CHALMERS | GÖTEBORG UNIVERSITY

PREPRINT 2007:8

Value of Information Analysis in Remedial Investigations

PÄR-ERIK BACK
LARS ROSÉN
TOMMY NORBERG

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

Preprint 2007:8

Value of Information Analysis in Remedial Investigations

Pär-Erik Back, Lars Rosén, Tommy Norberg

CHALMERS | GÖTEBORG UNIVERSITY

Department of Mathematical Sciences
Division of Mathematical Statistics
Chalmers University of Technology and Göteborg University
SE-412 96 Göteborg, Sweden
Göteborg, March 2007

Preprint 2007:8
ISSN 1652-9715

Matematiska vetenskaper
Göteborg 2007

Value of Information Analysis in Remedial Investigations

Pär-Erik Back[1], Lars Rosén[1,2] & Tommy Norberg[3]
Chalmers University of Technology, Göteborg, SWEDEN

Abstract

Site investigations of contaminated land are associated with high costs. From a societal perspective, just enough economic resources should be spent on investigations so that society's limited resources can be allocated to sustainable development in the best way. The solution is to design investigation programs that are cost-effective, which can be performed using Value of Information Analysis (VOIA). The principle is to compare the benefit at the present state of knowledge with the benefit that is expected after an investigation has been performed. Bayesian methods are used to calculate the expected change, i.e. the value of the investigation. A general framework for VOIA of site investigations is presented. The framework consists of seven modules: (i) the landuse scenario, (ii) the objective of investigation, (iii) a conceptual site model, (iv) a data collection module, (v) a prior information module, (vi) an uncertainty reduction module, and (vii) a decision model. The decision model is based on Bayesian risk-cost-benefit decision analysis. The result is an estimate of the value of an investigation program, and for specific problems, the optimal number of samples. Applications of the method illustrate that the objective of the investigation, the landuse, and the benefit of remediation have major impacts on the results. The main contributions of this work are: (1) a general framework for VOIA of site-investigations in remediation projects, (2) a tool-kit of VOIA models for practical application, and (3) a knowledge base of strengths and weaknesses of the methodology, including recommendations.

Introduction

During the history of mankind there has often been a conflict between development and environmental issues, long before the modern environmental movement was launched in the 1960s. A landmark in the protection of the environment was the introduction of the concept *sustainable development*, as defined by the Bruntland Commision (1987). According to their definition, sustainable development has three major dimensions: socio-cultural, ecological, and economical. This implies that development must consider environmental issues as well as economics. This view has also concretised in some national legislations. For example, economic considerations of environmental restoration are supported through the regulatory framework of the Swedish Environmental Code.

---

[1] Department of Environmental and Civil Engineering, Chalmers University of Technology, SE 412 96 Göteborg, SWEDEN

[2] Corresponding author, e-mail: lars.rosen@chalmers.se

[3] Department of Mathematical Sciences, Chalmers University of Technology and Göteborg University, SE 412 96 Göteborg, SWEDEN

One characteristic about the geological environment is that it is only visible at the ground surface, which only displays a tiny part of the total volume of geologic material. Characterisation of properties in the subsurface therefore requires investigations such as drilling, digging, excavation, geophysical investigations, and other costly investigation and measurement techniques. These methods are inevitably associated with large uncertainties. The uncertainties can be reduced by performing more investigations and collecting more data, but this will increase the costs. On the other hand, keeping the costs low implies large uncertainties, which may result in wrong decisions being made. Such decisions under uncertainty may lead to harm to humans and negative environmental effects.

The high costs involved in geo-related investigations are problematic because of society's limited economic resources. From a societal perspective, just enough resources should be spent on investigations; not more and not less. Spending more money is unsustainable in the sense that the resources could have contributed more to society solving a problem at a different location, or in other sectors of society. It is also unsustainable to spend less money because this will result in unacceptable risks. Thus there is a strong incentive to design investigation programs that are cost-effective.

Contaminated land is a type of problem that requires extensive investigation of the subsurface. The investigations include for example hydrogeochemical properties, subsurface transport properties, as well as geotechnical issues. Questions related to the risks for humans and the ecological systems have received increased attention during the last two decades. The Swedish EPA (2006) estimates that there are about 80,000 more or less contaminated sites in Sweden. The cost of remediating the 1,500 worst contaminated sites is estimated to 45 billion SEK (roughly 5 billion euros, or 6 billon USD).

The cost-effectiveness of a site investigation can be estimated by so called Value of Information Analysis (VOIA), or Data Worth Analysis (DWA). It can be performed at different levels of ambition and the highest level of ambition is based on risk-based decision analysis; see Freeze *et al*. (1992). The principle is to calculate the value of the investigation and compare it with the investigation cost. The approach requires that the total project economy is considered when a site investigation is planned, including economic valuation of risks. An investigation program should not be performed unless its value is bigger than its cost.

VOIA has been applied to contaminated land projects before, but only to a limited extent. In the last 15 years, VOIA has been applied to 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). Dakins *et al*. (1994; 1996) presented a decision framework for remediation of PCB-contaminated sediments, utilising Bayesian Monte Carlo methods, and Rautman *et al*. (1994) used a risk-based decision analysis approach to compare the reliability of different characterisation techniques for uranium contaminated soil. James *et al*. (1996b) presented a simple risk-based decision analysis framework for estimating the cost-effectiveness of a proposed sampling program for contaminated soil, whereas Kaplan (1998) described software with the purpose of locating sample points based on geostatistics and data worth in an iterative process. Norberg and Rosén (2006) presented a model for estimating the cost-effectiveness and the optimal number of samples for characterisation of soil in the remediation phase of a contaminated land project. Back (2006c) presented a model with similar objective, whereas Back *et al*. (2006) compared it with the model by Norberg and

Rosén (2006). Back (2006b) developed a model for early project phases based on hotspot sampling, including the uncertainties in hotspot detection.

Previous work indicates that there is a need for a more detailed evaluation of VOIA as a method to design cost-effective site investigations. The strengths and weaknesses of VOIA should be identified, the tool-kit of VOIA models for different phases of a remediation project needs to be expanded, economic valuation as a tool in VOIA needs to be evaluated, and the factors in a VOIA requiring special attention should be identified. All this requires that the theory is applied and thoroughly evaluated in case studies.

The aim of this paper is to improve the knowledge base for application of VOIA on remediation problems. The work aims at cost-effective site investigations in general, although the focus is on random sampling. The reason is that sampling is the most important collection method for hard data in contaminated land projects. Sampling theory is a basis for the work, since these principles apply to most types of investigations. The uncertainty in sampling is critical and must therefore be considered in the methodology.

The specific objectives are: (1) to develop a framework for VOIA on contaminated land problems, (2) to develop a tool-kit of VOIA models that can be used during different phases of remediation projects, and (3) to identify strengths and weaknesses of the methodology, including the factors that are most important for the result by means of applying the methodology.

