinates of observations into data processing is used to analyze a dataset. The data used is based on a sampling program developed using a randomized stratified sampling plan with modifications to remove, wherever possible, the potential effects of infrastructure. The nested sampling grid, covering an area 200 km by 200 km, was centered on the three smelters; Copper Cliff, Coniston and Falconbridge. The findings from the case study undertaken are that standard spatial interpolation techniques tend to overestimate the metal contents within the soil. The paper also highlights the potential for utilizing simulation and other geostatistical methods for providing the probabilistic distribution of metals within the soil and thereby improving the decision making process. Using the proposed methodology, differences up to 45% can be achieved when considering rehabilitation measures, thus showing the degree of risk the planner could face and how uncertainty could affect decisions.

Key Words: Geostatistics, Soil Characterization, Sudbury, Smelter, metal contaminants, spatial sample analyses.

Introduction

Sudbury has been home to mining, smelting and refining of nickel-copper ores since the late nineteenth century. The early methods of refining included the use roast yards in which layered sulphide ore with locally cut timber was ignited to heat the ore until the sulphide minerals ignited. The resultant nickel and copper concentrates were gathered for further refining (Winterhalder, 1995). The roasting process generated dense plumes of smoke, including sulphur dioxide (Freedman and Hutchins, 1980). Thus the local forests were denuded by felling for use as fuel and for construction of the railway, and also by the noxious gases emanating from the roast beds. Rapid, severe erosion of the barren soils ensued, resulting in exposure of the bedrock that has, in turn, been subject to intense acid weathering.

Estimates suggest that as much as $2.7 \times 10^4$ tonnes of SO$_2$ were emitted annually, together with many tonnes of heavy metal particulates (Holloway, 1917), at the peak of the ore roast yard era between about 1895 and 1928. In 1928 the use of open roast beds was forbidden by an order from the Ontario Legislature. The open roast beds were supplemented with more efficient smelter facilities with smoke stacks. The three smelters in the Sudbury region were located at Copper Cliff, Coniston and Falconbridge.

---

1 Paper was presented at Mining and the Environment IV Conference, Sudbury, Ontario, Canada, October 10-27, 2007.
2 jsavedra@mirarco.org. MIRARCO – Mining Innovation, Laurentian University, 935 Ramsey Lake Road, Sudbury, ON, P3E 2C6
3 gspiers@mirarco.org. Ibid.
4 pdunn@mirarco.org. Ibid.
In the mid 1970s, Ni and Cu emissions from the three smelters were estimated at 1100 tonnes per year (Cox & Hutchinson, 1981). The Coniston Smelter was decommissioned in 1972 when the “Super Stack” (381 m) was brought on line at CVRD INCO’s Copper Cliff smelter. Today all smelting within the region is carried out at the Copper Cliff (CVRD INCO) or Falconbridge (Xstrata Nickel) smelter. The CVRD INCO stack emitted 1.1 x 10^6 tonnes of SO2 and 1.2 x 10^6 tonnes in 1977 (Freedman and Hutchinson, 1980). Emissions from the Falconbridge smelter totaled approximately 2.0 x 10^5 tonnes in 1977, about 17% of the Copper Cliff total (Freedman and Hutchinson, ibid).

In addition to sulphur-rich emissions, large quantities of metal-containing particulate materials are vented through the stacks. In 1976 and 1977 metal emissions from the Copper Cliff stack amounted to 1.0 x 10^4 tonnes, a reduction from the total INCO emissions of 3.4 x 10^4 tonnes in 1970 (Freedman and Hutchinson, 1980). The particulate material emitted primarily comprises of iron oxides, with significant amounts of nickel and copper emissions. Smelter emissions from the Sudbury area smelters have decreased over the years. In 1995 total SO2 emitted from the INCO Copper Cliff and the Falconbridge smelters totaled 281,000 tonnes per year. Annual metal emissions, in 1995, were approximately; 140 tonnes Cu, 10 tonnes Zn, 87 tonnes Ni, 52 tonnes Pb, 10 tonnes Cd, and 48 tonnes As (OMOE, 1976). The Residual Discharge Information System (OMOE, 1996) places the annual total of suspended particulate matter emitted to the atmosphere for the Copper Cliff smelter at 7050 tonnes and for the Falconbridge smelter 1180 tonnes.

