Using the power of information of sparse data for soil improvement management

Author(s):
Schnabel, Ute; Tietje, Olaf; Scholz, Roland W.

Publication Date:
2002

Permanent Link:
https://doi.org/10.3929/ethz-a-004398105

Rights / License:
In Copyright - Non-Commercial Use Permitted
Using the Power of Information of Sparse Data for Soil Improvement Management

Ute Schnabel, Olaf Tietje, and Roland W. Scholz

July, 2002
Using the Power of Information of Sparse Data for Soil Improvement Management

Ute Schnabel, Olaf Tietje, Roland W. Scholz

Contents

Contents ............................................................................................................................................................................................................................................... 1

Abstract .................................................................................................................................................................................................................................... 2

Introduction .......................................................................................................................................................................................................................... 2

Data ........................................................................................................................................................................................................................................ 4

Methods .............................................................................................................................................................................................................................................. 7

Analysis of spatial dependency ...................................................................................................................................................................... 7

Ordinary Kriging ................................................................................................................................................................................................................. 8

Conditional Simulation ..................................................................................................................................................................................... 10

Uncertainty Assessment .................................................................................................................................................................................... 11

Results ........................................................................................................................................................................................................................................... 15

Spatial Estimation ........................................................................................................................................................................................................... 15

Uncertainty Assessment ................................................................................................................................................................................... 17

Consequences ..................................................................................................................................................................................................................... 19

Conclusion .................................................................................................................................................................................................................................... 20

Uncertainty Estimation ................................................................................................................................................................................................... 20

Decision Relevance ....................................................................................................................................................................................................... 21

Acknowledgements ............................................................................................................................................................................................................. 22

References ...................................................................................................................................................................................................................................... 23

1 Working Group: Environmental Decision-Making, Environmental Evaluation and Modeling
Abstract

For a sustainable management of natural resources, such as soil, the spatial distribution of the environmental impacts is a basic need for decision-making. However, for spatial interpolation in most cases only few data with a skewed distribution and uncertain information about soil contamination are available whereas decisions with high "correctness" are required. In order to assess the power of information of sparse data a site of 15 square km with 76 soil samples was investigated. The soil was cadmium contaminated predominantly due to airborne emissions from a metal smelter. A lognormal probability distribution was found to appropriately estimate the probabilistic distribution of the contaminant. The spatial interpolation compares lognormal anisotropic kriging and conditional simulation. The resulting overall uncertainty from data sampling, sampling preparation, analytical measurement, interpolation and numerical representation has been investigated. The uncertainty of spatial interpolation was analyzed as the major component due to coarse data sampling and spatial heterogeneity. It is shown that the uncertainty can be efficiently estimated by calculating the percentiles of the lognormal probabilistic distribution function. This procedure also allows the calculation of the local probability of exceeding legal threshold values. Although the estimated uncertainty of the local prediction of the cadmium concentration is rather high, the procedure yields what is often required by decision making: a qualified rough estimate of the contamination. Conclusively, predicting the probability of exceeding a site-specific threshold can be used to roughly delineate prior areas for soil improvement, remediation, or restricted area use, based on the decision makers probability requirement.

Key words: conditional simulation, lognormal ordinary kriging, uncertainty assessment, probability of occurrence.

Introduction

Land use management becomes challenging for decision-makers of local and regional planning authorities especially in case of extensive land use and widespread spatial distribution of contaminants (Roe and van Eeten, 2001). Long-term impacts on soil fertility (Grunewald, 1997; Gysi and others, 1991) and precautionary thinking (Fedra, 1998; Prato, 2000) additionally affects the value of land. The impacts may lead to a shortage in usable land and accordingly to competing interests in different land uses. Therefore the assessment of a regional soil contamination may become relevant even if there is no direct endanger of human health. Thus, many investigations have dealt with regional soil analysis, monitoring (see e.g. (Meuli and others, 1994; Schulin and others, 1994), spatial interpolation (see e.g. Meuli, 1997), material flux analysis (Baccini and von Steiger, 1993; Keller and others, 2000), and uncertainty assessment (Keller, 2000). In order to assess affected areas in a valid, reliable and economic way, a core issue is a spatial interpolation that is able to include a high range of uncertainty without being unreliable (Briassoulis, 2001). Sources of uncertainty are data sampling (Schulin and others, 1994) preparing and conducting measurements in the laboratory (Dubois and Schulin, 1993), spatial representation (Meuli and others, 1994; Meuli, 1997), spatial interpolation (Goovaerts, 2001) and subsequent modeling (Keller, 2000; Keller
and others, 2000). Generally the uncertainty – from sampling to modeling – appears to be high, especially when different sources of error gather throughout investigation. Thompson and Fearn (1990) discussed a relation between sampling and analytical cost and the appropriate quality of the data in the context of data assessment. They propose a fitness for purpose concept, which is “the property of data produced by a measurement process that enables a user of the data to make technically correct decisions for a stated purpose.” (Thompson and Fearn, 1990, p.271). The decision, whether the available data and its uncertainty supply sufficient information thus depends on the purpose of data use, and decision-makers have to judge the appropriateness of uncertainties as a substantive base for their decision (Ramsey and others, 1998; Thompson, 1995). It is rather the decision and its consequences that matter than the accuracy of prediction (Goovaerts and Meirvenne, 2001).

However available data are often sparse because of unsystematic and arbitrary sampling (see also Korre, 1999). In practice, additional and strategic sampling is cost extensive. As a consequence, local and regional planning authorities are facing the dilemma that few data and uncertain information on soil quality or soil contamination are given whereas decisions with high “correctness” are required. As Mowrer (2000) suggest, the implementation of probabilistic uncertainty distributions and probabilistic reasoning becomes more and more important and accepted in environmental management (Bouma and others, 1996; Goovaerts and others, 1997; Van Meirvenne, 2001).

Therefore an important question for regional soil management is how to estimate the probability that a legal threshold value will not be exceeded (Goovaerts and Meirvenne, 2001). Modeling of uncertainty by means of stochastic simulation is widely used (Gotway and Rutherford, 1993; Pan, 1997; Van Meirvenne, 2001). Goovaerts (2001) compared kriging-based estimation with simulation-based spatial estimation and concluded that simulation yields several advantages, in particular, providing a model of local uncertainty. He concluded that a parametric modeling approach yields better results for local uncertainty, whereas a non-parametric approach performed best for propagation of spatial uncertainty. Several applied geostatistical models address the amount of uncertainty from different interpolation techniques (Barabas and others, 2001; Bouma and others, 1996; von Steiger and others, 1996) and most of them focus on the use of non-linear methods, like disjunctive or indicator kriging.

