APPENDIX I
Comparison of Discrete (Grab) Sampling with Incremental Composite Sampling

Table I-1. Comparison of Discrete (Grab) Sampling with Incremental Composite Sampling

Parameter	Discrete Sampling	Incremental Composite Sampling
Assumptions about soil contamination relevant to sampling and analysis		
Soil homogeneity	Constituents of interest are homogeneously distributed within a soil matrix (like salt dissolved in water) whether at background concentrations or anthropogenic release. **Corollary: The concentration at two soil locations 0.1, 0.5, 1, 2, 5, 10 meters apart are approximately the same.**	Because constituents preferentially bind to certain soil particles, contaminants behave as solid particles (concentrated "nuggets") that are unevenly dispersed (heterogeneous) at spatial scales relevant to contaminant sampling and analysis. **Corollary: Concentrations may differ greatly even for grab samples taken near each other.**
Subsampling	Contaminant concentrations within a sample jar are the same at the top, middle, and bottom of the jar, so that any grab subsample represents the jar's average concentration. Stirring, cone-and-quartering, etc. are acceptable ways to "mix" samples.	Since contaminants are borne on particles, and particles segregate by size during sample transport and manipulation (such as stirring and weighing), a grab subsample can be strongly biased high or low as compared to the jar's average concentration. Appropriate sample processing and subsampling procedures are required to counter this bias.
Analytical mass	A concentration result will be the same no matter how much or how little soil from the jar is used to perform the analysis. The mass of subsamples is determined by laboratory convenience, the needs of instrumentation, and the desire for waste-reduction.	Smaller analytical samples are less precise than larger samples. Larger analytical subsamples are more likely to represent the true concentration within the jar than smaller subsamples. The size of the largest particle in the sample determines the appropriate analytical mass.
Grab sample reliability	Grab samples are the most appropriate way to collect field samples and subsample soil samples for analysis. Only one grab sample or subsample gives sufficiently accurate information on which to base a decision.	Grab samples can be trusted to provide accurate information <u>only</u> if there are enough of them to accurately measure variability and to provide an estimate of statistical confidence around the average calculated from the data.
Number of samples	The number of grab samples is determined by the available budget, and by non-scientific negotiations among the regulator, responsible party, and perhaps stakeholders, so that regulator and stakeholder "comfort" is achieved.	The number of increments is determined by: (a) using a conservative default shown to be sufficient for most situations (^{3}30 per an area of ½ acre or less), or (b) by calculation from the actual variability observed for those increments within the defined soil mass (the DU).

Table I-1. Comparison of Discrete (Grab) Sampling with Incremental Composite Sampling

Parameter	Discrete Sampling	Incremental Composite Sampling
Statistical data distribution of results	Usually not normal (lognormal, gamma, or nonparametric distributions), although too few discrete samples may be collected to reliably determine the distribution. High variability and non-detects are usually present in the data set. Non-normality and many non-detects can make statistical analysis of the data complex. Selection of an appropriate UCL calculation method can be unclear and controversial. UCLs can be unrealistically high. Conclusions may be uncertain because of the combination of too few samples and high variability.	Since each IS result is an estimate of the DU mean, assuming a normal distribution is justified unless the underlying variability is very high. Since there are usually only three replicates from which to calculate the DU mean and UCL, the options for calculating a UCL are limited to the Student's t and the Chebyshev. The UCL calculation is balanced by a high precision among the replicates against the penalty for only having three results. There are fewer (or no) non-detect results. Statistical analysis of the data can be easier and less subject to controversy. Unless the three replicates are exactly the same, the UCL will be higher than the highest replicate result, but the UCL (not the highest result) should be used as a conservative estimate of the DU mean.
Spatial resolution	Grab sample results are assumed to represent the actual spatial resolution present in the field. Grab sample results are often used to draw "contour lines" to delineate areas with different concentrations. Contour lines generated from high variability; low sample-density data sets uncertain.	No spatial resolution within the DU is possible unless a more complicated incremental design (involving SUs) is used. Alternatively, DUs could be made smaller. For "point" data purposes such as transects, very small SUs (1–4 square feet) can be represented by composite samples (³5 increments) to reduce the biasing effect of short-scale field heterogeneity. A composite sample needs to be processed and subsampled in the same way as an incremental sample.