This paper focuses on applying geostatistical techniques to metal contaminant analyses within soils. Geostatistics provides a set of statistical tools for incorporating spatial and temporal coordinates of observations in data processing (Goovaerts, 1999). Geostatistics allows for description and modelling of spatial patterns, prediction (kriging) at unsampled locations and assessment of the associated uncertainty, and stochastic simulation (Goovaerts, 1998). The paper is organized as follows; a review of the main concepts in geostatistics, then a discussion on the appropriateness of kriging and geostatistical simulation techniques when performing environmental risk assessments, application of the concepts discussed on the Sudbury regional soils dataset (Spiers, 2004) and finally some ideas that could serve as future research avenues.

Geostatistical Concepts

(Goovaerts, 1997) gives a treatment of geostatistics in very intuitive terms and with a correct use of the mathematical language to formalise concepts. (Chilès and Delfiner, 1999) offers a comprehensive treatment of the field, in a mathematically oriented presentation. The whole idea behind geostatistics is that natural processes can be represented trough the concept of a regionalized variable (Matheron, 1969, 1971). A regionalized variable (R.V.) is a function \( f(\vec{x}) \) of the point \( \vec{x} \). A R.V. is presented trough two contradictory aspects:

- a *random* aspect (high irregularity, unexpected variations from one point to other)
- a *structured* aspect (the R.V. should reflect on their own terms the structural characteristics of a regionalized phenomenon)

A good conceptualization of the concept of R.V. is obtained trough the use of *random functions* (R.F.). A R.F. can be defined as a mapping function that assign to each point of the space \( \vec{x}_0 \) a random variable \( Z(\vec{x}_0) \). Note that this concept accounts for the random aspect and under some reasonable hypotheses accounts also for the structured aspect of R.V.’s.
In order to perform statistical inference for the R.F. is usual to adopt some hypotheses\(^5\):

(i) Stationarity of the random function. This implies a constant mean
\[ m(Z(\bar{x})) = m(Z(\bar{x} + \bar{h})) \forall \bar{x}, \bar{h}, \]
and second order stationarity which implies the existence of the covariance function
\[ C(\bar{h}) = E\{Z(\bar{x})Z(\bar{x} + \bar{h})\} - m^2 \forall \bar{x}, \bar{h}, \]
with \( m \) the mean (constant because stationarity).

(ii) Ergodicity of the R.F., meaning that spatial averages on infinite domains converge to expectation in the model.

In practice, we measure dissimilarities through the experimental semi-variogram which is defined as (Wackernagel, 1992):

\[
\gamma^*(\bar{h}) = \frac{1}{2n_h} \sum_{\alpha=1}^{n_h} \left( Z(\bar{x}_\alpha + \bar{h}) - Z(\bar{x}_\alpha) \right)^2
\]

Traditionally within geostatistics, a theoretical model is fitted for the semi-variogram. There are some properties that a variogram function must satisfy; primarily it must be positive-definite (Chilès and Delfiner, 1999). This theoretical semi-variogram function is denoted by \( \gamma(\bar{h}) \) to differentiate from the experimental semi-variogram. Under the hypotheses above, (i) and (ii), we have:

\[
\gamma(\bar{h}) = C(\bar{0}) - C(\bar{h})
\]

The whole purpose of geostatistics is to provide a value of the attribute of interest at an unknown location \( Z(\bar{x}_k) \). Two different problems in geostatistics must be differentiated, namely estimation and simulation. When estimating, the problem that is being solved is to find the “best”\(^6\) value of the attribute of interest in a given unknown location. Simulation solves the problem of obtaining the “true”\(^7\) probability distribution of an unknown location. There are many techniques available to solve these two problems but it is not the purpose of this paper to review them in detail. Between all possible choices, the two most representative; Ordinary Kriging and Sequential Gaussian Simulation, are briefly reviewed.

