Seven Guidelines for Managing Uncertainty in Geoenvironmental Design

Robert B. Gilbert[1] and Travis C. McGrath[2]

ABSTRACT: Seven, practical guidelines for managing uncertainty in geoenvironmental design are presented. They range in scope from evaluating uncertainty to reducing it and limiting its impact on design performance. The guidelines are derived from theoretical considerations, but do not require an expertise of (or even a background in) the theory to be employed. Simple tools are provided and a design case history is analyzed to facilitate the implementation of these guidelines in practice.

INTRODUCTION

Uncertainty is significant in geoenvironmental design: site conditions vary spatially and with time; field-scale, long-term performance is not known for many engineered systems; and costs and schedules cannot be predicted with accuracy. There have been numerous attempts at applying decision theory and probability theory to account for uncertainty in geoenvironmental design. Specific applications include groundwater remediation (e.g., Sitar et al. 1987; Freeze et al. 1990; Massmann et al. 1991; Woldt et al. 1991), waste containment (e.g., Tang et al. 1994; Benson and Daniel 1994; Gilbert and Tang 1995b) and nuclear waste management (e.g., Roberds et al. 1991; Keeney and von Winterfeldt 1993). However, most of the theoretical work is confined to the world of research, and a large gap currently exists between uncertainty management in theory and practice. It is unreasonable to expect all practitioners to become experts in the theory so that they can apply it in practice. It is also unreasonable to expect that an expert will be involved in every geoenvironmental design project.

The goal of this paper is to present practical guidelines for uncertainty management that are derived from theory. Seven guidelines will be presented: three of the guidelines deal with evaluating uncertainty and its impact on geoenvironmental design, while the remaining four guidelines deal with

[1] Asst. Prof. of Civil Engrg., The University of Texas at Austin, Austin, TX 78712
[2] Grad. Res. Asst., The University of Texas at Austin, Austin, TX 78712

managing the uncertainty. Simple tools will be provided to facilitate the implementation of these guidelines in practice. A design case history will be used to demonstrate the guidelines and to illustrate key points. Appendix I provides background information for the case history; this information will be referenced in discussions throughout the remainder of the paper.

GUIDELINE 1

$\boxed{\text{Judgment is essential in evaluating and representing uncertainty.}}$

Uncertainty means that the outcome of a design (i.e., how it will perform) is not known at the outset. Although all uncertainty is ultimately due to a lack of knowledge, it is convenient in practice to consider two types of uncertainty: random uncertainty and systematic (or model) uncertainty.

Random Uncertainty

Random uncertainty results from variability with time and/or location. Examples include the daily precipitation at a site (varying from day to day), the hydraulic conductivity of a soil layer (varying with location), or the ratio between an actual construction cost and the estimated cost (varying from project to project). Because of this variability, there will be uncertainty in predicting new values (e.g., at future times or different locations).

A random variable model provides a practical tool to quantify variability. This model describes the frequencies with which different values of the variable will occur. For example, a random variable model for the ratio of actual cost to estimated cost in site remediation is shown as a probability distribution on Fig. 1a. This model was developed from data supplied by the U. S. Environmental Protection Agency for 102 Superfund sites (Davis 1996). A random variable model can be continuous (the variable can take on any value) or discrete as in Fig. 1a (the variable can only take on discrete values). In many instances, a discrete model is sufficient for practical purposes even though it may not be entirely realistic. There are several useful properties of a random variable model, as shown on Fig. 1a: the expected or mean value, μ (a measure of the central value for the variable), and the standard deviation, σ (a measure of the variability about the mean value). Since random uncertainty is due to variability, the standard deviation indicates the magnitude of random uncertainty.

Systematic (or Model) Uncertainty

Systematic (or model) uncertainty results from uncertainty in the random variable models used to represent random uncertainty. Rarely ever in geoenvironmental design will sufficient information exist to formulate the complete probability distribution for a random variable model. For example, the mean ratio between actual cost and estimated cost (Fig. 1a) is uncertain because it is based on data from only 102 sites.

Uncertainty in a random variable model is systematic in the sense that it does not vary (e.g., there is only one mean even though its exact value is not known).

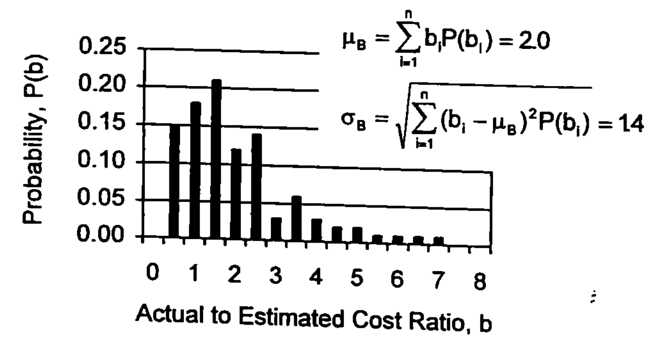
Systematic uncertainty is typically the dominant source of uncertainty in geoenvironmental design due to the scarcity of data. For example, remedial actions have been implemented at only a small percentage of the listed Superfund sites (Russell and Davis 1995). Also, design performance data from five years ago may not be relevant today due to improvements in technology or regulatory changes. Therefore, judgment is essential in evaluating systematic uncertainty.

Theory provides some guidance in the application of judgment. First, model uncertainty, like random uncertainty, can be represented using a probability distribution. For example, a probability distribution representing systematic uncertainty in the mean cost ratio is shown on Fig. 1b. Note that each possible mean cost ratio has an associated distribution accounting for random uncertainty. The probabilities in Fig. 1b do not represent the frequency of occurrence, but rather the likelihood that the mean is equal to different values.

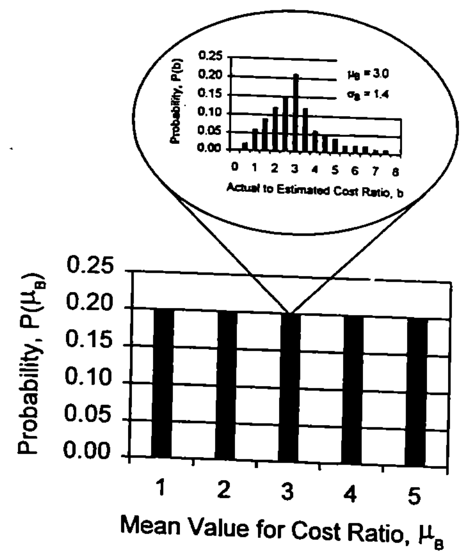
Second, the most important role of judgment is to identify all possible outcomes. The probabilities in Fig. 1b represent the likelihoods relative to all possible outcomes (i.e., a mean of 1, 2, 3, 4 or 5). Failure to consider a possible outcome (e.g., a mean cost ratio of 10) can lead to ineffective design decisions. For example, many pump-and-treat systems were implemented in the 1980s without recognizing their ineffectiveness at removing non-aqueous phase liquids (NRC 1994).

