

APPENDIX C

CHARACTERIZING VARIABILITY AND UNCERTAINTY IN THE CONCENTRATION TERM

C.0 THE CONCENTRATION TERM AND THE EXPOSURE UNIT

Incomplete knowledge of the concentration of one or more chemicals in various exposure media is often the major source of uncertainty in Superfund risk assessments. In any risk assessment, the derivation of the concentration term will reflect assumptions about: (1) properties of the contaminant, (2) the spatial and temporal variability in contamination, (3) the behavior of the receptor, and (4) the time scale of the toxicity of the chemical(s). This appendix expands upon concepts introduced in Chapter 5. This appendix does not provide detailed equations for performing calculations, but instead refers the reader to other Environmental Protection Agency (EPA) guidance documents in which both the recommended approaches and calculations are provided.

The concentration term is linked to the concept of an exposure unit (EU). For Superfund risk assessments, an EU is the geographical area in which a receptor is randomly exposed to a contaminated medium for a relevant exposure duration. Environmental sampling provides information about the contamination within and around an EU. Multiple EUs may be defined at a site based on the choice of a receptor, the exposure medium, and the nature of contact with the medium. For example, residential exposures to children may involve exposures via soil and dust ingestion both at the primary residence and recreational areas at a day care facility. Site-specific information regarding the activities of receptors should guide assumptions about the receptor's contact with exposure media.

Defining the EU is critical to the success of the remedial strategy, as it affects the calculation of the concentration to which receptors are exposed.

C.1.0 VARIABILITY IN PRA

In general, variability and uncertainty should be kept separate to the extent possible in any probabilistic risk assessment (PRA). For example, assume a one-dimensional Monte Carlo Analysis (1-D MCA) was developed to characterize variability in risk, but it combined a distribution for uncertainty in mean concentration with distributions for variability in exposure variables. The result would yield a single distribution for risk, however, each risk estimate would reflect both uncertainty and variability and distinguishing between the two would not be possible. Therefore, EPA's *Guiding Principles for Monte Carlo analysis* recommends against mixing distributions of variability and uncertainty in a 1-D MCA (U.S. EPA, 1997b) to avoid such ambiguities.

A fundamental concept in Monte Carlo analysis is that there is variability in exposure between receptors (inter-individual variability) as well as day-to-day variability for each individual (intra-individual variability). In most Tier 2 analyses (see Chapter 2), the goal of a 1-D MCA is to characterize inter-individual variability in exposure and risk. Typically, probability distributions for exposure represent variability (PDFv's) between individuals in the average value over the entire exposure duration. In this case, the exposure point concentration (EPC) should represent the average exposure concentration over the entire exposure duration. Because an EPC is calculated from a sample, there is uncertainty that

the sample mean equals the true mean concentration within the EU; therefore, to account for associated uncertainty, the 95% upper confidence limit for the mean (95% UCL) is generally used for Superfund risk assessments (U.S. EPA, 1992).

In a 1-D MCA, a point estimate for the EPC is combined with PDFv's for other variables to yield a probability distribution for risk. An alternative approach is to simulate long-term average exposures as a series of consecutive short-term exposure events. This approach is referred to as MicroExposure Event (MEE) Monte Carlo modeling, and is discussed in detail in Appendix D. In MEE modeling, the goal is to develop PDFv's for exposure variables that capture the event-to-event variability in exposures at the individual level. The concept of an averaging time still applies, but generally to a shorter time frame. For example, seasonal variability in exposure frequency might be expected among outdoor occupational workers so that different PDFv's are representative of inter-individual for each season. In this case, the EPC continues to represent an average concentration within the EU, but it would be linked to season-specific activity patterns. It may be important to develop two different weighted averages to reflect season-specific activity patterns and locations that are more frequently contacted in the summer compared with the winter, for example. As the time frame for the exposure scenario is shortened from the entire exposure duration, to a season, to a day, to an individual event, the concentration term should be reevaluated to assess the relevance of the assumption that concentrations contacted by the receptor are represented by the mean of the measured sample.

The following discussion introduces concepts of temporal and spatial variability as they apply to the estimate of the EPC for different exposure media and exposure scenarios. While the general rule of thumb applies to all Monte Carlo models—use a measure of the average concentration within the EU over the time frame of exposure—it is important to apply the site sampling data in a way that is consistent with the exposure scenario.

C.1.1 TEMPORAL VARIABILITY

Temporal variability in chemical concentrations may be an important consideration when developing a preliminary remediation goal (PRG) for any exposure medium (refer to Chapter 5 for a comprehensive discussion of using PRA to evaluate PRGs). For example, wind erosion may change chemical concentrations in surface soil over time; leaching may change concentrations in both subsurface soil and groundwater; and bioaccumulation may result in increasing concentrations in predatory fish with time. If possible, such factors should be considered early in the risk assessment process and included in the conceptual site model.

Development of the EPC normally will depend on the averaging time relevant to the exposure scenario and health endpoint of concern. In the shorter term, it may be unlikely that receptors are exposed throughout the entire EU due to temporal (and spatial) variability in the contaminant and inter-individual variability in activity patterns. Therefore, inter-individual variability in the EPC might be expected, and a distribution of EPCs may be developed to represent differences in exposure among the population. Variability in short-term exposure may be an important factor for assessing variability in acute toxicity. However, over time, short-term variability in the EPC will tend to smooth out and approach a long-term average concentration. A single estimate of the long-term average EPC may be reasonable to use in assessing risks to the receptor population. This is true regardless of the underlying distribution of the environmental sampling data (e.g., lognormal, normal, beta, etc.).

While most chemicals regulated by the Superfund program are based on concerns for chronic toxicity (e.g., lifetime cancer risk from exposure to a carcinogen for ten or more years), for some

chemicals, toxic effects occur with shorter exposure durations (e.g., nitrate in drinking water and methemoglobinemia in infants). Differences between acute and chronic health endpoints are important to consider for ecological receptors such as transient migratory species. Superfund guidance distinguishes between acute and chronic exposure to provide risk assessors the option of evaluating risk under different time frames. The EPC should be estimated within an EU during a period of time that has toxicological relevance for the exposed population.

☞ *The time scale of the concentration term should match the time scale of the toxicity criterion and exposure duration.*

C.1.2 SPATIAL VARIABILITY

Spatial variability in chemical concentrations is also an important property to consider when developing a PRG. Spatial variability arises from many factors, including the mechanism of contamination, physical and chemical dilution and transformation processes, and physical characteristics of the site (Cullen and Frey, 1999). Similarly, receptors may exhibit spatial variability in their contact with an exposure medium. In general, receptors are assumed to have equal access to all areas within an EU so that the concept of a long-term average concentration is applicable.

Often, the EPC is estimated without regard to the spatial patterns in contamination. The sampling design yields a measure of the variability in concentrations that is assumed to be representative of the receptor's contact with the exposure medium. However, even when the sampling design is representative (e.g., both are simple random samples within the EU), the concentrations may exhibit clear spatial patterns that could be used to reduce uncertainty in the EPC. Geostatistics (see Section C.5.2 and Appendix D) offers a wide range of techniques for incorporating spatial information into estimates of the EPC. These techniques are particularly useful when there is uncertainty in the representativeness of site sampling, due to a difference in scale between site sampling and the size of the EU, or the use of targeted sampling designs that oversample areas within an EU believed to contain the highest levels of contamination.

In point estimate risk assessments (Tier 1 of the PRA), the EPC is most often characterized by a point estimate of the mean concentration, typically given by the 95% UCL for the mean to account for uncertainty in the site characterization (U.S. EPA, 1992). Variability in concentrations is an important consideration for determining appropriate statistical methods used to estimate the 95% UCL. In addition, for some Monte Carlo models, a PDFv may be developed to determine the EPC for the exposure model. A PDFv for the EPC may be warranted in short-term exposure scenarios, particularly when the sampling density is relatively sparse in relation to the size of the EU (i.e., poor site characterization). For example, a risk assessment may include a future use residential scenario (e.g., currently the site is undeveloped) in which the EPC that is relevant to a potentially exposed population of children is the average concentration within a 0.5 acre lot. If the soil sampling yields 100 measurements, but a small subset of the samples (e.g., less than three) are available for any 0.5 acre area, the most appropriate measure of the average concentration for a hypothetical residence may be the maximum detected concentration or a single value from the PDFv in concentration among hypothetical receptors. In general, for any of the EU's that define a randomly located residence, the poor site characterization would be a source of uncertainty in both a point estimate and probabilistic risk assessment.

At the vast majority of sites, concentration data is the easiest data to obtain of all the exposure variables. In cases of poor site characterization, risk managers may opt to perform a point estimate risk assessment only using the maximum detected concentration and highly protective exposure assumptions.

In the scenario described above for 0.5 acre residential lots, it is possible that a residence would be located in an area in which the average concentration is represented by the maximum detected concentration in the sample. Should the risk manager opt for a Tier 1 point estimate risk assessment, the use of the maximum detected concentration of a chemical on the site should ensure the performance of a health-protective risk assessment within a smaller EU.

Consideration of variability is also warranted in short-term scenarios for ecological risk assessment (ERA) when the EU is much smaller than the site (see Section C.3.1.1). For example, the home range of the receptor populations may be relatively small in comparison to the spatial distribution of sampling locations (e.g., benthic invertebrates living in the sediment at the bottom of a river or soil invertebrates in a terrestrial habitat). In these cases, the receptor would be exposed to an area smaller than the sampling grid or measure of areal sampling density. A value from the PDFv that characterizes variability in the concentrations across a relatively large spatial scale may be used to define the EPC for a receptor population at a smaller scale. Again, risk assessors should take care in designing a 1-D Monte Carlo model when using a PDFv for the concentration term. It is inadvisable to mix a PDFv for the concentration term with PDFv's for other exposure scenarios when estimating risks within one EU. Use of the PDFv in this manner would incorrectly suggest that the mean concentration varied for each individual within the same EU according to the variability in concentration measured across a much larger area. A preferred approach is to use a PDFv to obtain a point estimate that represents the EPC, and then combine this point estimate with PDFv's for other variables in the Monte Carlo simulation to estimate risks in the small EU. If there are many EU's at a site, or if the boundaries of EUs are undefined, more advanced modeling approaches can be developed to efficiently run multiple scenarios. Methods for characterizing exposure point concentrations for ecological receptors are further discussed in Sections C.2 and C.3.

C.1.3 EXAMPLE OF TEMPORAL AND SPATIAL VARIABILITY

Exposure scenarios often require consideration of both temporal and spatial variability. The MEE might be used to assess temporal variability by simulating long-term intake as the sum of individual exposure events. The time step for MEE is an important consideration and will depend on the rate of change of the most rapidly changing exposure variable. In addition, there should be a correspondence between the time periods over which data were obtained and the time step used in the MEE model. For example, when a MEE is used for the risk assessment, the concentration term selected at each time period should match the "average" concentration within the EU appropriate for that particular time period. Assume that the receptor is a residential child, and the time period is a single day, and the child may contact only 1,000 square feet within the 0.5 acre (20,000 square feet) residential EU. The specific 1,000 square foot area may change with each day as the child chooses different areas in the yard to frequent. Hence, the variability in the sample may be a more appropriate measure of the concentration contacted by residential child receptor on a day-to-day basis than the long-term average within the 0.5 acre EU. Over the long-term, this receptor will be exposed to the entire EU and hence the average contaminant concentration within the 0.5 acre EU. Note that the day-to-day variability in concentration undergoes the familiar phenomenon of "regression to the mean" when considered over the long-term.

C.1.4 SPATIAL AND TEMPORAL VARIABILITY FOR DIFFERENT EXPOSURE MEDIA

C.1.4.1 VARIABILITY OF CONCENTRATIONS IN SOIL

Surface soil is subject to erosion by wind and surface water runoff. Over time, concentrations in surface soil may change, but generally at a slow rate relative to other media. The spatial variability of chemical contamination is most often due to the mechanism by which the contamination occurred. For example, particulate stack emissions will tend to fall in an even pattern downwind of the stack whereas over-application of pesticides and chemical spills can result in a patchy pattern of contamination.

Subsurface soil is not subject to wind erosion, so concentrations change mostly due to degradation processes or leaching of the contaminant to groundwater. At most Superfund sites, concentrations of chemicals in subsurface soil will remain relatively constant.

C.1.4.2 VARIABILITY OF CONCENTRATIONS IN GROUNDWATER

Exposure to groundwater contamination mostly occurs at a fixed point in space (e.g., the wellhead). Groundwater is subject to a variety of influences that can alter chemical concentrations within this medium such as aerobic and anaerobic biodegradation, volatilization, and absorption. Due to these influences, monitored natural attenuation is an appropriate remedy under certain site conditions. If a risk assessor wishes to use a measure of the long-term average of a concentration in groundwater, a hydrogeologist should be consulted.

C.1.4.3 VARIABILITY OF CONCENTRATIONS IN SURFACE WATER

Concentrations in surface water can be very dynamic. Streams are constantly flowing and the effects of mixing, dilution and evaporation can change the chemical concentrations in surface water over relative short time periods. Any sampling of surface water is truly a “snapshot” in time. The sampling methods used to characterize spatial and temporal variability of concentrations in surface water will have a direct effect on the uncertainty in estimates of the average concentration over both short and long time frames.

C.1.4.4 VARIABILITY OF CONCENTRATIONS IN SEDIMENT

In some situations, sediment may be considered a relatively stable medium, similar to soil. Alternatively, sediment may be physically moved by currents, tides, the movement of ships and other events. Trend analysis may be used to establish the long-term average sediment transport at a site. This information could provide the basis for choosing a representative “average” concentration in the sediment available to ecological receptors (Piest and Miller, 1975; Van Sickel and Beschta, 1983; Walling, 1983; Meade et al., 1990).

C.1.4.5 VARIABILITY OF CONCENTRATIONS IN FISH

Concentrations in fish may vary due to a change in the availability of food and environmental conditions. Factors that may be used to model population dynamics may include intensity of angler harvest, death/attrition of the population, and the introduction of a predator species or a more adaptive species. In risk assessments that include a fish ingestion exposure pathway, the activities of the angler may be a more important factor in determining the EPC than the changes in concentrations in fish over time. For example, an avid recreational angler may harvest fish from different locations within a lake and

consume fish of different sizes and species. In this way, with the consumption of contaminated fish, both the contaminated medium and the exposure point change throughout the exposure duration.

Unless, samples of fish are collected over time, knowledge of these factors will generally be unknown. Concentrations of bioaccumulative chemicals in territorial fish (e.g., largemouth bass) obtained in different locations will generally reflect the concentrations in the sediment in the individual's home territory. Concentrations of bioaccumulative chemicals in migratory fish will be more difficult to predict as the fish will contact areas with varying sediment and surface water concentrations.

C.1.4.6 EXAMPLES OF TEMPORAL AND SPATIAL VARIABILITY IN THE CONCENTRATION TERM FOR SELECTED EXPOSURE MEDIA

Whatever medium is considered in the development of EPCs, the risk assessor should be aware that the EPC embodies aspects of both the spatial distribution of contamination, the movement of the receptor, and possibly the contaminated medium within the EU. Table C-1 presents examples of sources of temporal and spatial variability in the concentration term based on both the contamination in selected exposure media and the receptor.