**Kriging versus Simulation**

**Kriging**

The idea behind kriging is to estimate the value at an unknown location \( \bar{x} \) by using the information of a set of samples \( \bar{x}_k \) by using a linear combination of those values (Rivoirard, 2003):

---

\(^5\) There are basically two lines of thought in geostatistics. Those who assume a set of stricter hypotheses and whose who do not. The methods differ slightly between those two schools, but the building blocks remain the same, with the more restrictive hypotheses used in this paper.

\(^6\) The use of “best” involves the concept of optimality of a given performance measure.

\(^7\) The use of “true” could seem abusive at first glance, is not possible to know the true distribution of the random variable \( Z(\bar{x}) \) (recall the definition of random function). By true we mean to reproduce the statistics for the stationary R.F.
Thus the weights in the summation above must be determined in order to minimize the error variance, with the unbiased condition\(^8\) being required in order to preserve the mean (constant due to stationarity assumptions). Minimization of variance produces the system (first order conditions of optimality):

\[
\begin{cases}
\sum_j \lambda_j C(\bar{x}_i - \bar{x}_j) - m = C(\bar{x}_i - \bar{x}) & \forall i \in \{1, \ldots, n\} \\
\sum_j \lambda_j = 1
\end{cases}
\]

This kriging variant is called **Ordinary Kriging**. Note that the mean is constant over the entire study area, but is not determined before performing the calculations, and is a characteristic of the random function. If the mean is calculated before, then the system varies slightly and the resulting kriging system is known as **Simple Kriging**. Finally, keep in mind that estimation by kriging is best in the least-squares sense because the local error is minimal. A shortcoming of the least-square criterion, however, is that the local variation of the Z values is smoothed. Another drawback is that the smoothing depends on the local data configuration; it is small close to the data locations and increases as the estimated location gets further from sampled locations. This uneven smoothing yields kriged maps that artificially appear more variable in densely sampled areas than in sparsely sampled areas. For all these reasons, interpolated maps should not be used for applications sensitive to presence of extreme values and their patterns of continuity, typically soil pollution data and physical properties (permeability, porosity) that control solute transport in soil. A better alternative is to use simulated maps which reproduce the spatial variability modeled from the data (Goovaerts, 1999).

**Sequential Simulation**

Sequential simulation algorithms are based on the following decomposition of any given probability distribution function (pdf):

\[
f(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}) = f(x_{\sigma(n)}|x_{\sigma(n-1)}, \ldots, x_{\sigma(2)}, x_{\sigma(1)}) \cdots f(x_{\sigma(2)}|x_{\sigma(1)}) \cdot f(x_{\sigma(1)})
\]

for every possible permutation function \(\sigma(\cdot)\) and \(|\) indicates conditioning (probabilistic conditioning).

To simulate a realization of a R.F., instead of using the joint pdf, the decomposition above can be used and generate an instance by visiting the nodes we want to simulate one by one. The value to be simulated at each node is drawn from a distribution with the pdf conditioned to the original data plus previously simulated locations. Each permutation provides with a different conditioning path and hence with a different realization. To ensure reproduction of the semivariogram model, each conditional cumulative distribution function (ccdf) is made conditional not only to the original \(n\) data but also to all values simulated at previously visited locations.

**Some Comments**

(Journel & Huijbregts, 1978) shows that the variance of a simulated location is twice the variance of the estimation for the same location:

\[8 \sum_k \lambda_k = 1\]
\[
\sigma^2_{\epsilon}(\bar{x}) = 2\sigma^2_{E}(\bar{x})
\]  

(6)

where \(\sigma^2_{\epsilon}(\bar{x})\) is the simulation variance and \(\sigma^2_{E}(\bar{x})\) is the estimation variance.

**Numerical Example**

**The Dataset**

The 2001 Sudbury Regional Soils Project was developed to allow preparation of an accurate geochemical map of the Sudbury smelter footprint to describe the impacts of anthropogenic metals on regional soils, waters, vegetation or sediments. The sampling program was developed using a randomized stratified sampling plan with modifications to remove, wherever possible, the potential effects of infrastructure. Previously existing "historical" sites sampled by INCO personnel on a regular basis since 1972 were included in the appropriate grid cell for this study to enable possible utilization of the historical data for time series analyses. The nested sampling grid, covering an area 200 km by 200 km, was centred on the three smelters, Copper Cliff, Conistone and Falconbridge. The centroid was in the vicinity of the Copper Cliff smelter.