Third, another role of judgment is to impose realistic constraints on the probability distribution. Some outcomes, such as a negative construction cost, may not be possible. Information theory provides guidance on appropriate probability distributions to represent systematic uncertainty within constraints imposed by judgment; an appropriate distribution is one that minimizes unintended biases (e.g., Tribus 1969). Suggested probability distributions for typical constraints are summarized in Table 1. As an example, a discretized version of a uniform distribution (Table 1) is used to model systematic uncertainty in the mean cost ratio (Fig. 1b) because upper and lower bounds can be assigned based on judgment.

Finally, judgment must be applied with care. Cognitive theory suggests potential problems in quantifying judgment into a probability distribution. The most common problems are underestimating the magnitude of uncertainty and imposing artificial biases (e.g., Capen 1976). Formal techniques for subjective assessments are available to minimize these effects (e.g., Roberds 1990). However, most of the problems can be controlled simply by being aware that they exist and clearly defining the parameters to be estimated.



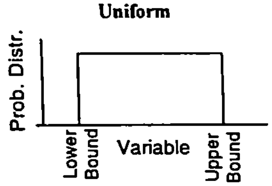
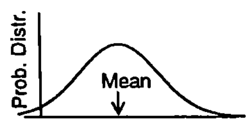
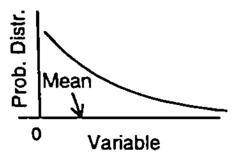
(a) Random Uncertainty in Cost Ratio



(b) Systematic Uncertainty in Mean Cost Ratio

Fig. 1. Random Variable Model for Ratio of Actual to Estimated Cost in Site Remediation

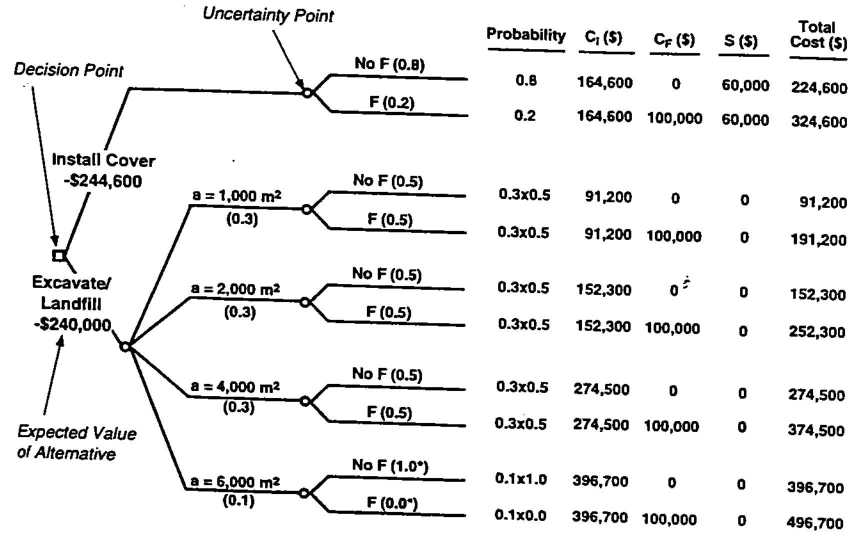
Table 1. Suggested Probability Distributions Given Constraints Imposed by Judgment

Constraints Imposed by Judgment	Suggested Probability Distribution	Significant Features
Upper and Lower Bounds	**Uniform** 	All possible values are equally likely Mean value at mid-point of lower and upper bounds Symmetrical distribution about mean value
Mean Value and Standard Deviation	**Normal** 	68 % of values are within ±1 standard deviation from mean 95 % of values are within ±2 standard deviations from mean Symmetrical distribution about mean value
Positive Values and Mean Value	**Exponential** 	63 % of values are less than mean Asymmetrical distribution

GUIDELINE 2

The outcome of a design (e.g., satisfactory or unsatisfactory performance) will be uncertain. The value of a design is related to the consequences of different outcomes weighted by the likelihoods that those outcomes occur.

Design constitutes a series of decisions that are made under uncertainty and, therefore, risk. A decision tree, as shown on Fig. 2 for the case history site at the completion of Phase I soil sampling (Appendix I), provides a useful tool for analyzing and discussing the design process. For each design alternative, different outcomes can occur. An outcome may be described by a single event (e.g., the cover performs satisfactorily and no future corrective action is required) or a series of events (e.g., 1,000 m^2 of contaminated soil is excavated to a depth of 0.5 m and there is no residual contamination or need for future corrective action).



F = Future corrective action required
a = Area of contaminated soil requiring remediation (to a depth of 0.5 m) based on sampling during excavation
C_I = Cost of implementation (construction, maintenance and monitoring)
C_F = Cost of future corrective action
S = Cost of restricting future land use at the site
* Entire work area is excavated - chance for residual contamination assumed negligible

Fig. 2. Decision Tree for Case History Site

Each design outcome has a likelihood of occurrence and an associated consequence. The likelihood of occurrence for an individual event is indicated by the values in parentheses on Fig. 2. As an example, the likelihood that 1,000 m^2 of contaminated soil will have to be excavated is equal to 0.3 based on existing data. The likelihood of occurrence for an outcome is obtained by multiplying the likelihoods together for all events leading to that outcome. The likelihood that 1,000 m^2 of soil will have to be excavated and that no residual contamination exists is 0.3x0.5 = 0.15 (Fig. 2). The consequence for each possible design outcome is expressed as a total cost on Fig. 2. Estimation of the likelihoods and consequences that are used in the decision tree is discussed in the next section.

The expected (or representative) value of a decision is obtained by weighting each possible consequence by its likelihood

$$\text{Expected Value} = E(\text{Value}) = \sum_{i=1}^{n} (\text{Consequence})_i \times p_i \quad (1)$$

where p_i is the probability that outcome i occurs, $(\text{Consequence})_i$ is the consequence associated with outcome i, and n is the number of possible outcomes. The expected value of a design alternative provides a convenient measure to compare the value of different designs. If benefits are positive and costs are negative, then the design with the maximum expected value will have the maximum value. In the decision tree shown on Fig. 2, the excavate alternative has the maximum (least negative) expected value. The expected value concept is also helpful in simplifying design decisions. Design outcomes that do not have large consequences or that are not likely are less important, and can possibly be neglected.

GUIDELINE 3

> **Design decisions should be based only on outcomes that are tangible and consequences that can be measured. Sensitivity analyses should always be conducted to evaluate how estimated quantities affect the value of different design alternatives.**

A major difficulty in design is quantifying the likelihoods and consequences of different design outcomes. The key to this process is defining outcomes that are tangible. If an outcome is not tangible, then it is impossible to assess its likelihood of occurrence (one would never know if the outcome had or had not occurred) and it is impossible to evaluate its consequences (consequences could never be linked directly to the outcome). While this guideline is seemingly obvious, it is frequently ignored in practice. For example, an increased risk of cancer in humans due to a site with groundwater contamination is not a tangible outcome, although it is commonly used as a design guideline (e.g., USEPA 1989). Even at the infamous Love Canal site, which is probably the most studied case of site contamination, direct links have never been drawn between environmental contamination and human health effects (e.g., Danzo 1988; Hoffman 1995). An alternative outcome that is tangible would be the contaminant concentration at an exposure point. The case history site (Appendix I) also provides an example of decision outcomes that are not tangible. An example set of tangible outcomes for this site is provided in Table 2.