Theoretical background

Basic approaches to sampling

Every remediation project is faced with the question of how much resources should be allocated to site investigations in order to reduce uncertainty. The answer to this question is surprisingly complex (Lindley, 1997) and there are different strategies that can be used. There are at least four basic approaches for designing a sampling program: (1) Minimise the uncertainty for a given sampling budget, (2) Minimise the cost of reaching a specified precision, (3) Follow regulations that specify the required sampling effort, and (4) Select the most cost-effective sampling program.

In addition to these four approaches there are others of limited importance, such as the rule-based adaptive sampling approach (Cox, 1999) and the haphazard sampling approach. The latter is based on the philosophy that ‘any sample will do’ (Gilbert, 1987). Traditionally, approach 1-3, or combinations of these, have been used to design sampling programs. However, all these approaches raise questions that are fundamental but very difficult to answer.

In the first approach there is a pre-specified sampling budget. Justified questions are why this specific amount of money should be allocated to sampling, and if this amount really will supply a result of sufficient quality.

In approach 2, the aim is to reach a pre-specified precision or uncertainty level. This approach is often suggested in guidelines or standards on sampling; see e.g. Cochran (1977) and CEN (2005). Obvious questions to ask are how we can define this specific level of precision, and why this particular level is desired. In addition, this approach can rise problems because many people are unaccustomed to thinking in terms of the amount of error that can be tolerated in estimates.

Approach 3 implies following regulations or guidelines that specify the required sampling effort in a quantitative way, e.g. how many samples one needs to collect. The approach can only be applied when regulations are very detailed. The natural question is: 'What is the basis for the specified sampling effort and is it sufficient, or necessary, for the actual problem?'

The objective of approach 4 is to find the sampling program that minimises the total cost in the long run, i.e. the expected total cost, or maximises the expected total benefit. Both sampling costs and other project costs and benefits are considered, deterministic as well as probabilistic. In some problems, however, the total number of possible sampling programs is extremely high and the computing time needed to find the best program is far too long. In such cases experts may be used to single out a small number of promising sampling programs. The best of these programs can then be found.

Approach 4 avoids many of the difficulties related to approaches 1-3. The basis for approach 4 is Value of Information Analysis (VOIA). One disadvantage of the methodology is that it can be rather complex. In addition, depending on the perspective of the decision-maker, questions like these could be raised: 'Why should we minimise the expected cost (or maximise the expected benefit)?' and 'Why should we care for other costs than sampling costs?' The key answer is that society[4] has limited economic resources, which means that money should be spent when required but not otherwise.

Decision analysis

VOIA is an application of Bayesian decision cost-benefit analysis. This kind of decision analysis is a systematic approach to decision-making under uncertainty. The question it tries to answer is how to select the best *alternative action* from a set of alternatives. Keeney (1982) defines decision analysis as "*a formalization of common sense for decision problems which are too complex for informal use of common sense*". Dakins *et al*. (1994) describe decision analysis as a technique to "*...help organize and structure the decision-maker's thought process, elicit judgements from the decision maker or other experts, check for internal inconsistencies in the judgements, assist in bringing these judgements together in a coherent whole, and process the information and identify a best strategy for action*". The complex problem must be analysed and formulated in a clear way, which requires a structured approach. In fact, one of the main advantages of a formalised decision analysis is that the logical and structured approach results in transparency. Keeney (1982) gives an excellent overview of decision analysis, including common misconceptions about its philosophy and usefulness.

The decision analysis problem is structured in a decision model. Such a model consists of the following parts (Martignon, 2001): (1) the decision options (alternative actions), (2) the relevant consequences (outcomes), (3) the probabilities (if the analysis is probabilistic), and (4) the decision rule (the aims of the decision-maker).

The *decision options* are the set of alternative actions from which the decision-maker will have to select one alternative. Each alternative is associated with one or more *outcomes*. The

[4] In this context, society includes both the problem-owner and the rest of society.

outcomes, or *consequences*, are often quantified as *payoffs*, especially when monetary costs or benefits are involved in the analysis. When future monetary costs and benefits are included it becomes necessary to calculate the *net present value*. There are usually a set of *events* that needs to be included in the analysis, in addition to the decision options. The decision problem can be structured in a *decision tree*, or *influence diagram*, to illustrate the chain of decision options, events, and outcomes.

In risk-based decision analysis, a probabilistic approach is used. Probabilities are assigned to uncertain events and are also used to model the uncertainty about the *true state of nature*. Risk-based decision analysis provides probabilistic results, such as the *expected cost* or *expected benefit*, by means of multiplying probabilities and consequences. *Risk* is defined as an expected cost.

It is common to apply the *Bayesian*, or subjectivist, view in decision analysis, which means that probabilities can and should be assessed by using subjective information (*soft data*), not just *hard data* (measurements) like in classical statistics. The subjective probabilities can later be combined with new data to reach an updated information state (Dakins *et al*., 1996).

A simple way to present the result of a decision analysis is to use a payoff matrix, or *payoff table*. The payoff table summarises the alternative actions and their corresponding payoffs, i.e. their stream of benefits and costs over a specified time horizon. The payoffs are calculated using Eq. 1, where Φ is an objective function that represents the expected total net present value of alternative *i*.:

$$\Phi_i = \sum_{t=0}^{T} \frac{1}{(1+r)^t} [B_i(t) - C_i(t)] \qquad (1)$$

where *B* represents the deterministic and probabilistic benefits [monetary units], *C* is the deterministic and probabilistic costs [monetary units], *T* is the time horizon, and *r* is the discount rate.

The basis for the decision is the decision rule, or decision criterion, of the decision-maker. Examples of decision criteria are (Hintze, 1994; Martignon, 2001; Norrman, 2004): the maximum expected utility (EU) criterion; the maximum expected monetary value (EMV) criterion; the maxiprobability criterion; the maximin or minimax criterion; the 'take the best' criterion; and the intuitive criterion.

The *maximum EU* criterion is common in decision-making. It implies that the decision alternative with the highest expected utility should be chosen. Utility is a concept representing the satisfaction, happiness, or wellbeing (Norrman, 2004). If utility is expressed in monetary units, the maximum EU criterion is identical to the *maximum EMV* criterion, i.e. the objective is to maximise the expected monetary value. Some support for the maximum EMV criterion can be found in the Swedish Environmental Code, which states that the benefit of protective actions should be compared with the costs for these actions. The maximum EU and EMV criteria takes both probabilities and consequences of different outcomes into account. The *maxiprobability* criterion on the other hand, ignores consequences. The VOIA analysis described in this paper uses the EMV decision criteria.

A pitfall common to decision analysis is the misconception that it provides a solution to the decision problem. However, it should be understood that decision analysis, like any type of analysis, only focuses on part of a problem (Keeney, 1982). Instead, its purpose is to produce insight to help decision-makers to make better decisions.

