Environ Geol (2007) 52:573–585
DOI 10.1007/s00254-006-0488-6

ORIGINAL ARTICLE

A model for estimating the value of sampling programs and the optimal number of samples for contaminated soil

Pär-Erik Back

Received: 20 April 2006 / Accepted: 24 August 2006 / Published online: 19 September 2006

Abstract A model is presented for estimating the value of information of sampling programs for contaminated soil. The purpose is to calculate the optimal number of samples when the objective is to estimate the mean concentration. A Bayesian risk–cost–benefit decision analysis framework is applied and the approach is design-based. The model explicitly includes sample uncertainty at a complexity level that can be applied to practical contaminated land problems with limited amount of data. Prior information about the contamination level is modelled by probability density functions. The value of information is expressed in monetary terms. The most cost-effective sampling program is the one with the highest expected net value. The model was applied to a contaminated scrap yard in Göteborg, Sweden, contaminated by metals. The optimal number of samples was determined to be in the range of 16–18 for a remediation unit of 100 m^2. Sensitivity analysis indicates that the perspective of the decision-maker is important, and that the cost of failure and the future land use are the most important factors to consider. The model can also be applied for other sampling problems, for example, sampling and testing of wastes to meet landfill waste acceptance procedures.

Keywords Contamination · Value of information · Data worth · Soil · Cost-effectiveness · Bayesian analysis

P.-E. Back (✉)
Department of Civil and Environmental Engineering, Chalmers University of Technology, 412 96 Göteborg, Sweden
e-mail: par-erik.back@geoinnova.se

Introduction

Contaminated soil and groundwater problems have received increased attention during the last decades. The human and environmental risks call for investigation of the degree of contamination at a site, and if required, remediation. The results from site investigations are typically associated with large uncertainty due to natural variability and lack of information.

Because of stiff competition on the market the consultant with the cheapest site-investigation often gets the assignment. A consequence is that it will be difficult to distinguish between ''nothing found'' because there was nothing there or ''nothing found'' because of a poor site-investigation (Bosman 1993). The latter may well be regarded as a success by the involved parties but may in fact lead to long-term human health problems or environmental effects. The opposite situation may also exist, i.e. huge quantum of data are collected at high cost based on demands from environmental authorities (James and Freeze 1993; LeGrand and Rosén 2000). Both these situations may thus result from bad data collection strategies.

Three strategies have traditionally been used to determine the size of the sampling effort: (1) to minimise the sampling cost for a specified level of accuracy or precision; (2) to minimise uncertainty for a given sampling budget; or (3) to respond to demands on sampling made by the legal authorities. There are also combinations of these strategies, such as the fitness-for-purpose approach (Thompson and Fearn 1996). The concept of Value of Information Analysis (VOIA), also named Data Worth Analysis (DWA) by some authors, constitutes a fourth alternative. In this approach, the decision to sample, or the design of the sampling program, is based on cost-effectiveness. A sampling program is regarded as cost-effective as long as the expected benefit associated with the new information is larger than the sampling cost. Another way of expressing this is that sampling is only justified if the sampling has potential to change the decision.

During the 1970s a number of applications of VOIA were reported for various water-related problems (Davis and Dvoranchik 1971; Maddock 1973; Gates and Kisiel 1974) and additional work for hydrogeological problems was reported during the 1980s (Grosser and Goodman 1985; Ben-Zvi et al. 1988; Reichard and Evans 1989). In the last 15 years, VOIA has been used for a number of groundwater contamination problems (Freeze et al. 1992; James and Freeze 1993; James and Gorelick 1994; Abbaspour et al. 1996; James et al. 1996a; IT-Corporation 1997; McNulty et al. 1997; Russell and Rabideau 2000).

Applications of VOIA on problems of contaminated particulate materials, such as soil or waste material, are more limited. Dakins et al. (1994, 1996) present a decision framework for remediation of PCB-contaminated sediments, whereas Rautman et al. (1994) apply a risk-based decision analysis approach to compare the reliability of different characterisation techniques for uranium contaminated soil. James et al. (1996b) present a simple risk-based decision analysis framework for remediation of contaminated soil, and Kaplan (1998) describes a software for the purpose of selecting sample locations based on the worth of new data. Back (2003) reviews the literature of VOIA for contaminated land and suggests a model for estimation of the data worth. Norberg and Rosén (2006) present a model for calculating the optimal number of samples for sites with contaminated soil.

In many of the models, sample uncertainty is ignored or is implicitly considered by geostatistical techniques with a complexity level that has restricted the application of VOIA for real-world contamination problems. In this paper a model is presented that explicitly includes sample uncertainty at a complexity level that can be applied to practical contaminated land problems with limited amount of data. All data uncertainties are considered, i.e. natural variability, sampling uncertainty, preparation uncertainty, and analytical uncertainty (Pitard 1993). Soft prior information is considered in the approach by applying Bayesian statistics.

This model is suited for soil sampling problems in the remediation phase of a contaminated land project and was developed with practical application for non-statisticians in mind. The principles of VOIA for sampling have been presented in shorter mathematical notation by Lindley (1997) with emphasis on the statistical aspects. An application of the model is presented for a sampling problem at the Wockatz scrap yard in Göteborg, Sweden, where the soil is contaminated by metals.

Methodology

Decision analysis framework

The VOIA is performed within the risk–cost–benefit decision analysis framework presented by Freeze et al. (1990). Two decision alternatives are considered in the analysis; remediation or no remediation of a remediation unit (RU). The decision problem is illustrated by a decision tree in Fig. 1.

An objective function Φ_i is applied for each decision alternative i (after Freeze et al. 1992):

$$\Phi_i = B_i - C_i - \gamma \cdot P_i \cdot \mathrm{Cf}_i \qquad (1)$$

where B_i is the benefit and C_i is the investment cost. The third term in Eq. 1 is the risk term, where P_i is the probability of failure, Cf_i is the cost of failure, and γ is a risk aversion factor. The risk aversion factor determines how the risk is weighted in the analysis and should be <1 for risk seeking behaviour, equal to 1 for risk neutral behaviour, and >1 for risk aversion. The approach in this paper is risk neutral and hence, $\gamma = 1$.

By collecting more information (sampling), the value of the risk term in Eq. 1 is reduced and thus the expected value of the objective function increases; the decrease in the risk term is solely due to a decrease in the probability of failure (Freeze et al. 1992). The expected increase in Φ_i is the value of the sampling program. Several different decision rules can be found in the literature (Norrman 2001), and the decision rule to maximise the value of the objective function is used in the model. This criterion will lead to the largest benefit, or lowest cost, in the long run.