Despite the common use of non-parametric methods, the use of a parametric approach promises a best (linear and unbiased) estimation. In case of lognormally distributed data ordinary lognormal kriging is easy to implement and – if a sensitive (back) transformation of data is provided - yields best results compared to other kriging methods (Papritz and Moyeed, 1999; Saito and Goovaerts, 2000). The main objective of this investigation is to analyze the power of information given by spatial interpolation of few measurements combined with a probabilistic reasoning for practical planning purposes on the basis of sparse data available for the case of Dornach (Canton Solothurn, Switzerland). The estimation of cadmium contaminants with lognormal kriging and conditional simulation models the local uncertainty distribution. A locally conditioned prob-
ability density function allows a preliminary spatial screening for areas of soil improvement treatment with the help of probability maps. The following questions will be addressed:

- How can distant, irregularly distributed measurement points be efficiently interpolated for environmental management purposes?
- What kind of uncertainty assessment is appropriate in practical applications and how can the uncertainty in an interpolation be assessed accordingly?
- What are the chances and limits of the application for practical planning purpose?

Data

76 soil samples were taken in an area of 15 square kilometers. Each sample is a composite sample from topsoil (20 cm) within a range of 10 square-meter (Wirz and Winstörfer, 1987). Soil samples were taken without a specific sample design, more or less randomly distributed, regarding local circumstances and trying to cover the supposed area of contamination (Figure 1).

The total heavy metal content of the soil was determined in a 2-molar HNO₃ extract due to the Swiss Ordinance Relating to Pollutants in Soil (VBBo, 1998) and measured with a flame-atom-spectrometer (AAS).
Table 1. Statistical properties of the data.

<table>
<thead>
<tr>
<th></th>
<th>Cd</th>
<th>ln Cd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count (N)</td>
<td>76</td>
<td>76</td>
</tr>
<tr>
<td>Geometric Mean</td>
<td>0.925397</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>1.428289</td>
<td>-0.07753</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>2.69425</td>
<td>0.797583</td>
</tr>
<tr>
<td>Variance</td>
<td>7.258982</td>
<td>0.636138</td>
</tr>
<tr>
<td>exp(average(ln(Cd)))</td>
<td>0.925397</td>
<td></td>
</tr>
<tr>
<td>exp(stddev(ln(Cd)))</td>
<td>2.220168</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>0.855</td>
<td>-0.15667</td>
</tr>
<tr>
<td>exp(median(ln(Cd)))</td>
<td>0.855</td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>7.359588</td>
<td>0.777339</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>59.66685</td>
<td>2.887344</td>
</tr>
<tr>
<td>Min</td>
<td>0.12</td>
<td>-2.12026</td>
</tr>
<tr>
<td>Max</td>
<td>23.3</td>
<td>3.148453</td>
</tr>
</tbody>
</table>

The data analysis in Table 1 reveals that the distribution of heavy metal measurements is highly skewed. For such distributions, the logarithmic transformation may be appropriate before interpolation. The hypothesis that the logarithmically transformed data follow a normal distribution was accepted by Kolmogorov-Smirnov-Test (p = 0.35, Table 2) and shown by a Q-Q plot (Figure 2).

Figure 2. Normal Q-Q Plot of the lognormal transformed cadmium data.
Table 2. One-Sample Kolmogorov-Smirnov Test (test distribution is normal, mean and standard deviation are calculated from data).

<table>
<thead>
<tr>
<th></th>
<th>Cd</th>
<th>ln Cd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count (N)</td>
<td>76</td>
<td>76</td>
</tr>
<tr>
<td>Normal Parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.4274</td>
<td>-0.0778</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>2.69345</td>
<td>0.79715</td>
</tr>
<tr>
<td>Most Extreme Differences</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absolute</td>
<td>0.314</td>
<td>0.107</td>
</tr>
<tr>
<td>Positive</td>
<td>0.304</td>
<td>0.107</td>
</tr>
<tr>
<td>Negative</td>
<td>-0.314</td>
<td>-0.061</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov Z</td>
<td>2.735</td>
<td>0.935</td>
</tr>
<tr>
<td>Asymp. Sig. (2-tailed)</td>
<td>0.000</td>
<td>0.347</td>
</tr>
</tbody>
</table>

Within the area of interest a large part of the cadmium concentration is below guide values of the *Swiss Ordinance Relating to Pollutants in Soil* (VBBo, 1998) (Table 3). Another large part of the measurements holds cadmium contents above the guide value, but below the remediation value. If the guide value is exceeded, a remediation of soil can be considered appropriate in order to sustain soil quality and soil fertility or if an endangerment of groundwater or other goods is possible. However, there is no general compulsory requirement to clean up. If the soil contamination is between guide value and remediation value the change of land use often provides a less expensive alternative to remediation. A remediation is legally obligatory, when the remediation value is exceeded.

Because Cadmium was blown out of a metal smelter (indicated as arrow in Figure 1, Dornach, Canton Solothurn, Switzerland) during the past century and as the wind comes mostly from the west, higher cadmium concentrations are expected east of the metal smelter (Hesske and others, 1998).

Table 3. Legal threshold values for remediation of cadmium in soil in Switzerland (VBBo, 1998).

<table>
<thead>
<tr>
<th></th>
<th>[ppm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guide Value / Richtwert</td>
<td>0.8</td>
</tr>
<tr>
<td>Trigger Value / Prüfwert</td>
<td></td>
</tr>
<tr>
<td>Use with probable oral intake</td>
<td>10</td>
</tr>
<tr>
<td>Use for food planting</td>
<td>2</td>
</tr>
<tr>
<td>Use for feed planting</td>
<td>2</td>
</tr>
<tr>
<td>Remediation Value / Sanierungswert</td>
<td></td>
</tr>
<tr>
<td>Agriculture and horticulture</td>
<td>30</td>
</tr>
<tr>
<td>Gardening / Haus- und Familiengärten</td>
<td>20</td>
</tr>
<tr>
<td>Playgrounds / Kinderspielplätze</td>
<td>20</td>
</tr>
</tbody>
</table>

Knowing soil parameters and contaminants behaving randomly on small scale (Webster, 2000), we have to consider small-scale variability within the range of less than a few 10 m for data analysis. Geiger and Schulin (1995) found high variability (n=50, average= 4.38 ppm, variance= 0.72 ppm², min=2.37 ppm, max=5.92 ppm) on a 40 m transect for cadmium in the same area.
Methods

