CHALMERS | GÖTEBORG UNIVERSITY

PREPRINT 2007:1

On the Cost-Effectiveness of Sampling Contaminated Soil: Comparison of Two Models

PÄR-ERIK BACK
TOMMY NORBERG
LARS ROSÉN

Department of Mathematical Sciences
Division of Mathematical Statistics
CHALMERS UNIVERSITY OF TECHNOLOGY
GÖTEBORG UNIVERSITY
Göteborg Sweden 2007

Preprint 2007:1

On the Cost-Effectiveness of Sampling Contaminated Soil: Comparison of Two Models

Pär-Erik Back, Tommy Norberg, Lars Rosén

CHALMERS | GÖTEBORG UNIVERSITY

Department of Mathematical Sciences
Division of Mathematical Statistics
Chalmers University of Technology and Göteborg University
SE-412 96 Göteborg, Sweden

Göteborg, January 2007

Preprint 2007:1
ISSN 1652-9715

Matematiska vetenskaper
Göteborg 2007

On the Cost-Effectiveness of Sampling Contaminated Soil: Comparison of Two Models

Pär-Erik Back*, Tommy Norberg† and Lars Rosén*

Summary

Remediation of contaminated soil is associated with substantial uncertainties and economic risks. Unfortunately, the extent of soil sampling is often determined without openly addressing these uncertainties and risks, leading to non-optimal use of available economic resources. Data worth analysis can be used to design soil sampling programs in order to manage these problems. Two different data worth models are applied at a site divided into remediation units and five different management alternatives are evaluated: (1) no action, (2) complete remediation without sampling, (3) to base the remediation decision solely on prior information, (4) to base the remediation on a pre-specified number of samples (traditional approach), and (5) to base the decision on the optimal number of samples calculated by the two models (optimal approach). The latter approach was found to be the most cost-effective alternative. The two models give similar results on site level but there may be differences for individual remediation units. It was found that the cost of failure, i.e. the cost associated with the wrong decision, has a strong impact on the result. This illustrates the importance the decision-maker's perspective has on the sampling design. The optimal sampling approach is more effective with a societal perspective compared to if a project perspective is applied. In the traditional approach there appears to be an implicit valuation of failure corresponding to a monetary cost of 1.5-2 times the remediation cost. If failure was valued higher, more samples would be required in traditional sampling.

Keywords

Contaminated soil, data worth, value of information, sampling, soil, economic valuation

* Department of Civil and Environmental Engineering, Chalmers University of Technology.
Address: Hugo Grauers gata 8, SE-412 96 Göteborg, Sweden.
Tel. +46 31 772 20 39
Fax. +46 31 772 19 42
E-mail: par-erik.back@geoinnova.se ; par-erik.back@geo.chalmers.se

† Department of Mathematical Sciences, Chalmers University of Technology and Göteborg University, Göteborg, Sweden.

Introduction

The Swedish Environmental Protection Agency (2005) estimates that there are approximately 40 000 contaminated sites in Sweden. The accumulated governmental cost for clean-up is so far approximately 2.5 billion SEK (1€ ~ 9.50 SEK), but the expected cost for remediation of the 1500 most contaminated sites is an additional 45 billion SEK. A major incentive for environmental clean-up is sustainable development. Sustainable development has three major dimensions; socio-cultural, ecological, and economical (Brundtland Commission, 1987). From a sustainability perspective, the large number of sites and the high costs make it necessary to consider the economic values, or benefits, of remediation in the management of contaminated sites. In addition, for Sweden economic valuation of environmental restoration is supported through the regulatory framework of the Swedish Environmental Code.

Due to complex hydrogeological and geochemical conditions at contaminated sites, it is usually not possible to obtain complete information and to characterise a site with a high degree of certainty. Investigations of contaminated areas are therefore typically associated with large uncertainties regarding *e.g.* type and extent of contamination and possible future contaminant spreading. These uncertainties transform into substantial economic risks associated with the remediation. In order to achieve cost-effective remediation, *i.e.* to select remediation alternatives that reach set environmental goals to the lowest total costs, these risks must be properly accounted for. Unfortunately, decisions regarding type and extent of investigations are often made without completely or openly addressing these uncertainties and risks. The implication of this is that the economical dimension of sustainable development is not fully recognized in the investigation phase, resulting in non-optimal investigations and consequently also non-optimal remedial designs. In order to comply with the goal of sustainable development there is a need for reliable methods for incorporating economic valuations in the investigation phase of remediation projects. Ignoring the economic dimension leads to two important types of mistakes: (1) remediation of sites or parts of sites where no remediation should have been made, and (2) no remediation where remediation should have been made. Both mistakes lead to non-optimal use of available economic resources; in the first case by spending money that could have been better used elsewhere or for other purposes, and in the second case by leaving an unnecessarily large residual risk to humans and ecological systems. The latter can also be translated into economic values due to health problems and lost ecosystem services.

Freeze et al. (1992), James et al. (1996), Back (2003), and Norberg and Rosén (in press) used decision analysis in the form of data worth analysis (often also called value of information analysis) for identifying cost-effective investigation strategies. By including the costs of the investigation program, the precision of the sampling technique, the contaminant spreading conditions, and the economical consequences of taking an erroneous decision at the site, the cost-effectiveness of a specific proposed sampling program can be evaluated. Jardine et al. (1996) used a decision analysis approach as described by Freeze et al. (1990) to the design of a performance monitoring network at a waste management facility overlying fractured bedrock. Barnes and McWhorter (2000) proposed a risk-cost approach for the design of soil vapour extraction system that takes into account uncertainties in the soil properties. The proposed system design approach is similar to the methodology first presented by Massmann and Freeze (1987) and expanded upon by Freeze et al. (1990). Angulo and Tang (1999) used the approach to design the most optimal groundwater detection monitoring system under uncertainty. The various methods for evaluating the value of sampling information use somewhat different models for the mathematical statistical calculations as well as for economic valuations. To further evaluate the use of these methods there is a need to compare them when applied to the same problem in order to assess the model uncertainties. Such studies are inherently cumbersome and therefore rarely done.

The purpose of this paper is to compare two different models described by Back (2003) and Norberg & Rosén (in press), denoted Model A and Model B respectively. The models have several similarities, but differ in some respect in the mathematical statistical calculations. The objective of both models is to estimate the data worth of sampling for classification of soil during the remediation phase of a contaminated site. Evaluation of the cost-effectiveness of sampling programs at earlier phases of remediation projects will require different models; see Nathanail (1998) and Back (submitted). Initially we give a general description of data worth analysis and then proceed into theoretical descriptions of the two models. Then we describe applications of the two models to the same problem with identical data, and evaluate the differences and similarities in both results and theory. We finalize this paper by giving conclusions considered important to further development and applications of data worth analysis at contaminated sites.

Methods

The problem

At the remediation phase, the contaminated site is divided into remediation units, RUs. A number of increments are collected in each RU, all located randomly within the RU (increments are portions of material collected with the only purpose to form a larger sample, composite sample, representing a larger volume of material). The increments are mixed to form a composite sample, which is sent to a laboratory for chemical analyses. The measured concentration is then compared to an action level (AL) and the RU is classified as either contaminated or not contaminated. Each RU is remediated completely or not at all. Failure is said to occur if the RU is not remediated when it in reality is contaminated, resulting in a failure cost. The questions to be answered are; what extent of sampling is required from a cost perspective, and what is the value of the sampling activity. Note that all costs in the presentation are expected costs.