Table 2. Example Set of Tangible Outcomes for Case-History Site

Original Outcomes Considered	Alternative Outcomes that are Tangible
Overall protection of human health and the environment Long-term effectiveness and permanence Reduction of toxicity, mobility or volume of contaminant Short-term effectiveness	No future corrective action required if mercury concentrations are kept below: 1) 20 mg/kg in the upper 0.5 m of on-site soils and sediments from drainage ditches; and 2) 2 µg/l in groundwater samples taken from wells completed in the shallow aquifer
Implementability	Minimize time to complete remediation
Cost	Minimize total expected cost

Once tangible outcomes are defined, there may still be considerable uncertainty in the likelihoods and the consequences associated with those outcomes. The most rigorous approach to account for this uncertainty is to develop full probability distributions representing both random and systematic uncertainties in these parameters. However, in general, uncertain likelihoods and consequences can be represented simply with estimates for their mean values in design decisions. The value for a design alternative is then approximated from (1) as follows

$$E(\text{Value}) \cong \sum_{i=1}^{n} \overline{(\text{Consequence})}_i \times \overline{p}_i \quad (2)$$

where $\overline{p}_i$ is an estimate for the mean likelihood and $\overline{(\text{Consequence})}_i$ is an estimate for the mean consequence associated with outcome i. Estimation of these mean values is addressed in the following subsections.

Estimation of Mean Likelihoods

Uncertainty in the frequency of occurrence for an outcome arises from a lack of performance data. Consider a case in which a given outcome has been observed to occur x times in n designs. An estimate for the mean likelihood of occurrence, $\overline{p}_i$, is given by $\overline{p}_i \cong (x+1)/(n+2)$ (Ang and Tang 1975). For example, if a design is implemented 5 times and performs successfully each time, then the mean likelihood of unsatisfactory performance, $\overline{p}_i$, would be estimated as 0.14 (i.e., $x = 0$ failures). Note that the actual frequency of poor performance may be much smaller, but the limited data set does not support a smaller estimate for the mean likelihood. Also note that if no data are available (i.e., $n = 0$), then the estimate for the mean likelihood is 0.5, which is consistent with a uniform distribution bounded by 0 and 1 (Table 1). For the design decision on Fig. 2, it is difficult to estimate the likelihood that future corrective action will be required if the soil is excavated (i.e., the likelihood that contaminated soil is left on-site, leading to future problems). Since the data needed to make this assessment do not exist, the likelihood is estimated to be 0.5.

Performance data will be rare in most instances. While a 0.5 value can serve as a starting point for the mean likelihood, performance modeling can also provide information on the likelihood. The performance model, which can be physical, empirical, analytical or numerical, predicts performance of the design under different conditions. For example, a numerical model of unsaturated flow is used commonly to predict infiltration rates through covers as a function of environmental conditions and cover properties. If probability distributions for the model input parameters are known, then the frequency of different outcomes can be estimated with this model (e.g., the frequency of different infiltration rates). There will be uncertainty in these calculated frequencies because of systematic uncertainties in the input parameters and in the performance model itself (i.e., how well the model represents reality).

A simple chart to account for model uncertainty in estimating $\overline{p}_i$ is provided on Fig. 3. Z is a measure of performance, defined such that Z<0 indicates unacceptable performance (e.g., Z = a maximum allowable, annual infiltration rate minus the actual infiltration rate through a cover). The mean value of Z, μ_Z, is assumed to be greater than zero (e.g., the cover will typically be designed so that the average, annual infiltration rate is less than a maximum allowable rate). The standard deviation of Z, σ_Z, represents the random uncertainty in Z (e.g., due to variability in precipitation between years). Model uncertainty (e.g., uncertainty in the model used to estimate the infiltration rate) is represented by σ_M. The curves on Fig. 3, which are developed assuming that Z is normally distributed (Gilbert and Tang 1995a), can be used to estimate the mean likelihood of unacceptable performance as a function of σ_Z/μ_Z and σ_M/σ_Z. For example, if the annual infiltration rate has a mean of 50 mm/yr and a standard deviation of 20 mm/yr while the allowable infiltration rate is 80 mm/yr, then $\mu_Z = 80 - 50 = 30$ mm/yr and $\sigma_Z/\mu_Z = 20/30 = 0.67$. If there is no model uncertainty ($\sigma_M = 0$), then the mean likelihood of unsatisfactory performance in a given year is 0.07 (Fig. 3). If $\sigma_M/\sigma_Z = 1.0$, then the mean likelihood increases to 0.14 (Fig. 3). The ratio σ_M/σ_Z will typically be greater than one, and may approach values as large as five.

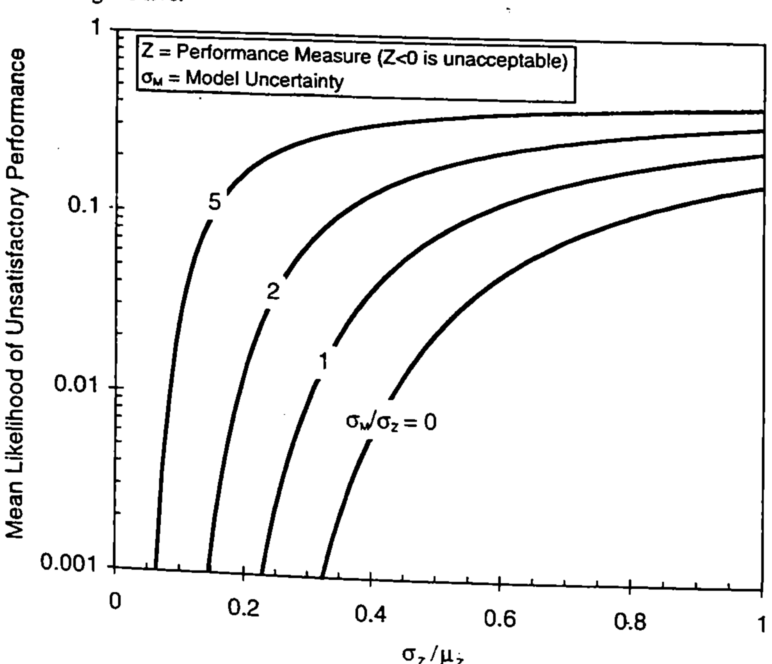


Fig. 3. Effect of Model Uncertainty on the Mean Likelihood of an Outcome

Estimation of Mean Consequences

The potential consequences of unsuccessful or successful designs include: design costs, construction costs, operation and maintenance costs, corrective action costs, legal costs, regulatory fines, environmental effects, human-health effects, worker safety, publicity, land re-use and good-will. There is uncertainty in estimating the consequences associated with a given design outcome (e.g., unsuccessful performance). The outcome may occur many years in the future or be drawn out over long periods of time. Also, an outcome may lead to multiple, incongruous costs and benefits that are difficult to quantify and combine into a single value.