Box 1. Prior analysis

Suppose a decision maker needs to choose between doing a complete remediation of a small potentially contaminated site or leaving it as it is. The site is presently posing a risk to the society which is estimated to $P(C)C_B$, where $P(C)$ denotes the (subjective) probability that the site is contaminated and thus in need of remediation, and C_B denotes the monetary benefits of remediating a contaminated site (C denotes the event that the site is contaminated). Let moreover C_R be the remediation cost. Then, cf eq (1),

$$\Phi_{\text{prior}} = \max(\Phi_0,\, \Phi_1)$$

where $\Phi_0 = 0$ (the null alternative has no benefits and no costs) and

$$\Phi_1 = B_1 - C_1 = P(C)C_B - C_R$$

So remediation is worth doing if $P(C) > C_R/C_B$. In the continuation of this example below in Boxes 2, 3 and 4, we will assume that $C_R = 0.3\,C_B$. Then,

$$\Phi_{\text{prior}} = \max(0,\, P(C) - 0.3)\, C_B$$

Uncertainty and error

It is not an easy task to define what uncertainty really is. The variety of types and sources of uncertainty, along with the lack of agreed terminology, can generate considerable confusion (Morgan and Henrion, 1990). Still, it is important to notice the difference between the terms uncertainty and error, although they are often used as synonyms. Terms like uncertainty, reliability, confidence and risk are probability-related and refer to *á priori* conditions, i.e. the situation *before* an event has occurred. Error, on the other hand, can only be measured *posteriori*, i.e. *after* an event has occurred. It is not possible to know what the errors will be before the event has occurred. However, uncertainty may be present even after an event has occurred if the error is not completely known.

Though not easily defined, it is important to distinguish between the different types and sources of uncertainty. Lacasse and Nadim (1996) divide uncertainties into two categories: (1) aleatory (inherent or natural) uncertainty, and (2) epistemic (lack of knowledge) uncertainty. The aleatory uncertainty is irreducible but epistemic uncertainty can be reduced by collecting more data.

It is quite common to distinguish between uncertainty in quantity (*parameter uncertainty*) and uncertainty about model structure (*model uncertainty*). Sturk (1998) classifies uncertainty in geological engineering problems into three classes: (1) inherent variability, (2) modelling uncertainty, and (3) parameter uncertainty.

Bedford and Cooke (2001) points out that uncertainty must be distinguished from ambiguity, which results from describing observations in an ambiguous language. It is impossible to remove all ambiguity but it can be removed to an acceptable level by linguistic conventions. Ambiguity must be satisfactorily removed before we can meaningfully discuss uncertainty.

Box 2. Posterior analysis

The decision maker is very uncertain about the probability $P(C)$, so her safest guess is $P(C) = 0.3$ (this is the probability where the decision uncertainty is maximal), in which case $\Phi_{\text{prior}} = 0$. This is not very satisfactory, so before deciding the decision maker wants to investigate the site in order to better "pin down" this probability. The standard data collection procedure is sampling followed by laboratory analyses. The laboratory analyses are usually reliable but the cost can be high. Denote by $P(D|C)$ and $P(D|C')$ the probabilities that it is truly and falsely discovered that the site is contaminated (D denotes the event that C is detected). Roughly, $P(D|C) = 0.4$ and $P(D|C') = 0.1$. Suppose C is detected, i.e, that D has occurred. Then, by Bayes formula,

$$P(C|D) = \frac{P(C)P(D|C)}{P(C)P(D|C) + P(C')P(D|C')} = 0.632$$

and $\Phi_{\text{posterior}} = 0.332\, C_B$. Similarly, $P(C|D') = 0.222$, so $\Phi_{\text{posterior}} = 0$ if D' occurs.

Efficiency and optimality

All investigation programs should be efficient and if possible also optimal. Minasny and McBratney (2002) define efficiency in three different ways:

1. Efficiency in terms of effort: $\text{Efficiency 1} = \dfrac{\text{quality of information}}{\text{effort}}$
2. Efficiency in terms of cost: $\text{Efficiency 2} = \dfrac{\text{quality of information}}{\text{cost of information}}$
3. Efficiency in terms of value of information:
$\text{Efficiency 3} = \text{value of information - cost of information}$

The *quality of information* can be expressed by the uncertainty, quantified by e.g. the standard deviation. The *effort* can be described by the required time to collect the information. The *cost of information* is the monetary price of acquiring the information. Efficiency 3 is denoted as the expected net value (ENV) in this paper.

In this paper, the terms 'efficient' and 'effective' are used synonymously. In contrast to an 'effective' investigation, the 'optimal' investigation is identified among all theoretically possible alternatives. Therefore, *an investigation can be effective but not necessarily optimal* because there may be other alternatives that are even more cost-effective (Figure 1). Another difference between the two concepts is the type of criterion used to define when an investigation is 'effective' or 'optimal'. The 'optimal' alternative is found were a specific property has an extreme value, for example where the expected net value (ENV) has a maximum or the expected cost has a minimum. An 'effective' alternative on the other hand must be defined based on a specific threshold, for example when the value of information is larger than zero. This implies that *an investigation can be optimal but not necessarily effective*.

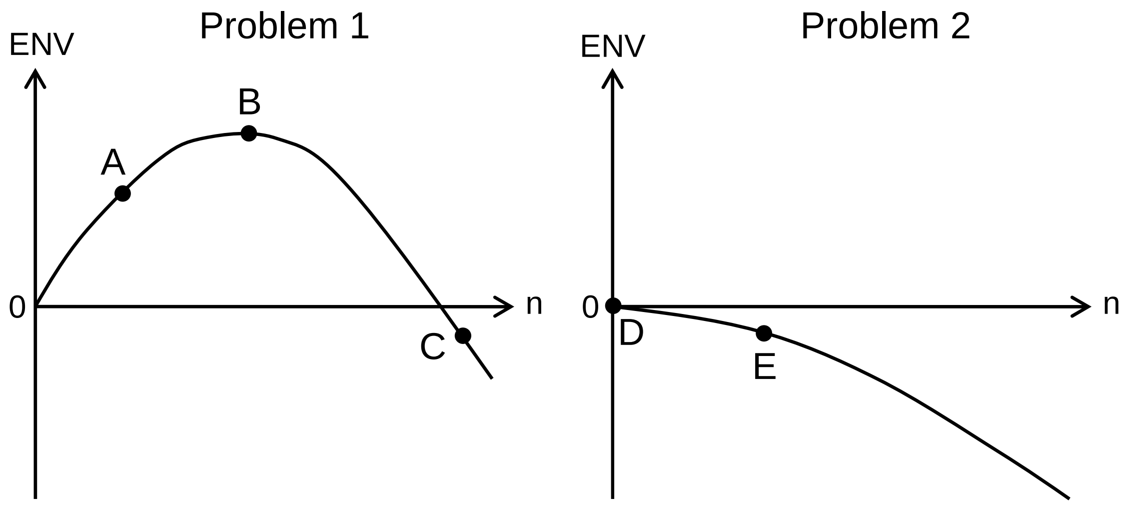

Figure 1 *Illustration of how the expected net value (ENV) may change as a function of the extent of sampling (n). Sampling program A and B are cost-effective (positive ENV) whereas C, D, and E are not. Program B and D are optimal for sampling problem 1 and 2, respectively. Note the program D implies doing nothing.*

VOIA can be used to identify *cost-effective*[5] investigation programs. Optimisation theory can be applied to identify the *optimal* alternative, but this is not further discussed in this paper. For specific problems however, VOIA can be used to calculate the optimal number of samples.