The value of data

For a general and thorough mathematical discussion of data worth in one RU; see Norberg and Rosén (in press). Here, the simplifying assumption that failure F is certain if the RU under consideration is contaminated (denoted C) and, conversely, that failure cannot occur if the RU is not contaminated. A further assumption is that failure cannot occur in a remediated RU.

By definition, the worth W_P of perfect data is the pre-posterior mean of Φ_0 - Φ_{1P}, where Φ_0 is the cost of the best prior alternative (*i.e.* the alternative having lowest expected cost), which is either remediate R or do nothing R', and Φ_{1P} is the cost of the best posterior alternative presuming knowledge of whether the RU is contaminated or not. That is,

$$\Phi_0 = \min(P(C)k_F, k_R)$$

$$\Phi_{1P} = \begin{cases} k_R & \text{on } C \\ 0 & \text{otherwise} \end{cases}$$

hence $E\Phi_{1P} = P(C)\, k_R$, and

$$W_P = \min(P(C)k_F, k_R) - P(C)k_R = \begin{cases} P(C)(k_F - k_R) & \text{if } P(C)k_F < k_R \\ (1-P(C))k_R & \text{if } P(C)k_F \geq k_R \end{cases}$$

(k_R and k_F denotes the remediation and failure cost, respectively). It is implicitly assumed that $k_F \geq k_R$, so $W_P \geq 0$.

Perfect data have no probabilities of error. Real data have. That is, typically the probabilities of an error of the first kind, $D|C'$, and of an error of the second kind, $D'|C$, are non-zero (D denotes the event that contamination is detected). It is assumed that contamination C and detection D are positively dependent, *i.e.*

$$P(C) \leq P(C|D)$$

This implies $P(C|D') \leq P(C)$. A necessary and sufficient condition for positive dependence is

$$P(D|C') + P(D'|C) \leq 1$$

Denote by W the worth of data under non-zero error probabilities. By definition, W is the pre-posterior mean of Φ_0 - Φ_1, where

$$\Phi_1 = \begin{cases} \min P(C|D)k_F, k_R) & \text{on } D \\ \min P(C|D')k_F, k_R) & \text{on } D' \end{cases}$$

It is not hard to see that

$$W = \begin{cases} 0 & \text{if } k_R \geq P(C|D)k_F \\ (P(C|D)k_F - k_R)P(D) & \text{if } P(C)k_F \leq k_R \leq P(C|D)k_F \\ (k_R - P(C|D')k_F)P(D') & \text{if } P(C|D')k_F \leq k_R \leq P(C)k_F \\ 0 & \text{if } k_R \leq P(C|D')k_F \end{cases} \tag{1}$$

This result also follows by straightforward manipulations from the general formula of Norberg and Rosén (in press).

A measurement may consist of several soil samples or increments taken at various locations in the RU. Clearly, it is in the long run worth performing the measurements if their worth W is bigger than their cost k_M, *i.e.* $W \geq k_M$. In conclusion we notice that the value W of new data depends on the probabilities $P(C)$, $P(D|C')$, $P(D'|C)$ and the costs k_R and k_F, and that the accompanying decision as to whether actually carry out the measurement in addition depends on its cost k_M.

Direct and indirect costs

Optimal sampling

If the worth W of new data is less than k_M, measurements are not cost effective and the optimal, *i.e.* most cost effective in the long run, decision is R if $P(C)\,k_F \geq k_R$ and R' otherwise. In the former case, the contribution of the RU to the project costs is direct and amounts to k_R (the cost will directly impact the project budget). In the latter case, the contribution is indirect and amounts to $P(C)\,k_F$ (this risk cost is sometimes referred to as *consumer's risk* (Gilbert, 1987)). From a project perspective, the direct costs are by economists referred to as internal costs, whereas the indirect costs are referred to as external.

Measurements are cost effective in the long run if $W \geq k_M$, *i.e.* the expected net value of the investigation, $W - k_M$, is positive. This is the case if new data have potential to change the best prior alternative. In this case, an RUs contribution to the project costs consists of k_M and either $P(C|D')\,k_F$ (if D' occurs) or k_R (if D occurs). The pre-posterior mean contribution is

$$k_M + k_R\,P(D) + \; P(C|D')\,k_F\,P(D')$$

The sum of the first two terms here will be regarded as a direct cost, and the remaining as an indirect cost. The total project cost is the sum of direct and indirect costs, and sampling costs.

The situation described here will be referred to as Optimal Sampling (OS), since it is most cost effective in the long run.

Traditional sampling

The result for OS will be compared to a situation in which sampling is performed independently of whether it is cost effective or not. With this type of "Traditional Sampling" (TS), an RUs contribution to the project costs consists of the measurement cost k_M and either $P(C|D')\,k_F$ (if D' occurs) or k_R (if D occurs). As in the foregoing section the pre-posterior mean contribution will be divided into a direct cost $k_M + k_R\,P(D)$ and an indirect cost $P(C|D')\,k_F\,P(D')$.

Prior knowledge

The results for OS will furthermore be compared to a situation in which the decision as to remediate or not depends only on the Prior Knowledge (PK). Here RUs contribution to the project costs is indirect and equals $P(C)\,k_F$ if this quantity is less than k_R. Otherwise, it is direct and equals k_R.

Total remediation and no action

The result for OS will also be compared to a situation in which remediation is always performed, to be referred to as Total Remediation (TR), and to a situation in which No Action (NA) is performed. In the former case an RUs contribution to the project costs is direct and equals k_R, while it is indirect and equals $P(C)\,k_R$ in the latter case.

Model A

Model A was presented in Back (2003), where an assumption is that contaminant data are normally distributed, which is usually not the case for contaminated land. Therefore, the model has been slightly modified to handle lognormally distributed data (Back, submitted). A prior probability density function (PDF) is defined based on existing data, available historical information about the site, and expert opinion. The prior PDF represents the expected mean of logtransformed concentrations in the RU, at a defined sampling support (Starks, 1986). Each measurement x_i is associated with uncertainty regarding the true mean concentration of the RU, specified by the coefficient of variation CV, *i.e.* the mean divided by the standard deviation (in untransformed concentration units). The estimated $\bar{y}$ of n logtransformed data y_i (i = 1,..,n) from the RU is assumed to be normally distributed around the population mean μ. The standard deviation σ_y of logtransformed data is calculated as (Strom and Stansbury, 2000)

$$\sigma_y = \sqrt{\ln(CV^2 + 1)}$$

and the corresponding standard error SE_y of the mean of logtransformed data is

$$SE_y = \sqrt{\frac{\ln(CV^2 + 1)}{n}}$$

Now, the distribution of the mean can be expressed as

$$\bar{y} \sim N\left[\mu\,;\,\frac{\ln(CV^2 + 1)}{n}\right] \tag{2}$$

Eq. 2 includes spatial variability within the RU and random measurement errors (potential bias is ignored). The probability of $\bar{y}$ falling below a logtransformed action level μ_{AL} is a function of the true logtransformed concentration, $p_{\bar{y}<\mu_{\mathrm{AL}}}(\mu)$. Now, it is possible to calculate the probability of falsely classifying the RU as uncontaminated when in fact it is contaminated,

$$P(D'|C) = \int_{\mu_{\mathrm{AL}}}^{\infty} p_{\bar{y}<\mu_{\mathrm{AL}}}(\mu) \cdot \frac{f_{\mathrm{prior}}(\mu)}{P(C)}\,d\mu \tag{3}$$

where D represents detection, *i.e.* the estimated mean exceeds AL, C represents the true state (contaminated), $f_{\text{prior}}(\mu)$ is the prior PDF, and P(C) is the prior probability that the RU is contaminated. The latter is estimated from the prior PDF.