Utility theory has been developed to manage the difficulties in quantifying consequences (e.g., Raiffa 1968). The premise of utility theory is to express all consequences, whether monetary or not, on a normalized utility scale so that they can be combined and compared. The advantage of the utility approach is its generality: different types of consequences can be handled (e.g., construction cost and aesthetic value) and different decision-maker preferences can be accommodated (e.g., a $10,000,000 loss may be much more significant to a small versus large company). The main limitation of utility theory is also related to its generality. Evaluation of utility values is subjective, which can affect the usefulness of this approach in practice. Many previous authors have discussed the problems in assigning utility values (e.g., Raiffa 1968; Clemen 1991) and in combining utility values from different decision makers (e.g., Merkhofer and Keeney 1987).

An alternative approach to utility that minimizes problems with subjectiveness is to express all consequences in measurable terms, such as economic terms as in Fig. 2. There are several advantages to this approach. It is easier to estimate consequences if they are expressed in concrete terms as opposed to abstract terms. Although tough issues such as aversion to large monetary losses and the intrinsic value of the environment still must be addressed, at least there is some basis for assigning an economic value in a capitalistic society. Also, it is easier to communicate consequences to decision makers and to the public if they are expressed in economic terms. Finally, if consequences can be measured, then the actual consequences of different outcomes can be determined over time after a design is implemented. In this way, actual consequences can be compared with estimated values to provide information for future designs. Therefore, it is recommended that design decisions be based not only on tangible outcomes, but also on consequences that can be measured.

Sensitivity Analyses

Even with tangible outcomes and measurable consequences, it can be difficult in practice to quantify likelihoods and consequences. Sensitivity analyses should always be performed to evaluate how uncertain quantities affect the value of different design options. For example, the following quantities are difficult to estimate in the decision tree shown on Fig. 2: the reliability of the cover system (i.e., the likelihood that it performs successfully in the long term); the cost of restricted land-use for the cover alternative, S; and the present-worth cost of future corrective action for either alternative, C_F (this cost is assumed to be equal for the two alternatives). The value of each design alternative has been evaluated as a function of these variables, and the results are shown on Fig. 4. Each curve for a given cover reliability represents combinations of S and C_F for which the two alternatives have equal expected costs. If the cover reliability is 0.8 and C_F is $100,000, for example, then the two designs have equal value for S equal to $55,400. For pairs of cost values above a particular cover-reliability curve, the excavation alternative minimizes the expected cost. For pairs below the curve, the cover alternative has the lower expected cost.

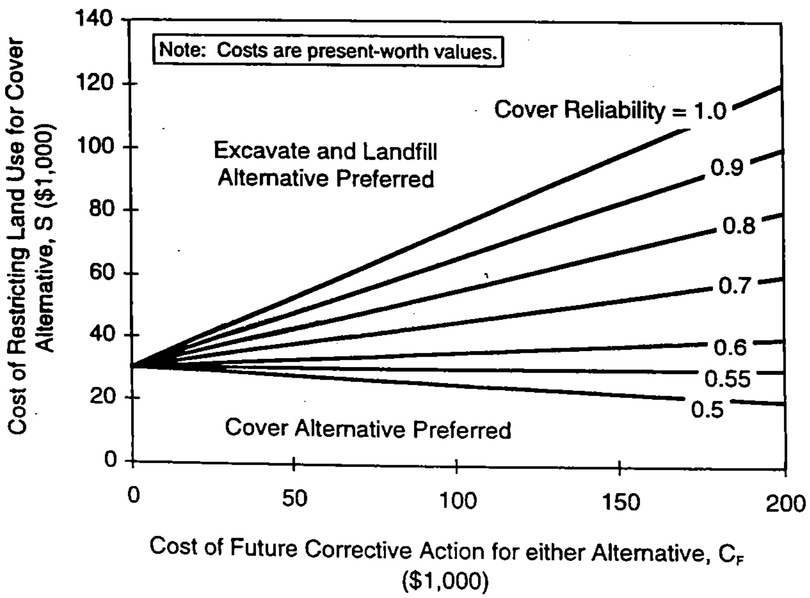


Fig. 4. Sensitivity of Decision to Uncertain Costs and Cover Reliability

There are several interesting conclusions that can be drawn from Fig. 4. The relative design values are insensitive to C_F if the cover reliability is equal to 0.55. At this reliability level, the excavate alternative will have the maximum value for all S values greater than $30,400. The sensitivity of the decision to the cost of future corrective action increases as the reliability level increases. The reliability of the cover is expected to be greater than 0.55. The primary mode of failure, gully erosion and subsequent contaminant outwash caused by a single, major storm event, is not very

likely. An upper-bound value for the reliability is 1.0. At this reliability level, the critical value for S is not much greater than $30,400 for small C_F values. However, as C_F approaches $200,000, the critical value of S approaches $120,000. Since a reasonable upper bound for C_F is $200,000, the excavate alternative will be preferred over the cover alternative if the cost of restricted land use is greater than $120,000, regardless of the cover reliability.

GUIDELINE 4

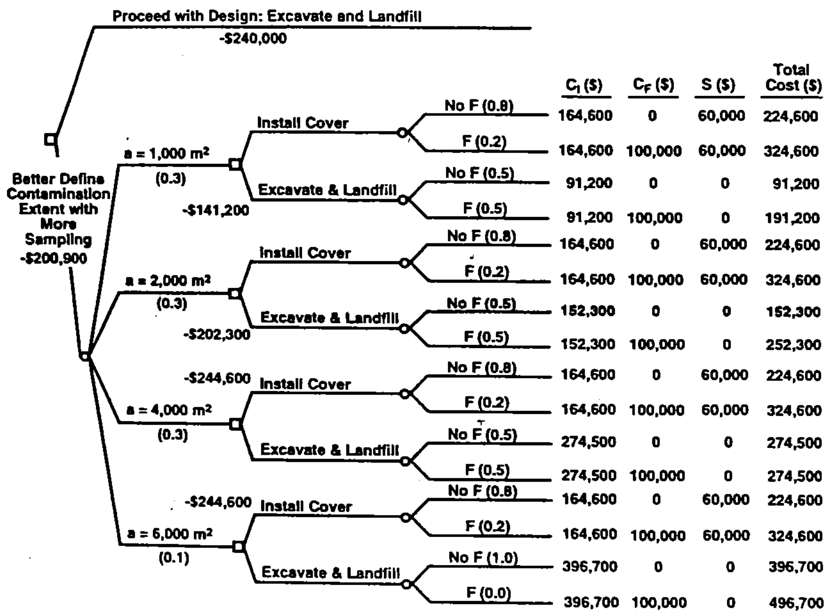
There are two approaches for managing uncertainty in design: (1) try to reduce the uncertainty by obtaining more information and (2) try to limit the impact of uncertainty on performance of the design.