Within the framework of geostatistical analysis the measured cadmium concentrations (see above) are assumed to be one realization of a stochastic process (Cressie, 1993; Goovaerts, 1997; Webster and Oliver, 2001) \( \{Z(x) : x \in D\} \), where \( D \) denotes the sample space for the investigated domain, in this case the area of Dornach (see Figure 1). Let \( x_1, \ldots, x_n \) denote the data coordinates and \( Z(x_1), \ldots, Z(x_n) \) indicate the logarithms of the measured cadmium concentrations of the 76 measurements. Then, the (logarithmic) concentration differences \( Z(x) - Z(x + h) \) between the sites \( x, x + h \in D \) depend on the distance vector \( h \) (Matheron, 1963). The dependency is expressed by means of the semivariance \( \gamma(h) \), which is defined as half the variance of concentration differences of all data pairs with the same distance \( h \). Detailed discussions of properties and characteristics of this function can be found in the geostatistical textbooks (such as Cressie, 1993, pp. 58ff). We analyze this spatial dependency in order to interpolate the cadmium concentrations using lognormal ordinary kriging and conditional simulation (see below).

Analysis of spatial dependency

The spatial dependency of the cadmium concentrations is empirically investigated by using standard geostatistical software (Geovariances, 1997). Due to the point source at the metal smelter and the wind direction a stationary stochastic process is not expected. The investigation includes the (supposed) total area of contamination. The semivariogram (Journel, 1989; Webster and Oliver, 2001) shows the semivariance against the distance, where the range \( r \) – the "spatial dependency" – is the maximum distance between sites beyond which the differences in cadmium concentrations are uncorrelated.

![Variogram map of the logarithmic transformed cadmium data. Each grid cell shows the semivariance of the lag vector (x,y). D1-D4 indicate the direction of the anisotropic variograms.](image)

\( \gamma(h) \)

\(< 0.35 - 0.38 \)

\(0.39 - 0.45 \)

\(0.46 - 0.50 \)

\(0.51 - 0.70 \)

\(0.71 - 0.79 \)

\(0.80 - 0.89 \)
The variogram map (see Figure 3) of the logarithmic transformed cadmium data displays the empirical semivariance depending on length and direction of the distance vector \( h \). It shows that in the southeast direction the range of “spatial dependency” is larger than in other directions. This refers to the dominant wind direction discussed in Hesske and others (1998) and Wirz and Winißtörf (1986). It indicates that an anisotropic semivariance could be used. We analyzed a spherical function (see Figure 3), which can be modeled as

\[
\gamma(h, \theta) = c_0 + c_1 \left\{ \frac{3h}{2\Omega(\theta)} - \left( \frac{h}{\Omega(\theta)} \right)^3 \right\} \quad \text{for } 0 < h < \Omega(\theta) \tag{1}
\]

\[
\gamma(h, \theta) = c_0 + c_1 \quad \text{for } h \geq \Omega(\theta), \gamma(0) = 0
\]

where \( \gamma(h, \theta) \) is the semivariance depending on \( h \), the lag, which is defined as Euclidian distance between 2 points in 2 dimensions, and on \( \theta \), the angle representing the direction of the lag. \( c_0 \) is the nugget effect and \( c_1 \) the sill and

\[
\Omega(\theta) = (A^2 \cos^2(\theta - \alpha) + B^2 \sin^2(\theta - \alpha))^{1/2},
\]

which defines the anisotropy by means of an ellipsoid with \( \alpha \), the direction in which the continuity is greatest, \( A \), the maximum diameter and \( B \), the minimum diameter, perpendicular to \( A \) (Webster and Oliver, 2001). This expresses the so-called geometric anisotropy, for which the range (and not the sill) depends on the direction. According to the wind direction discussed above the direction \( \alpha \) of the largest diameter was set to -65° (155° in geographical notation, see Figure 3).

The geometrically anisotropic semivariance is shown in Figure 3 for the directions D1 \( \theta = -65° \) to D4 \( \theta = 70° \) counterclockwise. As opposed to isotropic kriging, the presented geometric anisotropy is more appropriate for the spatial structure of the investigated cadmium concentrations. Other kinds of semivariance functions (like Gaussian or exponential) and the zonal anisotropy are discussed, for instance, in Armstrong (1998), Journel and Huijbregts (1978), Matheron (1973), Tietje (1993) and Goovaerts (2001). The spatial structure of the data presented above was analyzed and the semivariance defined by Equation (1) was considered the best description of it (e.g. compared to an isotropic variogram, data not shown).

**Ordinary Kriging**

Because kriging requires a normally distributed data sample, the current interpolation applies a lognormal kriging method, where the data are logarithmically transformed before and back transformed after kriging. Ordinary kriging interpolates by means of a weighted average

\[
Z^*(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i)
\]

where \( Z^*(x_0) \) is the estimated value at point \( x_0 \), \( Z(x_i) \) is the measured value at \( x_i \), and \( \lambda_i \) are the weights (summing up to 1), which minimize the kriging variance (see below) and thus depend on the semivariance of the data pairs (Cressie, 1993; Goovaerts, 1997; Journel, 1989). Because the variance is minimized the interpolation is expected to smoothen local variation. The weights
Using the Power of Information of Sparse Data for Soil Improvement Management

depend on the spatial distribution of the measured data and the anisotropic variogram model (Equation 1, Figure 3 and 4). Measurements near by are more correlated to the estimation point and thus carry a higher weight than more distant measurements. However, the nugget effect slightly reduces the weight of nearer points relative to the weight of distant points (Webster and Oliver, 2001). Anisotropic kriging increases the weight of sampling points along the preferred direction, thus emphasizes the local spatial structure, and promises a more appropriate interpolation, e.g. if a reduced nugget effect can be obtained.

The determination of the kriging weights is also influenced by the restriction to use a maximum of the 7 nearest points for the estimation (Equation 2). Distant measurements are not taken into account, if enough measurements are available within a shorter range. Moreover, the spherical semivariance model (Equation 1) implies that only measurements within a moving neighborhood around the estimation point are included. The moving (ellipsoidal) neighborhood is defined by the parameters of anisotropy, i.e. a rotation of $-65^\circ$ and a range of $A=900$ m and $B=500$ m in the directions of D1 to D4.