When Model A is applied to multiple RUs the prior PDFs are defined based on block kriging of logtransformed data. Previous measurements from earlier phases of the remediation project are utilised to construct a variogram model of the spatial correlation structure at the site. The expected mean in each RU is estimated by block kriging and the corresponding kriging standard deviation is a measure of the likelihood of different means. Consequently, the prior $f_{\text{prior}}(\mu)$ for each RU is defined by the block kriging mean and the kriging standard deviation.

By applying Bayes´ theorem, the pre-posterior failure probability $P(C|D')$ can be calculated:

$$P(C|D') = \frac{P(D'|C)\cdot P(C)}{P(D'|C)\cdot P(C) + P(D'|C')\cdot P(C')}$$

The pre-posterior probabilities for the other combinations of D, D', C and C' can be calculated in a similar manner. The data worth W is calculated in accordance with the presentation in previous sections. The optimal number of samples n_{opt} is the number with the largest positive difference between W and sampling cost.

Model B

Model B is a further development of the one that was presented in Norberg and Rosén (in press). A measurement consists of $n > 0$ independent soil samples taken at purely random locations in the RU. Denote by x_i the concentration of the ith sample and assume that log concentrations, $y_i = \log x_i$, vary according to a normal distribution with unknown mean μ and known standard deviation σ. The variation in the sample mean $\bar{y} = \frac{1}{n}\sum_i y_i$ is normal with mean μ and standard deviation $\sigma/\sqrt{n}$. Let μ_{AL} be a threshold, such that any RU should be considered contaminated if $\mu > \mu_{\text{AL}}$ and not contaminated otherwise. Typically, μ_{AL} is the logarithm of a site specific action level AL. Further, c be a threshold, such that contamination is considered detected if the mean $\bar{y} \geq c$. Then, as is shown in any basic statistics text, *e.g.* Rice (1995),

$$P(D|C') = 1 - \Phi\left(\frac{c-\mu}{\sigma/\sqrt{\text{n}}}\right) \leq 1 - \Phi\left(\frac{c-\mu_{\text{AL}}}{\sigma/\sqrt{\text{n}}}\right) \tag{4}$$

(Φ denotes the standard normal distribution function). The right hand side is the level of significance of the test of the null hypothesis H_0 : $\mu \leq \mu_{\text{AL}}$ *vs* the alternative H_1 : $\mu > \mu_{\text{AL}}$. Also,

$$P(D'|C) = \Phi\left(\frac{c-\mu}{\sigma/\sqrt{\text{n}}}\right) \leq \Phi\left(\frac{c-\mu_{\text{AL}}}{\sigma/\sqrt{\text{n}}}\right) \tag{5}$$

The sum of the bounds in equations (4) and (5) is 1. Thus if we use these bounding probabilities as our probabilities for an error of the first and second type, then C and D becomes independent and, necessarily, $W = 0$.

A way to circumvent this is to let C be the event $\mu \geq \mu_{\mathrm{AL}} + k_1\sigma$, while C' is the event $\mu \leq \mu_{\mathrm{AL}} - k_2\sigma$, for a suitable choice of $k_1, k_2 \geq 0$. We thus pretend that $\mu_{\mathrm{AL}} - k_2\sigma < \mu < \mu_{\mathrm{AL}} + k_1\sigma$ cannot occur. Then

$$P(D|C') \leq 1 - \Phi\left((c' + k_2)\sqrt{n}\right) \tag{6}$$

and

$$P(D'|C) \leq \Phi\left((c' - k_1)\sqrt{n}\right) \tag{7}$$

where $c' = (c - \mu_{\mathrm{AL}})/\sigma$ is the standardized threshold. The bounds in equations (6) and (7) will be used in the calculations of the data worth W. Since data worth depends monotonically on the two error probabilities (Norberg and Rosén, in press) the true worth is not smaller than the one we calculate.

The standard deviation σ is considered known above. Nevertheless it must be estimated. This presumes that some samples already have been taken and analysed. In this known set of data, we look for clusters of samples located close to each another with the property that the corresponding log concentrations could be assumed to vary normally around an unknown mean with the constant standard deviation σ. In each such cluster the sample variance is calculated and if there are more than one cluster, the calculated variances are pooled into one grand estimate s^2 of σ^2 having, say, v degrees of freedom.

The proper way to take into account the extra uncertainty when using the estimate s of σ instead of σ in formulae (6) and (7) for the probabilities of the two types of error, is to replace the normal cdf Φ with its more heavy-tailed Student t(v)-counterpart, denoted Φ_v. Thus,

$$P(D|C') = 1 - \Phi_v\left((c' + k_2)\sqrt{n}\right) \tag{8}$$

$$P(D'|C) = \Phi_v\left((c' - k_1)\sqrt{n}\right) \tag{9}$$

In order to specify a value for the prior probability of contamination, $P(C)$, and the two k's we need a guess z of the log concentration in the RU. This could be either the logarithm of the concentration in a known measurement or an interpolation of known values from nearby RUs. Let $z' = (z - \mu_{\mathrm{AL}})/\sigma$ be the corresponding standardized value. We then let

$$P(C) = \Phi_v(z')$$

The rationale for defining $P(C)$ in this way, is that $\Phi_v(z')$ is the confidence in the statement $\mu \geq \mu_{\mathrm{AL}}$.

Further, if $z' > 1$, we let $k_1 = z'$ and $k_2 = 0$ and if $0 \le z' \le 1$, we let $k_1 = z'$ and $k_2 = 1 - z'$. Symmetrically, if $z' < 1$, we instead let $k_2 = -z'$ and if $1 \le z' < 0$, we let $k_2 = -z'$ and $k_1 = 1 + z'$. That is, in principle, we discriminate between the two levels z and μ_{AL}. This does not work well if z is too close to μ_{AL}. We therefore introduced a bound in such a way that the distance between the two levels that we discriminate between is at 1.

Finally, we need to specify the critical level c above which contamination should be considered detected. For this, notice first that the data worth W depends continuously on the two error probabilities $P(D|C')$ and $P(D'|C)$, see (1), and that these in turn depend continuously on c, or rather the corresponding standardized level c', *cf* (6) and (7). Thus, W is a function of c'. The value chosen for c' is the value that maximizes W.

For each sample size n, we now may calculate the corresponding data worth W and the optimal sample size n_{opt} is the n that maximizes the difference of the data worth W and the measurement cost k_M, provided that this difference is non-negative. Otherwise n_{opt} is set to 0.