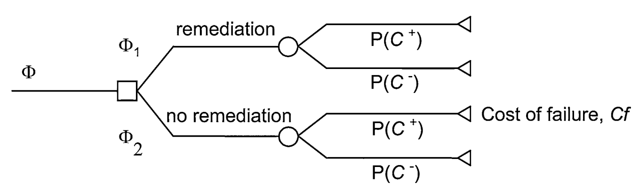

Fig. 1 Example of a decision tree for a contaminated land problem with two decision alternatives, remediation or no remediation, and two states, contaminated (C^+) or not contaminated (C^-) (after Freeze et al. 1992). Costs and benefits occur at the *triangular terminal nodes* (only the cost of failure is indicated)

Procedure of VOIA

The VOIA is carried out in a five-step procedure according to Fig. 2: (1) the sampling program, or a set of sampling programs, is defined; (2) the prior information about the mean concentration of the RU is transformed to a probability density function (PDF); (3) four different types of probabilities are estimated, based on the prior information and the proposed sampling program; (4) all involved costs and benefits are estimated; and (5) the value of information (VOI) and the optimal number of samples are estimated.

Step 1: sampling program

The sampling objective is to estimate the true mean concentration at the site, or a part of the site such as an RU. Arguments why the mean concentration is a suitable measure to be used in risk assessments are given by the US EPA (1992, 1996) and the Swedish EPA (1998). The spatial distribution of the contaminant is unknown but it is assumed that data are lognormally distributed at the sample scale, which is usually the case (US EPA 1992).

Two parameters need to be specified for the sampling program: the number of samples (n), and the sample uncertainty. Here, the sample uncertainty is defined as the uncertainty of a single concentration measurement in relation to the true mean concentration of the RU. The random part is quantified as a coefficient of variation, *CV* (relative standard deviation) and can be interpreted as the expected variability of measurements. It includes spatial variability, sampling errors, preparation errors, and analytical errors.

The systematic uncertainty is denoted ε_s, and is the expected systematic error. Note that the word *expected* is used, indicating that we are dealing with uncertainties prior to sampling; true errors on the other hand can only be estimated after sampling has been performed.

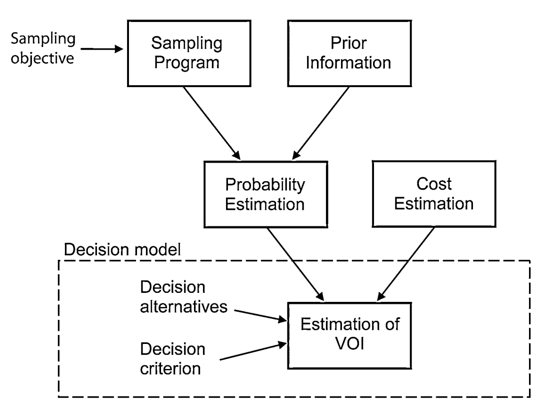

Fig. 2 The procedure for estimating the value of information for a sampling program

A design-based approach to sampling is applied (Brus and de Gruijter 1997). This means that no model of data is used for the inference of the mean concentration. The unknown value at any given location is considered as fixed, not random as in model-based approaches (Brus and de Gruijter 1997). The assumption of lognormally distributed data should not be confused with a model-based approach. This assumption is not the basis for the inference of population parameters, as in a model-based inference; instead the assumption is made in order to include sample uncertainty in a realistic way. Independence is created by locating the samples in a random manner in the RU, in order to derive at an unbiased estimate of the mean.

Contrary to what is commonly believed, samples do not need to be spatially uncorrelated when classical sampling theory is applied for estimation of the mean (Brus and de Gruijter 1997). Instead, independence is achieved when the sample locations are randomly selected. Gregoire (1998) states that spatial correlation is an irrelevant issue in design-based sampling. The misconception that classical approaches are incorrect when spatial correlation is present is widespread, see for example Back (2003), Einax and Kraft (2002), Provost (1984) and references made in Brus and de Gruijter (1997).

Step 2: prior information

Prior information is information that is available before the sampling is carried out. Here, prior information about the mean concentration of the RU is expressed by means of a probability density function (PDF). The PDF should reflect the belief of the likelihood of different mean concentrations of the RU. This PDF should not be confused with the PDF of the expected sample data.

Applying a Bayesian perspective suggests that prior estimation of the mean concentration can be based on soft information if no hard data is available. There are several possible sources of soft information, e.g. experience from similar sites, expert judgement, literature. Approaches for specifying a PDF from uncertain information is discussed by Hoffman and Kaplan (1999).

The prior PDF should reflect the mean concentration μ of logtransformed data. Any PDF can be used but a suitable distribution is the normal distribution, with the two parameters μ_N and σ_N defining the distribution. For practical applications it is easier to define the PDF by the median (m_u) and a reasonable maximum value (b_u) of the expected mean concentration (index u indicates untransformed units):

$$\mu_N = \ln(m_u) \tag{2}$$

σ_N is selected so that b is a specific percentile of the distribution (Fig. 3):

$$P_s = P_N(\mu < b) \tag{3}$$

where P_s is the specified percentile at b [decimal fraction], and P_N represents a probability based on the normal PDF. The proper value of σ_N is found by solving Eq. 3.

Step 3: probability estimation

According to Freeze et al. (1992) there are four types of probabilities to consider in the analysis, denoted as P[state], P[sample|state], P[sample], and P[state|sample]. State refers to the two alternative states: *contaminated* or *not contaminated*, i.e. μ > AL = C^+ or μ < AL = C^-. Sample represents *detection* or *no detection* of contamination, i.e. $\bar{y} > \mathrm{AL} = D^+$ or $\bar{y} < \mathrm{AL} = D^-$, where $\bar{y}$ is the mean of the logtransformed measurements. The action level AL can be a guidance value defining when remediation is required. The aim of the probability estimation is to update the prior probability P[state] to a preposterior probability P[state|sample], that together with P[sample] is required in the preposterior objective function. The updating is achieved with Bayes' theorem but requires the probability P[sample|state] derived subsequently.

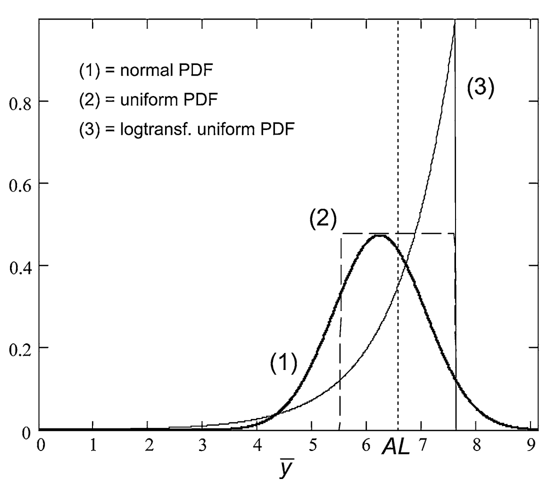

Fig. 3 Three alternative ways of specifying the prior information. In the application the normal PDF is defined so that b = ln(2,000 mg/kg) equals the 95-percentile, the uniform PDF so that $P(C^+) = P(C^-)$ = 0.5, and the logtransformed uniform PDF so that all mean values are below 2,000 mg/kg

Prior probabilities

The prior probabilities of the true state, P[state], are the probability of the RU being contaminated or uncontaminated, i.e. $P(C^+)$ or $P(C^-)$. These are estimated from the prior PDF as the area above and below the action level, respectively.

