ISM for Risk Assessment

Section 8 addresses the key issues risk assessors should consider when planning for the use of ISM data for risk assessment and provides guidance for using ISM data in risk-based decision-making. The focus here is the use of ISM in human or ecological risk assessments, so this section presumes familiarity with the basic concepts and practice of risk assessment and includes references for federal and state guidance documents as well as ITRC guidance (ITRC 2015).

This section provides risk assessors and other users with a go-to resource for using ISM data in risk assessment. Key concepts discussed are as follows:

- systematic planning considerations for risk assessment
    - laboratory analyses
    - EPCs and DUs/EUs
    - nature and extent
- calculation and use of ISM data for EPCs
- benefits, planning, and application of ISM background data
- ISM sampling for post-remediation comparison to risk-based goals
- communicating ISM-based risk assessment results

8.1 Introduction

ISM is a technically sound sampling approach to collecting data for a scientifically defensible risk assessment designed to assist in risk-based decision-making. Specifically, it provides an accurate estimate of the true mean concentration with a limited number of samples. Most agencies accept the use a 95% UCL on mean concentration for comparison of soil or sediment concentrations to screening values or ALs and for use in baseline risk assessments. A minimum of three ISM replicates are necessary to calculate the 95% UCL.

As described throughout this document, the use of ISM samples to characterize the soil within a DU can provide higher-quality data and fewer decision errors than conventional low-density discrete or composite sampling designs. Because ISM yields an estimate of mean concentrations within a DU, it is important to understand the appropriate spatial scale that was specified or implied in the development of the risk-based screening levels.

As discussed in Section 3.1, incorporating a risk assessment strategy into a project should include developing a CSM developed during systematic planning. The risk assessment strategy should also be designed to inform risk management. Communication with those involved in or affected by the risk assessment should start in the planning stage of the project as well. The regulatory jurisdiction overseeing the risk assessment should always be consulted in the study design for ISM sample collection. If the study goal is to complete a quantitative risk assessment, then the strategy for using ISM data should consider the quantity and quality of the data required for the risk assessment, as described in Section 3.3.2 and Section 6.1.3. Important attributes of a CSM related to ISM sampling strategy for risk assessment include the following:

- exposure scenarios
    - receptors (human and/or ecological, current and potential future land use)
    - pathways (direct contact – such as soil ingestion and dermal absorption, transport – such as wind dispersion, surface runoff, infiltration, and uptake by plants and animals)
- exposure media (soil/sediment)
    - area (source characterization, extent of contamination, and the exposure characteristics of potential receptors' activity patterns)
    - depth (activity pattern characteristics of ecological receptors or human soil disturbance activities, and the extent of contamination)

A key concept in risk assessment is the exposure area. Risk assessment generally requires estimating long-term (chronic) exposures, and on occasion, short-term (acute) exposures. A goal of the data collection plan for a risk assessment is to estimate exposure-based chemical concentrations in exposure media based on current or hypothetical future land use.

CSMs, as discussed in Section 3.1.2 of this document and in ITRC's risk guidance (ITRC 2015), are essential to the SPP and for completing well-designed risk assessments. The CSM describes the relationship between potential chemical sources, media, release mechanisms, fate and transport pathways, exposure pathways, exposure media, exposure routes, and current and future receptor groups. The CSM presents the current understanding of the project area but should be reevaluated and updated as new information is collected throughout the lifecycle of the project. The CSM helps identify data gaps and focus data collection efforts. Various styles of CSM can be useful, including narrative explanations supplemented

by one or more figures (pictorial or schematic flow chart) depicting current and potential future site land uses. The ISM sampling strategy should provide inputs for risk assessment study goals.

8.2 Systematic Planning for ISM Data Use in Risk Assessment

Sample design input for risk assessment study goals should provide sufficient data to evaluate all the media each exposure group (such as trespasser, resident, or terrestrial receptor) might encounter. ISM sample results can be collected to provide EPCs for soil or sediment exposure media.

Environmental data must be of the appropriate type, quantity, and quality to manage and evaluate uncertainty so that defensible decisions can be made. To ensure that the data obtained during environmental investigations are adequate for their intended purposes, it is strongly recommended that data collection activities be planned and developed through an SPP, as discussed in Section 3.1.

As described throughout this document, the use of ISM to characterize the soil within a DU can provide higher-quality data and fewer decision errors than conventional low-density discrete or composite sampling designs. In combination with well-considered investigation objectives as well as DU and SU designations, ISM samples can reduce the need for additional sample collection, increase the certainty of decisions, and reduce the time and money required to complete environmental projects. Although a project team may have an ISM strategy in mind during initial planning, a number of sampling and analysis options should be considered, and the sampling strategy selected should be an outcome of the SPP. For risk assessment, the sampling strategy must meet the input needs for the planned risk assessment.

Estimates of mean concentrations in environmental media are generally the appropriate statistic to compare to ALs and to use in risk assessments. ISM provides an estimate of the mean contaminant concentration in a defined volume (area and depth) of soil. An exposure area (also called an EU for ISM risk-based DU) is a geographic area over which a receptor is reasonably assumed to move randomly and is therefore equally likely to contact an environmental medium (for example, soil) at all locations. In addition to assumed activity patterns of receptor behavior, the boundaries of an EU can also be defined to account for knowledge of historical activity and the nature and extent of chemicals in environmental media (USEPA 1989a).