Model differences

The two models for calculating the sample size for which $W(n) - k_M(n)$ is maximal are clearly different, although both are based on the general theory for the value of data described above. The two models differ at least in the following aspects:

- The various ways in which $P(C)$ are calculated;
- The optimization of the critical threshold c in Model B has no counterpart in Model A;
- The various ways in which the error probabilities $P(D|C')$ and $P(D'|C)$ are calculated.

Case study

The Wockatz scrapyard

The two models are applied at the Wockatz scrapyard (Figure 1) situated along the river Göta älv in the central parts of Göteborg, the second largest city of Sweden. The scrapyard was operated between the 1930s and 1993. The soil consists of diverse filling material on top of a thick layer of glaciomarine clay. The thickness of the filling is about 1 m. Previous investigations at the site indicate substantial contamination of heavy metals and oil products. Here, the analysis will be based on cadmium. The adjacent river has a high protection value due to both the ecosystems in the river and in the river estuary. In addition, it serves as municipal water supply for about 700,000 inhabitants.

The landuse scenario for the site is residential with a mix of apartments and small offices. A site-specific guideline value for cadmium has been calculated to 2 mg/kg, based on the risk assessment model of the Swedish EPA (1997). This is the action level (AL) for the site. The site-specific guideline value is based on the assumption that groundwater is not used for consumption but that some level of protection of the groundwater is required because of the potential leaching of contaminants to the river.

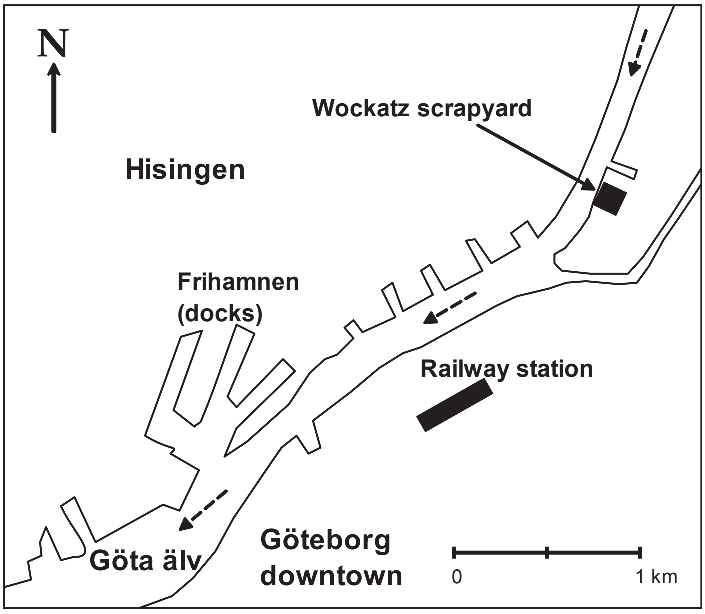

Figure 1. The Wockatz scrapyard in Göteborg, Sweden.

Management alternatives for the Wockatz scrapyard

In the remediation phase, the site is divided into 225 RUs, each of size 100 m^2 with a thickness of 1 m. Each RU is either remediated as a whole, or not at all. The remediation technique is excavation and deposition of the soil at a landfill.

Five different management alternatives for the Wockatz site are evaluated. The first one is *No Action*, *i.e.* no sampling and no remediation is performed. The second alternative is *Total Remediation* of all RUs without performing any sampling at all. As a third alternative, the decision to remediate an RU or not is based solely on *Prior Knowledge* about contaminant concentrations without performing any sampling. In this alternative, the most cost-effective decision is the one that minimises the total cost of the project.

The fourth alternative is a strategy that is frequently used in Sweden, here denoted as *Traditional Sampling*. This strategy implies taking four to five increments in each RU and mixing them to create a composite sample, which is analysed at a laboratory. The extent of sampling is thus the same in all RUs. The strategy is recommended by the Swedish EPA (1998) in situations when a low ambition is sufficient. A strategy with higher ambition is also presented by the Swedish EPA, with 4×5 increments collected in each RU, but this strategy is rarely applied due to the higher sampling costs. Finally, the *Optimal Sampling* alternative is evaluated. This approach implies collecting the number of increments that maximises the expected net value of the sampling program.

Remediation cost and sampling cost

Strategies two to five will result in remediation of at least some RUs and this cost is estimated to 200 kSEK per RU, including excavation, transportation, disposal at a landfill, and backfilling of the RU. The cost of field sampling is 20 kSEK per day and it is assumed that 20 increments can be collected in one day, including formation of composite samples. The number of increments that can be collected in a day is limited by the geological conditions (the diverse filling material that can be difficult to penetrate). The cost for laboratory analyses of a composite sample is 1 kSEK.

Failure cost

Failure will result in consequences that can be expressed as a monetary cost, *i.e.* the failure cost. This represents the additional cost that may evolve if and when the contamination is detected in the future (additional clean-up costs). It may also include potential long-term negative consequences for humans and ecological systems, due to exposure from undetected contamination in the ground. The failure cost must be estimated by careful consideration of the loss of value due to residual health and ecological risk. Valuations are made by using environmental economical valuation methods; see *e.g.* U.S. EPA (2000) or EFTEC (2006). It is evident that the latter is a very uncertain cost. For simplicity reasons we have in this paper normalised the failure cost by the remediation cost, using the variable k_F / k_R. We believe that also for practical purposes this normalisation may be useful. Detailed economic valuations are complicated and time-consuming, whereas reasonable estimations of the relative values of remediation cost and failure cost should be possible and sufficient in many cases, based on experience and detailed valuations made in similar situations.

Total cost

The expected total cost for each management alternative is calculated as the sum of expected sampling costs, expected remediation costs and expected failure costs. The remediation cost is a direct cost that a project manager will be facing (internal cost), whereas the failure cost is indirect (external; consumer's risk). The external costs do not impact the project budget for projects with a short timeframe.

Some costs that occur on site level are ignored in the application. These include costs for establishment of sampling equipment at the site, costs for evaluating the sampling data, and costs for evaluation of prior information. These costs are so low compared to remediation costs and failure costs that omitting them makes no difference for the conclusions.

Results

The results for the Wockatz site application are illustrated in Figures 2-9. The expected total costs for the five management alternatives are illustrated in Figure 2. The general pattern is similar for Model A and Model B, although Model B results in slightly higher cost for all alternatives except Total Remediation and Optimal Sampling. Notice however that the total expected cost when Optimal Sampling is performed is much lower for Model B than for Model A. This effect becomes more pronounced as

k_F / k_R gets larger. The main reason is that the critical threshold c in Model B is adapted to the failure and remediation costs. This makes the data value higher in Model B than in Model A, *cf* Figure 4.