Including sample uncertainty

As mentioned, μ is the true but unknown mean concentration (logtransformed) of the RU, which we try to estimate from a planned sampling program of primary samples i = 1,...,n. Each expected sample value x_i is associated with uncertainty regarding the true mean concentration of the RU, specified by the coefficient of variation, CV (random component of the expected error). The systematic component (expected bias) is expressed as ε_s for each x_i. It is assumed that all samples taken in the RU are associated with the same uncertainty. The estimated mean concentration $\bar{y}$ of n logtransformed sample values y_i (i = 1,...,n) is assumed to be normally distributed around μ. The assumption of normally distributed errors around the mean is commonly applied based on the Central Limit Theorem (Box et al. 1978) and is believed to be reasonable also for this problem, especially after logtransformation. The standard deviation σ_y is estimated as (Strom and Stansbury 2000):

$$\sigma_y = \sqrt{\ln(\mathrm{CV}^2 + 1)} \tag{4}$$

The corresponding standard error of the mean of logtransformed data is

$$\mathrm{SE}_y = \sqrt{\frac{\ln(\mathrm{CV}^2 + 1)}{n}} \tag{5}$$

The bias is expected to be the same for all samples and thus the bias in mean concentration is

$$\varepsilon_x = \varepsilon_s \tag{6}$$

If ε_x is quantified, a correction for the bias can be made so that $\bar{y}$ is normally distributed around ln (e^{μ} + ε_x). The distribution of $\bar{y}$ can be expressed as

$$\bar{y} \sim N\left[\ln(e^{\mu} + \varepsilon_x);\ \frac{\ln(\mathrm{CV}^2 + 1)}{n}\right] \tag{7}$$

High untransformed values exhibit a larger variation than low untransformed values because a constant coefficient of variation is used. This is reasonable because larger variability is expected for high concentrations than for low. However, after logtransformation the widths of the distributions for different $\bar{y}$ are seemingly equal.

The probability of $\bar{y}$ exceeding a logtransformed action level AL is

$$P(\bar{y} > \mathrm{AL}|\mu) = P(D^+|\mu) = p_{\bar{y}>\mathrm{AL}}(\mu) \tag{8}$$

where $p_{\bar{y}>\mathrm{AL}}(\mu)$ represents a probability derived from the normal distribution of $\bar{y}$ as a function of the true mean μ, according to Eq. 7. The function $p_{\bar{y}<\mathrm{AL}}(\mu)$ is calculated similarly. Theoretically, the high tail could exceed the highest possible concentration, but this is insignificant for most problems and is therefore ignored.

Weighing sample uncertainty with prior information

It is obvious that false classification of the RU may occur, especially close to the action level. The four different ways of classifying the RU correspond to the four probabilities P[sample|state]. The probabilities are estimated by weighing all possible distributions of $\bar{y}$ around μ with the prior PDF and normalising with the prior probability P[state], followed by integration over all possible values of μ larger, or lower, than the action level. The probability of falsely classifying the RU as uncontaminated is

$$P(D^-|C^+) = \int_{\mathrm{AL}}^{\infty} p_{\bar{y}<\mathrm{AL}}(\mu) \cdot \frac{f_{\mathrm{prior}}(\mu)}{P(C^+)} \mathrm{d}\mu \tag{9}$$

where $f_{\mathrm{prior}}(\mu)$ is the prior PDF. The other three probabilities P[sample|state] are calculated correspondingly. The probability of detecting contamination, P[sample], is calculated as

$$P(D^+) = P(C^-) \cdot P(D^+|C^-) + P(C^+) \cdot P(D^+|C^+) \tag{10}$$

Preposterior probabilities

The four preposterior probabilities P[state|sample] are estimated with Bayes' theorem. For example the probability of failure is estimated as

$$P(C^+|D^-) = \frac{P(D^-|C^+) \cdot P(C^+)}{P(D^-|C^+) \cdot P(C^+) + P(D^-|C^-) \cdot P(C^-)} \tag{11}$$

The other three preposterior probabilities are estimated correspondingly.

Step 4: cost estimation

Four types of cost- or benefit-related values must be specified: (1) benefits provided by the specific decision alternative, (2) costs induced by the specific alternative (remediation costs), (3) cost brought about by making the wrong decision (cost of failure), and (4) cost of the proposed sampling program. The first three values are required in the objective function (Eq. 1). The sampling cost is required for estimation of the expected net value of the sampling program. The principles for cost estimation are beyond the scope of this paper, but information on this issue can be found in e.g. National Research Council (1997) and in Hardisty and Özdemiroğlu (2005) for contamination problems.

The cost of failure (3) is of special importance because it is highly uncertain and has the potential to affect the result significantly (Back 2003). It is defined as the additional cost from making the wrong decision. Two types of failure can occur: (1) classification of the RU as contaminated when it is not, or (2) classification of the RU as "clean" when it is in fact contaminated. Flatman and Englund (1991) refer to these misclassifications as false positive and false negative, respectively. The false positive results in unnecessary remediation but no additional cost. The false negative is referred to as Consumer's risk by Gilbert (1987), and may result in long-term human or environmental negative effects. Such consequences are represented by the failure cost.

A reasonable assumption is that the failure cost depends on how much the true mean concentration exceeds the action level. This can be modelled by an asymmetric loss function, $f_{\mathrm{loss}}(\mu)$ (Flatman and Englund 1991). An effective failure cost, Cf, is estimated by integrating the loss function and the prior PDF normalised by the prior probability $P(C^+)$:

$$\mathrm{Cf} = \int_{\mathrm{AL}}^{\infty} \frac{f_{\mathrm{prior}}(\mu)}{P(C^+)} \cdot f_{\mathrm{loss}}(\mu)\, \mathrm{d}\mu \tag{12}$$

Step 5: estimation of the value of information

The value of information is estimated in a three-step procedure; (1) prior analysis, (2) preposterior analysis, and (3) calculation of the expected value of information (EVI). A decision tree for the prior analysis is presented in Fig. 1. The expected value of the prior objective function is

$$\Phi_{\text{prior}} = \max\left[\Phi_1; \Phi_2\right] \tag{13}$$

where Φ_1 and Φ_2 refer to the objective function for the two decision alternatives. The preposterior analysis is performed when the sampling program has been defined, but before the actual sampling is carried out. The expected value of the preposterior objective function is

$$\Phi_{\text{preposterior}} = P(D^+) \cdot \max\left[\Phi_1|D^+; \Phi_2|D^+\right] + P(D^-) \cdot \max\left[\Phi_1|D^-; \Phi_2|D^-\right] \tag{14}$$