The kriging variance $\sigma_{Z^*}^2$ is determined as the weighted sum of the semivariances, the estimation point, and the measured points (Equation 4). The kriging variance is then

$$\sigma_{Z^*}^2 = (x_0) = \sum_{i=1}^{n} \lambda_i \gamma(x_i,x_0) + \psi(x_0) \quad (3)$$

where $\gamma(x_i,x_0)$ is the semivariance between the data point $x_i$ and the estimation point $x_0$, the weights $\lambda_i$ are distributed in terms of the data model (see above) and under the requirement of unbiasedness, the results (see Equation 2) must sum up to one, and $\psi$ is the Lagrange multiplier used for minimization (Armstrong, 1998; Goovaerts, 2001).

After kriging the lognormal data must be backtransformed to the original scale. Several backtransformations have been proposed (Cressie, 1993, p.136; Saito and Goovaerts, 2000; Webster and Oliver, 2001 p.180). Many authors argue that lognormal kriging interpolation ought to yield the expectation of the original variable. In this case the backtransformation is conducted as the formula for calculating the expectation of the lognormal distribution

$$Z_r^* = \exp \left( Z^* + \frac{\sigma_{Z^*}^2}{2} \right) \quad (4)$$

where $Z_r^*$ is the estimated value in the raw scale, $Z^*$ is the estimated value and $\sigma_{Z^*}^2$ is the kriging variance, both on the logarithmic scale. For the kriging variance the transformation is,

$$\sigma_{Z^*}^2 = m_r e^{s_r^2} \left( 1 - e^{-\sigma_{Z^*}^2} \right) \quad (5)$$

where $m_r$ is the mean and $s_r^2$ the dispersion variance of the raw variable and $\sigma_{Z^*}^2$ is the kriging variance (on the log scale) (Geovariances, 1997).

As opposed to such backtransformations we are not interested in the estimation of the expectance of the Cadmium concentrations (as the raw variable). We want to derive a useful confi-
dence interval, which can easily be used for regional assessment. Therefore we follow the ap-
proach of Limpert and others (2001), who presented multiplicative confidence intervals for a mul-
titude of lognormal distributed data. We disregard the expectation of the lognormal distribution
and related arithmetic means, and calculate the *median* of the raw variable given the mean $\bar{Z}^*$ on
the log scale as (see Gut, 2001; Limpert and others, 2001; Stahel, 2000, p.136)

$$Z^*_M(x) = \exp(\bar{Z}^*(x))$$

(Median of the raw variable)

and the multiplicative standard deviation from the untransformed kriging variance $\sigma^2_Z$ on
the log scale as

$$\sigma(x) = \exp(\sigma^*_Z(x))$$

Please note, that $Z^*_M$ is the median (and not the expectation) of the raw variable and that
the multiplicative standard deviation (sometimes called geometric standard deviation) is calcu-
lated as the exponential of the standard deviation $\sigma^*_Z$ on the log scale and not as the exponential
of the variance on the log scale.

**Conditional Simulation**

In order to characterize the small-scale variability without a smoothing effect, a random re-
alization of a stochastic process is generated with the same geostatistical properties as the data.
The basic idea is to generate a multitude of such random realizations and characterize the distri-
bution at each estimation point through the average and standard deviation. The distribution is
Gaussian, so that the probability of an estimated value being less than a certain amount can be
easily computed. Because the original data (the Cadmium concentrations) are lognormally dis-
tributed, all calculations are conducted using the logtransformed data in order to obtain a Gauss-
ian distribution.

The stochastic simulation uses the turning bands method (Matheron, 1973; Tietje, 1993). The
turning bands method simulates realizations of a one-dimensional Gaussian process \{\(Y(tu)\): \(t \in R\}\) along the \(L\) lines \(L_1, \ldots, L_L\) through the origin defined by the unit vectors \(u_1, \ldots, u_L\). The lines have
all the same (arbitrary) origin (Gotway, 1994) and are uniformly distributed within the (in this
case) 2-dimensional space. If \(u_1, \ldots, u_L\) are unit vectors in the direction of \(L_1, \ldots, L_L\), the simulated value \(Z_s(x), x \in D,\) is an average of those values, which are generated at the orthogonal projection \(t_{ui}(x)\) of the terminal point \(x\) onto the line of direction \(u_i\) \(\{Y(t_{ui}(x)u_i) : i = 1, \ldots, L\}\). Finally the generated value at the point \(x\) is

$$Z_s(x) = \frac{1}{\sqrt{L}} \sum_{i=1}^{L} Y(t_{ui}(x)u_i)$$

(6)

A detailed discussion of the method can be found in Gotway (1994) and Journel (1974). Con-
ditioning the simulation now means to force the simulated surface to pass through the meas-
urements. Hence, the simulated value is obtained as (Journel and Huijbregts, 1978):

$$Z^{cs}_s(x) = Z^*(x) + [Z_s(x) - Z^*_s(x)]$$

(7)
Using the Power of Information of Sparse Data for Soil Improvement Management

where $Z^*(x)$ is the kriged value obtained form the data $Z(x_i)$, (measurements $i = 1, \ldots, n$), $Z_i(x)$ is the simulated value from Equation (6), and $Z_i^*(x)$ a kriged value obtained from $Z_i(x)$ using the realizations at the coordinates of the measurements. The term $[Z_i(x) - Z_i^*(x)]$ is not correlated with any data and is therefore independent of $Z^*(x)$. This term comprises the mere variation, i.e. if all measurements would be zero. The conditional simulation was conducted with 100 turning bands for each of 100 realizations at each gridpoint $x \in D$. A simulation post-processing averages the results of $i = 1, \ldots, 100$ simulations according to Equation 7 for each grid node

$$Z_{cs}(x) = \frac{1}{100} \sum_{i=1}^{100} Z_{cs}^i(x).$$

At each gridpoint $x \in D$ the mean is used to interpolate the cadmium concentration and the standard deviation

$$\sigma_{cs}^*(x) = \sqrt{\frac{1}{99} \sum_{i=1}^{100} \left( Z_{cs}^i(x) - Z_{cs}^*(x) \right)^2}$$

is used for the uncertainty assessment. In order to compare the results of the conditional simulation with the results obtained from log-transformed kriging and subsequent calculation of the median and the multiplicative standard deviation (see above), the same median calculations have been applied to the results of the conditional simulation:

$$Z_{cs}^M(x) = \exp\left( Z_{cs}^*(x) \right) \text{ (Median of conditional simulation of the raw variable)}$$

$$\sigma_{cs}(x) = \exp\left( \sigma_{cs}^*(x) \right) \text{ (standard deviation of conditional simulation of the raw variable)}$$

**Uncertainty Assessment**

Uncertainty gives an idea of the range of variability that might occur (Heuvelink, 1998) and of the difference between reality and the presentation of the reality. Uncertainty is often conceived of as the interval around the result of a measurement that contains the true value with high probability (Thompson, 1995).