The cells of the stratified sampling plan ranged in size, with the smallest cells being in the zones of historically known high smelter impact. The cells were 2, 4, 8 and 16 km square, respectively. The irregular 2 km grid cells extended approximately 8 km in each direction from the individual smelters. The 4 km grid cells extended 12 km out from the edge of the 2 km grid cells, with the 8 km grid cell area being squared off and designed to encompass the rest of the Greater City of Sudbury. The 16 x 16 km border grid was only one cell deep. The design of this survey will allow systematic expansion in any direction.

Within each cell one proposed sampling point was selected using a random lookup table in ArcView. A series of sampling exclusion zones were developed to minimize any resultant data bias from contamination by recent erosion, sedimentation or flooding, as well as from non-smelting anthropogenic activities such as effects of road construction, road salt drift, and railway traffic fugitive dust. These exclusion zones included: Industrial lands, such as tailings ponds, slag heaps, open pits

- Wetlands
- Lakes
- Rivers and streams
- 200 m of the centre line on primary roads
- 100 m of the centre line on secondary roads
- 100 m of railway lines
- 100 m of major utility lines.

In the 8 and 16 km outer cells, a replacement point was randomly reassigned if the initial sample location fell within an exclusion zone. The final number of sites sampled in the Regional Soils Project was 386.

Soil samples were collected according to standard protocol in the OMOE publication “Field Investigation Manual, Part 1, General Methodology” (OMOE, May, 1993). All sites for the survey were in stable landscape positions, with minimal evidence of erosion and with a full stable vegetation cover. Limed sites from the regional regreening program were not selected as core sites for the sampling program. In the wooded areas, sampling was conducted within a 10 m quadrant. The UTM coordinates of the stake at the quadrant corner were taken with a GPS unit and recorded on standard site description forms. Soil cores
were collected using Star Quality Soil Sampler that was cleaned between sample locations. According to OMOE procedure, a large “W” pattern was walked and cores collected along this pattern until a full sample of 30 cores was collected. Each core was sectioned from 0 to 5 cm depth, from 5 to 10 cm depth and from 10 to 20 cm depth. Cores were sectioned using stainless steel spatulas, with each depth interval being placed in a plastic bag.

The sampling at uniform depth increments as in this study may lead to variable mixing of the natural soils horizons, thus leading to a loss of any pedogenic overprint which could aid in understanding the natural geochemical and weathering processes responsible for translocation of soluble and colloidal materials.

**Basic Statistical Analysis**

The sampling points provided measurements of concentrations for multiple elements, with Nickel being the focus as it is most representative for the Sudbury area, given that the final objective of this paper is to compare assessment approaches rather than complete a comprehensive analysis of the available data.

Basic statistical analysis is always the first step in any geostatistical study. Usually, at this initial stage, we are interested in some general distribution characteristics for the attribute of interest. Almost always the histogram for the attribute of interest does not follow a known distribution. To fit our data to a specific distribution, formally speaking, statistical tests must be performed to check the validity of this assumption (e.g. Kolmogorov-Smirnov test). It is not uncommon to find the existence of very high values in the data. In classical statistics these high values are called outliers and usually some sort of action is taken to deal with them. In geostatistics outliers are very valuable, as they can be anomalous concentrations of mineral or, as in this study the presence of highly contaminated zones within the study area. Also the mean and the median are different, indicating presence of skewed data. This is classical behavior in earth sciences datasets. Sequential gaussian simulation requires normally distributed data, requiring that *Gaussian Anamorphosis*, a transformation of the data to make it follow a Normal distribution, be performed on this dataset.

**Variogram Fitting**

An examination of the basic statistics to provide some idea about the nature of the information allows progression to describe the spatial continuity of the data, a step crucial as almost all algorithms that can be applied depend on this description of the spatial continuity. New techniques are being developed to overcome variography, which improve the spatial continuity description by using multiple point statistics (Strebelle, 2002). Also because of the use of the two different techniques (kriging and sequential gaussian simulation), variography must be performed for both the original data and the normal scores data. To keep things simple, an omnidirectional variogram is fitted. It is possible to refine the analyses further and identify anisotropies. The variography is kept as simple as possible to illustrate the differences between geostatistical techniques.