The estimated value of the sampling program is EVI:

$$\text{EVI} = \Phi_{\text{preposterior}} - \Phi_{\text{prior}} \tag{15}$$

It is possible to estimate the value of a perfect sampling program, eliminating all uncertainty, without performing a preposterior analysis. This is the expected value of perfect information (EVPI) and represents an upper bound on EVI and the maximum exploration budget that can be justified without performing a preposterior analysis (Morgan and Henrion 1990; Freeze et al. 1992):

$$\text{EVPI} = \Phi_{\max} - \Phi_{\text{prior}} \tag{16}$$

$$\Phi_{\max} = P(C^+) \cdot \max\left[B_{i,1} - C_{i,1} - Cf_{i,1}\right] + P(C^-) \cdot \max\left[B_{i,2} - C_{i,2} - Cf_{i,2}\right] \tag{17}$$

After Klemeš (1977), the relative value of information (RVI) can be expressed as a function of the number of samples n in the sampling program:

$$\text{RVI}(n) = \frac{\text{EVI}(n)}{\text{EVPI}} \tag{18}$$

It can be interpreted as how efficient the sampling is in reducing the uncertainty. The expected net value (ENV) is a measure of the cost-effectiveness of the sampling program:

$$\text{ENV}(n) = \text{EVI}(n) - C_p(n) \tag{19}$$

where $C_p(n)$ is the sampling cost function. Here, the sampling cost is a function of an establishment cost (C_{estb}), a daily cost for labour and machinery (C_{day}), a per sample cost for laboratory analyses and other expenses (C_s), the number of samples that can be collected in 1 day (n_{spd}), and the number of samples in the sampling program (n):

$$C_p(n) = C_{\text{estb}} + k \cdot C_{\text{day}} + n \cdot C_s \text{ for } (k-1)n_{\text{spd}} < n \leq k \cdot n_{\text{spd}}; \quad k = 1, 2, \ldots \tag{20}$$

Equation 20 is applied when measurements are made on all primary samples. Composite sampling requires a slightly modified equation:

$$C_p(n) = C_{\text{estb}} + k \cdot C_{\text{day}} + n \cdot C_c + n \cdot C_s \text{ for } (k-1)n_{\text{spd}} < n \cdot n_{\text{inc}} \leq k \cdot n_{\text{spd}}; \quad k = 1, 2, \ldots \tag{21}$$

Here, n represents the number of composite samples, n_{spd} is the number of increments collected in a day, and C_c is the cost of forming a composite sample from n_{inc} increments.

The ENV can be used to identify cost-effective sampling programs and to compare the cost-effectiveness of different programs. The optimal number of samples, n_{opt}, is determined at the maximum of ENV. The ENV for the optimal number of samples is

$$\text{ENV}_{\text{opt}} = \text{EVI}(n_{\text{opt}}) - C_p(n_{\text{opt}}) \tag{22}$$

Only sampling programs with positive ENV are cost-effective, whereas negative ENV implies that the reduction in risk due to sampling cannot balance the sampling cost. In such cases $n_{\text{opt}} = 0$ and it is most cost-effective to base the decision on the prior analysis alone.

Application

Wockatz scrap yard

The Wockatz scrap yard is situated in the central parts of Göteborg, the second largest city in Sweden. It was operated between the 1930s and 1993. In the period 1950–60 metal scrap from shipyards was handled at the site. The adjacent river has a high protection value due to the ecosystems in the river and in the river estuary. In addition, the river serves as a municipal water supply for about 700,000 inhabitants.

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. The site has previously been investigated and results indicate substantial contamination of metals and oil products. No remediation of the site has been performed yet.

There are two landuse scenarios for the Wockatz site; residential and industrial landuse. In both scenarios the site needs to be remediated to meet the environmental standards. Here, focus is on the industrial landuse scenario but other landuses are considered in a subsequent sensitivity analysis.

The site is divided into remediation units (RUs) of size 100 m^2. Each RU is either remediated as a whole, or not at all. The remediation technique is excavation of contaminated soil. For illustrative purposes, only one RU will be analysed and one single contaminant is considered, i.e. zinc.

Sampling approach

A sampling exercised will be carried out with the objective to estimate the mean concentration in the RU. The RU is considered contaminated, and consequently remediated, if the mean exceeds an action level of 700 mg/kg, which is the Swedish generic guideline value for land with less sensitive use (Swedish EPA 1997). However, there is a possibility that the future landuse will be residential, which would require a lower action level.

The purpose of the analysis is to determine the optimal number of samples in the RU from a cost-effectiveness point of view. Samples are located in a random manner in the RU. For simplicity, composite sampling is not performed.

Prior information

Previous investigations at the site indicate a mean concentration of zinc less than 2,000 mg/kg with a median of about 500 mg/kg. A normal distribution (prior PDF) is defined on the logtransformed data, with the 95-percentile at 2,000 mg/kg. To study the effect of how prior information is defined, two additional prior PDFs are applied; a uniform PDF and a logtransformed uniform PDF (Fig. 3). For the uniform prior PDF, all logtransformed mean concentrations between a and b are assumed to be equally probable:

$$f_U(\mu) = \begin{cases} \frac{1}{b-a} & \text{for } a < \mu < b \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

Parameter b is defined as the 95-percentile of the normal PDF, and a is defined so that $P(C^+) = 0.5$, i.e. maximum prior uncertainty.

Now consider instead the case where all mean concentrations in untransformed units between $b_u = e^b$ and $a_u = e^a$ are assumed to be equally probable. The cumulative distribution function (CDF) for this case is

$$F(x) = \begin{cases} \frac{x}{b_u - a_u} & \text{for } a_u < x < b_u \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

Logtransformation of the x-axis yields:

$$F(\mu) = \begin{cases} \frac{e^{\mu}}{e^b - e^a} & \text{for } a < \mu < b \\ 0 & \text{otherwise} \end{cases} \tag{25}$$

and the logtransformed prior PDF is consequently:

$$f_{LU}(\mu) = F'(\mu) = F(\mu) \tag{26}$$

In the application, the parameters are selected as $b = \ln(2000)$ and $a = 0$.

Based on previous sampling, the variability in sample data is expected to be about 100% (CV = 1), including spatial variability within the RU and random errors in sampling and analyses. No systematic error is considered.

Probability estimation

In total, 12 different probabilities are estimated, i.e. all variants of the four types of probabilities related to *state* and *sample* (Freeze et al. 1992).