Framework for Value of Information Analysis

The origin of VOIA is decision theory and it has sporadically been applied to geo-environmental problems since the beginning of the 1970s. Cochran (1977) states that although it is not evident how frequently the value of information approach is capable to find complete solutions to decision problems, the method has value in stimulating clear thinking about the important factors in a good decision.

The basic idea of VOIA is simple: the value of additional information is the change in expected total cost (or benefit) caused by the new information. The value is estimated by analysing the uncertainty at the present state of knowledge, and comparing it with the reduced uncertainty that is expected when new information becomes available. The analysis is performed in three steps: (1) Prior analysis, (2) Preposterior analysis, and (3) Calculation of the value of information.

The *prior analysis* is based on the present state of knowledge and results in an expected total cost, or benefit. The result of the prior analysis is a value of the prior objective function Φ_{prior}, calculated by Eq. 1. The subsequent *preposterior analysis* is performed similarly, but is based on the information that is expected from the data collection program. This implies that the

[5] In general economic terminology, something (e.g. a policy option) is regarded as *cost-effective* when it achieves some specified objective (e.g. an environmental quality objective) at the least possible cost (Perman *et al*., 2003).

analysis is performed *after* ('posterior') the data collection program has been defined, but *before* ('pre') the data collection has taken place, and it results in a value of the preposterior objective function Φ_{prep}. In the third step, the Expected Value of Information (EVI) is calculated (Freeze *et al*. 1992):

$$\text{EVI} = \Phi_{preposterior} - \Phi_{prior} \qquad (2)$$

There is only a value of information if the investigation has the potential to change the decision. If the decision is already made, regardless of the outcome of the investigation program, there is no value of additional information. Note that the EVI does not consider the data collection cost C_p. To consider this cost, the Expected Net Value (ENV) is calculated (Freeze *et al*. 1992):

$$\text{ENV} = \text{EVI} - C_p \qquad (2)$$

Figure 1 illustrates how the ENV can behave. The first data to be collected usually contain the most information and have the highest value. As more data are collected, their marginal worth decreases (Dawdy, 1979), although there are exceptions to this rule for specific problems. The marginal worth is the slope of the ENV curve in Figure 1.

Box 3. Pre-posterior analysis

Before actually doing the investigation, the wise decision maker calculates its (pre-posterior) value. In doing so, she also needs the probability $P(D) = P(C)P(D|C) + P(C')P(D|C') = 0.19$. Then,

$$\Phi_{\text{preposterior}} = E\Phi_{\text{posterior}}$$
$$= 0.332 \cdot 0.19\, C_B = 0.063\, C_B = 0.21\, C_R$$

Hence

$$\text{EVI} = \Phi_{\text{preposterior}} - \Phi_{\text{prior}} = 0.21\, C_R$$

Assume that the investigation cost amounts to $C_P = 0.03\, C_R$ (sampling equipment, laboratory analyses and one day of labor). Then

$$\text{ENV} = \text{EVI} - C_P = 0.18\, C_R$$

Thus, the proposed investigation is clearly cost-effective.

VOIA can be performed at three different complexity levels (Figure 2). A Level 1 analysis is suitable for situations where it is impractical to assess the value in monetary terms, or when investigation costs are not of concern. In these cases, some surrogate of monetary value is often used (Dawdy, 1979). One such surrogate is reduction of uncertainty. This approach implies that cost-issues are not addressed at all; the analysis more or less resembles a traditional uncertainty analysis. A Level 1 analysis corresponds to the definition of 'Efficiency 1' by Minasny and McBratney (2002).

In a Level 2 analysis, the investigation costs are considered, but no other costs or benefits. The value of information is quantified as the quotient of uncertainty reduction and investigation cost. This corresponds to the definition of 'Efficiency 2' according to by

Minasny and McBratney (2002). Level 2 analyses for contaminated land problems have been demonstrated by McNulty *et al*. (1997) and the IT Corporation (1997).

With a Level 3 analysis, the monetary value of an investigation can be estimated, corresponding to 'Efficiency 3' by Minasny and McBratney (2002), i.e. the ENV. All costs and benefits in the remediation project are considered, both deterministic costs, benefits, and investigation costs, as well as probabilistic costs (risks). An advantage of this approach is that the measure of value is easier to understand and use than the more abstract measures in Level 1 and Level 2 analyses. In a Level 3 VOIA, all decision options, and possible events and their consequences, are considered already when the investigation is planned, so that the expected total benefit, or cost, of the project can be estimated.

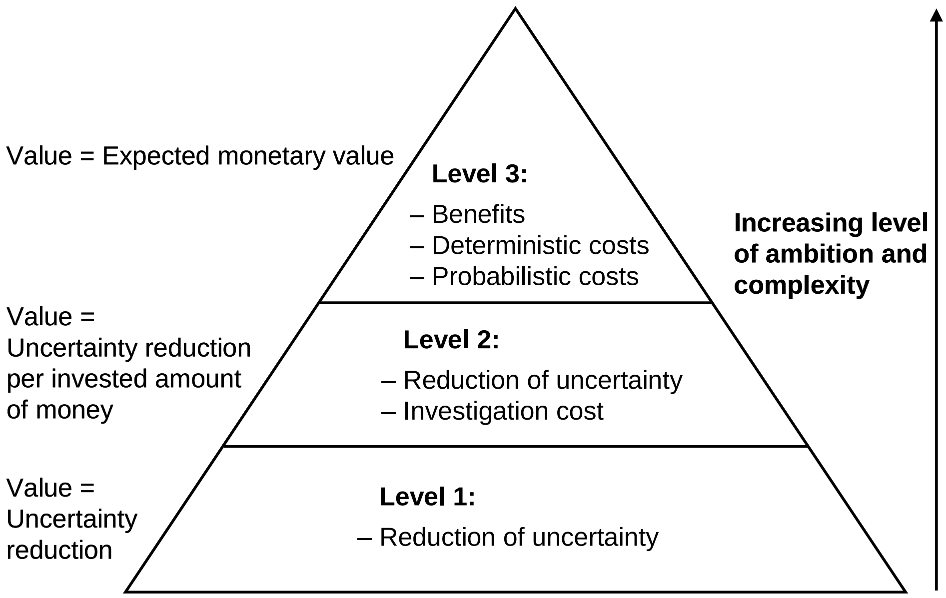

Figure 2. Three complexity levels of Value of Information Analysis for geo-environmental problems. Level 3 is the most complex one and represents the highest ambition.