Table C-1. Examples of temporal and spatial variability in selected media for the concentration term in common exposure scenarios.

| Factor | | Soil | Groundwater | Fish |
|----------------------|-------------|---|--|---|
| Temporal Variability | Contaminant | <ul style="list-style-type: none"> • none, if contaminant source is inactive • aerial deposition from ongoing source emissions affected by wind patterns • degradation over time • volatilization • migration to groundwater • radioactive growth and decay | <ul style="list-style-type: none"> • seasonal fluctuation in groundwater table • migration of contaminant plume • natural attenuation | <ul style="list-style-type: none"> • seasonal changes in species availability • bioconcentration • long-term changes in population dynamics • fish tissue concentrations linked to temporal variability in water and sediment concentrations • physical and chemical processes |
| | Receptor | <ul style="list-style-type: none"> • changes in activity patterns and behaviors over time (e.g., with age) | <ul style="list-style-type: none"> • none, fixed location at specific wellhead • changes in well location over time | <ul style="list-style-type: none"> • dietary preferences for fish species • cooking practices |
| Spatial Variability | Contaminant | <ul style="list-style-type: none"> • heterogeneity in concentrations over a small area and with depth, including presence of hotspots • heterogeneity in soil properties that influence bioavailability | <ul style="list-style-type: none"> • migration of contaminant plume, based on hydrogeology and source emissions (e.g., bulk flow or continuous source) | <ul style="list-style-type: none"> • migration of fish • changes in fish population structure |
| | Receptor | <ul style="list-style-type: none"> • daily activity patterns involve contact with different areas of the EU | <ul style="list-style-type: none"> • none, fixed location at specific wellhead • changes in well location over time | <ul style="list-style-type: none"> • change in recreational habits, and areas fished |

C.2.0 NONRANDOM EXPOSURES

As discussed in Section C.1.2, in the long-term it is generally assumed receptors exhibit random movement, such that there is an equal probability of contacting any area within the entire EU. Therefore, the long-term exposure concentration will most likely be the arithmetic mean of the concentration within the EU. However, in many situations, the assumption of random exposures in space may clearly be an oversimplification. People's behavior and preferences will cause them to access specific areas within an EU with greater frequency than others. The same is true in terms of ecological receptors with specific habitat preferences.

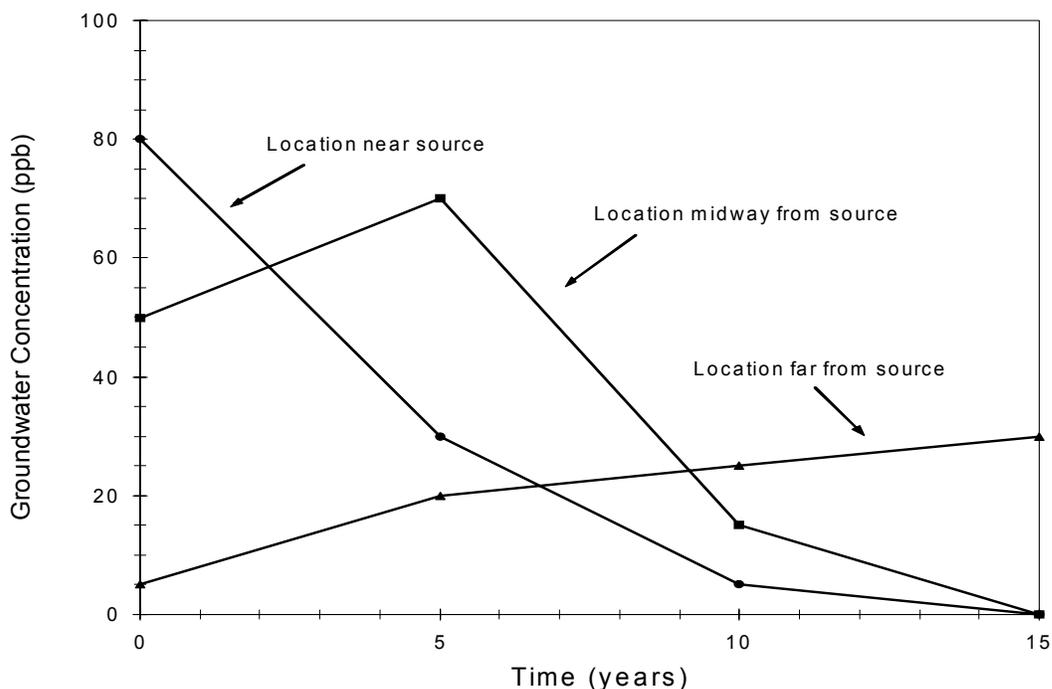


Figure C-1. Spatial and temporal variability in contaminant concentrations in groundwater.

For example, groundwater concentrations may show a large variation when sampled from wells in different locations (Figure C-1). Typically, residential receptors do not sample randomly from different wells, but draw chronically from individual wells. In such a case, the EU is a single wellhead. Fluctuations in the groundwater plume will depend on the hydrogeology of the site as well as the seasonal fluctuations in the water table. In this hypothetical example, concentrations are declining over time at distances nearest to the source, and concentrations are increasing as the plume moves farther from the source.

Incomplete information regarding the behavior patterns of people and environmental systems can be a large source of uncertainty in a risk assessment. Because of this, methods are being developed to model spatial relationships (between the contaminant and receptor) and nonrandom exposures. Recently, a quantitative technique to model nonrandom exposure has been proposed for ERA (Hope, 2000, 2001). Briefly, this technique divides the EU into smaller subunits and uses information about the attractiveness

of each subunit to assign a probability of the receptor occupying a given subunit for a period of time. Receptor movements are modeled stochastically and a time-weighted average of all the subunits provides a measure of the EPC. In some ecological risk assessments, telemetry data can be used to better characterize the areas of contamination that overlap with habitats of selected species. Hoff (1998) demonstrates an approach for American badgers (*Taxidea taxus*) in which telemetry data and geostatistical modeling provide an improved relationship between contaminant concentrations, tissue residues, and effects.

C.3.0 SOURCES OF UNCERTAINTY IN THE CONCENTRATION TERM

There are numerous potential sources of uncertainty in the estimate of the true mean concentration within an EU. As discussed in Chapter 5 (Section 5.1.1), sources of uncertainty can be grouped into four broad categories: sample data, location of the EU, behavior of the receptor, and from miscellaneous sources (e.g., physical and chemical processes). Development of an uncertainty distribution for the average concentration requires knowledge of the variability in chemical concentrations within the EU (unless distribution-free approaches are used), the toxicity of the chemicals, and the receptor's behavior. These distributions should be developed by risk assessors with the concept of the EU in mind. Differences in scale (e.g., small home range of an ecological receptor population relative to the site sampling design) can be a major source of uncertainty in ecological risk assessments. Methods for addressing such uncertainties in the concentration term are presented below. By incorporating these methods into the quantitative uncertainty analysis, risk managers may more effectively evaluate the importance of data-gaps and design subsequent rounds of site sampling to reduce the uncertainty in the EPC.

C.3.1 QUANTIFICATION OF UNCERTAINTY BASED ON THE SIZE OF THE EXPOSURE UNIT

Site characterization sometimes occurs before an EU has been defined. Therefore, an EU may be smaller than an entire site, equal to the site itself, or larger than the site. These three conditions lead to different conclusions and methods about the determination of the EPC. The most complex situation is when the EU is smaller than the site and the site can contain multiple EUs. For future scenarios in which the land use differs from the current land use, the difficulty in predicting the exact size and location of EUs necessitates accounting for the uncertainty in the EU.

Composite sampling is often used to maximize site information. However, it is important to note that the use of composite sampling influences the concentration term. If composite sampling is used exclusively at a site, the actual maximum concentration present or the best estimate of this maximum concentration will not be available. Depending on the time scale of the toxic effect or whether acute toxicity should be considered, this lack of knowledge of the maximum concentration present may be a large data gap. Risk assessors are urged to consider composite sampling and its ramifications for the concentration term.

C.3.1.1 WHEN THE EXPOSURE UNIT IS SMALLER THAN THE SITE

The size of the EU will be different depending on the length of exposure. A receptor can access a greater area if given more time. In almost all cases, the size of the EU for short-term exposure will be smaller than the EU for long-term exposure. Therefore, in addition to the uncertainty associated with sampling and analysis (which can be quantified with existing methods for calculating confidence intervals), there is uncertainty about the location of the EU within the site.

If contamination is evenly spread across the site, the location of the EU may not have any bearing on the EPC. In such a case, uncertainty may depend on the sample size or density of measurements within the EU relative to the entire site. In point estimate risk assessments, the concentrations of chemicals at the sampling location that poses the greatest risk may be considered as estimates of the EPC for this small EU. Using this “riskiest” sampling location as an estimate of the mean within an EU of unknown location accounts for both the uncertainty associated with limited sampling within a single EU and the uncertainty of the location within the site of the EU.

To express the uncertainty in location of the EU as a distribution, methods have been developed to place an EU of a given size randomly about a site (Burmaster and Thompson, 1997). A concentration term is developed for each of a large number of randomly located EUs. The distribution of these concentration terms will express the uncertainty in the location of the EU.

Risk assessors are cautioned to consider whether the statistical method used to estimate the EPC in an EU accounts for all sources of uncertainty in the concentration term. If only a few samples are used to characterize the average concentration within an EU, then the uncertainty in the EPC is large and should be presented in the risk characterization. These conditions may warrant additional sampling or the use of analytical methods that account for spatial variability within the entire site.

At some sites, geostatistical methods, pattern recognition, and geographical information systems (GIS) methods may provide additional insight and will aid in the development of the concentration term (see Section C.5.2). Although Table 3-1 shows several statistical methods for estimating both point estimates and distributions that encode uncertainty in the concentration term, a risk assessor’s understanding of these uncertainties should be conceptual as opposed to purely statistical.

C.3.1.2 WHEN THE EXPOSURE UNIT IS THE SAME SIZE AS THE SITE

In this case, the entire environmental data set within the site boundaries can be used for the determination of the concentration term. Assuming the EU occupies the entire site, then the source of uncertainty associated with knowing the average concentration within the EU is the sampling and analytical uncertainty.

C.3.1.3 WHEN THE EXPOSURE UNIT IS LARGER THAN THE SITE

In this case, the EU extends beyond the site boundaries. Therefore, the entire environmental data set within the site boundaries can be used for determination of the concentration term. However, an additional term in the exposure assessment may be needed to account for the fraction of the exposures that are expected to occur off site. Essentially, the contribution of the chemical concentrations measured on and off site are weighted by the fraction ingested or contacted in each area. Similarly, the term “area use factor” is used in ecological risk assessments to refer to the percentage of time or area an animal inhabits a contaminated area. An exposure scenario in which the EU is defined by the multiple locations that may be visited would be a common extension of this concept. One reasonable assumption regarding off site exposures is that the concentrations would be equal to the “background” concentrations. If this assumption is made, a site risk assessor should be consulted to determine appropriate methods for incorporating background concentrations into the risk assessment. Alternatively, additional sampling at off site locations would be needed to estimate the concentrations.

C.4.0 SUMMARY OF RECOMMENDATIONS FOR THE CONCENTRATION TERM

Table C-2 presents general guidelines for establishing a concentration term in various media based on exposure time and the size of the EU. These general guidelines along with site-specific exposure conditions are the driving factors in risk assessment decision making for establishing the concentration term.

Table C-2. Summary of factors that may be considered in developing an EPC.

| Medium | Exposure Time | Random | Non-Random | Size of EU relative to the site/sampling density | Recommendation (Human Health and Ecological) |
|--------------|---------------|--------|------------|--|--|
| Soil | Short-term | | X | small | HH - consider variability in concentration relative to the time scale of toxicity. ECO - time weighted average of smaller subunits. |
| Soil | Long-term | X | | variable | HH, ECO - consider uncertainty in the average concentration within an EU. |
| Fish | Short-term | | X | variable | HH, ECO - consider variability in sample concentrations relative to the exposure time. |
| Fish | Long-term | X | | variable | HH - consider uncertainty in the average concentration in consumed portion of fish. ECO - consider uncertainty in average concentration of whole fish. |
| Ground-water | Short-term | | X | small - single well head | HH - consider either the highest detected concentration or uncertainty around the concentration at the center of the plume as a measure of a single well and relate to the time scale of the toxic effect. ECO - not applicable |
| Ground-water | Long-term | | X | small - single well head | HH - consider variability among the higher concentration samples as a protective EPC. Alternatively, hydrogeologic modeling may be used to obtain a long-term average concentration in the most contaminated area. ECO - not applicable |

C.5.0 METHODS FOR ESTIMATING UNCERTAINTY IN THE MEAN CONCENTRATION

Confidence intervals (CIs) and UCLs are computed to characterize uncertainty in a parameter estimate. CIs can be computed for any parameter. The general method for estimating confidence intervals is presented in equation C-1.

$$CI = \text{parameter estimate} \pm (\text{critical value}) \times SE \quad \text{Equation C-1}$$

The parameter estimate is the estimated value for the unknown population parameter. The critical value is the number, *z*, with probability, *p*, lying to its right (for an upper critical value) or left (for a lower critical value). For a standard normal distribution (i.e., arithmetic mean=0, standard deviation=1), critical values are referred to as the *z-score* or *z-statistic*. These values are commonly given in statistics texts, and

may also be calculated using the Microsoft Excel function $Normsinv(p)$, where p corresponds to the probability lying to the right of the value. Distributions that characterize parameter uncertainty are sometimes referred to as sampling distributions. The standard error (SE) is the standard deviation of the sampling distribution for the parameter estimate. The confidence interval conveys two concepts: (1) an upper and lower confidence limit (for a 2-sided CI), and (2) a confidence level ($1-\alpha$), which gives the probability that the method yields an interval that encloses the parameter (Moore and McCabe, 1993). Methods for estimating SE vary for specific parameters. For example, the SE of a mean concentration may be calculated based on the sample variance and the sample size (due to Central Limit Theorem). Methods for calculating the SE for other parameters, such as the 95th percentile, are more complex, and may be estimated from a series of nested bootstrap simulations (Efron and Tibshirani, 1993; U.S. EPA, 2001a).

When comparing alternative approaches for quantifying parameter uncertainty, criteria that are important to consider include the variance of the original data set, and the bias and coverage of the CIs generated by each method. In statistics, a method is unbiased if the mean of the sampling distribution is equal to the true value of the parameter. Similarly, a method has accurate coverage if the probability p that a CI does not cover the true parameter is equal to the probability level used to construct the CI. For risk assessment, the most desirable method is one that deals well with high variance, yields CIs that are sufficiently wide (i.e., the CI does not underestimate the probability of enclosing the population parameter), and, more specifically, yields upper confidence limits that are not biased low. The choice of the most appropriate method will depend on the characteristics of the data set and a balance between two objectives: (1) the desire to be health protective and, therefore, have a low probability of underestimating the mean, and (2) a desire to be accurate, in the sense of choosing a method whose expected coverage equals the true coverage. As a general principle for quantitative uncertainty analysis, if alternative methods yield very different answers, it is helpful to explore the reasons for the differences. The objective is to explain why the estimates of the 95% UCL differ, and to determine if the differences are sufficiently great that they could alter the risk management decision or PRG. This information should be presented as part of the risk communication process associated with the scientific management decision points of the tiered process for PRA (see Chapter 2).