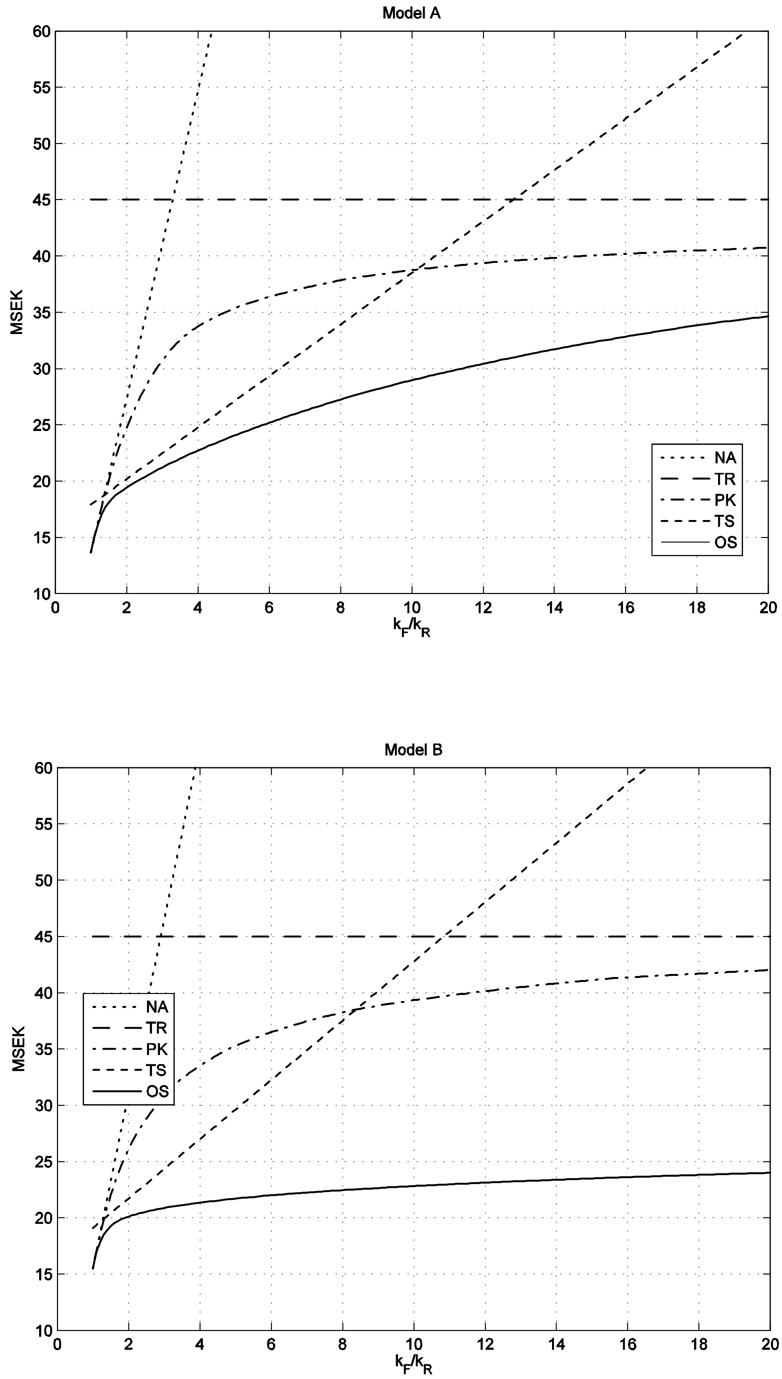

Figure 2. Expected total cost based on Model A and Model B for the five management approaches No Action (NA), Total Remediation (TR), Prior Knowledge (PK), Traditional Sampling (TS), and Optimal Sampling (OS).

The Optimal Sampling approach is by far the most cost-effective approach, at least for large values of k_F / k_R. Traditional Sampling results in slightly higher expected total cost than optimal sampling when k_F / k_R is about 1.5 to 2. At higher failure cost Traditional Sampling becomes much more costly. In fact, when k_F / k_R approaches 8 (Model B) and 10 (Model A) it is more cost-effective to base the remediation decision on prior information alone. Even for very low values of k_F / k_R (close to 1) prior knowledge may be more cost-effective to use than traditional sampling. The same applies to the No Action alternative but increasing k_F / k_R results in a rapid increase of the expected total cost for this approach. The alternative Total Remediation of the Wockatz site is, as anticipated, not a cost-effective alternative.

The cost reduction of applying Optimal Sampling instead of Traditional Sampling is illustrated in Figure 3. The curves for both models have minima at relatively low values of k_F / k_R; Model A at 1.8 and Model B at 1.5. This implies that Traditional Sampling is almost as effective as Optimal Sampling at low failure costs. In other words, if leaving contamination in the ground is not regarded as a large problem, then traditional sampling is sufficiently effective. On the other hand, if large human health or environmental values are lost when contaminated soil is not remediated, then Optimal Sampling is a much better alternative. The failure cost at the minima in Figure 3 is therefore an indication of the implicit valuation of failure in Traditional Sampling.

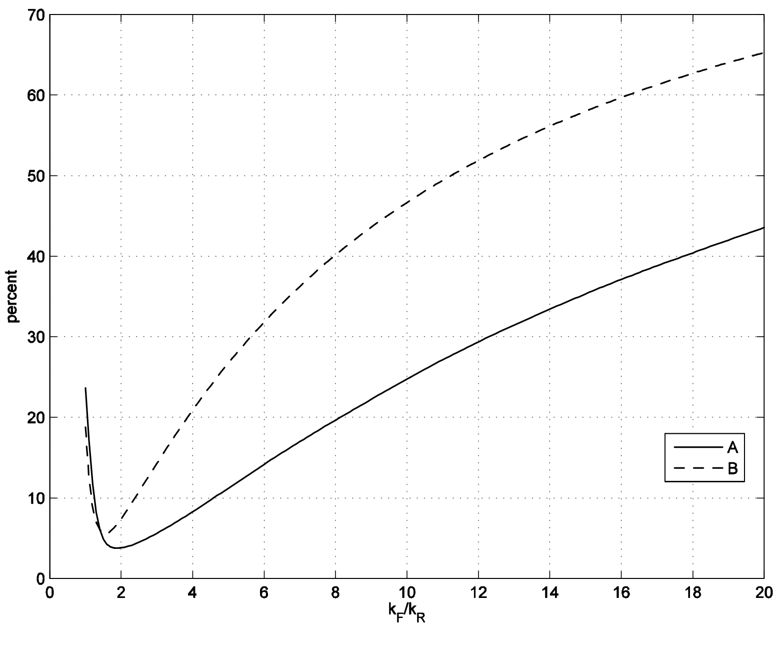

Figure 3. The expected project cost reduction of applying Optimal Sampling instead of Traditional Sampling for Model A and B.

The data worth of sampling is illustrated in Figure 4. The curve W_P represents the data worth of a perfect sampling exercise that eliminates all uncertainty. The maximum possible value of W_P is denoted lim W_P and is approached as k_F / k_R increases. The data worth for optimal sampling, W (n_{opt}), increases when k_F / k_R increases, up to a maxima followed by a decline. This decline occurs earlier for Model A than for Model B, since in the latter model the critical threshold c adapts itself to the failure and remediation costs. For comparison, the data worth of collecting five increments (Traditional Sampling) is indicated. The data worth is larger than the sampling cost, which indicates that the sampling exercise is cost-effective (it supplies more value than its cost).