A general framework for cost-effective site investigations is proposed in Figure 3 (Back, 2006a). It is based on the frameworks for Risk-Cost-Benefit decision analysis and DWA presented by Freeze *et al*. (1990; 1992). The VOIA framework was developed for Level 3 VOIA but can be applied also for Level 1 and Level 2 analyses. This makes the framework more general than the DWA framework presented by Freeze *et al*. (1992).

The framework consists of seven basic models, or modules, of varying complexity: (1) the Landuse Scenario, (2) the Objective of Investigation, (3) a Conceptual Site Model, (4) a Data Collection Module, (5) a Prior Information Module, (6) an Uncertainty Reduction Module, and (7) a Decision Model. In short, the principle of the framework is as follows. The Landuse Scenario influences the Conceptual Site Model, the Objective of Investigation, and defines what constitutes an unacceptable situation. The Objective of Investigation dictates what type of information the site investigation should supply and specifies the statistical parameter on which the decision should be made. The Conceptual Site Model describes the geological setting and forms the basis for parameter estimates of geological and hydrogeological properties in the other models.

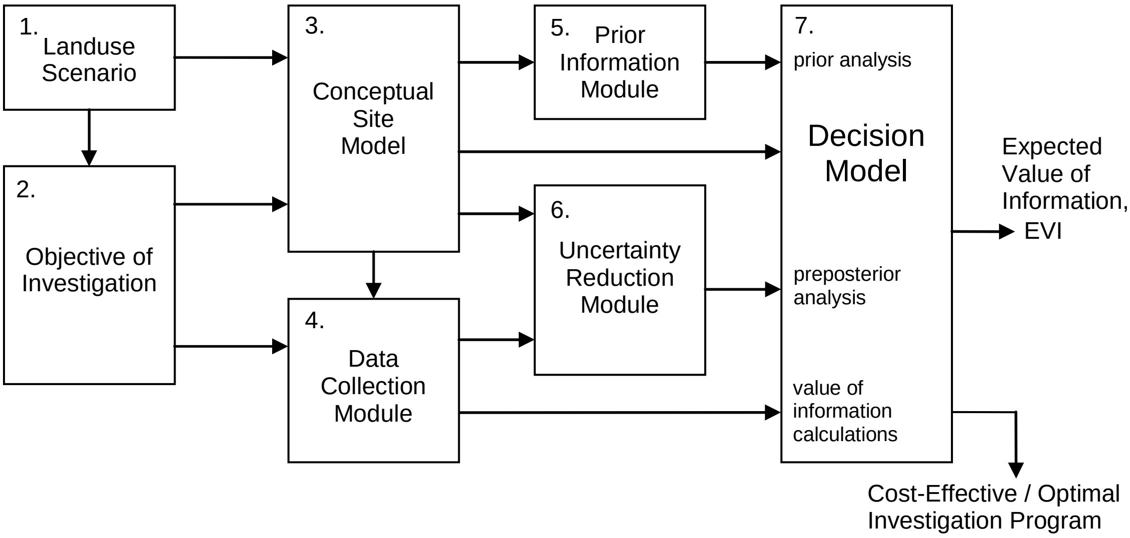

Figure 3. *A general framework for cost-effective site investigations based on Value of Information Analysis. The framework is structured based on the functions that are required in the analysis. The arrows indicate the main flows of information between the modules. The size of a box does not correspond to its importance.*

The Data Collection Module defines how and where data should be collected, the data uncertainty, and the cost of the investigation. In the Prior Information Module the prior knowledge about the site (contamination, transport properties etc.) is transformed into probability density functions (PDFs) and probabilities, to be used by the Decision Model. The approach is Bayesian in the sense that soft information is used for defining prior probabilities, which are later updated with new information. The purpose of the Uncertainty Reduction Module is to estimate the reduction of uncertainty that is expected when data from the site investigation becomes available. Finally, a Decision Model is applied to estimate the value of the site investigation. This is performed by comparing the results from a prior analysis and a preposterior analysis, see Eq. 2.

The framework was designed for site investigations in remediation projects, but in principle, it is general for all geo-environmental site investigation problems. A typical procedure for VOIA based the framework is as follows:

1. Define the problem and the investigations to be analysed.
2. Develop a VOIA model based on the presented framework.
3. Perform a VOIA with the model.
4. Choose the most cost-effective investigation program and perform the investigation.

Note that the design of a VOIA model is problem-specific to a large extent. Therefore, each investigation objective will require a separate VOIA model.

Box 4. Laboratory vs field-screening

As an alternative to laboratory analysis we will now calculate the value of a field-screening method, the results of which are not as reliable as laboratory analyses but its cost is lower. The prior analysis is still as in Box 1, but now $P(D|C) = 0.35$ and $P(D|C') = 0.15$ (cf Box 2). Assume also that $C_P = 0.01\, C_R$. Proceeding as in Boxes 2 and 3, we obtain

$$\Phi_{\text{preposterior}} = E\Phi_{\text{posterior}}$$
$$= 0.2 \cdot 0.21\, C_B = 0.042\, C_B = 0.14\, C_R$$

so

$$\text{EVI} = 0.14\, C_R \quad \text{and} \quad \text{ENV} = 0.13\, C_R$$

Thus laboratory analyses are more cost-effective than field-screening. In fact, even if the latter were gratis, laboratory analyses would be more cost-effective.

This, perhaps somewhat surprising fact, is due to the fact that it is assumed that exactly the same sampling programme is carried out in the two cases. Often, there is a fixed budget for sampling, and more samples are screened in the field. This may or may not compensate for the lacking reliability of the field-screening method.

Practical application in remediation projects

Application in different project phases

A remediation project is generally divided into a number of propject phases, defined slightly differently in different countries. The main questions are different in each project phase. Dta collection can take place in many phases of a remediation project but the type and objective of an investigation can be quite different. Generally, a range of different objectives are common in early phases of a project, such the preliminary study. The multiobjective nature of the preliminary investigation complicates VOIA; each investigation objective requires a specific VOIA model. To develop VOIA models for the preliminary investigation is therefore a challenge. An example of a VOIA model for early project phases is presented in Back (2006b).

Despite the problems of developing VOIA models for the preliminary study, VOIA has a large potential in this project phase. Traditionally, investigation programmes in early phases are designed largely using intuitive methods. However, as the problem complexity increases, the efficiency of intuitive appraisal decreases more rapidly than for formal analysis (Keeney, 1982). It may thus be more useful to analyse say 60% of a complex problem in an early phase of a project than 90% of a simpler problem in a later phase.

The application of VOIA has considerable potential in the main study phase, as this is the phase where investigation costs are the highest and the potential benefit of cost-effective investigations is therefore the greatest. In large remediation projects, the investigation in the main study is often performed in a step-wise manner. This provides an opportunity to use VOIA as a tool to decide when data collection should terminate; see the stopping rule approach below.