Table I-1. Comparison of Discrete (Grab) Sampling with Incremental Composite Sampling

Parameter	Discrete Sampling	Incremental Composite Sampling
Sample representativeness	The concentration result from a 1-g analytical sample grabbed from a 100-g sample jar can be assumed to represent the concentration for hundreds to thousands of kg of surrounding soil without a need for corroborating evidence. The volume and distance over which a single soil result will be extrapolated is determined by professional judgment, regulatory comfort, or whatever the grid size happened to be to accommodate the number of samples allowed by the budget. **Corollary: A SINGLE sample result provides actionable information for an undefined volume of soil.**	The true concentration for any large mass of soil is the same concentration that would be obtained by mathematically averaging the results of all potential analytical samples within that mass. The chance that any single analytical sample would be the same as (*i.e.*, represent) that true concentration is very small. **Corollary: The only way to estimate the true concentration of a soil mass is to mathematical or physically (via incremental sampling) obtain an average from an adequately large set of samples (or increments) taken from a pre-defined soil mass (the DU).**

Table I-1. Comparison of Discrete (Grab) Sampling with Incremental Composite Sampling

Parameter	Discrete Sampling	Incremental Composite Sampling
QC Results	The results of collocated field samples and laboratory duplicates can be ignored when using the associated analytical results for decision-making. This type of QC information is not used to improve sample collection and handling procedures or determine sample numbers when designing future sample collection efforts. **Corollary: A decision can be based on a single sample result without consideration of the result's uncertainty. Despite the high degree of sampling variability frequently measured by QC data, sample results are used "as is" and considered reproducible because that is what we've always done.**	The results of collocated field samples and laboratory duplicates provide valuable information about how much confidence can be placed in any single analytical result. These QC results are used to evaluate (and improve if necessary) the adequacy of current and future sample collection and processing procedures. **Corollary: Decisions should not be based on a single discrete sample unless the uncertainty in that result is estimated using QC data.** QC for Incremental Samples: Even decisions based on a single incremental sample may be uncertain unless there are sufficient QC data to measure sources of variability: (a) At least three independent replicate incremental samples from the same DU are collected to quantify precision over the entire measurement system. (b) At least three subsampling replicates are performed to quantify the precision of sample processing, subsampling, and analysis. (c) Laboratory control samples are used to measure analytical error. (d) From (a) and (b), the overall degree of field variability can be calculated to determine whether sufficient increments are being used. (e) From (b) and (c), the degree of sample processing and subsampling variability can be calculated to determine whether those procedures are sufficiently effective.

Table I-1. Comparison of Discrete (Grab) Sampling with Incremental Composite Sampling

Parameter	Discrete Sampling	Incremental Composite Sampling
Sample plans	It is sufficient for sampling plans to simply claim that "representative samples" will be collected. A description of what the data are supposed to represent or how the data will be used to make project decisions is not required for the sampling plan to be approved. After data are collected, "data review/data validation" does not include either a qualitative or quantitative determination of what the data represent.	Sampling plans must explain the intended project decisions that the data are to support. DUs are constructed so that the DU data are representative for the decision-making scenario(s). It can be concluded that incremental sample results represent the true DU concentration if: (a) sufficient increment density was used (default is less than or equal to 30 per an area of ½ acre or less) AND (b) less than or equal to three independent replicate incremental samples (from the same DU) agree.
Hot spots	A "hot spot" can be defined ("I'll know a hot spot when I see the data"), detected ("some unknown mass of soil is dirty"), or ruled out ("some unknown mass of soil is clean") using the result from a single grab sample. **Corollary: No forethought about defining hot spots is required before collecting the data.**	All concentration results represent an average concentration for some soil mass. The question is "What volume of soil is known to be represented by that result?" Without corroborating information, the only thing known for sure is that the analytical results represent the average concentration for a 0.5-, 1-, 10-, or 30-g subsample mass that is actually analyzed. The result cannot be assumed to represent the concentration for the sample jar, much less the concentration of some larger soil mass in the field. **Corollary: Reproducible detection of hot spots requires that project planning first define the volume and concentration of the soil mass that qualifies as a hot spot.**
Contaminant concentration data management and storage		

Table I-1. Comparison of Discrete (Grab) Sampling with Incremental Composite Sampling

Parameter	Discrete Sampling	Incremental Composite Sampling
Soil data documentation	Details about the procedures used to collect, handle, and analyze samples do not need to be stored in the database with the actual data. Sample ID, sample location, sample concentration, and the concentration units are enough for any future secondary uses of the data.	The following information should be stored in a database: • the spatial dimensions of the DU; • the number of increments making up the DU incremental sample; the sample support of the increments; the total mass of the incremental sample; the type of soil (sandy, clayey, etc.); how the sample was processed; • the particle size actually analyzed and what percentage of the total sample mass that particle size comprised; • the mass of the analytical sample and how it was prepared (grab or incremental subsampling); • and the QC data from which can be determined the magnitude of field variability, subsampling variability, and analytical variability. This will likely require attaching reports to the database.