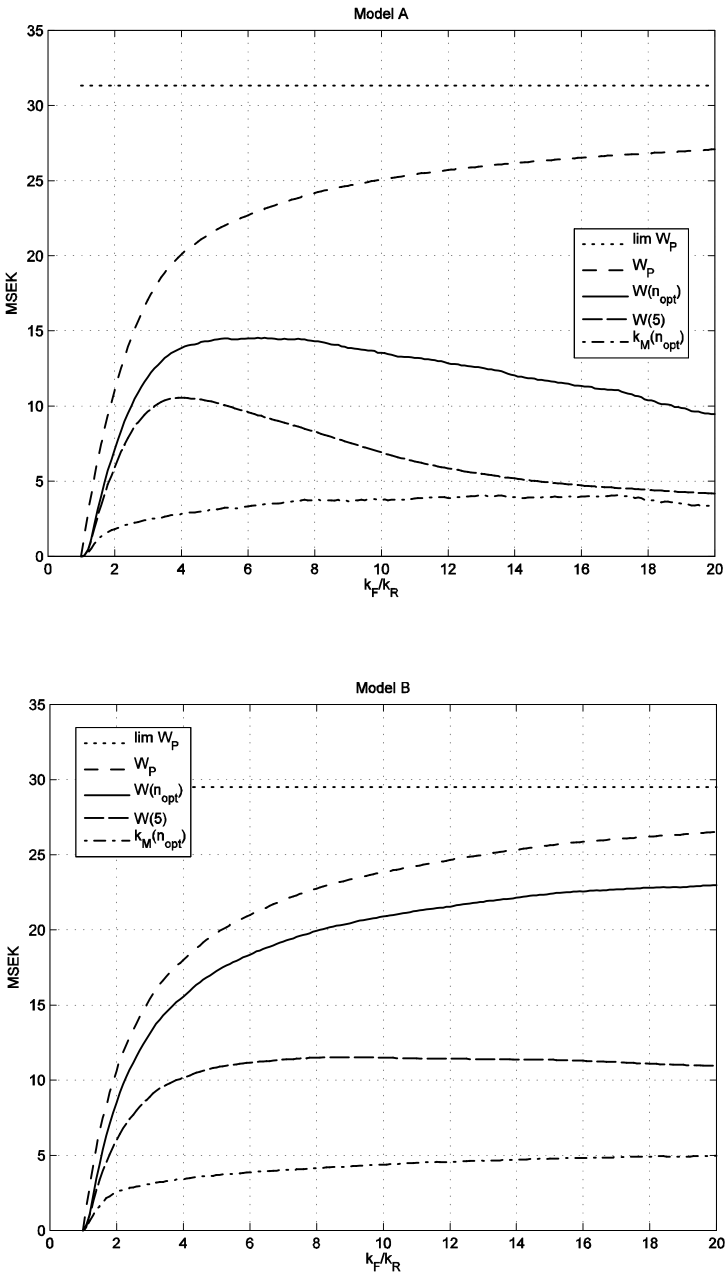

Figure 4. Data worth of sampling for Model A and Model B. The curve W_p represents perfect sampling, lim W_p the largest possible data worth, $W(n_{opt})$ the data worth for Optimal Sampling, $W(5)$ the data worth of collecting five samples in each RU (i.e. Traditional Sampling), and $k_M(n_{opt})$ the total sampling cost of collecting the optimal number of samples in each RU.

Figure 5 presents the direct and indirect costs for Optimal and Traditional sampling. The general behaviour is similar for the two models but not identical. For example, the indirect cost is lower according to Model B.

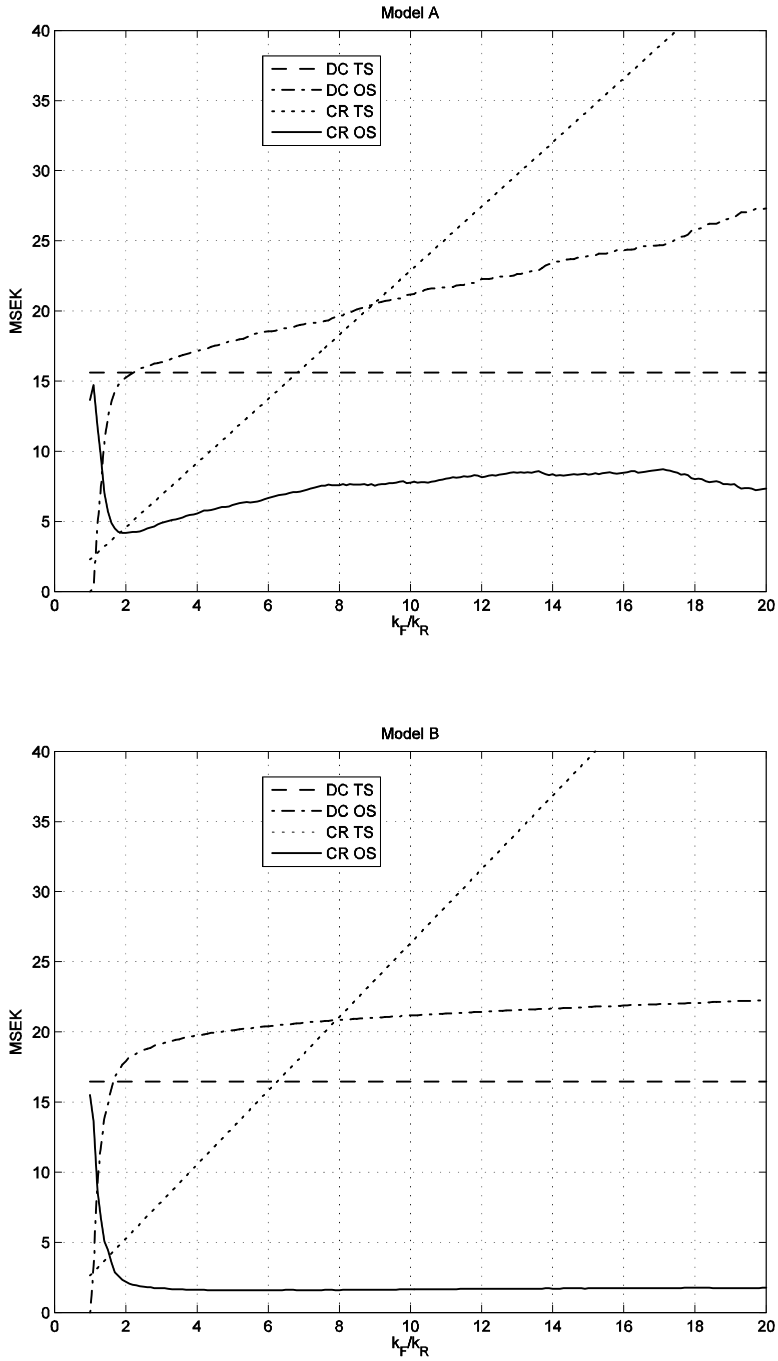

Figure 5. Direct costs (DC) and consumer's risk (CR) for Optimal Sampling (OS) and Traditional Sampling (TS) for Model A and Model B. Sampling costs are not included in DC.

Maps of the RUs at the Wockatz site are presented in Figure 6-8 for Optimal Sampling and Traditional Sampling. Black RUs indicate that the best option is to remediate directly without performing any sampling. This is because prior information about these RUs is so strong that sampling cost cannot balance the risk of leaving contamination (Back and Rosén, 2006).