The spatial heterogeneity of the Cadmium concentration is significant on different scales, from single point scale (local scale) with a characteristic length of some square meters, the practical scale of heavy metal field measurements to the scale of private property units (which is on the average on the scale of hectares in the investigation area of Dornach), and up to the "regional" scale of several square kilometers, the scale of geomorphological patterns and the reach of heavy metal air emissions. Scale questions have been addressed extensively within the 1990ies (see, e.g. Jarvis, 1995; Pahl-Wostl, 1995). The applied kriging interpolation includes a representation of variability on each of these scales. On the local scale and the practical scale of field measurements, the variability is expressed by the nugget effect and on the larger scale by the spatial variogram. In general, two types of spatial heterogeneity are considered. The spatial variability showing random effects within the soil, and the deterministic changes of heavy metal concentrations in space due to pollution patterns, land use, and geomorphology. Above, the kriging interpolation has been applied to make a deterministic estimate for the heavy metal concentrations. Then the correspond-
Due to the sparseness of the available data – especially considering the heterogeneity of the heavy metal concentrations in the soil – the interpolation includes a considerable amount of uncertainty. Therefore it is crucial, not only to estimate the local concentrations, but also their uncertainty. Moreover, because the uncertainty itself is being estimated and used to derive recommendations for specific site treatment options, an additional indicator for the “uncertainty of the uncertainty” is needed.

In order to assess the overall uncertainty of the interpolated concentrations we consider the following sources of error. Please note, that there is not only a deviation of the deterministic estimate, but also an influence on the implied uncertainty:

1. Field Sampling and Representation: Location and Mixing of soil samples: In our case 10 samples have been taken within a range of 10 times 10 m. In our geostatistical interpolation, this is represented by the nugget effect.

2. Sample preparation for analytical measurement: On the basis of literature data, the uncertainty is considered negligible with respect to the other, rather large uncertainties (see discussion below). It is notable that the Swiss Ordinance Relating to Pollutants in Soil (VBBo, 1998) prescribes to use HNO₃ extraction instead of extracting Cadmium with ‘aqua regia’. The kind of extraction has to be accounted for (qualitatively, because it cannot be calculated) in uncertainty assessment, when comparing with results from other countries, but it is ‘correct’ for Switzerland.

3. Analytical measurement: The uncertainty of the direct measurement (with the atom-absorption spectrometer, AAS) is also considered negligible with respect to the other, rather large uncertainties (see discussion below).

4. Interpolation: Due to the sparse measurements and the spatial heterogeneity, the interpolation introduces the main uncertainty. We try to estimate it by means of (lognormal) ordinary kriging and conditional simulation.

5. Numerical representation: The uncertainty due to numerical calculation and representation is neglected, because the large accuracy of the calculations and the small grid size introduce only a small error (2 orders of magnitude lower, data not shown).

Several authors discussed the contribution of different sources of uncertainty. Wächter (1997) showed from an interlaboratory test that the error induced by analyzing soil samples for heavy metal content varies from 10-30% (variation coefficient). The variation depends on the concentration. Low concentrations require a higher preciseness in the analysis and hence are more vulnerable to mistakes. This was verified by the comparative evaluation of European methods in soil sampling (CEEM project) (Desaules and others, 2001; Wagner and others, 2001). It demonstrated that the sampling and/or sampling preparation is the main sources of uncertainty and
even more depends on the analyzed element and the concentration level (Wagner and others, 2001b). Moreover, the investigation emphasized that especially for Cadmium the deviation from sampling is higher than the analytical deviations. From the results presented by Wagner and others (2001b, p. 87) we calculated an uncertainty factor of approximately 1.35 for sampling and sampling pre-treatment of different international samples at the same site. This uncertainty factor might also be caused by different sampling techniques (e.g. sampling depth, sieving), a high small-scale heterogeneity (about 40% see e.g. Wagner and others, 2001b, p.92) and different land use. The used reference sample (Desaules and others, 2001) yields variation coefficients for mixed samples of 8.1%, 7.2%, and 2.7% for forest, agriculture and grassland use. Von Holst (1997) also claims that the sampling itself causes the highest error followed by the sampling preparation. Wächter (1997) even infers that the error of the measurement instruments is negligible relative to the high error influence of sampling. Concerning the investigated area the accuracy of sample preparation and measurement (AAS) were assessed by different field studies for cadmium in topsoil (Table 4).

Table 4. Uncertainty of measurement preparation and analytical measurement (all measurements were done in the same laboratory).

<table>
<thead>
<tr>
<th>Number of probes</th>
<th>Variation coefficient [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-molar HNO₃ extract, flame-atom-spectrometer (AAS)</td>
<td></td>
</tr>
<tr>
<td>(Fröhlich, 1997)</td>
<td>12, one-fold measurement 1.6</td>
</tr>
<tr>
<td>(Wehrlen, 1998)</td>
<td>10, two-fold measurement 3.2</td>
</tr>
<tr>
<td>(Grünenfelder and Schmidli, 1998)</td>
<td>9, three-fold measurement 7.1</td>
</tr>
</tbody>
</table>

Theoretically, these contributions to the uncertainty are contained in the overall variance of the data. Because this might introduce a systematic error, an uncertainty factor of 1.35 is considered corresponding to the estimations of Wächter (1997) and Wagner and others (2001b, p. 87).

As the underlying data of this study are lognormally distributed, the uncertainty assessment is based on the logtransformed data. The statistical properties of the data are given in Table 1. Please note, that the statistical characteristics show, that the logtransformed data are not perfectly normal (see Figure 1). They show a skewness of 0.777 and hence there is a difference between the median (0.86 ppm) and the geometric mean (0.93 ppm) of the raw data. The Kolmogorov-Smirnov test (Table 2) shows, that the hypothesis of a lognormal distribution can only be rejected, if an error probability of more than 35% can be assumed. The hypothesis of a normal distribution can be rejected with an error probability of almost zero. Therefore the geometric mean and the multiplicative standard deviation (defined by ETH statisticians in (Limpert and others, 2001) can be used to estimate the characteristics of the distribution, especially for the estimation of the percentiles. The confidence limits for the geometric mean and the multiplicative standard deviation can be derived from the corresponding confidence limits of the log-transformed random variable Z. The confidence limits of the mean of Z can be estimated by
Ute Schnabel, Olaf Tietje, & Roland W. Scholz

\[ P\left(\bar{X} - \frac{S}{\sqrt{n}} t_{(1+\alpha)/2} \leq \mu \leq \bar{X} + \frac{S}{\sqrt{n}} t_{(1+\alpha)/2}\right) = \alpha \]  

(8)

with \(\bar{X}\) estimating the expectation, \(S^2\) estimating the variance, and \(t_{(1+\alpha)/2}\) being the \((1+\alpha)/2\) percentile of Students' t-distribution. The variance of \(Z\) can be assessed by

\[ P\left(\frac{n-1}{X_{(1-\alpha)/2}} S^2 \leq \sigma^2 \leq \frac{n-1}{X_{(1+\alpha)/2}} S^2\right) = \alpha \]  

(9)

with \(X_{(1±\alpha)/2}\) being the percentiles of the chi square distribution.