Thus far, this paper has addressed the evaluation of uncertainty and the analysis of its effects on design. The remainder of the paper will deal with the management of uncertainty in design. There are two basic approaches for managing uncertainty: (1) reduce the uncertainty and (2) limit its impact on design performance.

An example of uncertainty management through uncertainty reduction is shown on Fig. 5 for the case history site. A reduction of uncertainty in the area of contaminated soil (and therefore the volume assuming an average excavation depth of 0.5 m) through additional sampling may prove valuable. The benefit of knowing the area before deciding to excavate the soil is that the area may be larger than expected, in which case the cover option may be preferable. For example, if the entire 6,000 m^2 work area is contaminated, then the cover alternative has greater value than the excavate alternative (Fig. 5). The maximum possible value from this approach is obtained if the area is determined with certainty by the additional sampling (i.e., uncertainty is eliminated). The design value if uncertainty in the volume is eliminated is -$200,900, compared to a value of -$240,000 if no further sampling is performed (Fig. 5); therefore, the maximum possible value of additional sampling is $39,100.

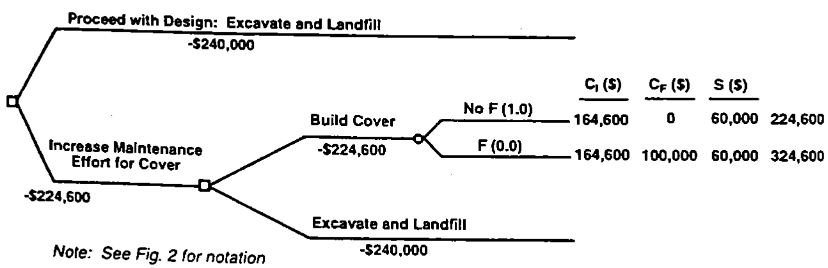
An example of uncertainty management through limiting its impact on design performance is shown on Fig. 6. The impact of uncertainty in long-term performance of the cover can be reduced by increasing the level of maintenance. In this case, the maximum possible value is obtained if the cover reliability is increased to 1.0. The design value for this approach is -$224,600, compared to the base case value of -$240,000 (Fig. 6); therefore, the maximum possible value of improved reliability is $15,400.

The value gained through uncertainty management depends on the increased benefits or reduced costs associated with the design, but also on the cost associated with the management. For example, the maximum $39,200 value associated with eliminating the uncertainty in the contaminated soil volume corresponds to approximately 200 soil samples analyzed for mercury content. It may not be possible to eliminate uncertainty in the contaminated area with 200 additional samples. As another example, it may not be possible to ensure successful cover performance by spending an additional present-worth amount of $15,400 on maintenance. Optimization of the value gained through uncertainty management is discussed in the following sections.



Note: See Fig. 2 for notation

Fig. 5. Value of Additional Sampling to Reduce Uncertainty



Note: See Fig. 2 for notation

Fig. 6. The Value of Limiting the Impact of Uncertainty

GUIDELINE 5

> **Uncertainty in an average (e.g., the average fraction of surficial soils contaminated with mercury) can generally be reduced with additional data. Uncertainty in a detail (e.g., the specific boundary of soil contaminated with mercury) is usually more difficult to reduce.**

Reduction of Uncertainty in Averages

Average values are commonly of interest in design decisions. Examples include the average contaminant concentrations in groundwater upgradient and downgradient from a site, the average volume of contaminated soils, the average hydraulic conductivity of a clay liner, the average shear strength along an interface in a containment system, and the average annual infiltration through a cover. Uncertainty in the estimated average for a variable is related to the amount of random variability in the variable (with time or location) and to the number of measured values (at different times or locations) as follows

$$\sigma_{\overline{X}} = \frac{\sigma_X}{\sqrt{n}} \tag{3}$$

where X is the variable, $\overline{X}$ is the average, $\sigma_{\overline{X}}$ is the standard deviation of the average (systematic uncertainty), σ_X is the standard deviation representing random variability in X and n is the number of independent measurements of X.

The simple relationship in (3) provides useful insight into optimizing the value of information for average values. First, variables with the largest variability (large σ_X) will have the greatest uncertainty in the average for a given number of measurements. The magnitude of σ_X is related both to real variations in X and artificial variations due to random errors in the measurement process. In general, investigation programs should focus more attention on variables with large σ_X values, although the focus also depends on the importance of $\overline{X}$ in the design decision and the cost of measurements. Second, there is a diminishing return in uncertainty reduction with additional measurements. The reduction in uncertainty can be measured by $\sigma_{\overline{X}}/\sigma_X$, which is equal to $1/\sqrt{n}$ from (3). Hence, the marginal reduction in the uncertainty with increasing n decreases as n increases, as shown on Fig. 7 (the curve labeled independent sampling).

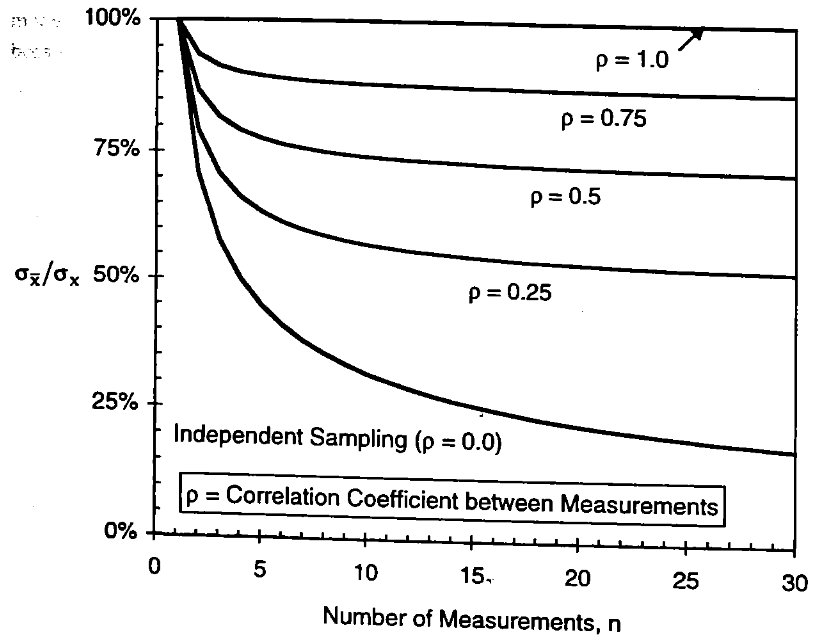


Fig. 7. Uncertainty Reduction versus Sampling Effort

The ability to reduce uncertainty in an average with measurements is affected by correlations between measurements. Measurements will be positively correlated if samples are spaced closely or if there is a systematic error in the measurement process (e.g., sample disturbance). For positively correlated measurements, (3) is modified as follows (e.g., Cochran 1963)

$$\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}}\sqrt{1+(n-1)\rho} \tag{4}$$

where ρ is the correlation coefficient between measurements (ρ ranges in magnitude from 0.0 for no correlation to 1.0 for perfect correlation). Each measurement is assumed to be correlated with every other measurement in (4), typical of a systematic error in the measurement process. As the correlation increases, the sampling efficiency decreases, as shown on Fig. 7. This reduction in efficiency occurs because positively correlated samples provide redundant information; the amount of

redundancy increases with increasing ρ. It is impossible to achieve a $\sigma_{\bar{X}}/\sigma_X$ ratio that is less than $\sqrt{\rho}$, regardless of the number of measurements (Fig. 7).