Several crucial decisions must be made when ISM projects are planned. As discussed in Section 3.1, sample design such as the number and size of DUs, the number of replicate ISM samples, and the number of increments for each ISM sample are guided by the problem formulation and study goals. When ISM sample design is chosen as an information input, the DU dimensions should be designed to address a specific study goal or multiple goals. Different goals may require different kinds of inputs. Risk assessments may necessitate multiple sampling objectives and strategies, such as when the CSM shows pathways for multiple receptors or exposure scenarios, which may require sampling EUs of different spatial scales based on different receptors' human and/or ecological activity patterns. The dimensions of each EU must be assessed for appropriateness of the exposure duration (long-term chronic exposures and/or short-term acute exposures), as well as specificity for each receptor under each current and potential future land use. The EU dimensions will depend on the appropriate area and depth pertaining to a receptor's potential exposure. As described in Section 8.2.2.1, the dimensions of EUs for human health and ecological assessments may differ.

8.2.1 Risk assessment considerations for ISM laboratory sample preparation

Laboratory processing of ISM samples prior to analyses for non-volatile compounds usually includes sample sieving and could include grinding of a sample to improve laboratory precision. These topics, as well as consideration of appropriate sample digestion/extraction methods, are discussed in the context of project planning in Section 3.1.5.3 and addressed in more detail in Section 5.2.

The representativeness of the sample after laboratory processing should be considered by the risk assessor with respect to each receptor evaluated in the risk assessment. A thorough discussion of this topic is beyond the scope of this document, but several key points are noted here:

- The surface area of particles per unit volume of soil or sediment is inversely related to particle diameter, and greater contaminant concentrations can sometimes be observed on fine particles relative to coarse particles.
- The appropriate particle size fraction for evaluating exposures in human and ecological risk assessment can vary as a function of the receptor and exposure route.
- Grinding of soil or sediment samples prior to analysis might produce analytical results that are not representative of environmental exposure conditions.
- Sample extraction methods should be representative of conditions associated with the exposure route (such as ingestion, inhalation, or dermal absorption) evaluated in the risk assessment.

8.2.2 DUs, EUs, and SUs

As described earlier in this guidance, an ISM DU is the smallest (horizontal) area and associated (vertical) depth of soil for which a decision will be made. An SU is either equal in size to – or is a subdivision of – the DU. An SU is comprised of at least 30 increments of soil collected to estimate the mean chemical concentrations for an area and depth of soil or sediment. For risk assessment, EUs represent the area over which a receptor could be exposed for the relevant exposure period corresponding to the particular receptor scenario. The relationship between DUs, EUs, and SUs, discussed in detail in Section 3.1, is described here in the context of risk assessment.

Environmental decisions are often based on the predicted risks from exposure to estimated mean concentrations of contaminants in a volume of soil. In some cases, a decision for additional investigation or remedial action might be made based on a comparison of ISM sample results to benchmark or screening levels, which are often risk-based. In other investigations, the estimate of the mean contaminant concentration provided by ISM samples might be used to calculate risk for human or ecological receptors. ISM results may also be used to estimate background concentrations, to assess the boundaries of source areas, or to evaluate the success of remedial activities. In each case, specifying the dimensions of EUs, SUs, and DUs is a critical component of the sampling design.

8.2.2.1 EUs

EU boundaries are based on exposure assumptions concerning the area and depth of soil where a receptor may be exposed over time. As discussed in Section 3.1, the primary types of DUs are those that are based on the known or expected locations and dimensions of releases (source area DUs or N&E DUs) and those based on the known or expected locations and dimensions of areas within which human or ecological receptors are randomly exposed (EUs). This discussion focuses on applying ISM to EUs.

During systematic planning to support risk assessment for contaminated soil and sediment, a primary question is, “Over what area and depth do samples need to be collected to provide data to represent the potential exposure of a receptor?” EUs based on exposure areas are a key tool in risk assessments and risk-based decision-making. For the purposes of this document, an *exposure area* is an area where human or ecological receptors could come into contact with contaminants in soil on a regular basis (refer to “Risk Assessment Guidance for Superfund, Vol. II” (USEPA 1989b)). Examples include residential yards, schoolyards, playgrounds, gardens, outdoor areas of commercial/industrial properties, or areas designated as exposure areas through other means (such as state laws).

EUs applicable to human receptors may not be readily applicable to ecological receptors, so when sites are evaluated for both human health and ecological receptors, multiple spatial scales may need to be considered for sampling. For example, EUs for human health evaluations may correspond with individual residential properties, while EUs established to address ecological receptors may correspond to the home range of an individual receptor or a population of receptors. This is a good example of why both the CSM and the second step in systematic planning (see Section 3.1.3, “What Types of Additional Information Do We Need?”) are very important – information needs for ecological and human exposure assessment can differ greatly. Other non-risk-based study goals might require different DU dimensions for purposes such as characterizing spatial patterns of contamination or providing data to evaluate remedial alternatives.

The primary use of data representing contaminant concentrations in an exposure area is to assess human or ecological risk. When study objectives involve risk assessment, the EU should ideally be based on the area where exposure is known or anticipated to occur. The size and placement of exposure areas depend on current use or potential future use of the site. However, if there are suspected areas of unique or elevated concentrations within an EU, it may be important to break the EU into two or more parts (see Example 2B in Section 3.1.6.2). Since risk assessment generally assumes long-term or chronic exposure, an underlying assumption of an EU is that the receptor spends an equal amount of time over all portions of the EU for that exposure period (meaning the receptor is randomly exposed across the entire EU).

In some situations, a standard-sized EU might not reflect the known or suspected movement of the receptor – that is, the human or ecological receptor prefers some areas over others. If there are subareas of elevated concentrations within an EU, then potential risks could be underestimated. In this case, the CSM and the resulting sampling plan should consider the suspected or actual movement of the receptor by the use of smaller EUs within which the receptor can be expected to move randomly. Swing sets and sandboxes in residential yards are the classic example for human exposure – in this scenario, a child is expected to spend more time on and around such play areas than in the remainder of the yard. Movement within such smaller areas is expected to result in equal exposure to all parts of the area, and therefore it meets the definition of exposure area. This concept is illustrated in Section 3.1.6.2 Example 2B.