We are modeling the uncertainty as the sum of variances induced by different causes. Hence the uncertainty \(U(x)\) (Equation 10) is a random field

\[ U(x) = \sigma^2_{\text{constant}} + V(x) \text{ for all } x \in D \]  

(10)

which includes a constant part \(\sigma^2_{\text{constant}}\) and a spatially variable part \(V(x)\). \(V(x)\) indicates the error due to interpolation, \(\sigma^2_{\text{constant}}\) indicates the uncertainty induced by soil sampling and measuring. It is considered as the sum of the uncertainties of soil sampling (sampling pattern, density, composite or spot sample), preparation of the soil sample (fragment removal, drying, grinding, sieving, splitting) and the measurement of cadmium content (flame-atom-spectrometer (AAS) itself (Muntau and others, 2001; Theocharopoulos and others, 2001).

\[ \sigma^2_{\text{constant}} = \sigma^2_{\text{sampling}} + \sigma^2_{\text{preparation}} + \sigma^2_{\text{measurement}} \]  

(11)

Due to the geostatistical theory (Armstrong, 1998; Webster and Oliver, 2001) the kriging variance is zero at measurement points and minimized (Equation 2,3) between the sampling points. The magnitude of the kriging variance depends on the distribution of the sampling points (note that the kriging characteristics, such as neighborhoods, anisotropy and the support have to be set carefully). A dense sampling clearly improves the local prediction, because closely spaced points provide more information about the predicted value and thus reduce the kriging error (Goovaerts, 1997; Meuli, 1997; Van Groenigen, 2000). Because the sampling in the above-described investigated area is sparse, large parts of the investigated area are expected to show a high kriging variance. The maximum of the kriging variance is related to the sill of the variogram and hence depends on the overall variance of the data.
Results

Spatial Estimation

The estimation of the spatial distribution of the cadmium concentration relies on the geostatistical analysis of the data. As expected, the variogram on the raw scale is not satisfactory in terms of spatial dependency, because the data are not normally distributed (data not shown). The isotropic and anisotropic variograms on the log scale are shown in 4 and 5.

### Figure 4
Anisotropic variograms in the directions D1-D4; their ranges are 900 m and 500 m, the nugget effect is 0.13 and the sill is 0.53.

### Figure 5
Isotropic variogram with a range of 813 m, a nugget effect of 0.3, and a sill of 0.332.

In both cases the sill equals approximately the data variance. Because the nugget effect is larger in the isotropic case, the anisotropic variogram is considered to better represent the spatial distribution of the measured cadmium concentrations. In the anisotropic case, the range is be-
between 900 and 500 meter, embracing the range of the isotropic model. The anisotropic model
(Figure 4) is computed for lag classes of 230 m and behaves discontinuous near the origin showing
a nugget effect of 0.125 (log scale). A spherical model fitted the empirical variograms (see table 5)
best in all cases.

Table 5. Geostatistical properties of the variograms, all data shown on log-scale.

<table>
<thead>
<tr>
<th></th>
<th>Isotropic variogram</th>
<th>Anisotropic variogram</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>spherical</td>
<td>spherical</td>
</tr>
<tr>
<td>Direction</td>
<td>omnidirectional</td>
<td>four directions, D1-D4</td>
</tr>
<tr>
<td>Range</td>
<td>813 m</td>
<td>-65° = 900m, -20°, +20°, +70° = 500m</td>
</tr>
<tr>
<td>Nugget effect</td>
<td>0.297</td>
<td>0.125</td>
</tr>
<tr>
<td>Sill</td>
<td>0.332</td>
<td>0.528</td>
</tr>
</tbody>
</table>

The differences between kriging of cadmium with isotropic and anisotropic variograms are
due to the different nugget effect. The larger nugget effect leads to a smoother isotropic kriging
interpolation (Webster and Oliver, 2001) with a smaller variance than in the anisotropic case. This
leads to an underestimation of the highest cadmium concentrations with the isotropic kriging.
The consequences for the kriging variance and the uncertainty assessment are discussed below.

Figure 6. Spatial distribution of Cadmium concentrations in top soil after conditional simulation. Dis-
played is the median of 100 realizations on the raw scale [Z^M] in the core area indicated in figure 1.

The estimated (interpolated) cadmium concentrations are shown in Figure 6. This is the re-
sult of the conditional simulation. The statistical properties of the interpolations (isotropic and
anisotropic) and the conditional simulation are shown in table 6. In order to avoid disturbing
boundary effects only part of the estimated (interpolated) cadmium concentrations are displayed.
The statistical properties (Table 7) have been calculated for the area indicated as ‘core area’ in Fig-
ure 1.
Using the Power of Information of Sparse Data for Soil Improvement Management

Table 6. Statistical properties of the interpolations (isotropic and anisotropic) and the conditional simulation.

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Isotropic estimation</th>
<th>Anisotropic estimation</th>
<th>Conditional Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counts (N)</td>
<td>76</td>
<td>11878</td>
<td>10802</td>
<td>10802</td>
</tr>
<tr>
<td>Geometric mean [e^μ ppm]</td>
<td>0.925</td>
<td>1.049</td>
<td>1.082</td>
<td>1.091</td>
</tr>
<tr>
<td>Standard deviation [e^σ ppm]</td>
<td>2.220</td>
<td>2.326</td>
<td>2.272</td>
<td>2.325</td>
</tr>
<tr>
<td>Minimum [ppm]</td>
<td>0.12</td>
<td>0.120</td>
<td>0.120</td>
<td>0.137</td>
</tr>
<tr>
<td>Maximum [ppm]</td>
<td>23.3</td>
<td>4.36</td>
<td>8.805</td>
<td>9.025</td>
</tr>
</tbody>
</table>

Table 7. Statistical properties for the core area of pollution indicated in Figure 1.