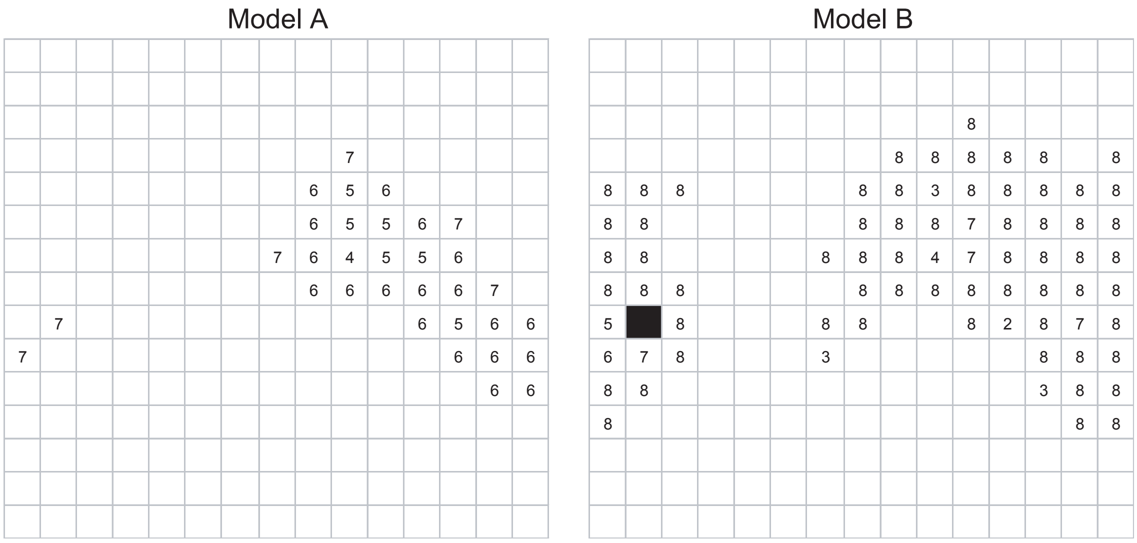

Figure 6. Best action alternative in each of the 225 RUs according to Model A and Model B when $k_F / k_R = 1.2$. Empty cells = no action, black cells = remediation, and numbered cells = sampling; optimal number of increments.

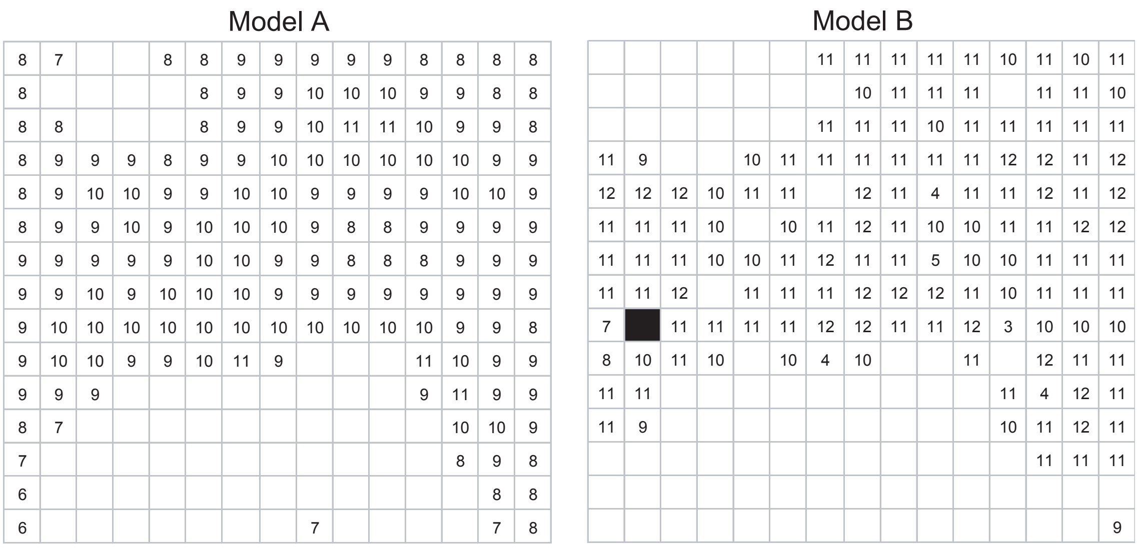

Figure 7. Best action alternative in each of the 225 RUs when the difference in expected total cost between Traditional Sampling and Optimal Sampling is at a minimum. This occurs at $k_F / k_R = 1.8$ for Model A and at $k_F / k_R = 1.5$ for Model B.

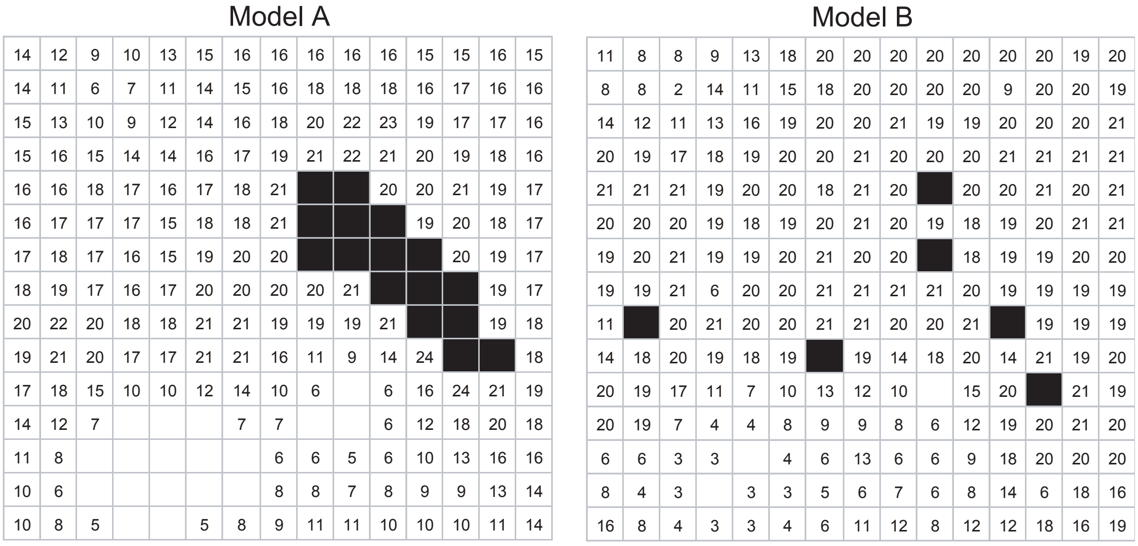

Figure 8. Best action alternative in each of the 225 RUs according to Model A and Model B when $k_F / k_R = 5$.

Slightly less sampling is required with Model A compared to Model B, most obvious at very low k_F / k_R -values (Figure 6), but also for larger values. At the k_F / k_R -value where Traditional Sampling best resembles Optimal Sampling (Figure 7), the general picture is similar for the two models, although the best option for individual RUs can be quite different. The cluster of black RUs for Model A in Figure 9 is probably a result of the smoothing effect accomplished by block kriging when prior information is defined. Model B exhibits a slightly larger variability in n_{opt} than do Model A; standard deviation of n_{opt} is 5.7 samples compared to 4.6 samples for Model A (RUs with $n_{opt} = 0$ not included). It is also evident that collecting only five increments in all RUs is far from optimal when failure cost is high.