An EU could be based on current land use or potential future land use. Site-specific information should be incorporated into the CSM so that exposure media and exposure dimension assumptions are clear. For example, if future residential exposure is going to be evaluated, then the EU should be designed and sampled at a scale consistent with the local residential

property size, and perhaps address the depth of soil that would be excavated for a foundation and used to regrade the area (bringing subsurface soils to the surface). Potential regrading during future development of the property should also be considered in systematic planning of the sampling design to focus proper sample collection in the soil horizon (depths) that may be encountered by future receptors. Some states define the area and depth to consider for a default residential property in the absence of existing residential boundaries, but the uncertainty section of the risk assessment should discuss EU assumptions, such as those associated with a hypothetical future residential EU property size and placement.

The depth and area of EUs should be defined consistently with the exposure scenario under consideration. In many such scenarios, the first few inches or centimeters of the surface is the appropriate sampling interval (recreational receptor). However, the depth and area considered acceptable for evaluation of surface or subsurface soil for exposures varies among agencies (ITRC 2007a) and risk receptor scenarios. Evaluation of risks posed by future excavation and spreading of deeper contamination to the surface could require EU depths many feet below the ground surface. The need for multiple, different vertical EUs depends on the CSM and its expected contaminant distribution with depth as well as on current and potential future land uses.

As discussed in Section 3.2, the number of samples needed from an exposure area (including duplicate and blank samples) depends on variability in chemical concentrations within the exposure area and the level of precision and accuracy required for the project's risk-based decision-making (Hartmann et al. 1993.).

8.2.2.2 SUs

As Section 2 notes, SUs are subdivisions of (or equal to) DUs from which separate ISM samples are collected. The boundaries of an SU indicate the coverage of a single ISM sample, thus, SUs define the scale of ISM sampling, whereas DUs define the scale of the decision(s) based on that sampling. These definitions allow for the possibility that ISM samples from several SUs composing a DU can be used collectively to make a decision on that DU. It is sometimes possible to later redefine SUs as DUs if the resulting scale and number of replicates meets project objectives.

The mean concentration over the entire DU is typically the basis for decision-making. If the DU is properly sized for risk assessment, then it may be comprised solely of replicates of a single SU or several SUs. Estimation of the DU concentration mean from individual and replicate SU samples is discussed in more detail in Section 3.2.8 and Section 6. When ISM samples are collected across the entire DU (where the DU is a single SU), replicates offer information on variability in the mean estimate. Replicates from a single SU or DU do not, however, provide any information on the spatial patterns of concentrations within that single SU or DU. If this information is desirable, an appropriate sampling design is to divide the DU into multiple SUs (or further divide an SU into smaller SUs) and take one or more ISM samples from each newly formed SU. Section 3.1.6.2 Example 2B, Figure 3-11b illustrates this concept of utilizing SUs within an EU for a residential yard, play area, and potential lead-based paint surrounding a home. With this approach, ISM samples from the SUs are not true DU replicates in that they are providing estimates of the mean for different subunits within the DU. Individually, they estimate the mean of a subarea (the SU), and collectively they can be used to estimate the mean of the entire DU. However, collecting three replicates from some SUs within the DU can provide information on spatial variability and help calculate both field variability and DU mean concentration. Dividing DUs into SUs could be used to answer the environmental question, "Is there evidence that remediation of an SU in the DU could lower the risk presented by exposure across the DU?"

Results for individual SUs within a DU should generally not be used to make decisions because, by definition, such SUs are at a smaller scale than appropriate for a decision. This is especially true if sampling involves only one ISM sample per SU (meaning no replicates were collected). Taking only one ISM sample from an SU should generally be done with caution because a single sample will not provide any information about the variability in ISM replicates for a particular SU to assess uncertainty in chemical concentrations within it. However, such an approach can be advantageous when sampling large areas. In this case, individual SU results are statistically treated in a manner identical to that of traditional discrete or composite samples. This application of SUs is discussed in more detail in Section 3.2.8.

SUs may also be used when there are multiple sampling objectives or sampling scales for a given volume of soil. For instance, when the same area is assessed for multiple receptors with exposure areas of different scales, then the various EUs will define the receptor-specific DUs and multiple SU sizes may be applied in that area to combine in appropriate horizontal and vertical dimensions for the various receptor-specific EUs at the site. For example, Receptor Scenario 1 may consist of a large DU = EU1 that is broken down into multiple SUs, where SU size is equivalent to EU2, which represents the DUs with a smaller spatial scale for Receptor Scenario 2.

8.2.3 Using ISM to evaluate the nature and extent of contamination

As described in Section 3.1.5, delineating the nature and extent of contamination is a common objective of environmental sampling in a remedial investigation. When considering the appropriate area (residential property, ecological receptor home

range, and so on) and depths (surface, or 0 to 1 ft; subsurface, or 1 to 10 ft) where receptors may be exposed to soil contamination, the integration of information about the nature and extent of contamination is essential for the risk assessor.