As discussed in Chapter 5, in Superfund risk assessment, the EPC is usually calculated as the 95% UCL for the mean to account for the uncertainty in estimating the average concentration within an EU. The 95% UCL is defined as a value that, when repeatedly calculated for randomly drawn subsets of size (n), equals or exceeds the true population mean 95% of the time. In other words, it is calculated and applied as a 1-sided confidence limit. The 95% UCL is one percentile on the probability distribution that characterizes uncertainty in the mean (i.e., the PDFu for the mean). It is equal to the 95th percentile of the sampling distribution for the mean. EPA's guidance on calculating the concentration term describes the rationale and methodology for selecting the 95% UCL as the point estimate for the concentration term (U.S. EPA, 1992).

Common methodologies for characterizing the 95% UCL for the arithmetic mean concentration include the following: (1) application of Equation C-1 using Student's t-statistic (for normal distributions), (2) Land method using H-statistic (for lognormal distributions) (Land 1971, 1975), and (3) bootstrap and Jackknife resampling techniques (Efron and Tibshirani, 1993). Details on these methods and on choosing an appropriate method are provided in the ORD/OSWER guidance bulletin, *Lognormal Distribution in Environmental Applications* (U.S. EPA, 1997a), and the more recent OSWER guidance bulletin, *Guidance on Calculation of UCLs at Superfund Sites* (U.S. EPA, 2001a). An overview of methods that may be used when data are not normal or lognormal is also provided by Schulz and Griffin (1999). It is the responsibility of the regional risk assessor to ensure that an appropriate method for

calculating a UCL or for developing an uncertainty distribution is chosen. Chapter 3 (Table 3-1) provides an overview of approaches for characterizing uncertainty in the concentration term in both 1-D MCA and 2-D MCA.

C.5.1 QUANTIFYING UNCERTAINTY WITHOUT INFORMATION ABOUT LOCATIONS OF SAMPLES AND RECEPTORS

Knowledge of both the sampling locations and the receptor's activity patterns with the EU can be used to derive a more representative estimate of the 95% UCL. If a risk assessor has access to an environmental data set without information about the sample locations, the risk assessor is forced to assume that the sample consists of a number of independent observations. The validity of this assumption depends on the unknown spatial variability of contamination at the site. The size and location of an EU, as well as the choice of a statistical method for estimating the distribution of uncertainty around the mean concentration will require often implicit (and possibly incorrect) assumptions about the spatial distribution of contamination. Similarly, if information regarding receptor activity patterns is unavailable, one must assume that any area within the EU is equally representative of potential exposures. The risk assessor is urged to explore the effects of these various assumptions and to make choices that are protective of human health and the environment.

C.5.2 QUANTIFYING UNCERTAINTY WITH INFORMATION ABOUT LOCATIONS OF SAMPLES AND RECEPTORS

In classical statistics, observations are assumed to be independent. This assumption is often invalid at contaminated sites where the method by which a chemical is released into the environment (e.g., deposition from airborne emissions; migration of contaminant plume from a point source) results in positive spatial autocorrelation. In other words, observations located next to each other tend to contain similar levels of contamination (i.e., redundant information) (Griffith and Layne, 1999). For example, the higher the spatial autocorrelation, the less incremental information is provided by adding observations in close proximity to existing observations. This decrease in the information content of a site sample is exacerbated by the tendency to choose sampling locations in the most contaminated areas rather than distributed at regular spatial intervals or specified using random sampling methodology.

At many hazardous waste sites, environmental sampling plans are designed with remedial actions rather than risk assessment in mind. Therefore, the risk assessor must establish a correspondence between the actual sampling locations and the locations a receptor would be expected to frequent. Geostatistics may provide information to establish this correspondence.

Geostatistics is a branch of spatial statistics that can be used to model spatial variability and parameter uncertainty. Geostatistics offers two fundamental contributions to risk assessment: (1) a group of methods to describe the spatial distribution of a contaminant in a quantitative fashion, and (2) the ability to maximize the information available in the data set (Deutsch and Journel, 1988; Isaacs and Srivastava, 1989).

Geostatistics is capable of using the information revealed by a correlation analysis of the data to estimate concentrations at unsampled locations. For example, geostatistics is able to use the spatial information contained in the data to model uncertainty in contaminant concentrations for areas where data are sparse, a situation commonly encountered in site assessment work. Using geostatistics, information from samples collected from outside an EU can be used to model the uncertainty in the mean concentration within an EU. Approaches that do not consider the geospatial information present in the

data are limited to the subset of samples within an EU. However, this ability to model uncertainty in areas where data are sparse is also limited, and a well characterized site is still the best path to understanding the risk at that site.

Geostatistical methods may be used to calculate a distribution of uncertainty in the mean of the concentration term for use in PRAs. In the past, geostatistics has not been widely applied to risk assessment, even though uncertainty in the exposure concentration is often a major source of uncertainty in risk estimates. Most risk assessors quantify uncertainty in the long-term average concentration without explicitly considering the spatial information present in data obtained from environmental sampling or knowledge of the receptor's movement and activities within the EU. When spatial information does not exist, the inherent assumption is that environmental sampling yields a data set that is representative of the spatial variability in concentrations encountered by a receptor. This assumption represents one source of uncertainty in the EPC. In addition, data collected outside an EU are often ignored in the analysis, even though they can provide a more comprehensive view of patterns of contamination across the site, including the EU of interest. Ignoring site-wide information may result in less informed estimates of risk and, therefore, less effective remedial designs (i.e., too little or too much remediation). In the past five years, with rapidly expanding software and hardware capabilities, some examples of the application of geostatistics can be found in exposure assessment and remedial design (e.g., Gomez-Hernandez, 1996; Goovaerts, 1996, 1997; Kriakidis, 1996; Ginevan and Splitstone, 1997; McKenna, 1997, 1998) as well as site assessment guidance (e.g., U.S. EPA, 2000).

A limit to applying geostatistics at hazardous waste sites is that the method is resource intensive and requires personnel experienced with the software and techniques. Risk assessors and risk managers should ensure that contractors and other personnel have the necessary capabilities before applying geostatistical methods to risk assessment or site cleanup. Geostatistics is a powerful tool, but it cannot incorporate quantitative knowledge regarding all sources of uncertainty. The risk assessor is cautioned to consider all possible sources of uncertainty as described in Chapter 5. As indicated previously, a full discussion of geostatistics is beyond the scope of this guidance, and interested readers are urged to consult the OSWER guidance document, *Guidance on Strategy for Surface Soil Cleanup at Superfund Sites* (U.S. EPA, 2001b).

EPA has produced several software packages used for geostatistical estimation. Among these are GEO-EAS and GEO-PACK. Expertise in geostatistics can be obtained from ORD/Las Vegas.

REFERENCES FOR APPENDIX C

- Burmaster, D.E. and K.M. Thompson. 1997. Estimating Exposure Point Concentrations for Surface Soils for use in Deterministic and Probabilistic Risk Assessments. *Human Eco. Risk Assess.* 3(3): 363-84.
- Cullen, A.C. and H.C. Frey. 1999. *Probabilistic Techniques in Exposure Assessment. A Handbook for Dealing with Variability and Uncertainty in Models and Inputs.* Plenum Press.
- Deutsch, C.V. and Journel, A.G. 1998. *Geostatistical Software Library and User's Guide, 2nd Ed.,* Oxford University Press, NY.
- Efron, B. and Tibshirani, R.J. 1993. *An Introduction to the Bootstrap.* Chapman and Hall, CRC Press.
- Ginevan, M.E. and D.E. Splitstone. 1997. Improving Remediation Decisions at Hazardous Waste Sites with Risk-Based Geostatistical Analysis. *Environ. Science Tech.* 31(2):92A-96A.
- Gomez-Hernandez, J.J. 1996. Issues on Environmental Risk Assessment. In: *Proceedings of the Fifth International Geostatistics Congress, Vol. 1.* Baafi, E.Y. and N.A. Schofield, eds.). Kluwer Academic Publishers, Dordrecht, 15-26.
- Goovaerts, P. 1996. Accounting for Local Uncertainty in Environmental Decision-Making Processes. In: *Proceedings of the Fifth International Geostatistics Congress, Vol. 2.* Baafi, E.Y. and N.A. Schofield, eds.). Kluwer Academic Publishers, Dordrecht, 929-940.
- Goovaerts, P. 1997. *Geostatistics for Natural Resources Evaluation.* New York: Oxford University Press.
- Griffith, D.A. and L.J. Layne. 1999. *A Casebook for Spatial Statistical Analysis.* Oxford University Press, NY.
- Hoff, D.J. 1998. Integrated Laboratory and Field Investigations Assessing Contaminant Risk to American Badgers (*Taxidea taxus*) on the Rocky Mountain Arsenal National Wildlife Refuge. Ph.D. Dissertation, Clemson University, Clemson, S.C.
- Hope, B.K. 2000. Generating Probabilistic Spatially-Explicit Individual and Population Exposure Estimates for Ecological Risk Assessment. *Risk Anal.* 20(5):575-590.
- Hope, B.K. 2001. A Case Study Comparing Static and Spatially Explicit Ecological Exposure Analysis Methods. *Risk Anal.* 21(6):1001-1010.
- Isaacs, E.H. and R.M. Srivastava. 1989. *An Introduction to Applied Geostatistics.* Oxford University Press, NY.
- Kriakidis, P.C. 1996. Selecting Panels for Remediation in Contaminated Soils via Stochastic Imaging. In: *Proceedings of the Fifth International Geostatistics Congress, Vol. 2.* Baafi, E.Y. and N.A. Schofield, eds.). Kluwer Academic Publishers, Dordrecht, 973-983.

- Land, C.E. 1971. Confidence Intervals for Linear Functions of the Normal Mean and Variance. *Ann. Math. Stat.* 42:1197–1205.
- Land, C.E. 1975. Tables of Confidence Limits for Linear Functions of the Normal Mean and Variance. In: *Selected Tables in Mathematical Statistics*, Vol. 3. American Mathematical Society, Providence, RI.
- McKenna, S.A. 1997. *Geostatistical Analysis of Pu-238 Contamination in Release Block D, Mound Plant, Miamisburg, Ohio*. SAND97-0270, Sandia National Laboratories, Albuquerque, NM.
- McKenna, S.A. 1998. Geostatistical Approach for Managing Uncertainty in Environmental Remediation of Contaminated Soils: Case Study. *Environmental and Engineering Geoscience*, 4(2):175-184.
- Meade R.H., T.R. Yuzyk, and T.J. Day. 1990. *Movement and Storage of Sediment in Rivers of the United States and Canada*. In: Wolman et al. (eds) *Surface Water Hydrology. The Geology of North America*. Geological Society of America, Boulder, CO.
- Moore, D.S. and G.P. McCabe. 1993. *Introduction to the Practice of Statistics*. W.H. Freeman and Company, NY.
- Piest R.F. and C.R. Miller. 1975. *Sediment Yields and Sediment Sources*. In: Vanoni V.A. (ed.) *Sedimentation Engineering*, American Society of Civil Engineers, NY.
- Schulz, T.W. and S. Griffin. 1999. Estimating Risk Assessment Exposure Point Concentrations when the Data are not Normal or Lognormal. *Risk Anal.* 19(4): 577– 584.
- U.S. EPA. 1992. *Supplemental Guidance to RAGS: Calculating the Concentration Term*. Office of Solid Waste and Emergency Response. Washington, DC. OWSER Directive No. 9285.7-081.
- U.S. EPA. 1997a. *Lognormal Distribution in Environmental Applications*. Office of Research and Development and Office of Solid Waste and Emergency Response. Washington, DC. EPA/600/R-97/006.
- U.S. EPA. 1997b. *Memorandum from Deputy Administrator Fred Hansen on the Use of Probabilistic Techniques (including Monte Carlo Analysis) in Risk Assessment, and Guiding Principles for Monte Carlo Analysis*. Office of Research and Development, Washington, DC. EPA/630/R-97/001. May 15.
- U.S. EPA. 2000. *Statistical Estimation and Visualization of Ground-water Contamination Data*. Office of Research and Development, Washington, DC. EPA/600/R-00/034.
- U.S. EPA. 2001a. *Draft Guidance on Calculation of Upper Confidence Limits for Exposure Point Concentrations at Superfund Sites*. Office of Emergency and Remedial Response, Washington, DC.

U.S. EPA. 2001b. *Supplemental Guidance for Developing Soil Screening Levels for Superfund Sites*. Peer review draft. Office of Solid Waste and Emergency Response. Washington, DC. OSWER No. 9355.4-24. March.

Van Sickle J. and R.L. Beschta. 1983. Supply-Based Models of Suspended Sediment Transport in Streams. *Water Resour. Res.* 19:768–78.

Walling D.E. 1983. The Sediment Delivery Problem. *J. Hydrol.* 65:209–37.

APPENDIX D

ADVANCED MODELING APPROACHES FOR CHARACTERIZING VARIABILITY AND UNCERTAINTY

D.0 INTRODUCTION

This appendix briefly describes the following advanced modeling approaches that can be used in probabilistic risk assessment (PRA) to characterize variability and uncertainty: two-dimensional MCA (2-D MCA), microexposure event analysis (MEE), geospatial statistics, and Bayesian analysis. Except for 2-D MCA, these approaches can also be applied to point estimate risk assessment. The application of many of these approaches will require access to expertise in specialized areas of statistics and, in some cases, specialized or even custom-designed computer software. The intent here is to introduce some of the basic concepts and terminology, as well as to provide references where the reader can find more exhaustive coverage of these topics.

D.1.0 EXPRESSING VARIABILITY AND UNCERTAINTY SIMULTANEOUSLY

A Monte Carlo analysis that characterizes either uncertainty or variability in each input variable (see Chapter 1) can be described as a one-dimensional Monte Carlo analysis (1-D MCA). A 2-D MCA is a term used to describe a model that simulates both uncertainty and variability in one or more input variables. All probability distributions that are used to describe variability in a PRA model have a certain degree of associated uncertainty. For example, suppose variability in soil concentration (ppm) is estimated using a normal probability density function (PDF) defined by a mean ($\mu_{\text{soil}}=5$) and standard deviation ($\sigma_{\text{soil}}=1$), and subjectively truncated (min, max) at (0, 50). Uncertainty in the parameter estimates can be represented in a PRA model by assuming both parameters are also random variables. To illustrate this concept, assume normal PDFs for *uncertainty* can be specified for both parameters. Uncertainty in the mean is described by the normal PDF with parameters ($\mu_{\text{mean}}=5$, $\sigma_{\text{mean}}=0.5$); similarly, uncertainty in the standard deviation is described by the normal PDF with parameters ($\mu_{\text{SD}}=1$, $\sigma_{\text{SD}}=0.5$). Model variables are represented in this manner when there is a compelling reason to believe that a unique probability distribution does not adequately describe one's knowledge of each variable in the model. A variable described in this way is called a second order random variable. Figure D-1 (Panel A) shows a collection of $n=20$ cumulative probability distributions (CDFs), each curve representing a unique set of (mean, SD) parameter estimates for the normal PDF for variability. Panel B shows the 90% *confidence interval*¹ based on 2,500 simulated CDFs. The 95% lower and upper bounds correspond to the distribution of 5th percentiles and 95th percentiles, respectively (i.e., CDF for 2,500 5th percentiles and CDF for 2,500 95th percentiles). The 90% credible interval (CI) for the 50th percentile is (3.4, 6.7).