Cost estimation

Each RU contains about 180 metric tons of soil ($10 \times 10 \times 1$ m^3 and density 1.8 ton/m^3) and the remediation cost is about 2.7 k SEK/ton based on current prices in Sweden for disposal of highly contaminated soil, including excavation, transport, and backfilling (10 SEK ~ 1.1 euros; ~ 1.3 USD). Because the cost of failure for an RU is associated with substantial uncertainty three different loss functions are evaluated. The first loss function is a step function with a constant value of 2 MSEK when $\mu >$ AL, see Fig. 4. This value is based on the perspective that misclassification of the RU will lead to large negative consequences, such as long-term human health effects, regardless of the size of the failure. The second loss function in Fig. 4

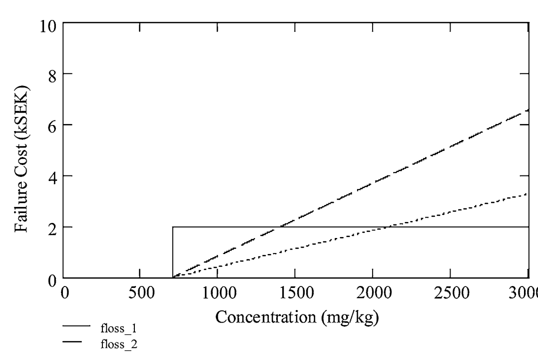

Fig. 4 Three different asymmetric loss functions representing different perspectives of failure, i.e. when the mean concentration exceeds AL (700 mg/kg)

Table 1 Result of value of information analysis of a proposed sampling program for a remediation unit at the Wockatz scrap yard

Result of VOIA	Normal prior PDF	Uniform prior PDF	Logtransformed uniform prior PDF
EVPI (kSEK)	328	250	175
Optimal no. of samples, n_{opt}	16	18	17
EVI at n_{opt} (kSEK)	258	175	114
ENV at n_{opt} (kSEK)	206	119	60
RVI at n_{opt} (%)	79	70	65

A step function of 2 MSEK is applied for the failure cost. The EVI, ENV, and RVI are listed for the optimal number of samples

assumes that the failure cost is zero at μ = AL and increases linearly to 2 MSEK at μ = 2AL. In the third loss function, 2 MSEK is reached at μ = 3AL. Note that the loss functions are linear in untransformed units but exponential after logtransformation.

The effective failure costs associated with the three loss functions were calculated with Eq. 12. With a normal prior PDF the results are: Cf_1 = 2.0 MSEK, Cf_2 = 1.97 MSEK, and Cf_3 = 0.98 MSEK. No economical benefits from remediation are considered in the analysis.

Three types of sampling costs are considered, see Eq. 20: An establishment cost of 5 k SEK, a labour and machinery cost of 15 k SEK per day, and a cost of 2 k SEK per sample. It is assumed that 20 samples a day can be collected.

Results

Results of the VOIA are summarised in Table 1. Although the three evaluated prior PDFs have quite

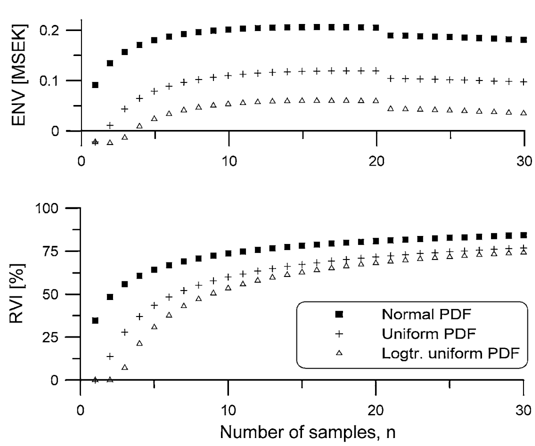

Fig. 5 Expected net value (*upper plot*) and relative value of information (*lower plot*) as functions of the number of samples for three different prior PDFs

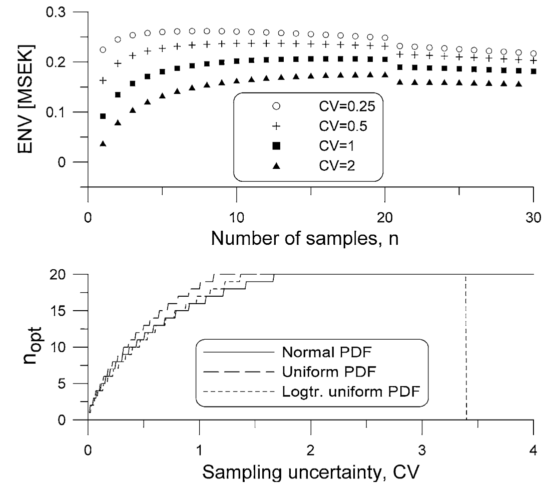

Fig. 6 Expected net value (*upper plot*) and the optimal number of samples (*lower plot*) as functions of uncertainty in sample data

different shapes there are only minor differences in the optimal number of samples in the RU. With the step function as loss function, the optimal number is in the range of 16–18, depending on how the prior information is defined. The optimal number of samples is slightly less with the second loss function in Fig. 4. The third loss function corresponds to a situation when the decision-maker regards failure as having lower consequences, resulting in an optimal number of samples in the range of 8–20.

The ENV and the RVI are illustrated in Fig. 5. Both are higher with the normal prior PDF than the other two PDFs. This is a result of the shapes of the prior PDFs in combination with the location of the action

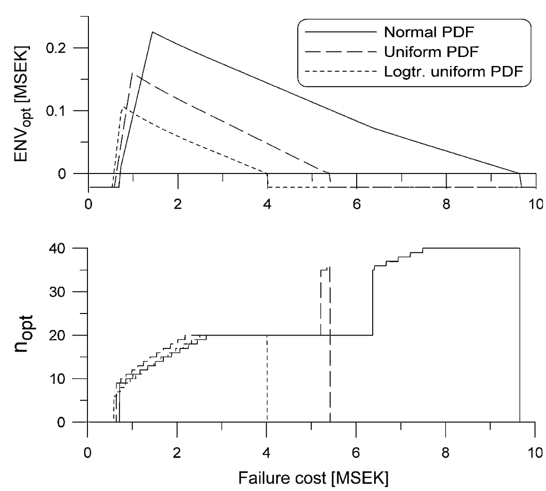

Fig. 7 The expected net value for the optimal number of samples (*upper plot*), and the optimal number of samples (*lower plot*), as functions of the cost of failure, for three different prior distributions

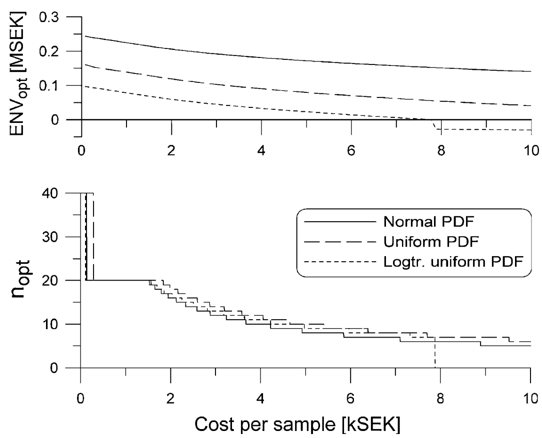

Fig. 8 The expected net value for the optimal number of samples (*upper plot*) and the optimal number of samples (*lower plot*) as functions of the cost per sample, for three different prior distributions

level. The expected mean concentration is defined as quite uncertain with these three prior PDFs. A more narrow PDF would result in a lower expected value of the sampling program because additional information is of less value when prior information is strong (Hammitt 1995).

The relative value of information reaches 65–80% at the optimal number of samples.