In the remediation investigation, the objective can be to classify soil into a specified number of contamination classes. Each contamination class is later treated differently in the implementation phases (different remediation techniques etc.). Classification of soil is a precisely defined sampling objective, which is an ideal situation for applying VOIA. Examples of VOIA for remediation investigations of soil are presented in Back (2006c) and Back *et al*. (2006).

There is also a potential for VOIA of investigations during the implementation of the remediation activities, especially for the investigation objective of determining when contaminant concentrations have been reduced to acceptable levels. The VOIA model in Back (2006c) can be applied relatively easily to this problem if the remediation goal is based on the mean concentration. VOIA can also be used for the design of monitoring programmes. Its potential is greatest for monitoring programmes with precisely defined objectives, and where the consequences are easily determined. One such example is a monitoring programme with the objective of determining whether degradation of a contaminant takes place or not, and where lack of degradation has monetary consequences.

Toolkit for Value of Information Analysis

The successful use of VOIA in remediation projects requires a toolkit of VOIA models for the various sampling objectives in the different project phases. Every investigation objective requires a unique VOIA model. Ferguson and Abbachi (1993) formulate a hope that sampling models "*...will be treated as tools in the site-assessor's toolbag – to be used when appropriate, to be modified if necessary to suit the circumstances of a particular site, or to be left in the toolbag when other tools are more effective*". This also applies to VOIA models.

Because the value of information is problem-specific, variants of the models are required so that a model that matches a specific objective exactly can be applied. For example, the search-sampling model in Back (2006b) was developed for a random sampling design, but systematic sampling would require a modified model. Similarly, the model in Back *et al*. (2006) was developed for composite sampling, where a composite sample is formed from increments located by simple random sampling. Systematic sampling would require a slight modification of the model.

Discussion

VOIA is based on decision analysis and there are many useful by-products of such analyses (Keeney, 1982). Here, four different types of application of VOIA for site investigations are described:

1. Evaluation of cost-effectiveness
2. Stopping rule
3. Optimal number of samples
4. Indirect valuation of benefit

In addition to these, it is also possible to perform some kind of qualitative VOIA without statistical calculations and economic valuations, based solely on experience and expert opinion.

Evaluation of cost-effectiveness

The expected net value (ENV) can be used to compare and identify the most cost-effective data collection program among a set of alternatives. This type of analysis is likely to have the largest potential in the preliminary study and the main study of remediation projects. For example, it can be applied to determine which type of data collection technique is the most cost-effective in the preliminary investigation: sampling/laboratory analyses of soil samples, or field-screening methods. Evaluation of the cost-effectiveness of site investigations is demonstrated in Back (2003), Back (2006b), Back (2006c) and Back *et al*. (2006).

Stopping rule

VOIA can also be used to determine when sufficient amount of data have been collected, i.e. VOIA is used as a stopping rule (Freeze *et al*., 1992). This type of application requires an iterative site investigation strategy. Figure 4 illustrates the principle. First, a prior analysis is performed to identify the best decision option (alternative action) based on existing information. Secondly, the cost-effectiveness of the sampling program is calculated, after a preposterior analysis. Sampling will stop if the sampling program is not cost-effective. In that case, the best alternative action according to the prior analysis is selected. On the other hand, if the sampling program is cost-effective it is carried out. The new information from the sampling is used to update the present state of knowledge in a posterior analysis (Bayesian updating). Then, a new investigation program is evaluated by a preposterior analysis to determine its cost-effectiveness. This iterative loop continues until no additional sampling is cost-effective.

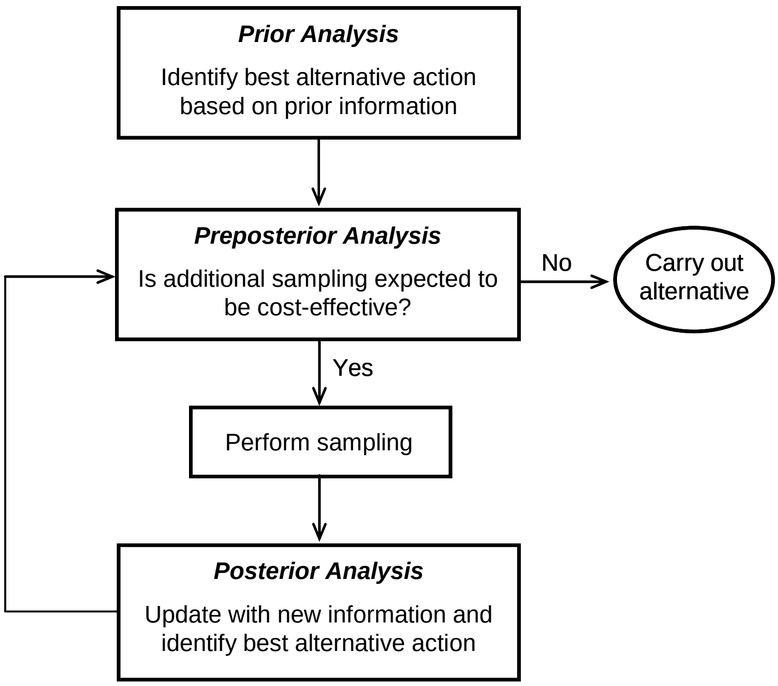

Figure 4. Value of information analysis applied as a stopping rule in an iterative manner (after James and Gorelick, 1994).

Application of the stopping rule approach is natural if sampling is performed several times at a site in an iterative manner. Its largest potential for application is in the main study of large remediation projects.

Optimal number of samples

The value of a sampling program depends on the number of samples. There are two resulting effects from increasing the number of samples: (1) the EVI increases towards a limiting value, and (2) the investigation cost increases. Because the ENV is the difference between EVI and investigation cost the change in ENV depends on if the increase in investigation cost is larger or smaller than the increase in EVI. The ENV reaches a maximum at a specific number of samples: the optimal number of samples. The principle is illustrated in Figure 1.

Estimation of the optimal number of samples by means of VOIA is demonstrated in Back (2003), Back (2006b), Back (2006c) and Back *et al*. (2006). This type of analysis has a large potential for application in remediation projects.

Indirect valuation of benefit

The framework for VOIA offers an interesting possibility to estimate how the benefit of remediation is valued when 'traditional' site investigations are designed. The assumption is that there must be some underlying philosophy, or perspective, that forms the basis for how many samples it is recommended to collect during a site investigation. Such recommendations can be found in sampling standards, reports from environmental agencies etc. The number of samples that is collected during a 'traditional' site investigation is usually regarded as sufficient by the involved parties, if the recommendations are followed. This implies that these 'traditional' site investigations are regarded as optimal, or at least cost-effective, by the involved parties, otherwise some other sampling program would have been used instead.

It is possible to analyse these 'traditional' data collection programs and make an inverse calculation to estimate the benefit of remediation indirectly. The principle is to find the benefit-value that makes the 'traditional'sampling program best resemble the optimal sampling program; see Figure 5. First, an optimal sampling program is designed with a VOIA model[6]. Next, the same model is applied to the 'traditional' sampling program. The difference in expected benefit, or cost, between the two sampling programs is plotted as a function of k_B/k_R, i.e. the ratio between the benefit of remediation[7] and the remediation cost. The k_B/k_R-value where the difference is at a minimum gives an indication of how the benefit of remediation is valued in 'traditional' sampling.