¹Note that the term "credible interval" may be more appropriate than "confidence interval" given that the range is based on subjective as well as statistical considerations. Brattin, Barry, and Chiu (1996) provide additional examples of uncertain PDFs that illustrate this concept.

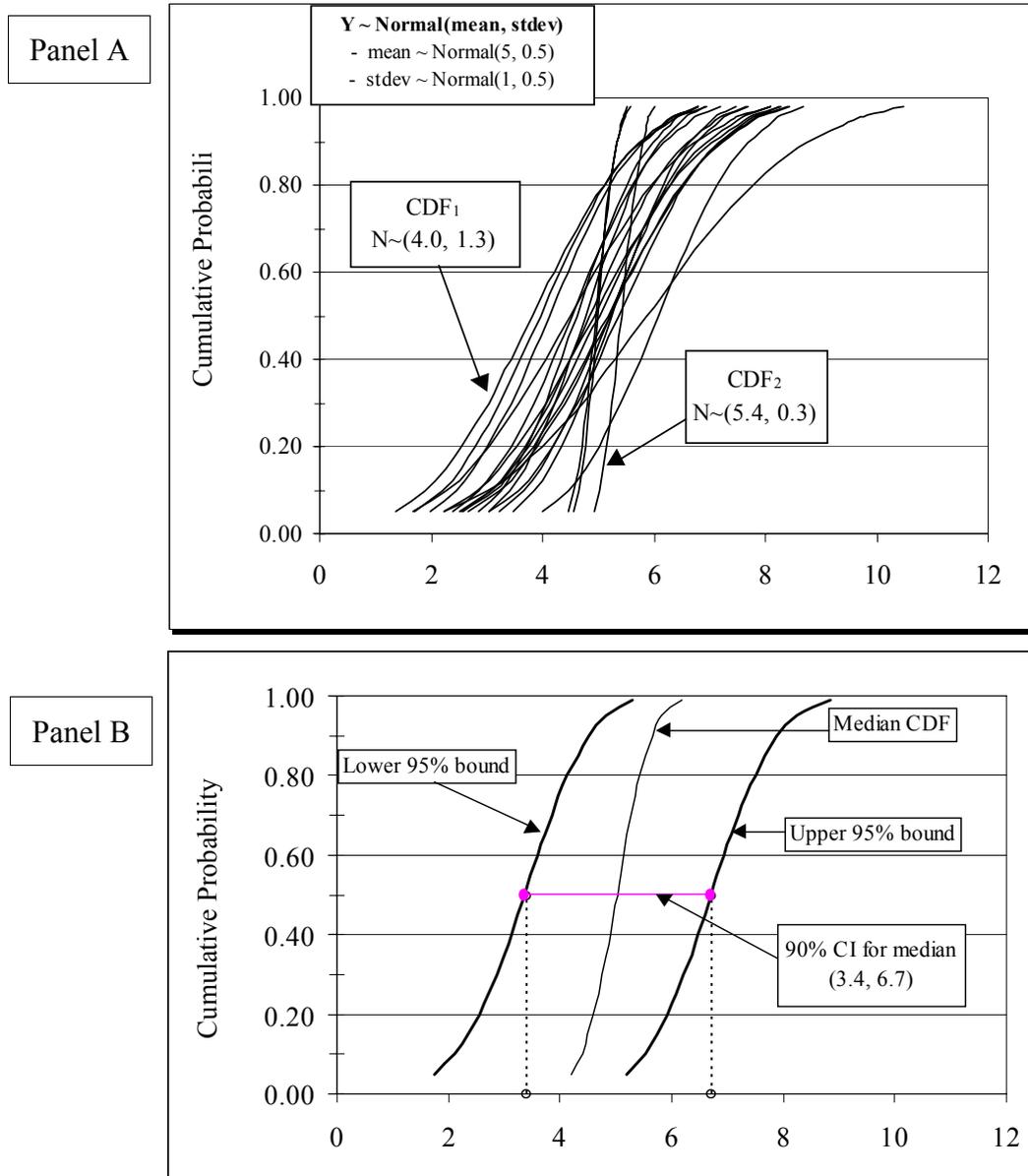


Figure D-1. Panel A shows a family of 20 CDFs for a hypothetical random variable, Y (e.g., concentration in units of ppm), characterized by a normal PDF where both the mean and SD are also random variables representing uncertainty in the parameter estimates: Mean \sim Normal(5, 0.5), SD \sim Normal(1, 0.5). Each CDF represents a single simulation of $n=2500$ iterations using a unique set of parameters. For example, CDF₁ represents $N \sim (4.0, 1.3)$ while CDF₂ represents $N \sim (5.4, 0.3)$. Panel B shows the “90% credible interval” for the CDF based on 2,500 simulations, each simulation using $n = 2500$ iterations (i.e., a 2-D MCA with 2,500 outer loop iterations and 2,500 inner loop iterations). Lower, median, and upper bounds represent the simulated 5th, 50th, and 95th percentiles, respectively. The 90% confidence interval for the estimate of the 50th percentile is: {3.4, 6.7}.

EXHIBIT D-1

DEFINITIONS FOR APPENDIX D

Bayesian Statistics - A specialized branch of statistics that views the probability of an event occurring as the degree of belief or confidence in that occurrence.

Geospatial Statistics - A specialized branch of statistics that explicitly takes into account the georeferenced context of data and the information (i.e., attributes) it contains.

Frequentist - A term referring to classical statistics in which the probability of an event occurring is defined as the frequency of occurrence measured in an observed series of repeated trials.

Image Analysis - A technique in geostatistics used to restore a degraded image or interpret images that have been contaminated by noise or possibly some nonlinear transformation.

Kriging - A geostatistical method of spatial statistics for predicting values at unobserved locations.

Likelihood Function - A Bayesian term referring to a probability distribution expressing the probability of observing a piece of new information given that a particular prior belief is true.

Location Tag - The spatial coordinates of a sampling location (e.g., longitude, latitude).

Microexposure Event Analysis (MEE) - An approach to modeling exposure in which long-term exposure of an individual is simulated as the sum of separate short-term exposure events.

Point Pattern Analysis - A technique in geostatistics of restricting the analysis to location information, ignoring attribute information, addresses two location problems: (1) describing points according to spacing, and (2) describing points according to density.

Posterior Distribution - A Bayesian term referring to a probability distribution that has been updated with new information.

Prior Distribution - A Bayesian term referring to the hypothesized, expected, or calculated probability distribution for an event prior to the collection of new information.

Spatial Autocorrelation - The tendency of data from locations that are relatively close together to be geographically correlated.

Thiessen (Voronoi) Polygon Analysis - A method of spatial statistics in which an area is subdivided into subregions, or polygons, in order to predict values at unobserved locations.

Time Step - A modeling term used to describe the time interval within which variable values do not change.

Two-Dimensional Monte Carlo analysis (2-D MCA) - Separate representation of variability and uncertainty in an MCA, usually accomplished using nested computation loops.

In the example shown in Figure D-1, the mean and standard deviation for soil concentration were allowed to vary independently. Thus, a distribution could be defined by a combination of a low mean and a high standard deviation, high mean and low standard deviation, or any other combination in between. The assumption of independence of variable parameters may not be valid in all cases. It may be unreasonable to assume that a high mean soil concentration would occur with a low standard deviation. An alternative assumption would be that the standard deviation of the mean is a constant proportion of the mean (i.e., a constant coefficient of variation). Correlations between parameters should be considered in the design of the PRA. One approach that is especially useful for characterizing relationships between the slope and intercept of a simple linear regression is to specify the bivariate normal distribution for the parameter estimates.

D.2.0 TWO-DIMENSIONAL MONTE CARLO ANALYSIS (2-D MCA)

Two-dimensional MCA is an approach for computing risk (or hazard) when combining distributions that represent variability and uncertainty. In 2-D MCA, distributions representing variability and uncertainty are sampled using nested computational loops (Figure D-2). The inner loop simulates variability by repeatedly sampling values for each variable from their defined probability distributions. With each circuit of the outer loop, new parameter values for each variable are selected, and the inner loop sampling is repeated. The result is a collection of inner loop simulations, one for each parameter value selected. If the inner loop samples 5,000 times, and the outer loop samples 1,000 times, then each

variable is sampled 5,000,000 times and 1,000 simulated probability distributions of risk are generated from the PRA model. These probability distributions can be analyzed to estimate the distributions for specific risk estimates. For example, confidence limits on the estimate of specific risk percentiles can be simulated using 2-D MCA (Figure D-3).

Simulation Logic for 2-Dimensional MCA

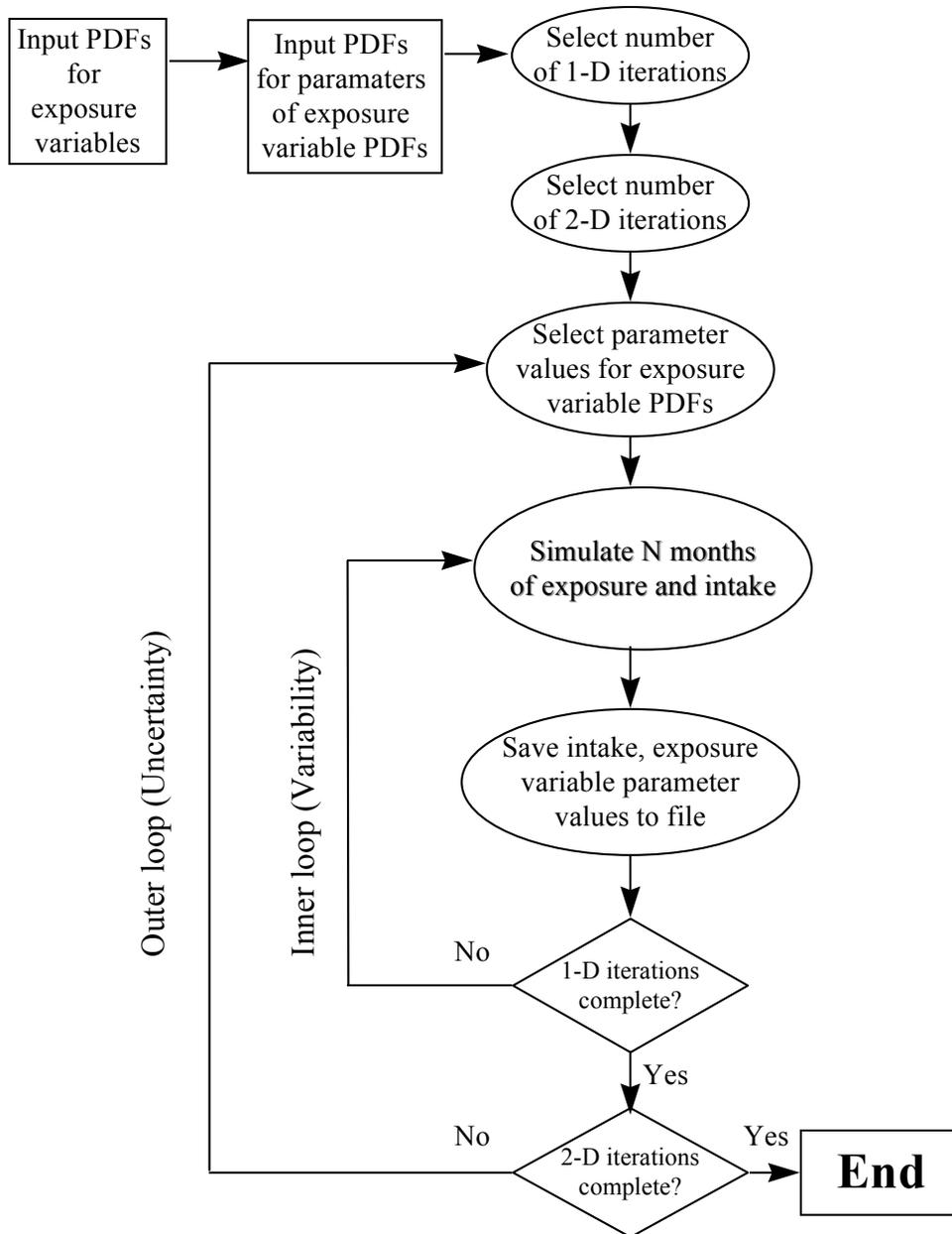


Figure D-2. Diagram showing of a 2-D Monte Carlo model in which the variability and uncertainty dimensions are computed in nested loops. In this example, values for exposure variables in the inner loop represent monthly averages.

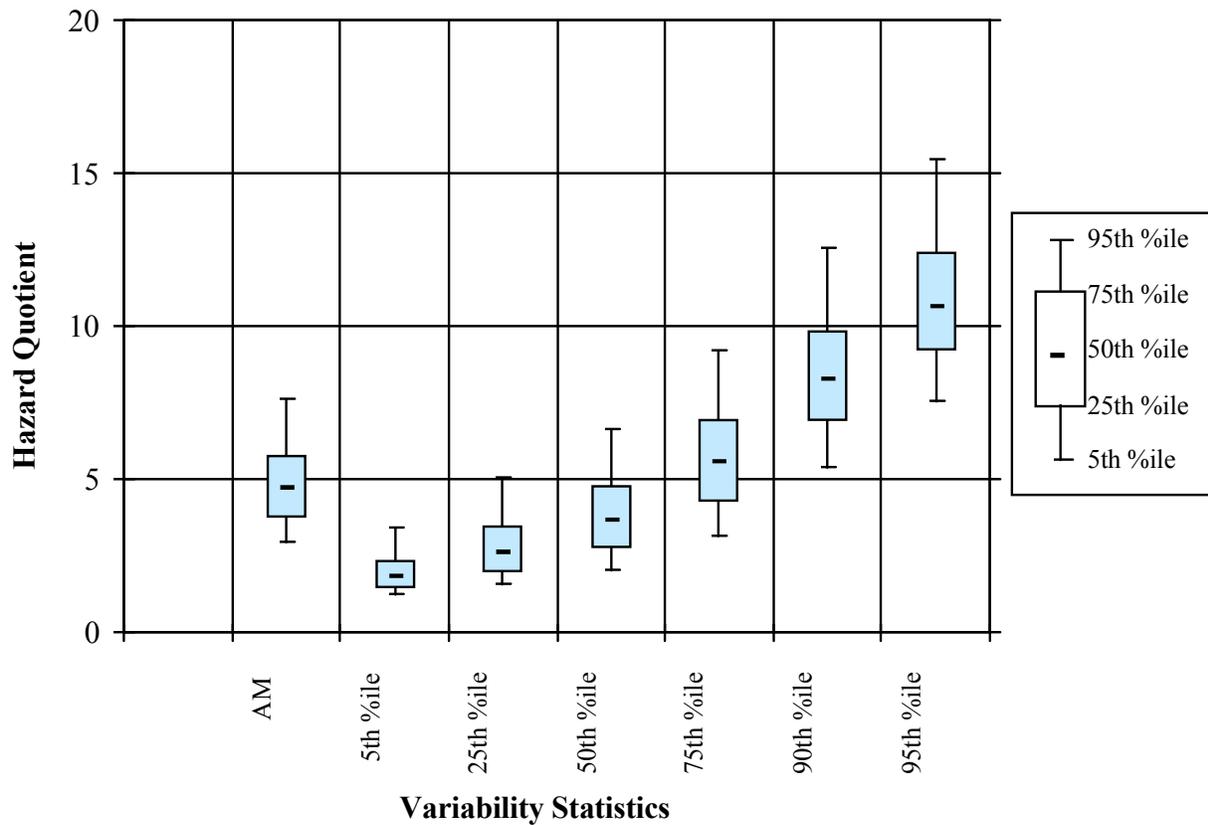


Figure D-3. Output from a 2-D MCA showing the estimated mean Hazard Quotient (HQ) and the 90% confidence interval for the arithmetic mean (AM) and selected percentiles of the HQ distribution. The 95th %ile HQ would be the reasonable maximum exposure (RME) risk estimate. The simulation suggests that there is a 95% probability that the RME HQ (95th percentile) is below 16.