<table>
<thead>
<tr>
<th>Anisotropic modeling</th>
<th>Core area</th>
<th>Total area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cadmium estimated</td>
<td>Cadmium simulated</td>
</tr>
<tr>
<td>Counts (N)</td>
<td>3870</td>
<td>10802</td>
</tr>
<tr>
<td>Geometric mean [e^μ ppm]</td>
<td>1.025</td>
<td>1.027</td>
</tr>
<tr>
<td>Standard deviation [e^σ ppm]</td>
<td>1.954</td>
<td>1.972</td>
</tr>
<tr>
<td>Minimum [ppm]</td>
<td>0.283</td>
<td>0.284</td>
</tr>
</tbody>
</table>

Uncertainty Assessment

The overall uncertainty assessment is based on the statistical properties of the data and their interpolation (see Table 6). We are not interested in the mean and the variance of the estimations, but in the assessment of percentiles. The rationale is illustrated with Figure 7, a south north transect (easting 614240m). The singular realization is one of 100 turning bands simulations showing a considerable variability and randomly high cadmium concentrations. The 50% percentiles for lognormal kriging and conditional simulation are estimated as the exponential of the mean values on the logarithmic scale. The other percentiles are calculated accordingly as the exponential of the percentiles on the log scale \( \exp(Z_{\log}^*(x) \pm \phi(p)\sigma_{\log}(x)) \), with the 16% and 84% percentile of the normal distribution \( \phi(16\%) = -1 \) and \( \phi(84\%) = 1 \), according to Equations 8 and 9, and Limpert and others (2001). The standard deviation on the log scale is obtained as the standard deviation of 100 realizations (conditional simulation) or the root of the kriging variance (lognormal kriging).
Figure 7. A south-north transect (easting 614240 m) with the results of one singular realization and some percentiles of the conditional simulation and lognormal kriging.

The kriging variance as used in Figure 7 measures the estimation error of kriging (Cressie, 1993; Webster and Oliver, 2001). Hence the percentiles yield the probability and confidence limits, if only the error due to interpolation is accounted for (see above number 4). The numerical representation error is neglected. Due to the discussion above the total uncertainty to be accounted for is 

$$\sigma_u^2(x) = \sigma_{\log}^2(x) + \ln(1.35)^2$$

which is used to estimate the spatially distributed uncertainty shown in Figure 8. Here the probability of exceeding a threshold of $C=2$ ppm is estimated by solving

$$P(Z^* > t) = 1 - \phi\left[\ln C - Z_{\log}^*(x)/\sigma_{\log}(x)\right]$$

for the probability $p_x$ of the standard normal distribution $\phi$.

Figure 8. Spatial distributed probabilities to exceed the threshold value $C=2$ ppm.
Consequences

The consequences of the interpolation combined with uncertainty assessment have to be related to the Swiss Ordinance Relating to Pollutants in Soil (VBBo, 1998) (Table 3). It contains the three thresholds values related to a) a long-term fertility goal (guide value), b) the concentration which indicates that a more detailed investigation is required (remediation value), and c) the concentration beyond a remediation is prescribed (remediation value).

Similar to Figure 8, where the probability of exceeding 2 ppm (trigger value for food and feed planting) is presented, the probability of exceeding any other threshold can be shown. As can be seen from Figure 7, there might be some local cadmium concentrations, which exceed the 84% percentile. Therefore the local administration should determine, which percentiles are going to be used for the assessment. The presentation of one singular realization might help to understand the meaning and the consequences of the results, which are stated as: “there is a 16% probability that a specific cadmium concentration is exceeded.” This is helpful, because - even for experts - probabilities are often difficult to assess (Scholz, 1987; Tietje and Scholz, 1996).

After having defined adequate probabilities of exceeding a threshold value, it becomes possible to delineated corresponding areas, which are subjected to more intensive investigations, like additional measurements or recommendations to the public. Our results show that in general the uncertainty from spatial interpolation is a factor of 2 (in the smaller region) or 2.3 (in the whole region of interpolation) (see table 6), the remaining uncertainty has been estimated as a factor 1.3. Hence assuming that these uncertainties are uncorrelated, the overall uncertainty is

\[ f = \exp \sqrt{\sigma^2_{\text{log}} + \sigma^2_{\text{log}}} \]

The accuracy of factor 2, postulated by Fresenius and others (1995), can be achieved only within one standard deviation \( \phi(p_x) = 1 \), and the demand of 97.5% accuracy leads to a uncertainty factor of about 4.2. A combination of the spatially distributed percentiles with other data of a GIS, such as land use classifications, might lead to valuable decision material, e.g. when additional measurements or at some sites even possible remediations are to be decided.
Conclusion

The aspiration of the paper is to generate a "correct" spatial assessment of soil contamination for a medium sized area (~ 10 km² - 100 km²) based on sparse data. Taking into account different sources of uncertainty from data sampling to the spatial representation of the interpolation results, we acknowledge that statements such as The probability that the soil contamination at this point exceeds 2 ppm cadmium is less than 60%, in our opinion, seems to be the only statements that can be given in a "correct" way, although they are itself uncertain (because, e.g., the median is estimated by the geometric mean). The method to gain the required spatial information, demonstrated for the prediction of cadmium concentration in top soil, shows that even sparse data can be used for first assessment of contaminated land. A logarithmic transformation of the data and a backtransformation after interpolation according to Limpert and others (2001) can be easily used to obtain the percentiles of the probabilistic distribution function and to predict the conditional cumulative density function including the uncertainty due to sparse sampling.

In order to represent the spatial distribution of the Cadmium concentrations, a lognormal ordinary kriging (Papritz and Moyeed, 1999; Matheron, 1963) and a conditional simulation approach (Goovaerts, 2000; Journel, 1974) were applied. The ordinary kriging approach could be improved by the application of an anisotropic variogram. The anisotropic variogram shows varying distances of autocorrelation in different directions. This reduced the nugget effect and thus decreased the smoothing effect of kriging (Goovaerts, 2001). In order to get an idea of the possible variations of the cadmium concentration, several three dimensional stochastic realizations were produced and a conditional simulation was applied. The conditional simulation (as the stochastic realizations are forced to hit the measured data) yielded also an estimation of the local uncertainty. The geometric mean of all realizations approximated the geometric mean of the (lognormal) ordinary kriging estimation. Moreover, the standard deviation of all the realizations approximated the standard deviation of the (lognormal) ordinary kriging, i.e. the square root of the kriging variance (see Table 6).