Single-factor sensitivity analysis by graphs

Graphical plots are simple and illustrative tools for studying the sensitivity of model results to a specific input variable. A multitude of combinations of input

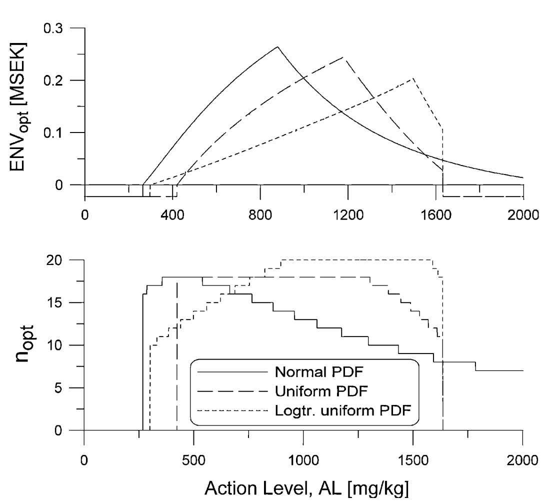

Fig. 9 The expected net value for the optimal number of samples (*upper plot*) and the optimal number of samples (*lower plot*) as functions of the action level, for three different prior distributions. The action level 350 mg/kg represents residential landuse, whereas 700 mg/kg represents industrial landuse

and output data can be analysed for this multi-dimensional and non-linear model but only a few are illustrated here, in addition to the different evaluated prior PDFs and loss functions. Four variables were analysed by graphs: (1) the sample uncertainty, (2) the failure cost, (3) the cost per sample, and (4) the landuse.

The effect of spatial variability and uncertainty in sample data is illustrated in Fig. 6. Increasing CV reduces the ENV and consequently the optimal number of samples increases. This is obvious since a less precise sampling will supply less information.

The effect of failure cost and sampling cost is illustrated in Figs. 7 and 8. The optimal number of samples increases when the cost of failure increases. If the failure cost is close to or lower than the remediation cost, sampling are of little or no value, and the optimal number of samples is zero. This is because the economic risk of leaving the contaminant in the ground is low and consequently it is more cost-effective to leave the site without sampling and remediation. On the other hand, if the failure cost is very high the most cost-effective decision may be to remediate without considering to spend money on sampling ($n_{opt} = 0$).

Sampling cost has the opposite effect compared to failure cost. The optimal number of samples is strongly affected by the extra cost that is induced when an additional day of fieldwork is required. This is evident in Figs. 6, 7, and 8 as plateaus at multiples of $n = 20$.

The sensitivity for different landuse scenarios is illustrated by applying a range of different action levels (Fig. 9). There is an almost linear increase in the EVI when AL increases, within a certain range and up to a maximum. As illustrated, the optimal number of samples can be very sensitive to the landuse. Only a small change in AL can reduce n_{opt} from 15–20 down to zero. Such threshold effects indicate that landuse is a major factor to consider. The action level affects the probability $P(C^{+})$ and if this probability is very low or high, the prior information is so strong that more data have little potential of changing the decision. Therefore, the optimal number of samples will be zero in such cases, and the decision to remediate or not can be based on the prior information alone.

Sensitivity analysis by Monte Carlo simulation

Monte Carlo simulation was performed in order to study the relative effects of the input variables. Table 2 lists the variables that were defined as stochastic, together with assigned PDFs. The rule for selecting ranges of the stochastic variables was to provide reasonable estimates of uncertainties in the Wockatz application. Generally, standard deviations of 10–25%

Table 2 Uncertainty in input variables for sensitivity analysis performed by Monte Carlo simulation

Input variables	Unit	Type of PDF	Mean	Std. dev.
Most likely mean of prior PDF, m_u	mg/kg	Normal	500	50
95 percentile of prior PDF, b_u	mg/kg	Normal	2,000	200
Action level, AL_u	mg/kg	Bimodal[a]		
Sampling uncertainty, CV		Normal	1	0.1
Number of samples per day		Poisson[b]		
Establishment cost, C_{estb}	MSEK	Normal	0.005	0.001
Cost per day, C_{day}	MSEK	Normal	0.015	0.0015
Cost per sample, C_s	MSEK	Normal	0.002	0.0002
Remediation cost, C_{rem}	MSEK	Normal	0.5	0.05
Cost of failure, Cf	MSEK	Normal	2	0.5

[a] 350 and 700 mg/kg with probability 0.5, respectively

[b] Rate = 20

of the mean were used. The uncertainty in future landuse was modelled by applying two different action levels, each with a probability of 50%, corresponding to residential and industrial landuse.

No correlation between input variables was considered because this would violate the purpose of the simulation: to identify the most important input variables. However, it is likely that some variables are correlated, for example, failure cost and landuse (simulations taking correlation into account were performed but were of little value and are therefore not presented). The analysis was limited to the normal prior PDF and 20,000 runs were made in the simulation.

The non-linear nature of the sampling problem is evident from Figs. 5, 6, 7, and 8, and therefore rank-order correlation was used to evaluate the sensitivity. The Spearman rank correlation coefficients between the 10 input variables and the output variables are illustrated in Fig. 10. For the optimal number of samples the three most important variables are the failure cost, the number of samples that is collected in 1 day, and the cost per sample.

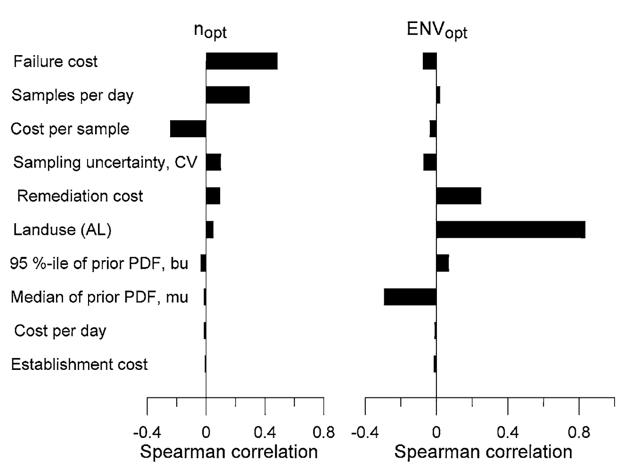

Fig. 10 Result of sensitivity analysis: Spearman correlation between input variables and the calculated optimal number of samples (*left*) and the expected net value (*right*)

Interestingly, three different variables have a dominating effect on the ENV compared to the optimal number of samples. By far the most important one is the landuse, but m_u and the remediation cost are also important. Failure cost seems to be less important for the net value than for the optimal number of samples. In addition, sample uncertainty (including spatial variability within the RU) has some effect on both n_{opt} and ENV_{opt} but to a minor extent.