D.3.0 MICROEXPOSURE EVENT ANALYSIS

The standard dose equation generally used in Superfund site risk assessments represents exposures averaged over a specified time period that is relevant to the health endpoint of concern (Equation D-1). If the risk assessment is directed at assessing life-time risk to humans, the averaging time used in Equation D-1 would generally be 70 years (i.e., estimated average human lifetime), and the calculated chemical intake would generally represent the life-time average daily dose (LADD). Where information is available to characterize variability on a smaller time scale than life-time, an alternative expression of dose that accommodates such variability may be desirable.

Concentrations in various environmental media can be expected to vary over time. For example, wind erosion may change chemical concentrations in surface soil. Leaching may change concentrations in both subsurface soil and groundwater. The change in the concentration term is most readily apparent when considering anglers harvesting fish. If an angler consumes a large amount of fish from a single location (e.g., a specific lake, pond, or river), then the average chemical concentration in the fish consumed by that angler can be expected to be similar to the average of the chemical concentration of fish in the population. However, if an angler consumes fish only occasionally, or harvests fish from different locations, there will be considerably more uncertainty in the concentration term. In addition, a harvesting angler may consume varying amounts of fish over the period of the exposure duration due to changing tastes, changes in the fish population size or other factors.

Daily activity patterns, food intake, soil ingestion and other behavioral factors are measured in a time period of less than a year. The extrapolation of these short term results to the chronic exposure situation is a source of uncertainty. Exposure events are real but unknowable, whereas data regarding the nature and magnitude of these events is known but its application to a real world situation is uncertain. Microexposure event analysis (MEE) attempts to explicitly quantify this uncertainty. Figure D-5 presents the general approach for MEE analysis. (Price et al., 1996, 2000). MEE modeling provides an alternative to the standard time-averaging approach represented by Equation D-1. In the MEE approach, long term intake is viewed as the sum of individual exposure events (Equation D-2). Implementing the MEE approach in a PRA requires dividing the exposure duration into short epochs, or time steps, within which the values assigned to exposure variables remain constant, but are allowed to vary from one time step to the next. In a PRA model, exposure variables are adjusted at each time step by selecting values from the probability distributions representing each variable (Figure D-4). Discussion of the implementation of

Standard Time-Averaging

$$\text{DOSE} = \frac{C \times \text{IR} \times \text{EF} \times \text{ED}}{\text{BW} \times \text{AT}} \quad \text{Equation D-1}$$

Microexposure Event Modeling

$$\text{DOSE} = \frac{1}{\text{AT}} \sum_{j=1}^{\text{ED}} \frac{1}{\text{BW}_j} \sum_{i=1}^{\text{Events}_j} C_{ij} \cdot \text{IR}_{ij} \quad \text{Equation D-2}$$

C = Concentration; I = exposure event; j = year of life
 IR = Intake Rate
 EF = Exposure Frequency
 ED = Exposure Duration
 BW = Body Weight
 AT = Averaging Time

MEE analysis in risk assessment and its merits and limits can be found in Wallace et al. (1994), Price et al. (1996), Slob (1996), and Buck et al. (1997).

In MEE modeling, the time step becomes an important variable, with associated uncertainty. The time step should be selected based on information available to describe how exposures change over time. For example, a model of a moving plume of solvents in groundwater might suggest that chemical concentrations in a given location are dropping by between 16 and 25% quarterly. Several rounds of sampling may support this prediction. This rapid decline in concentrations suggests that an appropriate time step might be one quarter (i.e., three months).

On the other hand, where risk is being assessed for metals, dioxin, or PAHs in soil, the concentrations might be expected to change much more slowly, if at all, and the basis of the time step might be the increase in age and corresponding changes in behavior of the receptor. The time step may be global; that is, one time step may apply to all variables in the model. In this case, the same number of random values would be selected for each exposure variable in a Monte Carlo simulation. A more complex model may use different time steps for different variables, requiring some probability distributions to be sampled more often than others. The selection of a value for a time step implies that the value represents the average value for that variable during the time step.

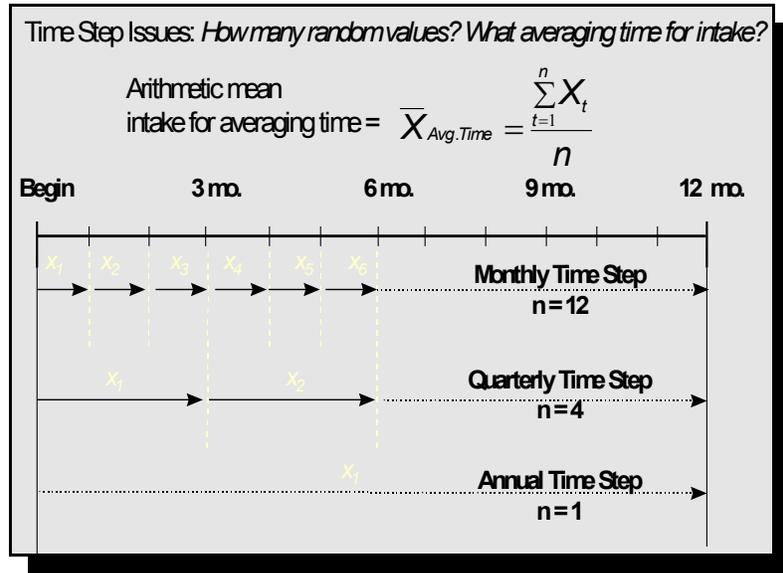


Figure D-4. Time Step for MEE.

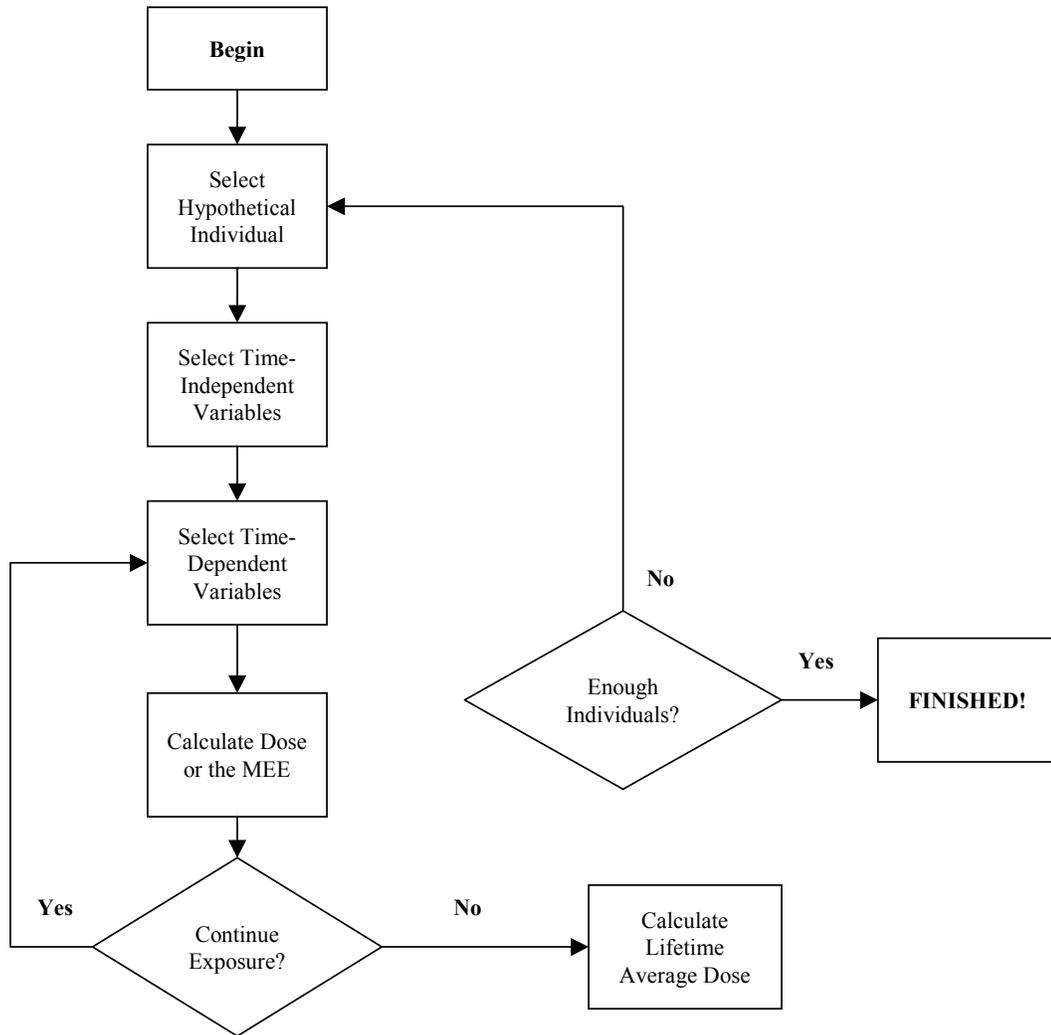


Figure D-5. Flowchart showing general approach for Microexposure Event (MEE) analysis.

Two important issues related to time step should be considered in implementing the MEE approach in PRA models. The first is the relationship between the length of the time step and the number of times random values are generated from a defined probability distribution. As the time step decreases, more time steps are needed to simulate exposures over a specified duration. For example, given a time step of one year and an exposure duration of 30 years, each random variable will be sampled 30 times (once per year); for a time step of one month and an exposure duration of 30 years, each random variable would be sampled 360 times (i.e., 12 months/year x 30 years). The Central Limit Theorem indicates that as n increases, the distribution of sample means is approximately normal, and the standard deviation of the sample distribution is inversely proportional to the square root of n . Thus a highly skewed input distribution (e.g., lognormal) may tend to become less skewed with increasing n (Figure D-6). A biased estimate of the RME risk in a PRA model may result if an inappropriately small or large time step is used in the model. This emphasizes the importance of having an empirical basis for selecting the time step and of exploring the time step as a variable in a sensitivity analysis of the model.

The second issue related to the time step concerns temporal correlations. Is it reasonable to assume that random values selected for consecutive time steps are completely independent? For example, consider body weight. The body weights of an individual measured at different times would be expected to show positive temporal autocorrelation; that is, body weight is likely to be similar (but not constant) from one time step to the next. For example, if an individual weighs 60 kg during one month, it is unlikely that they will weigh 80 kg the next month. If this scenario is accepted, then body weight should not be allowed to vary independently from one monthly time step to the next in the model. At shorter time steps, temporal correlation becomes more likely as a result of temporal autocorrelation. For example, one can expect a higher correlation between body weights on an individual measured on two successive days (one-day time step) than between weights measured at the midpoint of two successive years. Approaches to simulating temporal correlations in probabilistic models might include fixing an individual within a percentile range of a distribution (e.g., randomly assigned quartile) or using randomly assigned fluctuations (e.g., $BW_t = BW_{t-1} \pm x$).

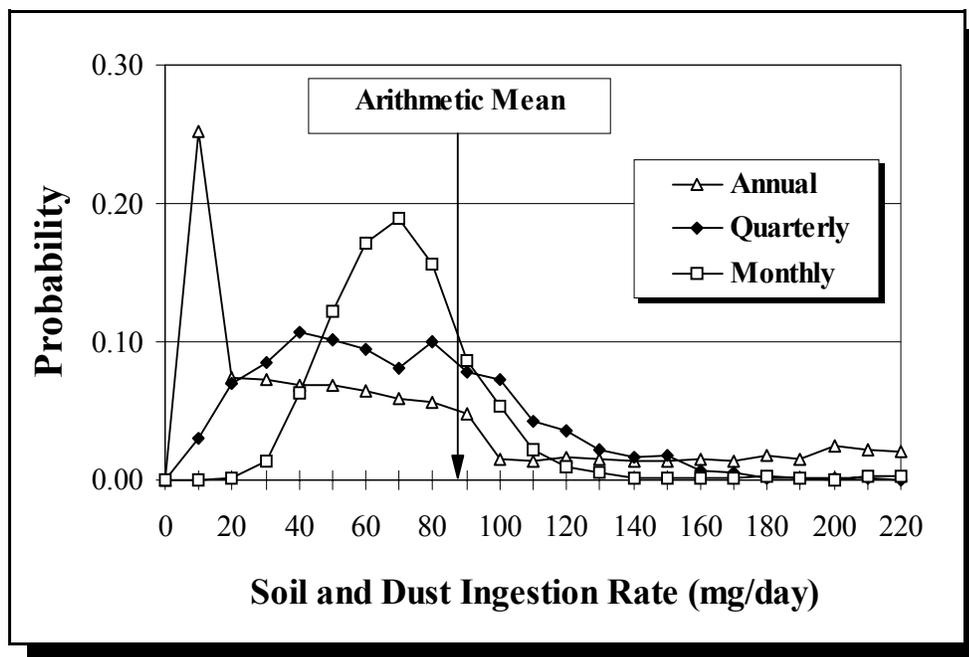


Figure D-6. Hypothetical example showing the effect of model time step on the probability distribution for soil and dust ingestion rate in children over a 1-year period. Number of samples (n) needed to simulate exposures: Annual (1), Quarterly (4), Monthly (12).

D.4.0 GEOSPATIAL STATISTICS

Spatial statistics is a specialized branch of statistics, falling under the heading of multivariate statistics, that explicitly takes into account the georeferenced or locational tagged context of data. Generally, environmental samples collected at Superfund sites have this geolocational information. By acknowledging the geography of site chemicals, information about the spatial distribution of contamination can be incorporated into an exposure assessment. In addition, knowledge about a receptors home range or patterns of movement may also be incorporated into the definition of the exposure unit (see Appendix C, Section C.2.0). Explicitly accounting for spatial relationships may lead to a more accurate estimate of the confidence limits for the arithmetic mean concentration. Geospatial statistics quantifies the spatial autocorrelation (Exhibit D-2) of sample measurements and allows for the exploration of the spatial distribution of exposure and risk using techniques of map generalization. By recording locational tags for each sample, information about spatial patterns within an exposure unit (EU) can be exploited to estimate both pre- and post-remediation exposure and risk.