In many cases, only samples that are located close together will be correlated. As an example in the case history (Appendix I), mercury measurements in surficial soils were highly correlated within a radius of several meters. In these cases, the effective correlation coefficient for the group of n samples will be smaller than the correlation coefficient between closely spaced samples. When there is a correlation of ρ between n_ρ pairs of samples and no correlation between the remainder of the samples, the effective correlation coefficient for the entire sample group, ρ_{eff}, can be approximated by the following (e.g., Gilbert 1987)

$$\rho_{eff} = \left(\frac{2n_\rho}{n(n-1)}\right)\rho \tag{5}$$

For example, if twenty measurements are made (n = 20) and pairs of adjacent measurements are correlated (n_ρ = 19) with ρ = 0.8, then ρ_{eff} is equal to 0.08. Journel and Huijbregts (1978) and Vanmarcke (1983) provide more detailed treatments for spatially correlated data.

Additional unknowns will also limit the effectiveness in reducing $\sigma_{\bar{X}}$ with measurements. The effect of additional unknowns can be accounted for approximately by replacing σ_X in (3) with an effective value, σ_{eff}, that is larger to reflect the additional uncertainty. As an example, if the standard deviation of X is not known (rarely ever will σ_X be known if the mean of X is not known), then σ_{eff} is approximated by $\sigma_{eff} \equiv s\sqrt{(n-1)/(n-3)}$ where s is an estimate for σ_X based on the measured data (Ang and Tang 1975). Consider the case where $s \approx \sigma_X$ and $n = 5$; σ_{eff} is approximately 1.4 times greater than the estimate for σ_X due to the additional uncertainty. The difference between σ_{eff} and σ_X decreases with increasing n.

An application of these concepts to the case history site demonstrates their usefulness. Mercury-contaminated soil is distributed across the site in concentrated areas (or hot spots) that correspond to areas where mercury was temporarily stored. It is not possible to determine the location, size or frequency of these hot spots based on the limited site history information and Phase I sampling. The average fraction of contaminated area (relative to the 6,000-m^2 work area) will provide information about the expected volume of contaminated soil. Reduction in the uncertainty about this average fraction is shown on Fig. 8 as a function of the number of additional samples. These curves were generated by numerically simulating a random pattern of contamination across the site, and then sampling from the simulated site according to a fixed grid. Curves are shown for two different average hot-spot diameters. Note that the curves follow the same general trend as that shown on Fig. 7 for the simple model. Also, note that the sampling efficiency decreases for smaller sample spacings because adjacent samples are correlated (i.e., they may come from the same hot spot).

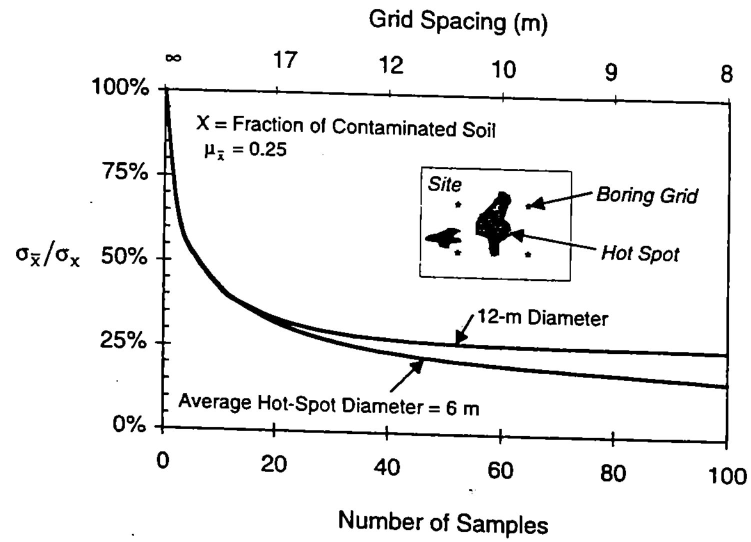


Fig. 8. Numerical Simulation Results for Estimating the Fraction of Contaminated Soil

Reduction of Uncertainty in Details

Details are more useful pieces of information than averages in many design decisions. For example, information about specific hot spots (i.e., their location and size), not the average fraction of contaminated soils, is necessary to develop an excavation plan and accurately estimate the cost of clean-up. While averages can generally be inferred from measured data, details are more difficult to infer. An example is the average mercury concentration (an average) versus the concentration at a specific location in the soil (a detail). Uncertainty in the average can be reduced by taking measurements from anywhere on the site as long as they provide representative samples. However, the only way to reduce uncertainty in the concentration at a specific location is to take a measurement from at (or near if concentrations are correlated) the location. If the location is not known, as is the case in trying to define the boundaries of mercury contamination, then significant effort may be required to reduce uncertainty in the detail.

The case history site provides valuable lessons in the difficulties associated with reducing uncertainty in details. First, a common approach is to try to reduce the number of required samples by identifying the locations of hot-spots based on site history, visual markings, or other information. The Phase I sampling program was designed around these identified hot spots. However, based on subsequent results from the Phase II sampling effort, only 15 percent of the hot spots were identified and 85 percent were undetected in Phase I because of its biased design. Therefore, care should be exercised in applying judgment (as discussed in the first section of this paper). In general, a regular grid of sample locations will be the most effective. Second, the 181 additional samples obtained in Phase II were largely ineffective at reducing uncertainty in the boundaries of contamination because the average sample spacing was greater than the typical hot-spot sizes. Therefore, this additional sampling effort had limited value because it was not designed to extract the information needed for design.

GUIDELINE 6

Redundant designs (i.e., systems that are not highly dependent on the performance of an individual component) can reduce the impact of uncertainty on performance. The value added through redundancy depends on the degree of dependence between the performance of individual components and that of the system.

One approach to reduce the impact that uncertainty has on design performance is to include redundancy in the design. Redundant designs incorporate systems that do not depend heavily on any single aspect or component. For example, a composite liner (i.e., a compacted clay layer overlain by a geomembrane) forms a redundant containment system. Leakage across the compacted clay is primarily controlled by its hydraulic conductivity, while leakage across the geomembrane is generally governed by defects in the geomembrane (e.g., Giroud and Bonaparte 1989). Since the hydraulic conductivity of the clay is essentially independent of the size and frequency of holes in the geomembrane, the reliability of the composite liner is expected to be greater than that for either a single clay liner or a single geomembrane liner (i.e., the likelihood of excessive leakage should be less for the system than for either component alone).