To compile the CSM, understanding the nature and extent, as well as the fate and transport, of contamination is also crucial in determining both where receptors may be exposed to elevated levels of contamination and the complete exposure pathways for all current and potential future receptors. Both the horizontal and vertical extent of contamination must be understood to determine where and what analyses are required from what media for estimating exposures to receptors in the risk assessment process,. For example, for future land use and potential redevelopment where soil mixing may redistribute subsurface soils to shallower depths or the surface, data on the nature and extent of contaminants from deeper soil are needed to evaluate the potential risks from direct contact with soil. Similarly, characterizing the lateral extents are critical for determining potential receptors (for both current and future human and ecological receptors), particularly if the lateral extents go beyond the property boundary on the site to an area where land uses may be different. EU locations and dimensions (breadth and depth) are commonly informed by the nature and extent of contamination, so characterizing the nature and extent is an iterative process that should involve the entire project team so that they obtain the proper data for determining the concentrations of each chemical in each media that could potentially contribute to an unacceptable risk. Evaluating the nature and extent of contamination with ISM is not intrinsically different than doing so via discrete or traditional composite samples. Regardless of sampling technique, however, the sampling design derives from the CSM and the study questions in the SPP.

Sampling to characterize the nature and extent addresses a complementary study goal to the risk assessment study goal (which is to determine EPCs). Typically, soil samples are initially collected at biased locations based on site history or physical features to determine if contamination was released and if chemical concentrations might exceed risk-based or regulatory criteria. After the nature of the contamination is determined (that is, what chemicals have been released), delineation of the extent can be evaluated and used to inform other components of the investigation, such as the potential migration of contaminants (for example, soil contamination at depths merging with shallow groundwater flowing to a surface water body). This data is usually intentionally biased:

- The data will not likely be useful for determining average concentrations over an exposure area.
- The data may over- or underestimate the mean concentration in the area/volume of soil or sediment depending on the proportion of samples collected from the highest area versus the lowest concentration locations (such as in delineating the extent of the impacted area).

One advantage of ISM is that the data from SUs used to define nature and extent are not unduly biased by intense sampling from areas of highest concentrations within the source area or lowest concentrations at the perimeters, thus delineating the extents of contamination. As discussed in Section 8.2.2.1, the dimensions of an EU cannot always be known with certainty. Therefore, understanding the spatial distribution of contamination is important to provide confidence that potential exposures have not been underestimated. ISM results from multiple SUs can be used to estimate EPCs for different possible EUs, providing useful information for risk management. Use of ISM in this manner provides better coverage and a more representative estimate of mean concentrations in a DU, which reduces uncertainties in the risk assessment EPC.

8.2.3.1 Misconception regarding small areas of elevated concentrations.

As discussed in Section 2.5.2, Section 2.5.3.2, Section 3.1, Section 3.2, and Section 7.3, because ISM has superior spatial coverage, there is a higher probability that the resulting DU mean will capture the effects of areas with higher concentrations. Strategies are available for use with ISM to safeguard against missing significant small areas of elevated contaminant concentrations within a DU that can have the potential to change a sample concentration from below to above the decision threshold if they are captured in their proper spatial proportions by an ISM sample.

ISM sample results provide better estimates of the true population mean of the whole DU, and because of the increased spatial coverage, increased sample support, and rigorous sample processing, ISM can in fact find small areas of elevated contaminant concentrations when the sizes and concentrations of concern are defined upfront. See "***Evaluating the potential presence of subareas of elevated contaminant concentrations with ISM***" (within Section 3.1.5.2) to address these types of planning considerations. A DQO study goal could be to not miss significant small areas (horizontal and depth) of elevated contaminant concentrations above risk-based concentrations within an EU because the EU could be comprised of several SUs designed to meet the "small area" volume requirement. It is in the systematic planning phase that project teams must define and designate what concentration and what volume, surface area, or mass are significant to their decision-making. The size and concentration of a *significant* small area of elevated contamination can be established using an Excel spreadsheet tool if the critical condition of a mature CSM is met. For an example and more details on this concept, see the White Paper (Crumbling 2014).

Statistically based sampling designs can be developed to determine whether localized areas of higher soil concentrations exist, even if the locations of such subareas within a larger site are unknown. As discussed in more detail in Section 3.1.5.2, a free software program called VSP is available to determine increment spacing for the DU grid so as not to miss sampling from a significant small area of elevated concentrations within the DU. The spacing of increments (and thus the number of increments needed to fill the DU's area) can be set to have a desired statistical probability of increments being collected from within an area of defined size for incorporation into the field sample. In this case, if the size of a potential subarea of elevated concentration is specified, sampling can be conducted to determine whether one or more such areas exist within a DU with an objective degree of confidence and scientific defensibility (see Section 3.1.5.2). For more details on this concept and the VSP tool, see the see the White Paper (Crumbling 2014). Users are strongly encouraged to fully understand and consult the additional details on VSP designs as well as the inherent assumptions and limitations that are available in the VSP help files some of which are noted in Section 3.1.5.2.

Although somewhat subjective, not statistically rigorous, and less scientifically robust, increment spacing within a DU (or the sampling density) can provide some confidence that small areas of elevated contaminant concentration within a DU are not obscured by either increasing the number of replicates or increasing the number of increments per replicate. Nonetheless, this process may define the size of a small area of elevated concentration within a DU that is *observable* via ISM. As an example, consider a ¼-acre residential EU. If three replicates of 30 increments are collected, they correspond to three increments (one from each replicate) from an area of approximately 18 ft by 20 ft, which results in a sampling density of one increment from each 10 ft by 12 ft area. If 70 increments are collected instead, this translates to three increments from an approximately 12 ft by 13 ft area, or a sampling density of one increment for each 6.5 ft by 8 ft area. In these examples, the small areas of 10 ft by 12 ft or 6.5 ft by 8 ft of elevated concentrations would be represented in ISM samples of 30 or 70 increments, respectively, from the residential lot. If the area sizes and concentrations are significant, sufficient increment sampling density can guard against missing a small area of elevated contaminant concentration and underestimating the true DU concentration.