In the past five years, with rapidly expanding software and hardware capabilities, some examples of the application of geostatistics can be found in exposure assessment and remedial design (e.g., Gomez-Hernandez, 1996; Goovaerts, 1996, 1997; Kriakidis, 1996; Ginevan and Splitstone, 1997; McKenna, 1998; Hope, 2000; 2001) as well as site assessment guidance (e.g., U.S. EPA, 2000).

Several important risk assessment issues are closely linked to geospatial statistics, as described in Exhibit D-3. Geospatial statistics comprises:

- *spatial autoregression*
- *geostatistics*
- *point pattern analysis*
- *image analysis*

The first three of these subjects can contribute to spatial statistical support of site risk assessments. The key concept linking all three is spatial autocorrelation, which refers to covariation among samples for a single chemical, or the tendency of data from locations that are relatively close together to be geographically correlated. By analogy, classical statistics treats soil samples as though they are balls, each having a battery of attributes, that can be placed into an urn for statistical analysis; geospatial statistics treats soil samples as though they are clusters of grapes,

EXHIBIT D-2

POSITIVE SPATIAL AUTOCORRELATION

- Locations with a high value of Y tend to be surrounded by nearby high values of Y.
- Locations with a medium value of Y tend to be surrounded by nearby medium values of Y.
- Locations with a low value of Y tend to be surrounded by nearby low values of Y.

EXHIBIT D-3

EXAMPLES OF RISK ASSESSMENT ISSUES LINKED TO GEOSPATIAL STATISTICS

- Sampling tends to disproportionately represent “hot spots” (i.e., a relatively large portion of a data set with a small sample size (n) tends to be concentrated at “hot spots”).
- The upper confidence limit (UCL) for the arithmetic mean exposure concentration (e.g., chemical concentrations in soil) depends on the sample size.
- Additional sampling may be needed, especially to better define the spatial patterns or the extent of contamination.
- There is uncertainty about locations not sampled at a site, as well as uncertainty regarding the representativeness of neighboring samples in nearby EUs.

with the branchy stems representing locational tags. Concentrations located on the same “branch” will be more strongly correlated than concentrations on different branches.

How is Geostatistics Different from Classical Statistics?

In general, geostatistics provides information beyond that provided by classical statistical techniques for at least two reasons. First, in classical statistics, observations are assumed to be independent. This assumption is often invalid at contaminated sites where the method by which a chemical is released into the environment (e.g., deposition from airborne emissions; migration of contaminant plume from a point source) often results in positive spatial autocorrelation (see Section D.4.1). In other words, observations located next to each other tend to contain similar levels of contamination (i.e., redundant information). For example, the higher the spatial autocorrelation, the less incremental information is provided by adding observations in close proximity to existing observations. This issue is compounded when the sample locations have been preferentially determined (e.g., “hot spot” sampling) rather than distributed at regular intervals or specified using random sampling methodology.

Second, geostatistics is able to use the geospatial information contained in the data to model uncertainty in contaminant concentrations for areas where data are scarce, a situation commonly encountered in site assessment work. Using geostatistics, information from samples collected from outside an EU can be used to model the uncertainty in the mean concentration within an EU. Approaches that do not consider the geospatial information present in the data are limited to the subset of samples within an EU.

D.4.1 CORRELATION AND SPATIAL AUTOCORRELATION

Several simple bivariate statistical approaches may be used to introduce the concept of spatial autocorrelation. Consider two variables, X and Y. For positive correlation there is a tendency for high values of X to be paired with the high values of Y, medium values of X to be with the medium values of Y, and low values of X with the low values of Y. The tendency is in the opposite direction for negative correlation; high values of X tend to be paired with low values of Y, and so on. Spatial autocorrelation, which virtually always is positive, directly parallels these definitions, but is written in terms of a single variable as shown in Exhibit D-2.

Just as the bivariate relationship between two variables, X and Y, can be portrayed by a scatter plot (Y versus X), the spatial autocorrelation relationship can be portrayed for a single variable, Y, (e.g., Y versus Y). A good example is the Moran scatterplot, which plots the sum or average of nearby values of Y versus Y. This plot is most effective when Y has been converted to z-scores. As shown in Figure D-7 and Section D.4.2, scatter plots can be used to illustrate some important issues related to sample size.

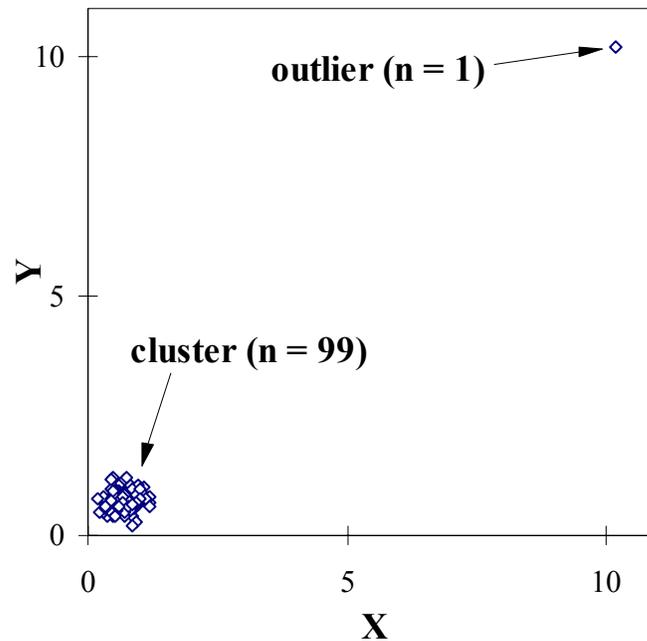


Figure D-7. Effect of an outlier on measured correlation: $r=0.956$ with outlier ($n=100$), whereas $r=0.086$ excluding outlier ($n=99$ clustered points).

If no soil samples were collected at a site ($n=0$), there is no information about the chemical concentrations in soil, and any guess may be considered an estimate. However, if the chemical concentration of a single sample ($n=1$) is measured, some information is obtained that partly restricts this estimate. As each additional independent sample is taken, more information is obtained, and the restriction on the estimate becomes more binding. If the same location is selected repeatedly for sampling, then the repeated measures, which may vary through time, will tend to be highly positively correlated; part of the information obtained from each sample is the same, and should not be counted more than once in estimating the site-wide soil concentration. Similarly, if immediately adjacent locations are sampled, the measures will often tend to be highly positively correlated (spatial autocorrelation). Once the first sample is taken, each additional sample provides only a fractional increment of new information about the site in its entirety.

D.4.2 EFFECTIVE SAMPLE SIZE (n^*) AND DEGREES OF FREEDOM

Repeated measures can result in data clustering, which can be illustrated in a scatter diagram. Because two points determine a straight line, if $(n-1)$ points cluster together on a scatter diagram while a single additional point occurs far away from this cluster (i.e., an outlier), then the resulting bivariate correlation will be very high (see Figure D-7). This situation alludes to the notion of effective sample size (n^*): the n^* is no longer equal to the number of observations (n), but rather is dramatically reduced by the presence of inter-observational correlation. For the example shown in Figure D-7, n^* is slightly greater than 2 rather than 100 (i.e., n).

Spatial autocorrelation plays an analogous role in georeferenced data. If a sampling network is arranged as a 25-by-25 square grid (one sample point per grid cell), and superimposed over a large site so that a very large distance separates nearby sample locations, then essentially zero spatial autocorrelation should be present in the geographic distribution of the concentrations of any given chemical. Concentrations will appear to be haphazard across the site, rendering the effective sample size as $n^*=625$. If the distance between nearby locations on the sampling mesh is decreased so that the spatial correlation is only $r=0.050$, then the effective sample size decreases to $n^*=514$. The effect of reducing the inter-sample distance on spatial autocorrelation and n^* for a 25-by-25 grid is shown in Exhibit D-4. If r increases to 1, then n^* reduces to 1. Therefore, obtaining a measure of latent spatial autocorrelation is essential to estimating n^* ; this in turn is critical to determining confidence limits

| r | n* |
|-------|-----|
| 0.000 | 625 |
| 0.050 | 514 |
| 0.539 | 64 |
| 0.957 | 3 |
| 1.000 | 1 |

for estimates of mean concentrations, which are sensitive to sample size. The UCL for the mean will be biased *only* when very high levels of spatial autocorrelation are present; this is because the Student-t statistic used to estimate the UCL (assuming a normal distribution) changes very little as the degrees of freedom (related to sample size) increases above 10; part of the difference between n and n^* is offset by an inflation of the variance.

The concept of effective degrees of freedom is important in exposure assessment because high positive spatial autocorrelation can bias the estimate of the UCL concentration if geospatial statistics are not considered. This should be of particular concern when specific locations at a site are intensively sampled (e.g., suspected “hot spots”), and other locations are relatively undersampled. Accordingly, the design of the sampling network itself can be evaluated from the perspective of geospatial statistics in order to ascertain the quality of sample information. The ideal sampling network should provide geographic representativeness, should be roughly uniformly distributed over a site, and is best implemented as a stratified random sampling design; that is, the site is partitioned into geographic stratum (e.g., EUs), and then a random sampling of points is selected within each strata. In practice, sample designs may need to focus on objectives that are in conflict with the above ideals. For example, intense sampling of suspected “hotspots” may be necessary at some sites, at the expense of a more representative spatial coverage of the site. In such cases, several statistical techniques are available for assessing the statistical benefit (in terms of reducing uncertainty) of additional sampling at undersampled locations.

D.4.3 ASSESSMENT OF ADDITIONAL SITE SAMPLING

Thiessen Polygons. In addition to calculating nearest neighbor statistics, the adequacy of a sampling network can be assessed by Voronoi (i.e., Thiessen polygon) surface partitioning, a popular approach used in mapping intra-site geographic distributions. This procedure divides a site into a mutually exclusive set of polygons, each polygon containing a single measured concentration. Each polygon has the unique property that any location within the polygon is closer to the polygon’s sample location than to any other sample point (Clifford et al., 1995). The concentration measured at the sample point in the polygon is assigned to the entire area of the polygon. The intensity of sample points on a surface can be measured by Equation D-3 mean inverse polygon areas:

$$SI = \frac{1}{m} \sum_{i=1}^m A_i^{-1} \quad \text{Equation D-3}$$

where SI is a measure of the sampling intensity, A_i is the area of the i^{th} polygon, and m is the number of interior polygons (those not along the edge of the site); $m < n$. The variance of the sampling intensity can be expressed by Equation D-4:

$$SI_{\text{Variance}} = \frac{1}{m-1} \left[\sum_{i=1}^m A_i^{-2} - \frac{1}{m} \left(\sum_{i=1}^m A_i^{-1} \right)^2 \right] \quad \text{Equation D-4}$$

If the sampling network is uniform (i.e., polygon areas are equal), the variance will be essentially zero. The variance will increase as the network deviates from uniform. This measure can be used to assess whether or not additional samples will improve the spatial coverage.

☞ Sampling locations that would yield a dramatic reduction in the variance should be given priority for future sampling efforts.

Thiessen polygons can be used to develop area-weighted estimates of the arithmetic mean concentration ($C_{\text{soil,w}}$) according to the following general equation:

$$C_{\text{soil,w}} = \sum_{i=1}^n C_i \frac{A_i}{A_T} \quad \text{Equation D-5}$$

where C_i is the concentration in the i^{th} polygon, A_i is the area of the i^{th} polygon in the EU, and A_T is the total area of the EU. The weight for each measurement is essentially the ratio of the area of each polygon to the total area of the site. Clifford et al. (1995) applied this approach to an ecological risk assessment of the burrowing owl with the following simplifying assumptions: habitat range is circular, size of EU is constant (75 ha) although location may vary, and organisms spend equal time in all portions of their habitat. Given these assumptions, a nonparametric bootstrap method can be used to determine the approximate 95% UCL for the mean concentration (see Appendix C). Using Monte Carlo analysis, $C_{\text{soil,w}}$ can be estimated for different locations of the EU according to Equation D-5, and confidence limits can be generated from the multiple bootstrap estimates. Burmaster and Thompson (1997) demonstrate a similar approach in which the EU (with constant area but random rectangular dimensions) is overlaid on the Thiessen polygon surface and 95% UCL for the mean is calculated from the bootstrap sample.

Linear Regression. Another diagnostic is found in the linear regression literature. The locational tag coordinates (e.g., longitude, latitude) can be converted to z-scores (say z_u and z_v) for the following calculation:

$$Y = \frac{1}{n} + \frac{z_u^2 + z_v^2 - 2r_{uv}z_u z_v}{(n-1)(1-r_{uv}^2)} \quad \text{Equation D-6}$$

where Y is a measure of the sampling network, r_{uv} is the correlation between the coordinate axes, and n is the number of samples. Any sampling location (z_u, z_v) in which $Y > 9/n$ may be considered too isolated in the sampling network. Additional sampling locations would be positioned closer to it to improve the overall coverage of the sampling network.

D.4.4 MAP GENERALIZATION

Another important application of geospatial statistics to risk assessment is that of map generalization, which draws on the subjects of geostatistics and spatial autoregression. Techniques developed for both topics exploit spatial autocorrelation in order to produce a map.

Kriging and Semivariograms. Geostatistics may employ kriging, which yields statistical guesses at values of a chemical at unsampled locations based on information obtained from sampled locations. Kriging assumes that the underlying geographic distribution is continuous, evaluates spatial autocorrelation in terms of distance separating sample points, and employs a scatter diagram similar to the Moran scatter plot to portray this relationship (i.e., the semivariogram plot: half the squared difference between measured concentrations for two sampled locations versus distance separating these two locations). The best-fit line to this scatter of points is described by one of about a dozen equations (semivariogram models).

Many different kriging approaches can be applied to quantify the spatial relationships among geographic attributes within an exposure unit. For example, site-specific chemical concentrations may be correlated with geologic information, such as glacial deposits, soil characteristics of core samples, and attributes that represent favorable habitats for ecological receptors. This information can be used to expand the available data and improve estimates of chemical concentrations at unsampled locations by employing a technique called co-kriging.

Thiessen Polygons and Spatial Autoregression. Spatial autoregression assumes a discretized surface, uses the Thiessen polygon surface partitioning to construct a Moran scatter plot, and can be used to estimate values at selected points with a regression-type equation. Theoretically, the exponential semivariogram model relates to the conditional autoregressive model, and the Bessel function semivariogram model relates to the simultaneous autoregressive model; in practice, though, the spherical semivariogram model often provides the best description of a semivariogram plot. Regardless of which approach is taken to map generalization, one relevant contribution of these two subjects is the following observation:

☞ Including positive spatial autocorrelation results in more accurate variance estimates; this in turn yields more accurate estimates of the 95% UCL for the mean concentration.

D.4.5 IMPLEMENTATION ISSUES RELATED TO GEOREFERENCED DATA

Estimation of parameters, for either geostatistical or spatial autoregressive models, cannot be achieved with ordinary least squares (OLS) techniques; nonlinear least squares must be used. While OLS provides unbiased regression coefficients, these estimates are not necessarily sufficient (i.e., they do not summarize all of the information in a sample pertaining to the population), efficient (i.e., the standard errors often are incorrect), and consistent (i.e., the asymptotic sampling distribution concentration will not be at the parameter value). In other words, OLS essentially uses the wrong degrees of freedom in its calculations, as described in Section D.4.2. Two additional complications of georeferenced data that do not appear in other types of data are (1) spatial autocorrelation might be directional (i.e., directional dependency); and (2) variance might be nonconstant over space as well as over the magnitude of the dependent variable, Y (e.g., chemical concentration). Several statistical approaches, which are beyond the scope of this guidance, are available for analyzing these potential sources of bias in the exposure concentration estimates (Isaaks and Srivastava, 1989; Cressie, 1991; Griffith, 1993; Ginevan and Splitstone, 1997).