**Estimation and Simulation**

There is a fundamental difference between simulation and interpolation techniques. Recall that interpolation methods arise when minimizing the error variance, and hence the smoothing affect that can be appreciated as in figure 1-(a).
By construction, simulation techniques reproduce the “true” variability of the data. “True” variability in this context means the actual spatial variability represented through the covariance model (or equivalently the variogram) is reproduced. A different model for spatial continuity will imply reproduction of a different spatial structure, hence the necessity of a good description of spatial continuity. This is the most important step in every geostatistical study and all the experience and interpretation goes here. Twenty (20) simulations were performed, by construction all of them are equiprobable realizations of the same random function. For clarity only one is shown 1-(b). Note that simulations are not smooth, but all of them honor the same data and are based on the same description of spatial continuity and their appearance is quite similar. Obviously there are some small differences, due to the nature of the simulation algorithm that in each realization selects a different random path to go through the nodes (see figure 1-(b)).

Some Warnings

Validation. To check the validity of any simulation or estimation effort, the moments of the target distribution (given by the original data) is compared with the statistics of the generated data (estimated or simulated). The final objective is to reproduce those statistics as accurately as possible. Nevertheless, that does not guarantee that the spatial distribution will be represented. In order to overcome these limitations new approaches to geostatistics have been developed in recent years where spatial continuity is described by means of multiple point statistics (Strebelle, 2002).

Variography. Variogram fitting is always the most delicate aspect in every geostatistical study. Good variogram fitting is an art that develops in time and usually needs to be performed by experienced and qualified personnel. Recall that as sequential simulation algorithms usually work under the assumption of Gaussianity of the sample data, new variograms must be fitted to transform the data. Simulation using the original variogram will not reproduce the spatial variability of the sample data.

Use of the Information

The Kriging mapping generated is different from the simulated maps. In order to make a reasonable comparison the average of all the realizations (E-type estimate) could be taken for and compare with the kriging map (figure 2).
Each cell in the grid has 2000 by 2000 meters. If we consider that rehabilitation will affect only the first layer of soil up to 5 cm, then each cell contains a volume of 200000 m$^3$. Between several alternatives for remediation, we are going to consider phytoremediation, because of its cost advantage (Schnoor, 1997), similar analysis can be done is other remediation technique is chosen.

<table>
<thead>
<tr>
<th>Type of treatment</th>
<th>Cost m$^3$</th>
<th>Time required (months)</th>
<th>Additional factors/expense</th>
<th>Safety Issue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixation</td>
<td>90-200</td>
<td>6-9</td>
<td>Transport/excavation</td>
<td>Leaching</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Long-term monitoring</td>
<td></td>
</tr>
<tr>
<td>Landfilling</td>
<td>100-400</td>
<td>6-9</td>
<td>Long-term monitoring</td>
<td>Leaching</td>
</tr>
<tr>
<td>Soil extraction, leaching</td>
<td>200-500</td>
<td>8-12</td>
<td>5000 m$^3$ minimum</td>
<td>Residue Disposal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Chemical recycle</td>
<td></td>
</tr>
<tr>
<td>Phytoextraction</td>
<td>15-40</td>
<td>18-60</td>
<td>Time/land commitment</td>
<td>Residue Disposal</td>
</tr>
</tbody>
</table>

Suppose now that a site will be declared contaminated if the Ni level exceeds 250 µg/g. Assume that the cost per cubic meter is the lowest possible from Table 1. If only the E-type estimate is used we see that we need to invest at least $1,227,000,000; if we compare with the kriging estimate ($1,197,000,000) we get at least a $30,000,000 difference. The big question now is; is this useful? As a first order approximation it is useful information, but like the kriging map, the E-type map has the drawback of not representing correctly the risk associated with the phenomenon. A risk profile for each cell in the map rather than a deterministic answer is more interesting. Taking into account that the simulations are an equiprobable set of realizations of the same random phenomena, for each node in the map the probability of being above the threshold of interest can be calculated (in the case of our example the threshold was set to 250 µg/g). The construction of this probability map is very simple, requiring a count of the number of times a given cell is above the threshold value and dividing by the number of total cases. The probability map is compared with the ordinary kriging estimation in figure 3.