In addition to (or as an alternative), the SPP DQOs can be used to define the allowable variation in replicates that will be acceptable (the MQO), as discussed in Section 3.3.2.1 and Section 6.4.2. Reproducibility is a hallmark of a scientifically defensible study, and reproducibility among replicates can be used to uncover one or more small areas of elevated contamination within a DU. ISM includes QC procedures designed to measure overall sampling and analysis precision, including the collection of field replicates. High variability between replicates demonstrated by a high RSD is an indication of either a localized small area of elevated concentrations within the DU or unequal distribution of the COPC due to the nature of the source that resulted in the release (such as munitions) or the actual nature of the COPC (such as propensity to form nuggets or hydrophobicity). See Section 3.2.4.2 text and Table 3-3, which classify heterogeneity of increments in terms of low, medium, and high CV of replicates. If results for the replicates do not agree, one reason may be that the number of increments collected was not adequate to representatively include areas of non-random higher concentration scattered throughout the DU. If other causes of data variability can be ruled out by QC data, disagreement among field replicates is an indication that more increments may be needed to manage the heterogeneity caused by small areas of elevated concentration. Examples in Section 3.2.5.2 demonstrate that large disagreement in replicate concentrations is a clear sign of extreme heterogeneity, most likely manifested as small areas of elevated concentration within the DU. Refinement of the study design to include smaller DUs and/or more increments will shed light on the cause and uncover whether the underlying reason is indeed a small area of elevated contamination within the original DU sampling design.

Examples that integrate the use of ISM sample data collected to investigate nature and extent, as well as support the estimation of EPCs for risk assessment, are included in Section 3.1.6 and Appendix A. Topics include ISM application to different types of releases, such as widespread theoretically homogeneous contamination (for example, agricultural field pesticide applications) and spatially heterogenous contamination from one or more potential source release areas.

8.3 EPCs from ISM Data for Human and Ecological Risk Assessment

The calculation of a receptor's average daily dose is based in part on the average concentration of a contaminant in an exposure medium, which means that the reliability of the chemical concentration data used to develop EPCs is very important. Samples collected by ISM provide reliable data with less uncertainty than those historically collected with discrete or composite sample methodology. Screening level risk assessments using ISM data are as valid as (and have less uncertainty than) those using the maximum concentration from discrete or other traditional sampling approaches.

There are various approaches to defining DUs, but the focus should be on ensuring that the data collected will aid in making the decision associated with the DU area. The approach selected should be consistent with the understanding of the site reflected in the CSM and should support the objectives of the investigation. Human health or ecological EUs – and the spatial distribution of contamination – should provide the basis for designating DUs for risk assessment.

8.3.1 Scenario-specific EUs

The CSM depicting current and potential future scenario-specific receptors is used to define scenario-specific EUs for ISM risk-based study questions. Assumed current and future human and ecological activity patterns associated with exposure media are the primary basis for defining the boundaries of a scenario-specific EU. The concentration of a soil contaminant an individual would be exposed to from long-term random exposure within an EU is, in fact, the average concentration. In other words, it is because a human or ecological receptor is assumed to contact soil across the EU in a random manner over time that we use the average to represent the exposure concentration. This concept can be extended to consideration of exposure to soil as a function of depth as well as area. If, for example, an exposure model states that the activities of humans (or burrowing animals) might reach a certain depth, then the average soil concentration from the ground surface to that depth is of interest. Practically, we rarely know with a high degree of confidence what (future) exposure patterns are going to be. As part of our sample design assumptions, we might state that humans could excavate soil to a depth corresponding to a basement, but we do not necessarily know they will, or what the exact location and volume of the excavation will be. As discussed in Section 8.2.2.1, for future exposure scenarios, the EU should be designed and sampled at a scale consistent for each of the receptors (such as the local residential property size for a future residential scenario). For future receptor scenarios, the depth of soil that would be excavated for a foundation and used to regrade the area (bringing subsurface soils to the surface) should also be considered. This is why it is important to consider both source areas (that is, the known or inferred spatial pattern of contamination) and exposure areas in developing DUs to meet information inputs for study goals in sampling designs (see Section 3.1). As described in Example 2 in Section 3.1.6, if there are specific areas that might receive higher use in an exposure area, such as a play area, that area could be evaluated as a smaller and separate SU or an N&E DU, particularly when localized contamination is suspected. Sections 3.1.6.2 and 3.1.6.3 provide Examples 2B and 3, respectively, for identifying EUs under different conditions and for different study goals. These examples address both human health and ecological risk assessments and integrate ISM sampling for investigations addressing source areas and EUs. The use of ISM to evaluate the spatial extent of contamination is also described in these examples.

8.3.2 Use of ISM replicate data in risk assessment

The number of ISM field replicates required for a scientifically defensible risk assessment is flexible and based on the DQOs established for the study questions in the SPP. A single ISM sample (singlet) does not provide a CI for assessing uncertainty in the mean or provide evidence that the common assumption of relatively low heterogeneity within the DU is met. A minimum of three replicates are needed from a DU to estimate a 95% UCL, and they may be required from some SUs even when a DU is comprised of multiple SUs, depending on study objectives.

> ***In order to calculate a 95% UCL on a mean, at least three replicates are needed from the DU.***

The objective of ISM is to provide a reliable estimate of the mean contaminant concentration in a DU, recognizing that any individual ISM sample may over- or underestimate the mean to some degree. This sampling error may be attributed to a variety of factors, as discussed in Sections 2.3 through 2.5. One objective of systematic planning for most sampling designs is to minimize the major sources of error in both the field and the laboratory. ISM standardized protocols minimize sampling errors, and replicates provide a measurement of the cumulative field and laboratory errors. In addition, replicates provide a measurement of the heterogeneity within the SU, with data from replicates used to assess data quality and CSM assumptions, as described in Section 6. In practice, the estimated variance is often viewed as an overall measure that includes the contribution of many sources of error. Section 3.2.4 provides an overview of the attributes of ISM samples that support robust estimates of a 95% UCL, including superior physical site coverage, relatively low variability, and a large sample mass.