D.5.0 EXPERT JUDGMENT AND BAYESIAN ANALYSIS

Up to this point in RAGS Volume 3: Part A, risk has been characterized as having a population probability distribution with parameters (e.g., mean, standard deviation) that can, theoretically, be estimated from observation. In theory, risk estimates could be derived by repeatedly measuring risk in subsets of the population of interest (e.g., repeated measurements of site-related cancer risk). The unstated expectation, or goal, is that the PRA model will accurately simulate this *real* risk distribution. This approach derives from a *classical* view of probability. The *classical* or *frequentist* view defines the probability of an event as the frequency with which it occurs in a long sequence of similar trials. From the *frequentist* perspective, the probability of having a flipped coin land *heads-up* is given by the frequency distribution of heads-up results derived from repeated similar trials of coin flips. For real-world decisions such as those informed by Superfund risk assessments, there is uncertainty that the sample data are representative of the population (see Chapter 1, Section 1.2.4).

Bayesian View of Probability. A Bayesian perspective on probability allows distributions to be constructed based on the judgment of an expert in the field. The subjectivist or Bayesian view is that the probability of an event occurring is the degree of belief a person has in the occurrence. Probabilities can be assessed by experts using scientific knowledge, judgment, data, past experience, and intuition. Different people may assign different probabilities to an event, and a single individual may assign different probabilities to the same event when considered at different times. The consequence is that probabilities become conditional and the conditions must be explicitly stated (Howson and Urbach, 1989; Morgan and Henrion, 1990; Ott, 1995; Sivia, 1996). These conditional probabilities can, of course, be updated with new information.

Using the coin flip analogy above, a Bayesian perspective might be that, based on experience with coins, assuming that most coins are *fair*, and that a fair coin would be expected to land heads-up half the time, the expected probability of the tossed coin landing heads-up is 0.5. If the outcome of repeated trials was different from the expected, the Bayesian approach would be to update the probability based on the new data. In the coin flip example, both the Bayesian and frequentist approaches will arrive at the same conclusions, because the outcome is amenable to rigorous experimentation. Where the two approaches can be expected to differ is in the assignment of probabilities to events that cannot be rigorously measured; for example, the probability of a site-related cancer risk, or the probability of a child ingesting a specific amount of soil.

The subjective judgment of experts is, therefore, an important tool in the Bayesian approach to risk assessment. For example, the input distributions for a PRA may be based upon the judgment of one or more experts who rely upon estimates from the literature, data from experimental studies, and any other information they consider relevant. Even when formal elicitations of expert opinion are not done, the final selection of the form and parameters of the input distributions usually involves some subjective judgment by the analyst. One of the challenges of incorporating judgments from experts or lay people is that there can be overconfidence bias (i.e., people tend to underestimate their uncertainty). There is a rich literature about the protocol for conducting expert elicitations and using the results to support decisions (Lichtenstein and Fischhoff, 1977; Morgan and Henrion, 1990; Shlyakhter and Kammen, 1992). Elicitation of expert judgment has been used to obtain distributions for use in risk assessments (Morgan and Henrion, 1990; Hora, 1992; U.S. EPA, 1997;) and in developing air quality standards (U.S. EPA, 1982).

In addition to providing input distributions for PRAs, Bayesian analysis allows the current state of knowledge, expressed as a probability distribution, to be formally combined with new data to reach an updated information state. The distribution expressing the current knowledge is the *prior distribution* and may be the output of a PRA (Figure D-8). An appropriate *likelihood function* for the data must also be formulated. The likelihood function is based upon an understanding of the data gathering process and is used to determine the probability of observing a new set of data given that a particular risk estimate is true.

EXHIBIT D-5

COMPONENTS OF BAYES THEOREM IN PRA

- Input probability distributions for exposure (or toxicity) based on available data or expert judgment
- Prior probability distribution for risk based on input probability distributions (output from PRA)
- New data
- Likelihood function, expressing the probability of observing the new data conditional on prior risk estimates
- Posterior (updated) probability distribution for risk

Once the prior distribution is determined, the new data values are collected, and the likelihood function is assumed, Bayes theorem (Exhibit D-5) provides a systematic procedure for updating the probabilistic assessment of risk. The updated information state is called the *posterior distribution* and reflects the reduction in uncertainty arising from the new information.

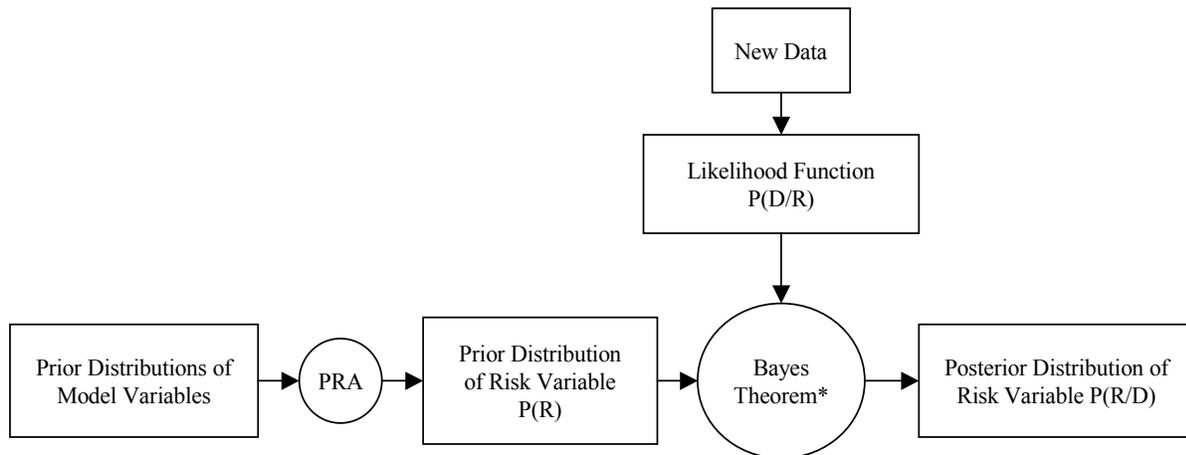


Figure D-8. Conceptual model of Bayesian Monte Carlo analysis. A PRA simulation yields a prior distribution of risk based on probability distributions for input variables. Given new data for an input variable, and a likelihood function for risk, Bayes Theorem (Eq. D-7) can be used to generate a posterior distribution of risk. The expression $P(D/R)$ refers to a conditional probability, “the probability of D , given R ”. Conditional probabilities can be thought of as relative frequencies, where R is the information given, and D is the event being computed when a particular value of R occurs.

$$\text{Bayes Theorem}^*: \quad P(R_i/D) = \frac{P(D/R_i) P(R_i)}{\sum_{j=1}^N P(D/R_j) P(R_j)} \quad \text{Equation D-7}$$

- D = new data
- R_i = i^{th} risk prediction associated with new data
- R_j = j^{th} risk estimate simulated from PRA model
- N = number of risk estimates from the PRA model

For example, suppose a model is available to relate soil tetrachlorodibenzodioxin (TCDD) concentrations at a site with serum concentrations of TCDD. A probability distribution of soil concentrations is created based upon expert judgment and a limited amount of site specific data. Using the model, the soil concentrations can be associated with a distribution of serum TCDD concentrations (P°), the prior distribution). New site-specific data (D) are subsequently collected on serum TCDD concentrations in order to reduce uncertainty in the risk estimate. Assume that it is known that serum TCDD concentrations generally follow a lognormal distribution and that the best estimate of the parameters of this distribution come from the prior distribution on serum TCDD. This creates the likelihood function ($P(D|R)$). Using Bayes Theorem, the new data are used to form a revised distribution of serum TCDD. This is the posterior distribution ($P(R|D)$).

Bayesian Monte Carlo analysis. In the past, the use of Bayesian analysis was limited by the degree of mathematical complexity involved. Using Monte Carlo analysis to carry out the PRA, rather than mathematical equations to describe the distributions, allows the calculations to be done much more easily. This variation on traditional Bayesian methods is called Bayesian Monte Carlo analysis (Patwardan and Small, 1992; Dakins et al., 1996). In the TCDD example discussed above and illustrated in Figure D-7, the required calculations are carried out for each of the N iterations of the Monte Carlo analysis (I and j go from 1 to N).

Bayesian Monte Carlo analysis is appropriate in several situations. If a model has been created and a distribution developed using PRA, new information may be incorporated without the need to repeat the entire analysis. This information could be on one of the uncertain parameters of the model or on the model output variable. Similarly, a generalized risk model with generic parameter distributions may be used for a Superfund risk assessment with the model predictions fine-tuned using data from a particular site of interest. Finally, after a distribution is developed, the amount of uncertainty that exists may be too large for the risk manager to make a decision. In this case, the risk manager might seek out new information that would refine the analysis and decrease the uncertainty.

Bayesian Monte Carlo analysis can also be combined with techniques from decision analysis to help determine the type and quantity of data that should be collected to reduce uncertainty. Decision analysis is a technique used to help organize and structure the decision maker's thought process and identify a best strategy for action. To determine the appropriate action, one defines the range of possible decisions, evaluates the expected value of the utility or loss function associated with each decision, and selects the decision that maximizes the expected utility or minimizes the expected loss.

Decision analysis provides a quantitative approach for evaluating the benefits of including an expanded assessment of uncertainty and the subsequent benefits of reducing this uncertainty.

Value of Information. Value of information (VOI) analysis involves estimating the value that new information can have to a risk manager before that information is actually obtained (Clemen, 1996). It's a measure of the importance of uncertainty in terms of the expected improvement in a risk management decision that might come from better information. Examples of VOI quantities are the expected value of including uncertainty (EVIU), the expected value of sample information (EVSII), the expected value of perfect information (EVPI). Calculation of these quantities can be done using mathematical methods, numerical integration (Finkel and Evans, 1987), or Monte Carlo techniques (Dakins, 1999)

Value of information calculations require the specification of either a utility or a loss function. A loss function states the losses associated with making different types of decision errors including both direct monetary costs and losses associated with other consequences. Loss functions take various forms depending on the risk management situation (Morgan and Henrion, 1990).

Expected Value of Including Uncertainty. The expected value of including uncertainty, EVIU, is a measure of the value of carrying out a PRA. It's the difference between the expected loss of a decision based on a point estimate risk assessment and the expected loss of the decision that considers uncertainty (Figure D-9). If uncertainty in a risk assessment has been estimated using Monte Carlo techniques and a loss function has been specified, the EVIU can be easily calculated. First, the management decision from the point estimate assessment is determined. The loss from making this decision is calculated for each iteration of the Monte Carlo, each time assuming that the risk estimate from that iteration is true. The expected loss is the average of these individual losses. The expected loss for the PRA is determined by calculating the expected loss for a full range of management decisions and selecting the decision with the lowest expected loss. The EVIU is calculated by subtracting the loss associated with the PRA from that associated with the point estimate risk assessment.

Expected Value of Sample Information. The expected value of sample information is the difference between the expected loss of the decision based on the PRA and the expected loss of the decision from an improved information state. As such, the EVSI is a measure of the value that may result from the collection and use of new information (Figure D-9). Calculation of the EVSI involves a technique called preposterior analysis and is somewhat more complicated.

This type of analysis is termed "preposterior" because it involves the possible posterior distributions resulting from potential samples that have not yet been taken. For each replication from the Monte Carlo simulation, the predicted value from the model is used to randomly generate a set of K data points. Each set of data points is then used to calculate the posterior probabilities for the N Monte Carlo simulated values. These posterior probabilities are then used to obtain the optimal answer to the management question at this new level of uncertainty by selecting the decision that minimizes the expected loss over all possible management decisions.

This procedure is repeated for each of the N replications of the Monte Carlo analysis resulting in N posterior distributions, N management decisions, and N associated expected losses. Because each of these outcomes is equally weighted, the expected loss associated with the state of uncertainty expected to exist after the data collection program is carried out is simply the average of the N expected losses. The EVSI is the difference between the expected loss based on the results of the PRA and the expected loss from the updated information state.

Expected Value of Perfect Information. The EVPI is the difference between the expected loss of the decision based on the results of the PRA and the expected loss of the optimal management decision if all uncertainty were eliminated. In actual application, no research plan or data collection program can completely eliminate uncertainty, only reduce it. The EVPI is an upper bound for the expected value of efforts to reduce uncertainty and so provides the ultimate bound on what should be spent on research and data collection efforts.

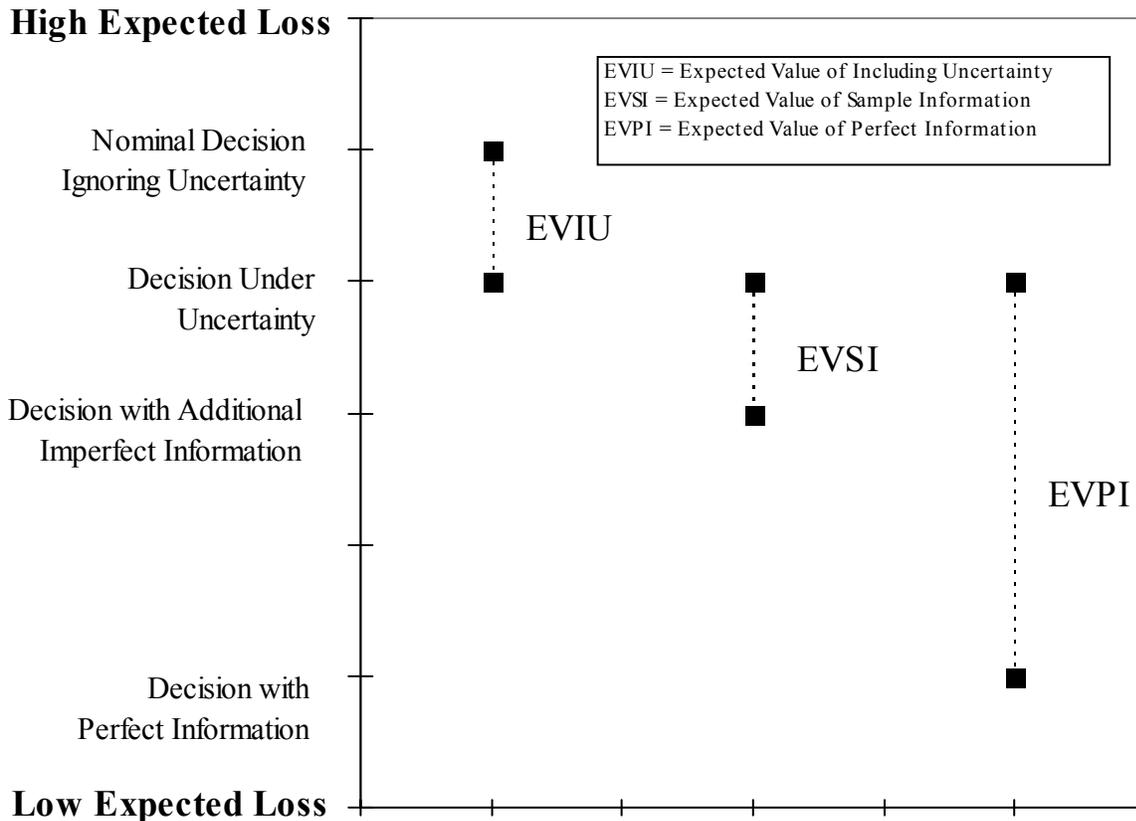


Figure D-9. Expected Loss associated with various types of information incorporated into a generic uncertainty analysis. The x-axis reflects different categories of value of information (VOI) quantities. The y-axis reflects the increasing Expected Loss with increasing uncertainty.