The following redundancy factor provides a useful measure to quantify system redundancy

$$RF = \frac{\text{Likelihood that a Component Performs Unsatisfactorily}}{\text{Likelihood that the System Performs Unsatisfactorily}} \qquad (6)$$

where RF is the redundancy factor. The magnitude of RF indicates the gain in reliability due to redundancy: RF = 1.0 indicates a non-redundant system, while RF >> 1.0 indicates a highly redundant system. For example, consider a liner that is designed with a maximum allowable leakage rate of 1,000 lphd. For a single geomembrane liner, the likelihood of exceeding this rate is high due to the potential for holes in the geomembrane: $P(Q_{FML} > 1{,}000 \text{ lphd}) = 0.9$. For a compacted clay liner with a hydraulic conductivity less than $1.0\text{x}10^{-9}$ m/s, the likelihood is smaller: $P(Q_{CCL} > 1{,}000 \text{ lphd}) = 0.2$. Finally, the likelihood for the composite liner is very small: $P(Q_{COMP} > 1{,}000 \text{ lphd}) = 0.001$. These likelihoods are estimated based on an analysis of laboratory and field data (Gilbert 1993). If performance of the clay liner is assumed to be independent of the geomembrane liner, then the redundancy factor is estimated as follows

$$RF = \frac{P(Q_{FML} > 1{,}000 \text{ lphd} \cup Q_{CCL} > 1{,}000 \text{ lphd})}{P(Q_{COMP} > 1{,}000 \text{ lphd})} = \frac{0.9 + 0.2 - (0.9)(0.2)}{0.001} = 920 \quad (7)$$

where $\cup$ indicates that either or both of the outcomes occur. Hence, this composite liner system exhibits substantial redundancy. The value of this redundancy depends on how the increased reliability, say from 0.8 for a single compacted clay liner to 0.999 for a composite liner, affects the value of the design (e.g., Fig. 6). If this value is greater than the cost of a geomembrane liner, then the composite liner provides a useful alternative to manage uncertainty in the performance of the clay liner.

The magnitude of RF depends on the degree of dependence between the performance of individual components and that of the system. In this composite liner, the frequency and size of defects in the geomembrane are unrelated to the hydraulic conductivity of the clay liner. Further, the performance of the composite liner system is not very sensitive to either the defects or the hydraulic conductivity. An example of a less redundant system is a composite liner consisting of a geosynthetic clay liner (GCL) overlain by a geomembrane. Since both the GCL and the geomembrane are susceptible to damage during construction and operation, RF is estimated to be about an order of magnitude smaller.

GUIDELINE 7

Flexible designs (i.e., systems that can be modified during and after construction in response to unexpected conditions), if implemented properly, are very effective at adding value.

Flexible designs are those that can be modified in response to unexpected conditions during and after construction. This approach, which has been referred to as the Observational Method in geotechnical engineering (Peck 1969) and proposed for geoenvironmental engineering (e.g., Holm 1993), is very effective at improving the value of a design. Its only potential drawback, besides the cost required to implement a flexible design, is the possibility that the design will not be flexible enough to accommodate unanticipated conditions (note that unanticipated conditions are those considered impossible at the design stage, while unexpected conditions are those

considered possible but not typical). However, this drawback applies to all design alternatives and is not peculiar to flexible designs.

The case history site provides examples of both inflexible and flexible design alternatives. The excavate-and-dispose option is inflexible because there is no flexibility available if the contaminated volume of soil is unexpectedly large. An example of a similar design that was unsuccessful because of inflexibility is the PPI Superfund site (Acar et al. 1995). An excavate-and-dispose option was selected but had to be abandoned, at considerable expense and time delay, because of the release of volatile organics during excavation. Conversely, the cover alternative is more flexible in that it can be inspected and maintained continually over its lifetime.

The keys to successful implementation of a flexible design are to (1) design a monitoring program that will reduce the important uncertainties over time and (2) have contingency plans available to manage unexpected outcomes. These concepts have already been addressed in this paper in the context of new designs, and the same principles apply to flexible, evolving designs.

CONCLUSIONS

Seven, practical guidelines for managing uncertainty in geoenvironmental design are presented and discussed in this paper. While these guidelines are derived from theoretical considerations, all of the guidelines could be classified as common sense (a result that lends some credibility to the theory). There are two major conclusions that we have drawn from applying the theory of uncertainty management to practice. First, there is a tremendous need for more, higher quality performance data (both technical and economic) for geoenvironmental systems. Much of what has been compiled is never analyzed (e.g., it is only compiled to satisfy regulatory requirements) or does not provide useful information as feedback to the design process. An integrated, national effort should be initiated to compile and analyze performance data to improve the value of future designs. Second, the most important aspect of uncertainty management is communication. All parties involved in a design need to be aware of the uncertainties, the potential design outcomes (favorable and unfavorable), and the potential consequences associated with those outcomes. This communication is essential, whether or not a formal approach is adopted to manage uncertainty.

APPENDIX I. CASE HISTORY SITE DESCRIPTION

The case history site is a 14,000-m^2 residential plot (Fig. 9) located in southeast Texas. Soil within a 6,000-m^2 work area has been contaminated with mercury from a processing operation. The owner has not divulged the details of the operation to investigators, so the number, locations and sizes of the former storage piles must be inferred from physical evidence (e.g., disturbed soil).

The site is generally flat and level. The site stratigraphy, from on-site well logs, is 1.5 to 2.0 m of silty clay overlying 1.5 to 7.5 m of clayey, fine sand, which grades into clay at a depth of about 9 m. This clay layer extends to a depth of approximately 40 m, below which lies the primary-use aquifer for the local population. Laboratory tests indicate that the surficial clay is overconsolidated and that its hydraulic conductivity in laboratory samples is approximately 2.5×10^{-6} cm/s. No secondary features have been observed in the surficial clays, and no mercury has been detected in the groundwater within the upper layer of fine sand.

A two-phase sampling program was conducted to determine the extent of soil contamination at the site. In Phase I, 16 boreholes were drilled down to a maximum depth of 9 m and mercury-test samples were taken intermittently to a maximum depth of 3.5 m (no mercury was detected below a depth of 1.5 m). The boreholes were placed where contamination was believed to be most likely based on visual evidence of waste transport and storage and soil disturbance. However, the results from the first phase did not sufficiently delineate the extent of contamination, so a second sampling phase was conducted. In Phase II, 181 soil samples were collected from across the entire site to a maximum depth of 0.3 m. Mercury concentrations in the samples ranged from non-detect (<0.12 mg/kg) to 1,760 mg/kg. A health-risk assessment for the site indicated that the soil should be remediated to a mercury concentration below 20 mg/kg.

Three remedial alternatives were originally considered at the site. The first, a monitoring-only alternative, was removed from consideration because it did not satisfy applicable or relevant and appropriate requirements. The second, containment of the work area with a cover, and the third, excavating the contaminated soil for disposal in an existing off-site landfill, were compared in a qualitative discussion of pros and cons. The following criteria were applied to compare the alternatives: overall protection of human health and the environment; long-term effectiveness and permanence; reduction of toxicity, mobility or volume of contaminant; short-term effectiveness; implementability; and cost. The excavation alternative was selected.