There is an important limitation regarding the previous conclusions. Both the relationship between landuse (AL) versus n_{opt}, and failure cost versus ENV_{opt} are non-monotonic for the interval covered by the sensitivity analysis, i.e. an increase in the *x*-variable does not result in a strict increase or decrease in the dependant variable (Figs. 9 and 7, respectively). In Fig. 7, low failure cost exhibit positive correlation with ENV_{opt}, in contrast to high failure cost. This results in a small negative correlation on average, according to Fig. 10, although the correlation can be strong over short cost intervals. Therefore, it is likely that the importance of failure cost for estimating ENV_{opt} is larger than Fig. 10 suggests for specific problems. The same applies to the importance of landuse for estimation of n_{opt}. These limitations are well-known for this type of sensitivity analysis when applied to non-monotonic relationships (Decisioneering 2001).

Conclusions and discussion

A model for value of information analysis is presented for the sampling objective of estimating the mean concentration where sample data is lognormally distributed. The model has limitations regarding some aspects of practical application. Only one sampling objective is considered: estimation of mean concentration. This limits the applicability of the model in the early phases of a contaminated land project when questions are broad and sampling is performed with a range of different objectives in mind. The analysis is also performed on a single chemical substance, whereas in reality several contaminants need to be considered. In addition, some contaminants may require different remedial actions, which complicate the analysis. Such methodological limitations can be overcome by further development.

The model is limited to a single and isolated subarea of the site, such as an RU. Usually, consecutive sampling of several RUs is performed, which will affect the sampling cost function. In addition, the consequences of failure are not restricted to a single RU but will also depend on failure at site-level. To consider such aspects a model at site level is required.

The estimated optimal number of samples is purely statistical, i.e. n_{opt} is the best preposterior option in the long run. However, in a specific project posterior analysis may reveal that n_{opt} was in fact not the best option. This situation is impossible to avoid because of the inherent randomness and variability in geoenvironmental problems.

A rough rule of thumb presented by the US EPA (1992) says that less than 10 samples result in a poor estimate of the mean, 10–20 samples give a better estimate, and 20–30 samples result in a fairly consistent estimate of the mean. With a cost-effectiveness perspective on sampling fewer samples may be enough. In other situations, typically when the mean is close to the action level, it can be cost-effective to collect larger number of samples.

At the Wockatz scrap yard the optimal number of samples in an RU was estimated to be in the range of 16 to 18. This is significantly more than the number of samples that are commonly used in Sweden for classification of 100 m^2 RUs. The curve of ENV as a function of sample size is quite flat (Fig. 5) and there is a range of sample sizes with only minor differences in ENV. This occurs because the sampling cost is low compared to the other costs and almost the same as the increase in the value of the sampling program. This implies that not much value is lost if a non-optimal sampling program is selected. A higher sampling cost would result in a more pronounced optimum of the ENV curve.

A consequence of the flat ENV curve is that the optimal number of samples can be highly sensitive to the cost for additional days of fieldwork, which is indicated by the plateaus in Figs. 6 and 8. A conclusion is that under such circumstances, it is cost-effective to assign a full day to sampling, given that the sampling cost function adequately describes how the costs evolve.

Previous studies of decision analysis for contaminated land problems indicate that the failure cost estimate has a large impact on the result (Back 2003; Norrman 2004). The findings in this study indicate that other variables should not be neglected. However, it must be emphasised that decision analysis problems are highly unlinear and problem-specific (Lawrence 1999). Therefore, conclusions based on one single case study should not be generalised. Figures 6, 7 and 8 can give some insight that moderate changes in assumptions may yield quite different results.

As demonstrated, the value of a sampling program, and the optimal number of samples, is not only a function of physical properties, but also at least as much depend on the objective and perspective of the decision-maker. Subject to the same sampling problem two decision-makers may arrive at quite different conclusions about the optimal sampling effort, and both may be right based on their perspective. At least four important aspects related to the objectives and perspectives of the decision-maker will be discussed: (1) the sampling objective, (2) the consequences of failure, (3) the risk aversion of the decision-maker, and (4) the landuse scenario for the site.

The presented model is based on the sampling objective to estimate the mean concentration. A different objective, e.g. sampling for “hot spots”, could yield a very different result and would require a different model. It is quite common that sampling plans are developed with several sampling objectives in mind, although this is rarely stated explicitly. An important development would be to apply VOIA to real-world multiple objective problems. This would allow the value of information to be estimated for a wide range of situations where today only highly subjective and qualitative assessments can be made concerning the cost-effectiveness of investigations. The presented model is one step in this direction.

Of importance for the result is the perspective of misclassification of the RU (failure). If the perspective is that there are only minor consequences, then sampling is of course of little value. Since the consequences are likely to be assessed differently by the consultant, the decision-maker, other stakeholders, and the environmental authorities, it is obvious that their different perspectives will affect the estimated value of the sampling program. This is a reality that is easily overlooked but with a model like the presented one, such differences in perspective become evident and can be explicitly analysed.

The risk aversion may vary among the involved parties but a risk neutral perspective is reasonable as a starting point. The landuse scenarios may also vary but are usually defined beforehand. However, a landuse scenario is always coupled to a timeframe, implicitly or explicitly, and there may be different opinions about the long-term landuse.

Several of the input variables to the model are to some extent coupled to the geological conditions at the site, for example the sampling uncertainty (CV) and the number of samples that can be collected in 1 day (n_{spd}). The latter is of special importance in Sweden where glacial till dominates the landscape and sampling can be difficult due to the coarse and densely packed soil. An interesting question is to what extent the geological variability and uncertainty affects the result. This is a difficult question to answer because the extent to which geological conditions affect the individual variables is not known. A detailed answer would require each variable to be modelled by a subset of new variables where the geology is one of many. However, from the earlier discussion a reasonable conclusion is that the basis for economic valuation is more important for the outcome of the analysis than the geological uncertainty for this sampling problem. This in agreement with the conclusions of other authors, see James et al. (1996b).

Important aspects of decision analysis, as applied in this paper, are that the problem must be analysed and structured in a transparent way, and that critical questions are openly addressed. Many real-world problems would benefit from this structured and transparent analysis, making communication between involved parties easier. Decision analysis will pinpoint the critical questions and the ones most relevant for the problem can be identified. Efforts can thus be directed to the most important issues.

The presented model can be applied to a range of contaminated land problems but is especially suited for detailed investigations of RUs in the remediation phase of a project. It is applicable for problems where the decision alternatives are to remediate or to do nothing, and it effectively answers the commonly asked question of how many samples one should collect by applying a cost-effectiveness approach. This simple question is found to be embarrassingly difficult to answer even for statisticians (Lindley 1997).

The application of this model is not restricted to contaminated land problems. It can be applied to other sampling problems where cost-effectiveness is important. One such possible application is sampling and testing of wastes to meet landfill waste acceptance procedures.

Acknowledgments The author gratefully acknowledges support from Stena Metall AB and Renova AB to the Forum for Risk Investigation and Soil Treatment (FRIST) research centre at Chalmers University, as well as the support from the Post-Graduate School for Natural Materials in Environmental Science and Cultural Conservation (NMK) at Göteborg University.