Uncertainty Estimation

The aim of the interpolation was not only to assess the spatial distribution of the Cadmium concentration but also to estimate the corresponding uncertainty. Due to the probabilistic distribution function of the data, it is good practice to apply a lognormal transformation before interpolation (Journel and Huijbregts, 1978; Kravchenko and Bullock, 1999; Saito and Goovaerts, 2000). Regarding the representation of the uncertainty many other investigations applied indicator kriging (Van Meirvenne, 2001; Webster and Oliver, 2001) or disjunctive kriging (Papritz and Moyeed, 1999; von Steiger and others, 1996), because of the complicated estimation of the local probabilistic distribution function when applying a traditional backtransformation (Geovariances, 1997; Webster and Oliver, 2001, p.19). Due to the fact that the local probabilistic distribution function on the raw scale is skewed the estimation of mean and standard deviation on the raw scale is not convenient. Instead, applying the concept of Tietje (1998) and Limpert and others (2001) the uncertainty on the raw scale is estimated by means of the geometric mean and the multiplicative stan-
standard deviation. With this approach the uncertainties involved in the interpolation can be reliably and conveniently estimated. Hence this procedure can be applied for uncertainty assessment of lognormally and spatially distributed data.

The uncertainty assessment shows, that the main part of the overall uncertainty is due to the interpolation and the spatial sparseness of the data, although it is very difficult to assess the overall uncertainty (Quevauviller, 1995; Ramsey, 1998; Theocharopoulos and others, 2001) of the final interpolation results. The second important uncertainty is due to the sampling and sampling preparation. Because corresponding data are hardly to obtain (see e.g. CEEM-Project, (Wagner and others, 2001)) we roughly estimated this as 35% of the local Cadmium concentration using quantitative (von Holst, 1997; Wächter, 1997; Wagner and others, 2001b) and qualitative results (Ramsey, 1998; Tiktak and others, 1999). Other sources of uncertainty, such as the analytical measurement and the numerical representation have been determined to be negligible against the first and second sources of uncertainty described above. The large amount of uncertainty arises throughout the whole area; it is smaller only in the vicinity of the measurements. This indicates that the variability of the data determines the final uncertainty. But, if so, why can we conclude that the uncertainty estimation is not far too high due to the sparse data? The estimated logarithmic standard deviation of the data does only marginally change if an additional measurement would be included (increasing from 2.22 up to 2.38 if a cadmium concentration between 0.1 and 20 would be additionally measured). Thus the resulting overall uncertainty is not too high to prevent the identification of regions with a higher risk of contamination.

**Decision Relevance**

The requirement of Fresenius and others (1995), to determine the uncertainty within a range of ‘factor 2’, is fulfilled within a confidence of 68% in our study. However the desired reliability and confidence of 90 to 95% confirms the hypothesis that further measurements should be conducted. The interpolation combined with the uncertainty assessment is sufficiently reliable to delineate potentially contaminated sites pictured in “maps of probability of occurrence”. For example, the probability to exceed legal threshold values helps to be aware of failures and necessities in sustainable soil improvement – depending on the decision-makers probability requirements. With in a GIS environment they can be linked with economic or social criteria and therefore contribute to an evaluation of remediation options. For the case considered in this study the following conclusion are evident from the interpolation and the uncertainty assessment:

- Cadmium contamination above legal remediation value is not likely on any side.
- A small area is contaminated above trigger value with a probability of more than 90 %
- Cadmium contamination above guide value is widespread with more than 50 % probability
- High probabilities for cadmium contamination above 0.8 ppm are delineated mainly around the source
- Only a few areas are not contaminated according to threshold values
To sum up the area is contaminated with cadmium between guide and trigger value of *Swiss Ordinance related to Soil* (VBBo, 1998) (Table 3). According to law, a remediation is not necessary, but there are some areas, where at least a change of land uses has to be considered by local authorities. However the results of the proposed method show widespread areas, where soil quality is affected by cadmium contamination. These areas might be considered for soil improvement measures, especially where land is used for agriculture or gardening. Since the contamination from different sources like sewage sludge or fertilizer is ongoing (although on a low level and mainly on agricultural land), decision maker should bear in mind that the cadmium concentration might still be accumulating (Keller, 2000) and any kind of soil improvement will help to sustain the multi-functional use of soil.

**Acknowledgements**

The work reported in this paper was undertaken in the context of the Integrated Project Soil of the Swiss Priority Program Environment, which follow the aim to elaborate a concept for gentle soil remediation of heavy metal contaminated sites in 1997-2001. We are grateful to Martin Fritsch, who initiated the project and guided it until 2000, and to the colleagues at UNS and within the IP Soil for the helpful discussions. The project was funded by the Swiss National Science of Foundation (Project Number 5001-44758) and the Swiss Federal Institute of Technology (Project Number 0048-41-2609-5).
References


Goovaerts, P., and V. Meirvenne. 2001. Delineation of hazard areas and additional sampling strategy in presence of a location specific threshold, in A. Soares, J. Gomez-Hernandez, and J. Froi-


Tietje, O. 1998. Note about lognormal transformation in spatial interpolation methods. Personal communication


UNS-Working Paper 1 (Out of Print)

UNS-Working Paper 2 (Out of Print)

UNS-Working Paper 3

UNS-Working Paper 4

UNS-Working Paper 5 (Out of Print)

UNS-Working Paper 6 (Out of Print)

UNS-Working Paper 7 (Out of Print)

UNS-Working Paper 8 (Out of Print)
UNS-Working Paper 9 (Out of Print)

UNS-Working Paper 10 (Out of Print)

UNS-Working Paper 11

UNS-Working Paper 12 (Out of Print)

UNS-Working Paper 13 (Out of Print)

UNS-Working Paper 14 (Out of Print)

UNS-Working Paper 15

UNS-Working Paper 16 (Out of Print)

UNS-Working Paper 17

UNS-Working Paper 18

UNS-Working Paper 19 (Out of Print)

UNS-Working Paper 20

UNS-Working Paper 21

UNS-Working Paper 22 (Out of Print)

UNS-Working Paper 23 (Out of Print)

UNS-Working Paper 24

UNS-Working Paper 25

UNS-Working Paper 26 (Out of Print)

UNS-Working Paper 27

UNS-Working Paper 29

UNS-Working Paper 30

UNS-Working Paper 31

UNS-Working Paper 32

UNS-Working Paper 33