Each ISM replicate provides an estimate of the true mean for that SU. As such, the distribution of ISM replicate results is related to but conceptually different from the distribution of discrete samples. The two approaches may share the same grand mean but can be expected to have different estimates of variance (see Section 4.2.1 and Figure 4-3 in (ITRC 2012)). For ISM, the mean of replicates is analogous to repeated trials of discrete sampling (the mean of the means, or grand mean), and the SD is analogous to the SE for the mean in discrete sampling. Even the most comprehensive sampling protocols will introduce some degree of sampling error, so it is possible that a single ISM replicate result can be well above or well below the true mean. The magnitude of the under- or overestimate depends on the overall heterogeneity of the underlying distribution, which increases as heterogeneity increases; this is why, ideally, at least three replicates of each ISM SU should be collected. As described in Section 3.2, both statistical simulation studies and case studies with ISM data support the need for three or more replicates.

8.3.2.1 One ISM result: pitfalls and high uncertainty.

For sites where there is a regulatory requirement to calculate a 95% UCL, at least three replicates should be collected within a DU. For sites where there is no regulatory requirement to calculate a 95% UCL, it is important to understand the potential for decision errors if a decision is to be informed by a single ISM result. Two critical components to a decision error are the likelihood of underestimating the mean and the magnitude of the underestimation, which correlates to the direction and magnitude of the uncertainty in the risk estimate. Section 3.2.5 discusses decision errors, and the 2006 G-4 document beginning on page 63 (USEPA 2006b) further describes estimating decision errors.

Section 6.2.1.1 provides a summary of cautions for drawing conclusions based on comparison of a single ISM result to an AL. Statistical information presented in Section 3.2 provides important concepts about each ISM result. As discussed in Section 6.2.1.1, a single ISM sample gives no information for estimating uncertainty in the mean concentration, which greatly limits the scientific defensibility of this approach. A single ISM result might support a risk-based decision when the estimated mean concentration is much greater than or much less than an AL. In this situation, the ISM sample might provide confirmation of what may have already been strongly suspected – that the DU clearly passes or fails. Obviously, as discussed in Section 6.2.1, uncertainty about making the right decision increases as the ISM sample result gets closer to the AL. The risk assessment uncertainty analysis should clearly discuss the large uncertainty if only one ISM sample is used in risk-based decision-making. Use of one ISM sample rather than three replicates for estimating risks has a very high level of uncertainty that may be inconsistent with many state or agency needs for risk-based decision-making. The uncertainty associated with making decisions with only one ISM sample may make this approach unacceptable to regulators for use in risk assessment.

8.3.3 Calculation of EPCs and 95% UCLs with ISM data

This section discusses the various EPC estimates that may be used in a risk assessment conducted with ISM data. Topics include the applicability of ISM EPCs for risk assessment use, similarities and differences between ISM and discrete EPCs, considerations for ensuring adequate spatial coverage for the EPC, the importance of calculating statistically sound EPCs with ISM, ensuring that the EPC (95% UCL, see Section 3.2.4) calculated from the ISM sample means encompasses the “true mean” of the soil population (see Section 3.2.4 on coverage of the 95% UCL), and circumstances when multiple small DU results may be combined to obtain an EPC for a larger DU. An example of using “weighted means” with multiple SUs is also discussed (see also Section 6.2.2).

The EPC is intended to be a conservative estimate of the mean concentration that the receptor is in contact with over daily exposure (see Section 3.2 and Section 3.3.2). It implies that the receptor’s exposure is assumed to be spatially equal in all areas of the EU throughout the exposure period. Project objectives may specify that the estimate of the mean concentration provided by ISM sampling must be health protective, meaning there is a low chance of underestimating the actual mean concentration within the EU. Recall that a 95% UCL of the arithmetic mean is an upper bound estimate of the mean concentration in a given environmental medium. Use of the 95% UCL is health protective because there is only a 5% chance that the mean is underestimated, assuming that appropriate statistical and sampling methods are used such that the coverage is not less than a 95% probability of encompassing the true mean. It is important to recognize that the likelihood of underestimating the mean from any sampling method (discrete, composite, or ISM) increases as the distribution of concentrations becomes more positively skewed (see Section 3.2). Traditionally, with discrete samples, the concern for underestimating the mean has been addressed by specifying an acceptable level of uncertainty (often 5%) and a method for calculating an estimate of the mean with that level of confidence (such as a 95% UCL on the mean). A similar approach is used with ISM data, as discussed below.

For those accustomed to working with 95% UCL values from discrete datasets, there are some important differences with 95% UCLs from ISM data. As discussed in Section 3.2.4, calculation of a 95% UCL for ISM data requires a minimum of three ISM samples, which is generally fewer samples than is required for discrete datasets to yield reliable 95% UCL values. The first ISM Team built an ISM UCL calculator in an Excel spreadsheet file that has been updated since then with an improved modeling procedure. Further information about the updated ISM 95% UCL calculator can be found in Section 3.2.4. Additional ISM replicates above the minimum of three increases the performance of the mean estimate, thus providing a 95% UCL closer to the actual mean; although this also increases the cost, it may be necessary if the site is relatively heterogeneous and worthwhile if the result is anticipated to be close to a level of concern. Alternatively, the CSM may need revision and multiple smaller DUs established if the inherent assumption that there are no significant spatial patterns of contamination within the DU is suspected of being incorrect.