References

Abbaspour KC, Schulin R, Schläppi E, Flühler H (1996) A Bayesian approach for incorporating uncertainty and data worth in environmental projects. Environ Model Assess 1996(1):151–258

Back P-E (2003) On uncertainty and data worth in decision analysis for contaminated land. Department of Geology, Chalmers University of Technology, Göteborg, Sweden

Ben-Zvi M, Berkowitz B, Kesler S (1988) Pre-posterior analysis as a tool for data evaluation: application to aquifer contamination. Water Resour Manage 1988(2):11–20

Bosman R (1993) Sampling strategies and the role of geostatistics in the investigation of soil contamination. In: Contaminated Soil '93, 4th international KfK/TNO conference on contaminated soil. Kluwer, Berlin

Box GEP, Hunter WG, Hunter JS (1978) Statistics for experimenters. An introduction to design, data analysis and model building. Wiley, New York

Brus DJ, de Gruijter JJ (1997) Random sampling or geostatistical modelling? Choosing between design-based and model-based sampling strategies for soil (with Discussion). Geoderma, vol 80, pp 1–24

Dakins ME, Toll JE, Small MJ (1994) Risk-based environmental remediation: decision framework an role of uncertainty. Environ Toxicol Chem 13(12):1907–2915

Dakins ME, Toll JE, Small MJ, Brand KP (1996) Risk-based environmental remediation: Bayesian Monte Carlo analysis and the expected value of sample information. Risk Analysis 16(1):67–29

Davis DR, Dvoranchik WM (1971) Evaluation of the worth of additional data. Water Resour Bull 7(4):700–207

Decisioneering (ed) (2001) Crystal Ball 2000 users manual. Denver

Einax JW, Kraft J (2002) Small-scale variability of metals in soil and composite sampling. Environ Sci Pollut Res 9(4):257–261

Flatman GT, Englund EJ (1991) Asymmetric loss functions for Superfund remediation decisions. In: Proceedings of the business and economic statistics section of the American Statistical Association

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):574–288

Freeze RA, Massman J, Smith L, Sperling T, James B (1990) Hydrogeological decision analysis: 1. A framework. Ground Water 28(5):738–266

Gates JS, Kisiel CC (1974) Worth of additional data to a digital computer model of a groundwater basin. Water Resourc Res 10(5):1031–2038

Gilbert RO (1987) Statistical methods for environmental pollution monitoring. Van Nostrand Reinhold, New York

Gregoire TG (1998) Design-based and model-based inference in survey sampling: appreciating the difference. Can J For Res 28(1998):1429–2447

Grosser PW, Goodman AS (1985) Determination of groundwater sampling frequencies through Bayesian decision theory. Civ Eng Syst 2(December):186–294

Hammitt JK (1995) Can more information increase uncertainty? Chance 8(3):15–27, 36

Hardisty PE, Özdemiroğlu E (2005) The economics of groundwater remediation and protection. CRC, Boca Raton
Hoffman FO, Kaplan S (1999) Beyond the domain of direct observation: How to specify a probability distribution that represents the ''state of knowledge'' about uncertain inputs. Risk Anal 19(1):131–234
IT-Corporation (1997) Value of information analysis for corrective action unit no. 98; Frenchman Flat, US Department of Energy, USA
James BR, Freeze RA (1993) The worth of data in predicting aquitard continuity in hydrogeological design. Water Resourc Res 29(7):2049–2065
James BR, Gorelick SM (1994) When enough is enough: the worth of monitoring data in aquifer remediation design. Water Resourc Res 30(12):3499–2513
James BR, Gwo J-P, Toran L (1996a) Risk-cost decision framework for aquifer remediation design. J Water Resourc Plann Manage 122(6):414–220
James BR, Huff DD, Trabalka JR, Ketelle RH, Rightmire CT (1996b) Allocation of environmental remediation funds using economic risk–cost–benefit analysis: a case study. Groundwater Monitoring and Remediation (Fall), pp 95–205
Kaplan PG (1998) A tutorial: a real-life adventure in environmental decision making, The Center for Risk Excellence, US Department of Energy
Klemeš V (1977) Value of information in reservoir optimization. Water Resourc Res 13(5):837–850
Lawrence DB (1999) The economic value of information. Springer, Berlin Heidelberg New York
LeGrand HL, Rosén L (2000) Systematic makings of early stage hydrogeological conceptual models. Ground Water 38(6):887–293
Lindley DV (1997) The choice of sample size. Statistician 46(2):129–238
Maddock T (1973) Management model as a tool for studying the worth of data. Water Resourc Res 9(2):270–280
McNulty G, Deshler B, Dove H (1997) Value of information analysis; Nevada Test Site, USA
Morgan MG, Henrion M (1990) Uncertainty, a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, Cambridge
National Research Council (1997) Valuing ground water. Economic concepts and approaches. National Academy Press, New York
Norberg T, Rosén L (2006) Calculating the optimal number of contaminant samples by means of data worth analysis. Environmetrics (in press)
Norrman J (2001) Decision analysis under risk and uncertainty at contaminated sites. A literature review. Department of Geology, Chalmers University of Technology, Göteborg
Norrman J (2004) On Bayesian decision analysis for evaluating alternative actions at contaminated sites. Department of GeoEngineering, Chalmers University of Technology, Göteborg
Pitard FF (1993) Pierre Gy's sampling theory and sampling practice. CRC, Boca Raton
Provost L (1984) Statistical methods in environmental sampling. Environmental sampling for hazardous wastes. In: Schweitzer GE, Santolucito JA, American Chemical Society, Washington DC, pp 79–96
Rautman CA, McGraw MA, Istok JD, Sigda JM, Kaplan PG (1994) Probabilistic comparison of alternative characterization technologies at the Fernald Uranium-in-Soils Integrated Demonstration Project. Waste Management 94, Tucson, Arizona
Reichard EG, Evans JS (1989) Assessing the value of hydrogeological information for risk-based remedial action decisions. Water Resourc Res 25(7):1451–2460
Russell KT, Rabideau AJ (2000) Decision analysis for pump-and-treat design. Groundwater Monitoring and Remediation, Summer, pp 159–268
Strom DJ, Stansbury PS (2000) Determining parameters of lognormal distributions from minimal information. AIHAJ 61(November/December):877–280
Swedish EPA (1997) Development of generic guideline values. Model and data used for generic guideline values for contaminated soils in Sweden. Stockholm
Swedish EPA (1998) Requirements for site remediation. Guidelines for practical achievement of acceptable residual concentrations and quantities—methods and quality aspects. Stockholm
Thompson M, Fearn T (1996) What exactly is fitness for purpose in analytical measurements? Analyst 121(March):275–278
US EPA (1992) Supplemetal guidance to RAGS: calculating the concentration term. Intermittent Bulletin, volume 1, number 1, Washington D.C., Office of Solid Waste and Emergency Response, Hazardous Site Evaluation Division
US EPA (1996) Soil screening guidance: technical background document. Office of Solid Waste and Emergency Response, Washington