Conclusions

The main conclusions of this study are:

- Two different models for estimating the expected monetary value of sampling was compared. The sampling objective was estimation of the mean concentration in remediation units (RUs) for classification of contaminated soil in the remediation phase of contaminated land projects. Both models can be applied for estimating the optimal number of increments in an RU.
- The basic philosophy behind both models are the same but, based on existing information, the two models use different ways to estimate the parameters that control the data worth. The main reason for the different results of Models A and B is the fact that in Model B there is a critical threshold c that adapts itself to the failure and remediation costs. Different methods of calculating prior probabilities and failure probabilities also contribute to the difference.

- Optimal Sampling, *i.e.* maximising the expected net value of the sampling, was found to be the most cost-effective approach of the five evaluated ones, as anticipated. Its cost-effectiveness, in relation to other approaches, is higher at high failure costs.
- In the Traditional sampling approach there appears to be an implicit valuation of failure, corresponding to a monetary cost of about 1.5 to 2 times the remediation cost. This could be an indication of how seriously society considers the issue of leaving contaminated soil untreated, for this and similar contaminated sites.

The findings in this paper have general implications on soil sampling at the remediation phase concerning direct (internal) and indirect (external) costs and the perspective of the decision-maker. When k_F is large, the direct cost for Optimal Sampling is higher than for Traditional Sampling, according to Figure 5. This implies that a decision maker that is only concerned with the direct cost of a remediation project would prefer the Traditional Sampling approach, because it results in the lowest costs that directly impact the project budget. On the other hand, from a societal point of view the Optimal Sampling approach will result in the lowest costs in the long run because the sum of direct and indirect costs are much lower than for the Traditional Sampling approach. This illustrates the importance the decision-maker's perspective has on the sampling design. Assuming that the environmental authority has a societal perspective and that the developer of a site has a more restricted perspective, one would expect that they would arrive at quite different sampling designs.

At values of k_F / k_R close to 1 the direct cost of Optimal Sampling is lower than for Traditional Sampling because several RUs do not require sampling with the optimal approach, which reduces sampling cost. In situations where the correlation distance is short, less than the length of an RU, a different situation can be expected. Under such conditions, more RUs will require sampling (less empty cells in Figure 6-8) because information about contaminant concentrations in one RU cannot be utilised in another cells to the same extent. Consequently, the optimal and traditional sampling approaches will show more similarities at low values of k_F. This is believed to be a typical situation for many sites contaminated by diverse filling material with short correlation distance.

A number of simplifications must of course be made when sampling models are developed, and the models in this paper are no exception. One limitation in the models is that the total expected cost for Optimal Sampling is somewhat overestimated because in reality there is a limit of the number of increment that can be economically collected in an RU. Above this limit it is less expensive to excavate the whole RU, homogenise the soil, and finally collect a representative sample that is sent to the laboratory. The overestimation implies that the difference between optimal sampling and traditional sampling in reality is even more pronounced than Figure 2 indicates, especially for large values of k_F / k_R when n_{opt} is large.

A common sampling problem during the remediation phase concerns classification of RUs in several classes, depending on contamination level and waste acceptance criteria at potential landfills. This problem can be modelled by applying several different action levels, and different remediation and failure costs. The methodology can also be extended to handle different sampling designs, *e.g.* systematic sampling, systematic random sampling, and stratified sampling. Instead of collecting only one composite sample in each RU, several could be collected, each one formed by a specified number of increments. A range of such adjustment of the methodology could easily be implemented for specific applications.

References

Angulo M, Tang WH. 1999. Optimal Ground-Water Detection Monitoring System Design under Uncertainty. *Journal of Geotechnical and Geoenvironmental Engineering*, **125**(6): 510-517.

Back P-E. 2003. On Uncertainty and Data Worth in Decision Analysis for Contaminated Land. Licentiate of Engineering Thesis, Chalmers University of Technology, Göteborg, Sweden; 129.

Back P-E. Submitted. A Model for Estimating the Value of Sampling Programs and the Optimal Number of Samples for Contaminated Soil. *Environmental Geology*.

Back P-E, Rosén L. 2006. Economic Valuation of Contamination Risks and its Effect on Site Investigations. In *VALDOR 2006 (VALues in Decisions On Risk)*. Andersson K (Editor), Congrex: Stockholm, Sweden; 67-74.

Barnes DL, McWorther DB. 2000. Design of Soil Vapor Extraction Systems Under Conditions of Uncertainty. *Ground Water*, **38**(2): 209-217.

Brundtland Commission. 1987. Report of the World Commission on Environment and Development. A/42/427, United Nations.

EFTEC. 2006. Valuing Our Natural Environment. NR0103, Department for Environment, Food and Rural Affairs, London.

Freeze RA, Bruce J, Massman J, Sperling T, Smith L. 1992. Hydrogeological Decision Analysis: 4. The Concept of Data Worth and Its Use in the Development of Site Investigation Strategies. *Ground Water*, **30**(4): 574-588.

Freeze RA, Massman J, Smith L, Sperling T, James B. 1990. Hydrogeological Decision Analysis: 1. A Framework. *Ground Water*, **28**(5): 738-766.

Gilbert RO, 1987. Statistical Methods for Environmental Pollution Monitoring. Van Nostrand Reinhold: New York; 320.

James BR, Huff DD, Trabalka JR, Ketelle RH, Rightmire CT. 1996. Allocation of Environmental Remediation Funds Using Economic Risk-Cost-Benefit Analysis: A Case Study. *Groundwater Monitoring & Remediation,*(Fall 1996): 95-105.

Jardine K, Smith L, Clemo T. 1996. Monitoring networks in fractured rocks; a decision analysis approach. *Ground Water*, **34**(3): 504-518.

Massman J, Freeze RA. 1987. Groundwater Contamination from Waste Management Sites: the Interaction Between Risk-Based Engineering Design and Regulatory Policy; 1, Methodology. *Water Resources Research*, **23**(2): 351-367.

Nathanail CP, Ferguson CC, Tucker P. 1998. The Use of Bayesian Statistics and Prior Information to Design Sampling Strategies. In *Eight International Congress, International Association for Engineering Geology and the Environment*. Moore D and Hungr O (Editors), Balkema, Rotterdam: Vancouver; 369-376.

Norberg T, Rosén L. In press. Calculating the Optimal Number of Contaminant Samples by means of Data Worth Analysis. *Environmetrics*, Accepted 2005.

Rice JA, 1995. Mathematical Statistics and Data Analysis. Duxbury Press.

Starks TH. 1986. Determination of Support in Soil Sampling. *Mathematical Geology*, **18**(6): 529-537.

Strom DJ, Stansbury PS. 2000. Determining Parameters of Lognormal Distributions from Minimal Information. *AIHAJ*, **61**(November/December 2000): 877-880.

Swedish EPA. 1997. Development of Generic Guideline Values. Model and data used for generic guideline values for contaminated soils in Sweden. Report 4639, Stockholm.

Swedish EPA. 1998. Requirements for Site Remediation. Guidelines for practical achievement of acceptable residual concentrations and quantities - methods and quality aspects. Report 4808, Stockholm.

Swedish EPA. 2005. Website http://www.naturvardsverket.se. 14th March 2005.

U.S. EPA. 2000. Guidelines for Preparing Economic Analyses. EPA 240-R-00-003.