With discrete sample datasets, the maximum concentration observed is sometimes used as the EPC if it is less than the calculated 95% UCL. However, with both discrete and ISM data, the maximum concentration observed may still underestimate the population mean if the sample size is low or the population is highly variable. As discussed in Section 3.2.6.1, the calculated 95% UCL value should always be used as the EPC with ISM samples, even if it is higher than any of

the individual ISM results. This situation is not uncommon, particularly when the number of replicates is small. In fact, with three replicates, the 95% UCL will always exceed the highest individual ISM result.

As discussed in Section 3.2.4, two methods for calculating the 95% UCL from ISM data are recommended: Student's-t and Chebyshev. The choice of method depends on the known or anticipated shape of the probability distribution of contaminant concentrations in the DU. Section 3.2 presents the more common 95% UCL equations and decision criteria for selecting the approach that is most likely to provide the desired coverage of the arithmetic mean concentration. An Excel tool (ISM 95% UCL Calculator) is available from the ITRC webpage to facilitate these calculations. USEPA's ProUCL 5.1 software, which implements a wider range of 95% UCL calculation methods, is generally not recommended for use with ISM datasets consisting of fewer than 8 replicates. Furthermore, USEPA guidance on the use of ProUCL 5.1 (USEPA 2015) cautions that at least 10 to 15 observations are needed before relying on bootstrap resampling techniques to estimate 95% UCLs.

The concepts of coverage, CI widths, and accuracy of the 95% UCL are discussed in Section 3.2. In risk assessment, each of these are a measure of uncertainty in the EPC with coverage being the frequency that the 95% UCL is expected to equal or exceed the mean. Typically, a 95% UCL is used for risk assessment, with the resulting uncertainty of a 5% chance that the true mean is above the EPC. As discussed in Section 3.2 and in ISM-1, Section 4 (ITRC 2012), simulation studies have demonstrated that the degree of skewness of the underlying distribution of increments affects whether a 95% UCL method can provide a coverage of 95%. For unknown distributions or highly skewed distributions with CV greater than 1, the Chebyshev 95% UCL method is capable of calculating an EPC with 95% coverage. While we have no way to demonstrate the accuracy of an EPC, the CI width provides a measure of the precision. The more precise the replicates, the lower the RSD and the narrower the CI. A narrower CI reduces uncertainty in the EPC and risk assessment. A small mean-to-95%UCL width is desirable when the goal is to confidently estimate the true DU mean. The degree of variability (that is, the range of data values), expressed as the SD, is a common measure of variability and is used to calculate the 95% UCL – less variability (a lower SD value) gives a narrower mean-to-95% UCL width. Tolerance for the mean-to-95%UCL CI width (see Section 3.2.4.3) can be defined in the DQOs during systematic planning (Table 3-3). The 95% UCL may provide an unreliable estimate of exposure if the dataset is from too few increments or replicates and/or is highly variable (see Section 3.2.2). On occasion, it may be desirable to undertake an additional phase of investigation with redesigned DUs and/or more increments per DU to achieve lower RSD, a narrower CI, and more confidence in the EPC and risk estimates.

8.3.3.1 Combining SUs, EUs, or DUs for 95% UCL calculation of EPC

On occasion, there might be a desire to combine information from multiple SUs, EUs, or DUs into a single larger EU or DU area. Recall that DUs for risk assessment are EUs, although DUs for other study questions may also be investigated at a site. Three types of situations are described here:

- Each SU or EU has three replicate ISM samples with either the same or different spatial coverage.
- Three replicates from one or more random SUs or EUs and a singlet ISM is sampled from all other SUs or EUs.
- A random subset of SUs or EUs is sampled from a very large CSM-equivalent DU or project area.

Cautions on combining SUs or EUs into larger EUs or DUs relate to the increased uncertainty in the EPC and risk estimates. If there is a spatial trend across SUs such that some SUs are more similar to each other than others, then ISM samples collected from each SU would not be independent, meaning sample independence is violated. An example of a spatial trend across SUs is if a release occurred primarily on the top of a hill but migrated down certain parts of that hill, then the SUs comprising the preferential path downhill and the SUs outside the preferential downhill path would each have similar spatial trends but would be different from each other. The consequence of violating the statistical assumption of independence is that the variation will be underestimated, resulting in an underestimation of the EPC and risks. Uncertainty in the EPC and risks is also higher for the situation in bullet 3 above, due to a lack of sampling across the entire DU or large project area.

***ISM three replicates per SU or EU*.** When each ISM SU or EU has three replicates, and there are multiple SUs in an EU or multiple EUs within a larger DU, then a mean can be used to provide an estimate of the 95% UCL EPC across a receptor-specific EU. If the SUs or EUs are all the same size (area and depth) with the same increment density, then a standard 95% ISM UCL can be calculated using the ISM 95% UCL calculator described earlier in Section 8.3.3 and in Section 3.2.4. The variable sizes of each SU, EU, or DU can be taken into account by using a weighted mean. Weighted means take into account the spatial scales of contamination or receptor-specific activity patterns. Section 6.2.2 provides a detailed discussion and equations for calculating weighted means, There are two primary situations when calculation of a weighted mean might occur for risk assessment:

- The CSM of a site anticipates different expected levels of contamination in different areas within a larger area that we would like to define as an EU. Each of those subareas might be investigated as a separate DU for site characterization and then combined to define a single EU. Section 3.1.6, Example Set 3, provides an example

CSM with multiple potential source areas within an EU.

- For ecological or human health risk assessment, we might need to consider a variety of sizes of EUs to accommodate multiple receptor scenarios because different receptors' exposure areas are of different spatial scales. For example, if the area of a pocket mouse habitat is a quarter of that of a muskrat, which is an eighth of that of a fox, we might need to sample SUs defined for pocket mice but then combine the SUs into EUs for receptors with larger home ranges.