Indirect valuation of the benefit of remediation was performed in Back (2006b) and Back *et al*. (2006). Results indicate that the benefit of remediation (or cost of failure) is regarded as relatively low; in the order of 1.2 to 3 times the remediation cost based on the two case studies (both are regarded as typical Swedish contaminated land problems). The rather low benefit can be interpreted in at least three different ways:

[6] Two different VOIA models are used in Figure 5, with slightly different results; see Back *et al*. (2006).

[7] Benefit of remediation k_B can be replaced by failure cost k_F if the analysis is based on a No Risk baseline condition, as in Figure 5.

1. The benefit of remediation resembles the relatively low risks to humans and the environment at contaminated sites in Sweden. Low risks imply low benefits when a site is remediated.
2. The benefit of remediation resembles the decision-maker's attitude towards contamination risks: Leaving some contamination in the ground is not regarded as a big problem, as long as some action is taken at the site.
3. The precision in traditionally designed sampling is overestimated, which results in a situation where the decision-maker falsely believes that all contamination has been located when a traditional but insufficient sampling program is carried out. This insufficiency is revealed in a VOIA, and the indirectly estimated benefit of remediation is therefore low.

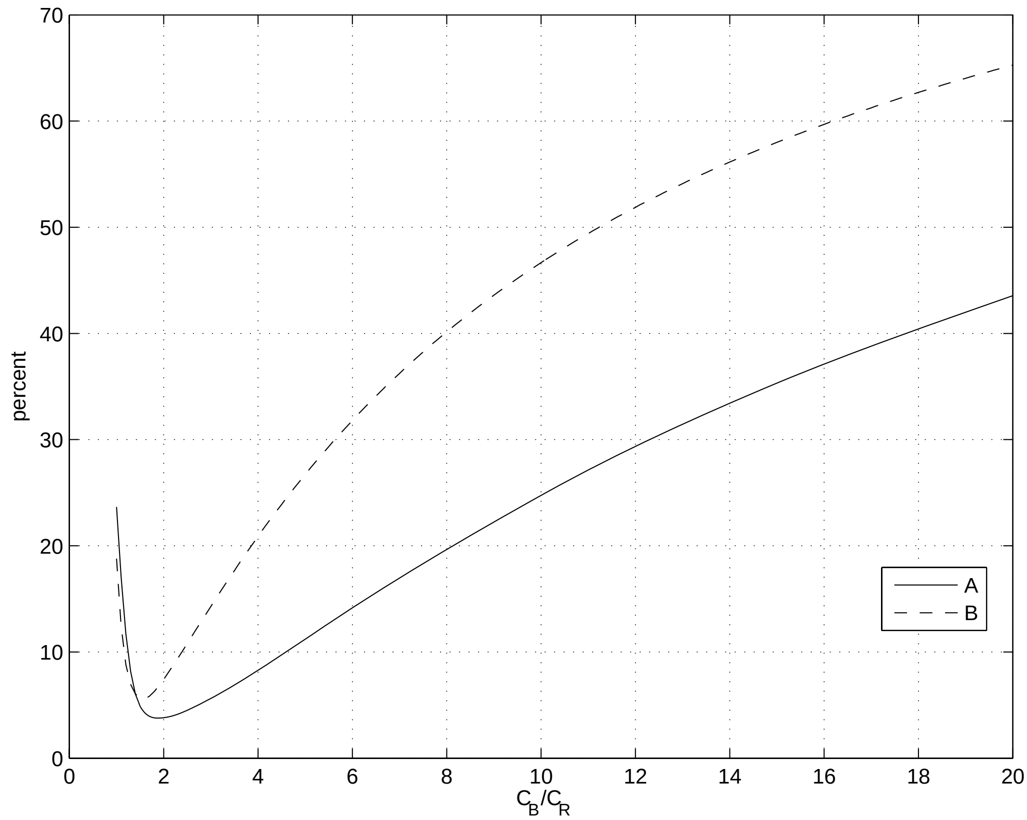

Figure 5. *The difference in expected project cost (percent) by using an optimal sampling program instead of a 'traditional' one. The difference is lowest when k_F/k_R is about 1.5 – 2 for both VOIA model A and B, i.e. the failure cost (k_F) is approximately 1.5 – 2 times the remediation cost (k_R).*

It should be stressed that this type of analysis only can give a rough indication of the benefit of remediation. Still, it is an interesting way to communicate the underlying value basis of 'traditional' sampling. A similar and interesting application is for communicating the different perspectives of the involved parties and their effect on site investigations (Back and Rosén, 2006). One such example of application of VOIA is to communicate to clients the implications of small sampling budgets.

Conclusions

The main conclusions of this work (Back, 2006a) are:

- A framework for VOIA is presented and successfully applied on both early and late phases of remediation projects.
- The main strength of the methodology is that the methodology promotes clear thinking and forces the decision-maker to reflect on issues that otherwise would be ignored, which leads to a better understanding of the problem. The main weakness is the complexity of VOIA models.
- The potential for application of VOIA depends on in which phase of a remediation project the investigations are performed, and the size and complexity of the project. It is easier to apply the methodology during the late phases, when investigation objectives are few and precisely defined, than during early phases. Applying the methodology in small projects requires pre-developed and simplified VOIA models.
- VOIA is best suited for problems that can be broken down to one single quantifiable key issue, such as exceeding or not exceeding a threshold. Data collection in remediation projects aim at supplying information regarding a range of key issues, at least in early phases of a project. Therefore, VOIA models of early phases of a remediation project are bound to be complex.
- The value of information is expressed as an EMV. Use of an expectation is reasonable from a point of view of maximising utility in the long run, but in a single project the decision will only be taken once, and so there is no 'long run' (Bedford and Cooke, 2001). Therefore, an ideal situation is to use VOIA to design sampling strategies on a national level, so that the 'long run' condition is fulfilled. This will maximise the benefit to society, and thus promote sustainable development. Organisations or companies responsible for a large number of contaminated sites may also benefit from using VOIA, for the same reason.
- In very large and complex projects it may be difficult or impossible to properly model how site investigations are coupled to the total project cost. In such cases, a simplified VOIA can be performed and the value of an investigation could be quantified in another way than as EMV. One such unit of value is the quotient of uncertainty reduction and investigation cost.
- Because value of information is problem-specific, general conclusions are difficult to draw. However, the three most important factors in a VOIA were found to be: (1) the objective of the investigation, (2) the landuse, and (3) the benefit of remediation.
- The benefit of remediation is a difficult variable to estimate, because it depends strongly on the perspective of the decision-maker and how the economic valuation is performed. However, the ratio between the benefit and the cost of remediation was found to be a useful variable for studying its effect on the result.