When a PRA has been carried out using Monte Carlo techniques, the expected loss associated with perfect information is calculated by determining the expected loss for each iteration of the Monte Carlo, assuming that the correct management decision, if that iteration were true, is made. As always, the expected loss is the average of these losses, and the EVPI is calculated by subtraction.

Uses of Value of Information in Risk Assessment. VOI analysis has many benefits for risk managers. First, VOI analysis makes the losses associated with decision errors explicit, balances competing probabilities and costs, and helps identify the decision alternative that minimizes the expected loss. VOI analysis can help a decision maker overcome a fear of uncertainty by developing a method to handle it. If the losses associated with making a poor decision are unclear, small uncertainties can take on major importance. Conversely, if the losses associated with different risk management decisions are similar, little additional effort need be expended to continue to consider the alternatives.

In addition, VOI analysis helps prioritize spending on research. It provides insights into how resources could be spent to achieve the most cost-effective reduction in uncertainty by identifying which sources of uncertainty should be reduced, what type of data should be obtained, and how much data is

needed. Finally, VOI analysis may help decision makers explain the rationale for their decisions to the public and help the public understand the multiple objectives considered in managing risks.

Expected Loss is usually greatest when uncertainty in risk estimates is ignored. For example, by quantifying uncertainty in risk (e.g., 2-D MCA, Bayesian Monte Carlo analysis) a risk manager may determine that the cleanup level associated with the 90th percentile of the risk distribution (rather than the 95th percentile) is adequately protective. Quantifying uncertainty may also result in lower expected loss when more soil remediation is required due to the losses associated with possible under-remediation, e.g., cost of additional sampling or lost revenue due to failure to meet land use requirements. The expected loss may be further reduced by collecting additional soil samples, which would presumably reduce uncertainty in estimates of mean exposure point concentrations. The expected loss may be minimized by obtaining "perfect" information (i.e., no uncertainty); however, as shown in Figure D-9, EVPI spans a wide range of expected loss because the value associated with reducing uncertainty may be tempered by costs associated with additional sampling and analysis. In practice, risk assessors consider this issue when deciding to obtain additional samples for site characterization.

The decision to obtain additional information in order to reduce uncertainty should be made on a site-specific basis, taking into account the potential impact that reducing uncertainty may have on the overall remedial decision. Important questions to consider include: (1) Are the risk estimates sufficiently sensitive to an exposure variable that collecting further data will reduce uncertainty? and (2) Are the confidence limits on the 95th percentile risk estimate sufficiently wide that reducing uncertainty may alter the cleanup goal? An example of decision framework applicable to PRA is presented in Figure D-10. The framework has three tiers. Tier 1 includes the point estimate approach and an assessment of the need for PRA. In Tier 2, the EVIU is calculated and, if warranted, a PRA is conducted. In Tier 3, the value of additional information is assessed and Bayes Theorem would be used to incorporate the new information and update probability distributions.

Limitations of These Techniques. Figure D-10 illustrates situations where Bayesian analysis and value of information quantities may not be helpful. For example, if point estimate risk assessment is selected as the appropriate method, these techniques do not apply. In addition, as site-specific data become available that are increasingly comprehensive and representative of the population of interest, Bayesian Monte Carlo analysis and the Monte Carlo analysis using the classical (frequentist) methods will approach the same result. This is because the site-specific data are incorporated into both approaches. To be representative and comprehensive, the data set must be sufficiently large, randomly selected, and represent the full range of variability that exists in the population (e.g., temporal, spatial, inter-individual). However, data sets are rarely perfect, often too small, suffer from relatively high sampling and/or measurement errors, or don't represent the entire population variability over time, space, age, gender, or other important variables. If the data cannot be assumed to describe the population distribution sufficiently well, then PRA will help to more fully develop the entire range of the population distribution and the Bayesian Monte Carlo analysis will act to refine the model estimates.

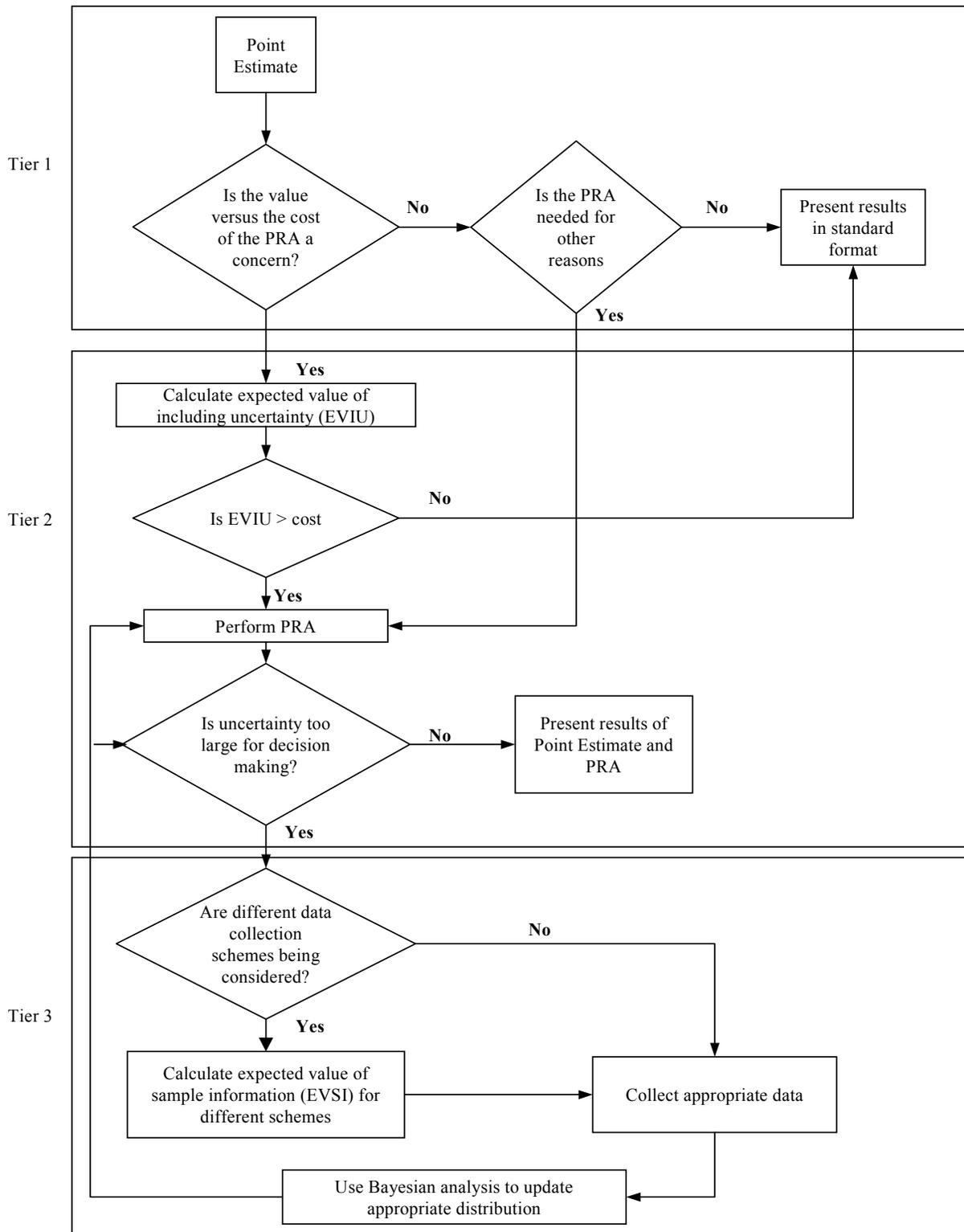


Figure D-10. Conceptual model for evaluating the expected value of including uncertainty in a Bayesian Monte Carlo analysis.

In order to carry out VOI calculations, a loss function must be assumed. Definition of the loss function may be complex due to multiple decision goals and/or multiple decision makers and may be difficult to capture in an equation. Finally, for Bayesian analysis and the calculation of the EVSI to be helpful, one or more sources of new data must exist. In addition, some information must be available about these data since a likelihood function describing its probability distribution must be assumed.

REFERENCES FOR APPENDIX D

- Brattin, W.J., T.M. Barry, and N. Chiu. 1996. Monte Carlo Modeling with Uncertainty Probability Density Functions. *Hum. Eco. Risk Assess.* 2(4):820–840.
- Buck, R.J., K.A. Hammerstrom, and P.B. Ryan. 1997. Bias in Population Estimates of Long-Term Exposure From Short-Term Measurements of Individual Exposure. *Risk Anal.* 17:455–466.
- Burmester, D.E. and K.M. Thompson. 1997. Estimating Exposure Point Concentrations for Surface Soils for Use in Deterministic and Probabilistic Risk Assessments. *Hum. Eco. Risk Assess.* 3(3):363–384.
- Clemen, R.T. 1996. *Making Hard Decisions: An Introduction to Decision Analysis*. Duxbury Press, Pacific Grove, CA.
- Clifford, P.A., D.E. Barchers, D.F. Ludwig, R.L. Sielken, J.S. Klingensmith, R.V. Graham, and M.I. Banton, 1995. An Approach to Quantifying Spatial Components of Exposure for Ecological Risk Assessment. *Environ. Toxicol. Chem.* 14(5):895–906.
- Cressie, N. 1991. *Statistics for Spatial Data*. Wiley, New York, NY.
- Dakins, M.E., J.E. Toll, M.J. Small, and K.P. Brand. 1996. Risk-based Environmental Remediation: Bayesian Monte Carlo Analysis and the Expected Value of Sample Information. *Risk Anal.* 16:67–79.
- Dakins, M.E. 1999. The Value of the Value of Information. *Human Eco. Risk Assess.* 5(2):281–289.
- Finkel, A.M. and J.S. Evans. 1987. Evaluating the Benefits of Uncertainty Reduction in Environmental Health Risk Management. *J. Air Pollut. Control Assoc.* 37:1164–1171.
- Ginevan, M.E. and D.E. Splitstone. 1997. Improving Remediation Decisions at Hazardous Waste Sites with Risk-Based Geostatistical Analysis. *Environ. Sci. Technol.* 31(2):92A–96A.
- Gomez-Hernandez, J.J. 1996. Issues on Environmental Risk Assessment. In: *Proceedings of the Fifth International Geostatistics Congress*, Vol. 1. (Baafi, E.Y. and N.A. Schofield, eds.). Kluwer Academic Publishers, Dordrecht, 15–26.
- Goovaerts, P. 1996. Accounting for Local Uncertainty in Environmental Decision-Making Processes. In: *Proceedings of the Fifth International Geostatistics Congress*, Vol. 2. (Baafi, E.Y. and N.A. Schofield, eds.). Kluwer Academic Publishers, Dordrecht, 929–940.
- Goovaerts, P. 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, NY.
- Griffith, D.A. 1993. *Spatial Regression Analysis on the PC: Spatial Statistics Using SAS*. Association of American Geographers. Washington, DC.
- Hope, B.K. 2000. Generating Probabilistic Spatially-Explicit Individual and Population Exposure Estimates for Ecological Risk Assessment. *Risk Anal.* 20(5):575–590.

- Hope, B.K. 2001. A Case Study Comparing Static and Spatially Explicit Ecological Exposure Analysis Methods. *Risk Anal.* 21(6):1001–1010.
- Hora, S.C. 1992. Acquisition of Expert Judgment: Examples From Risk Assessment. *J. Energy Eng.* 118:136–148.
- Howson, C. and P. Urbach. 1989. *Scientific Reasoning: The Bayesian Approach*. Open Court, LaSalle, IL.
- Isaaks, E. and R. Srivastava. 1989. *An Introduction to Applied Geostatistics*. Oxford University Press, Oxford.
- Kriakidis, P.C. 1996. Selecting Panels for Remediation in Contaminated Soils via Stochastic Imaging. In: *Proceedings of the Fifth International Geostatistics Congress*, Vol. 2. (Baafi, E.Y. and N.A. Schofield, eds.). Kluwer Academic Publishers, Dordrecht, 973–983.
- Lichtenstein, S. and B. Fischhoff, 1977. Do Those Who Know More Also Know More About How Much They Know? *Organizational Behavior and Human Performance* 20:159.
- McKenna, S.A. 1998. Geostatistical Approach for Managing Uncertainty in Environmental Remediation of Contaminated Soils: Case Study. *Environ. Engin. Geosci.* 4(2), Summer, 175–184.
- Morgan, G.M. and M. Henrion. 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press, NY.
- Ott, W.R. 1995. *Environmental Statistics and Data Analysis*. CRC Press. Boca Raton.
- Patwardhan, A. and M.J. Small. 1992. Bayesian Methods for Model Ancertainty Analysis with Application to Future Sea Level Rise. *Risk Anal.* 12:513–523.
- Price, P.S., C.L. Curry, P.E. Goodrum, M.N. Gray, J.I. McCrodden, N.H. Harrington, H. Carlson-Lynch, and R.E. Keenan. 1996. Monte Carlo Modeling of Time-Dependent Exposures Using a Microexposure Event Approach. *Risk Anal.* 16:339–348.
- Price, P.S., J.Y. Young, C.F. Chaisson. 2000. *Assessing Aggregate and Cumulative Pesticide Risks Using LifeLine™* Version 1.0. A Report Submitted to U.S. EPA Science Advisory Panel. August 31.
- Shlyakhter, A.I. and D.M. Kammen. 1992. Sea-level Rise or Fall? *Nature* 253:25.
- Sivia, D.S. 1996. *Data Analysis: A Bayesian Tutorial*. Clarendon Press. Oxford.
- Slob, W. 1996. A Comparison of Two Statistical Approaches to Estimate Long-Term Exposure Distributions from Short-Term Measurements. *Risk Anal.* 16: 195–200.
- U.S. EPA. 1982. *Air Quality Criteria for Particulate Matter and Sulfur Oxides*. ECAO, Office of Research and Development. EPA/600/8-82/029.

U.S. EPA. 1997. *Exposure Factors Handbook. Update to Exposure Factors Handbook*. Office of Research and Development, NCEA. EPA/600/8-89/043, May 1989. August.

U.S. EPA. 2000. *Statistical Estimation and Visualization of Ground-water Contamination Data*. Office of Research and Development, Washington, DC. EPA/600/R-00/034.

Wallace, L.A., N, Duan, and R. Ziegenfus. 1994. Can Long-term Exposure Distributions be Predicted From Short-Term Measurements? *Risk Anal.* 14:75–85.