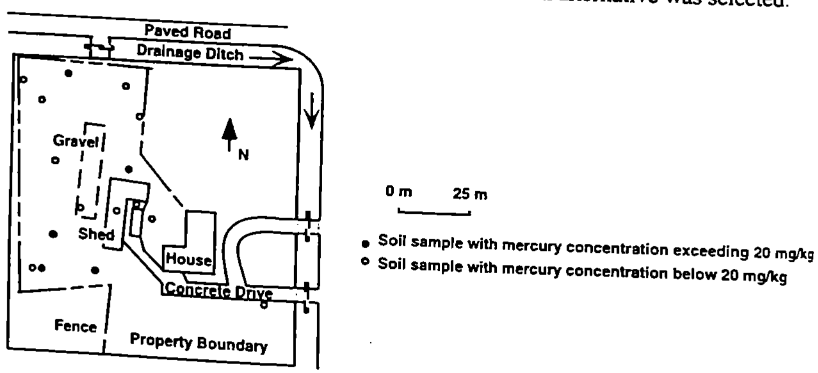


Fig. 9. Site Plan for Case History Site

APPENDIX II. REFERENCES

Acar, Y. B., Taha, M. R., and Constant, W. D. (1995), "The PPI Superfund Site: Remedial Measures and Alternatives," *Proc.,* Geoenvironment 2000, ASCE, 1684-1699.

Ang, A. H-S. and Tang, W. H. (1975), Probability Concepts in Engineering Planning and Design, Vol. 1 - Basic Principles, John Wiley & Sons, New York.

Benson, C. H., and Daniel, D. E. (1994), "Minimum Thickness of Compacted Soil Liners: I. Stochastic Models," *Journal of Geotechnical Engineering*, ASCE, Vol. 120, No. 1, 129-152.

Capen, E. C. (1976), "The Difficulty of Assessing Uncertainty," *Journal of Petroleum Technology*, Vol. 28, 843-850.

Cawlfield, J. D. and Wu, M-C. (1993), "Probabilistic Sensitivity Analysis for One-Dimensional Reactive Transport in Porous Media," *Water Resources Research*, Vol. 29, No. 3, pp. 661-672.

Clemen, R. T. (1991), Making Hard Decisions: An Introduction to Decision Analysis, Duxbury Press, Belmont, CA.

Cochran, W. G. (1963), Sampling Techniques, 2nd Ed., John Wiley & Sons, Inc., New York.

Danzo, A. (1988), "The Big Sleazy: Love Canal Ten Years Later," *The Washington Monthly*, September, pp. 11-17.

Davis, K. L. (1996), Joint Institute for Energy & Environment, Knoxville, TN, personal communication.

Freeze, R. A., Massmann, J., Smith, L., Sperling, T. and James, B. (1990), "Hydrogeological Decision Analysis: 1. A Framework,"*Ground Water*, Vol. 28, 738-766.

Gilbert, R.B. (1993), "Performance Reliability of Landfill Lining Systems," Ph.D. Dissertation, University of Illinois, Urbana, Illinois.

Gilbert, R. B. and Tang, W. H. (1995a), "Model Uncertainty in Geotechnical Reliability and Decision Analyses," Intl. Conf. on Applied Stat. and Prob., Paris, 109-114.

Gilbert, R. B. and Tang, W. H. (1995b), "Reliability-Based Design of Waste Containment Systems," ASCE Specialty Conf., Geoenvironment 2000, 499-513.

Gilbert, R. O. (1987), Statistical Methods for Environmental Pollution Monitoring, Van Nostrand Reinhold Company, New York.

Giroud, J. P. and Bonaparte, R. (1989), "Leakage Rates Through Liners Constructed with Geomembranes - Part I. Geomembranes," Geotextiles and Geomembranes, Vol. 8, No. 1, 27-67.

Hoffman, Andrew J. (1995), "An Uneasy Rebirth at Love Canal," *Environment*, Vol. 37, No. 2, 4-16.

Holm, L. A. (1993), "Strategies for Remediation," Geotechnical Practice for Waste Disposal, D. E. Daniel, Ed., Chapman and Hall, New York, New York.

Journel, A. G. and Huijbregts, C. J. (1978), Mining Geostatistics, Academic Press, London.

Keeney, R. L., von Winterfeldt, D. (1993), "An Analysis of Strategies to Manage Nuclear Waste from Power Plants," Decision and Risk Analysis Report 93-1, Institute of Safety and Systems Management, Univ. of Southern California, Los Angeles, CA.

Massmann, J., Freeze, R. A., Smith, L., Sperling, T. and James, B. (1991), "Hydrogeological Decision Analysis: 2. Applications to Groundwater Contamination," *Ground Water*, Vol. 29, 536-548.

Merkhofer, M. W., and Keeney, R. L. (1987), "A Multiattribute Utility Analysis of Alternative Sites for the Disposal of Nuclear Waste," *Risk Analysis*, Vol. 7, No. 2, 173-194.

NRC (1994), Alternatives for Ground Water Cleanup, National Research Council, National Academy Press, Washington, D.C.

Peck, R. B. (1969), Ninth Rankine Lecture: "Advantages and Limitations of the Observational Method in Applied Soil Mechanics," *Geotechnique*, 19(2), 171-187.

Raiffa, H. (1968), Decision Analysis: Introductory Lectures on Choices Under Uncertainty, Addison-Wesley, Reading, MA.

Roberds, W. J., "Methods for Developing Subjective Probability Assessments," *Transportation Research Record*, No. 1288, 1990, National Research Council, Washington, D. C., 183-190.

Roberds, W. J., Haerer, H. A. and von Winterfield, D. (1992), "Decision Management for the Hanford Environmental Dose Reconstruction Project," *Proc.*, Intl. High Level Radioactive Waste Management Conf., Las Vegas.

Russell, M. and Davis, K. L. (1995), "Resource Requirements for NPL Sites: Phase II Interim Report," Joint Institute for Energy & Environment, Knoxville, TN.

Sitar, N., Cawlfield, J. D., and Der Kiureghian, A. (1987), "First-Order Reliability Approach to Stochastic Analysis of Subsurface Flow and Contaminant Transport," *Water Resources Research*, Vol. 23, No. 5, 794-804.

Tang, W.H., Gilbert, R.B., Angulo, M. and Williams, R.S. (1994), "Probabilistic Observational Method for Settlement-Based Design of a Landfill Cover," *Proc.*, Settlement '94, ASCE Geotech. Special Publ. No. 40, College Station, Texas, 1573-1589.

Tribus, M. (1969), Rational Descriptions, Decisions and Designs, Pergamon, New York.

USEPA, (1989), Risk Assessment Guidance for Superfund, Volume I, Human Health Evaluation Manual; Part A; Interim Final, EPA/540/1-89/002, U. S. Environmental Protection Agency, Washington, D.C.

Vanmarcke, E. (1983), Random Fields: Analysis and Synthesis, The MIT Press, Cambridge.