Figure 3: (a) Probability Map (b) Kriged values above 250 µg/g (c) E-Type values above 250 µg/g.
The differences between (a), (b) and (c) in figure 3 are quite interesting. Whereas the probability map (figure 3a) indicates the likelihood of a given cell of being above the threshold level, kriging and e-type estimates only provide a deterministic answer. Also note the differences between the kriging estimate and the e-type estimate (figure 3, (b) and (c)): the kriging estimate is smoother and with less “holes” than the e-type estimate, this is direct consequence of the geostatistical technique used for each case.

At a risk level of 50% (the probability of having a contaminated cell is 50% or more) The experimental results show that between ordinary kriging and the classification based on simulation there exist a difference of 72 cells, with ordinary kriging being more conservative. In terms of costs this is translated on a lower bound for the difference of $216,000,000, or approximately 18% of the cost initially estimated. The use ordinary kriging to assess the amount of money required for remediation could lead to a serious overestimation. Furthermore, if we fix a risk level of only 25%, i.e. if the probability of having a contaminated cell is 75% or more, we find that misclassification increases to 184 cells and the corresponding cost difference is at least $552,000,000, or approximately 46% of the cost initially estimated. This simple example shows how different conclusions can be reached by using varying geostatistical techniques. Also, in this particular example, ordinary kriging overestimates the cost when considering some confidence level for decisions. This result implies that decisions based in ordinary kriging estimates must be considered with care.

Conclusions

When dealing with uncertainty one can choose to use several different geostatistical approaches. As better characterization of uncertainty is required, these more sophisticated techniques become necessary. Also, increased computational power makes the application of more sophisticated models feasible. Geostatistics has been demonstrated as a potential tool for site remediation assessment. But this approach is not free of obstacles which need to be taken into consideration. The use of the random function model simplifies the problem but at the cost of imposing unrealistic hypotheses on the phenomenon under study. Recent developments in geostatistics try to alleviate these strong assumptions, but in the meantime a certain behavior of the random field model is still required. Variogram fitting requires experience and could affect seriously the results that can be obtained. Interpolation methods are simple and can be applied with minimum computational effort. This simplicity, however, comes with the problem of producing results that do not replicate correctly the true spatial variability. Simulation techniques correctly address interpolation technique limitations, but given the theoretical requirements of simulation methods, the amount of required effort is increased by working in the normal score space for the data (trough Gaussian Anamorphosis).

For the case study investigated, interpolation techniques overestimate the requirement of resources for reclamation. This is not expected to be always the case and really depends on the particular problem under consideration. The use of different geostatistical techniques such a cokriging, indicator kriging, etc. provides new possible alternatives for performing the same task (Goovaerts, 1997), (Rivoirard, 2003), (Wackernagel, 1992) and (Chilès and Delfiner, 1999). Special emphasis has been put in multiple point geostatistics during the last decade and these new techniques seems to be promising in capturing better descriptions of spatial continuity than previous classical geostatistical techniques (Strebelle, 2002).

Finally, geostatistical simulation gives a good approximation of the true probabilistic distribution of the phenomenon under study and can be used with additional tools for decision making. There are several alternative methods ranging from the simple scenario approach and “what if” analysis to more complex mathematical models for resource allocation. Correct assessment of the underlying uncertainty can transform decision making processes and provide useful tools for efficient analyses. The approaches discussed in this paper should be investigated further to help maximize rehabilitation effectiveness within budget constraints.
References


Spiers, G. (2004). Metal Levels in the Soils of the Sudbury Smelter Footprint. A Report To: Safety, Health and Environment, INCO Ltd., Ontario Division, Copper Cliff, ON. P0M 1N0 and Falconbridge Ltd., Sudbury Smelter Business Unit, Falconbridge, ON. P0M 1S0. Submitted: July 12, 2004