- The framework for VOIA can be used also for other types of geo-environmental problems than contaminated land. Potential applications include geotechnical investigations, investigation of properties in rock mechanics, design of sampling programs for waste, design of environmental monitoring programs etc.

Further development of VOIA for remediation projects can proceed in four different directions: (1) Modelling of the full complexity of remediation problems, by developing computerised tools (software tool-kit) for the various investigation objectives, (2) Development of strategies on how to make simplifications in sampling problems, and how VOIA models can be simplified without loosing their strengths, (3) Development of alternative VOIA approaches, such as the Level 2 VOIA presented in this thesis, and (4) Development of a framework for qualitative VOIA based on experience and expert opinion.

It is likely that remediation projects would benefit from all these strategies, depending on the complexity of the problems and the size of the project. A combination of strategy one and two is recommended. Practical application of VOIA on a regular basis in remediation projects requires that the tool-kit of VOIA models is expanded, and that correct simplifications are made.

Suggested review persons:
Thomas Öberg, Högskolan i Kalmar (thomas.oberg@hik.se)
Mitchell Small, Carnegie Mellon University (ms35@andrew.cmu.edu)
Sverker Molander, Chalmers tekniska högskola (sverker.molander@chalmers.se)
Tom Purucker, University of Tennessee (SADA-group) (sada@tiem.utk.edu)

References

Abbaspour KC, Schulin R, Schläppi E and Flühler H. 1996. A Bayesian approach for incorporating uncertainty and data worth in environmental projects. Environmental Modeling and Assessment, 1996(1): 151-158.

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.

Back P-E. 2006a. Value of Information Analysis for Site Investigations in Remediation Projects. PhD Thesis No. 2551, Chalmers University of Technology, Göteborg.

Back P-E. 2006b. A Model for Estimating the Value of Sampling in the Early Phases of Contaminated Land Investigations. Submitted to *Journal of Environmental Engineering* 2006-11-14.

Back P-E. 2006c. A Model for Estimating the Value of Sampling Programs and the Optimal Number of Samples for Contaminated Soil. Accepted for publication in *Environmental Geology* 2006-08-24.

Back P-E, Norberg T and Rosén L. 2006. On the Cost-Effectiveness of Sampling Contaminated Soil: Comparison of Two Models. Submitted to *Environmetrics* 2006-07-12.

Back P-E and 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; 68-75.

Bedford T and Cooke R. 2001. Probabilistic Risk Analysis. Cambridge University Press: Cambridge.

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

CEN. 2005. Characterization of waste - Sampling of waste materials - Framework for the preparation and application of a Sampling Plan. EN 14899:2005, European Committee for Standardization, Brussels.

Cochran WG. 1977. Sampling Techniques. John Wiley & Sons: New York.

Cox LA. 1999. Adaptive Spatial Sampling of Contaminated Soil. Risk Analysis, 19(6): 1059-1069.

Dakins ME, Toll JE and Small MJ. 1994. Risk-Based Environmental Remediation: Decision Framework an Role of Uncertainty. Environmental Toxicology and Chemistry, 13(12): 1907-1915.

Dakins ME, Toll JE, Small MJ and Brand KP. 1996. Risk-Based Environmental Remediation: Bayesian Monte Carlo Analysis and the Expected Value of Sample Information. Risk Analysis, 16(1): 67-79.

Dawdy DR. 1979. The Worth of Hydrologic Data. Water Resources Research, 15(6): 1726-1732.

Freeze RA, Bruce J, Massman J, Sperling T and 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 and 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.

Hintze S. 1994. Risk analysis in foundation engineering with application to piling in loose friction soils in urban situations. TRITA-JOB, Report No. 94/2 Thesis, Royal Institute of Technology, Stockholm.

IT Corporation. 1997. Value of Information Analysis for Corrective Action Unit no. 98; Frenchman Flat. DOE/NV/13052-T1, U.S. Department of Energy, USA.

James BR and Freeze RA. 1993. The Worth of Data in Predicting Aquitard Continuity in Hydrogeological Design. Water Resources Research, 29(7): 2049-2065.

James BR and Gorelick SM. 1994. When Enough is Enough: The Worth of Monitoring Data in Aquifer Remediation Design. Water Resources Research, 30(12): 3499-3513.

James BR, Gwo J-P and Toran L. 1996a. Risk-Cost Decision Framework for Aquifer Remediation Design. Journal of Water Resources Planning and Management, 122(6): 414-420.

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

Kaplan PG. 1998. A Tutorial: A real-life adventure in environmental decision making. U.S. Department of Energy, The Center for Risk Excellence. 2000-08-21. http://www.doe.gov/riskcenter/

Keeney RL. 1982. Decision Analysis: An Overview. Operations Research, 30(5): 803-838.

Lacasse S and Nadim F. 1996. Uncertainties in characterising soil properties. In Uncertainty in the Geologic Environment: From Theory to Practice. Schackelford CD, Nelson PP and Roth MJS (Editors), American Society of Civil Engineers: Madison, WI; 49-75.

Lindley DV. 1997. The choice of sample size. The Statistician, 46(2): 129-138.

Martignon LF. 2001. Principles of Adaptive Decision-making. In Rational Decision-making in the Preservation of Cultural Property. Baer NS and Snickars F (Editors), Dahlem University Press: Berlin; 263-275.

McNulty G, Deshler B, Dove H and IT Corporation. 1997. Value of Information Analysis; Nevada Test Site. CONF-970335-37, USA.

Minasny B and McBratney AB. 2002. The efficiency of various approaches to obtaining estimates of soil hydraulic properties. Geoderma, 107(2002): 55-70.

Morgan MG and Henrion M. 1990. Uncertainty, A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press.

Norberg T and Rosén L. 2006. Calculating the optimal number of contaminant samples by means of data worth analysis. Environmetrics, 17(7): 705-719.

Norrman J. 2004. On Bayesian Decision Analysis for Evaluating Alternative Actions at Contaminated Sites. PhD Thesis No. 2202 Thesis, Chalmers University of Technology, Göteborg.

Perman R, Ma Y, McGilvray J and Common M. 2003. Natural Resource and Environmental Economics. Pearson Education Ltd: Harlow, UK.

Rautman CA, McGraw MA, Istok JD, Sigda JM and Kaplan PG. 1994. Probabilistic comparison of alternative characterization technologies at the Fernald Uranium-in-Soils Integrated Demonstration Project. In Waste Management -94. Tucson, Arizona; 2117-2124.

Russell KT and Rabideau AJ. 2000. Decision Analysis for Pump-and-Treat Design. Groundwater Monitoring & Remediation,(Summer 2000): 159-168.

Sturk R. 1998. Engineering Geological Information - Its Value and Impact on Tunnelling. PhD Thesis, Royal Institute of Technology, Stockholm.

Swedish EPA. 2006. Official website. Swedish Environmental Protection Agency. 2006-10-10. http://www.naturvardsverket.se