When these considerations are incorporated in the initial planning stages, they can be addressed by using a stratified sampling design. Within each stratum (smaller SU or EU), it may be appropriate to use ISM, but appropriate systematic planning is needed to ensure that the ISM data from the different strata can be combined for a larger EU. If there are three replicate ISM samples in each SU or EU, a weighted mean could be calculated as described in Section 6.2.2. While the ISM samples in this case are not true replicates of the mean throughout the EU in the sense that they provide information on different portions of the EU, they can provide an unbiased estimate of the mean for the whole EU.

Table 6-3 provides numerical examples of this calculation, where data from two scenario-specific EUs are combined to derive a 95% UCL for a larger scenario-specific EU. In these examples, an elementary school is divided into two EUs representing different play areas: EU1 is the kindergarten playground, and EU2 is the playground for older children. A maintenance worker has contact with both of the smaller EUs, and a separate EU is constructed to reflect the exposure of this worker. If it is assumed that, on average, a maintenance worker spends equal time in EU1 and EU2, then the replicates from each EU can be weighted equally, yielding the results shown in the "Equal weight" row of Table 6-3 for the maintenance worker's EU. Alternatively, it may be assumed that a maintenance worker's exposure is proportional to the respective areas of the two smaller EUs, and the equations from Section 6.2.2 can be used to generate summary statistics for the proportionally weighted combined area, presented in the "Proportionally weighted" row of Table 6-3. The weighting factors applied to each EU should sum to 1.0, which is achieved by dividing each area by the sum of the two areas.

You can click here to download the updated ISM 95% UCL Calculator (ISM 95% UCL Calculator) for a combined EU from several smaller EUs.

By design, ISM sampling provides an estimate of the average concentration within an EU. As noted above, if areas of elevated concentrations or other spatial patterns of contamination are suspected, these subareas can be evaluated as separate N&E DUs and then combined to define a single EU. This type of ISM design also supports risk management decisions for individual DUs that might contribute to risks within a larger EU.

The same methodology for calculating a weighted average 95% UCL described above for multiple surface soil SUs, EUs, or DUs could be used to combine a surface SU or EU with one or more corresponding subsurface SUs or EUs. The only slight difference would be that the weighting term would reflect the proportion of the total soil volume within the vertically integrated EU.

Three replicates from one or more SUs. If one or more SUs or EUs are represented by a single ISM sample, and one or more other SUs or EUs are represented by at least three replicates, the following methods are options for calculating the 95% UCL for the larger scenario-specific EU:

- One option, discussed in Section 3.2.6.2, is to use pooled variances from SUs or smaller scenario-specific EUs with three replicates that are applied to calculate 95% UCLs for the singlet SUs or EUs. This method is appropriate for CSM-equivalent SUs and EUs where a statistical test that compares variances demonstrates that the differences in variances are not significantly significant (that is, at the 95% level of confidence).
- Another method for computing a 95% UCL could employ the random selection of one replicate result from each SU or EU with multiple replicates (for example, the first replicate) after establishing in the SPP how the replicate used in calculating the 95% UCL would be randomly selected, as described in Section 3.3.2.
- If a similar variation is expected across the SUs or smaller scenario-specific EUs, an alternative option discussed in Section 3.3.1 involves collecting from one of the SUs or EUs at least three replicates, applying the measured RSD to each of the SU or EU results, and proceeding with the weighted 95% UCL calculation discussed above.
- Section 3.3.2 gives additional details on these and various possibilities to structure the statistical analysis of scenario-specific EUs composed of smaller SUs or EUs.

Random sampling subset of SUs or EUs to characterize a very large EU or project area. Occasionally, a risk-based study question may be regarding a very large receptor-specific EU. A brief discussion of using ISM data from multiple SUs or smaller scenario-specific EUs within a larger EU when a singlet ISM sample is obtained for each SU or smaller EU is provided in Section 8.2.2.2, and details pertaining to statistical evaluation of these data and a calculation of a 95% UCL for the very large EU is provided in Section 3.2.8.1. A broader discussion for approaches to evaluate risks or compliance with risk-based concentrations from large EUs follows.

In some cases, a large EU can be divided into equally-sized multiple SUs or smaller scenario-specific EUs, and only a portion of the SUs or EUs need to be sampled to provide either a mean of the entire EU or determine if the means for each SU or EU will likely be less than a risk-based benchmark. While the grids within the EU should be equally spaced and contiguous (see Section 3.2.8.1, Figure 3-25), the SUs sampled do not have to be contiguous. A feature consistently required for sampling a random subset of SUs or smaller EUs is a CSM-equivalent study area supported by a mature CSM.

When the risk assessment is evaluating a large EU, and the area is a CSM-equivalent EU, an approach for sampling a random subset of SUs is described in Section 3.2.8.1. The example is an 80-acre EU for a farmer that is divided into 80 equal-sized 1-acre SUs. A minimum of 10 randomly selected SUs are each sampled with three replicate ISM samples from at least one of the SUs and a singlet ISM sample from the remaining of the randomly selected SUs. Using a minimum of 10 SUs in the sampling design allows for calculating the 95% UCL with ProUCL. Three replicates from one of the SUs provides a measure of the variability between replicates to verify the accuracy of CSM equivalency across the EU. Increasing the number of SUs sampled can reduce the width of a CI (and the magnitude of the 95% UCL) because the number of samples factors into the 95% UCL calculation; likewise, a determination of compliance (95% UCL ≤ threshold) may be sensitive to the choice of number of SUs. Thus, although sampling a subset of 10 SUs is statistically sufficient for calculating a 95% UCL via ProUCL, the variability in ISM results between the SUs should be established in the planning and documented in the DQOs. Furthermore, the spatial coverage of the large EU may be too small and create an unacceptable amount of uncertainty in the risk assessment for broad use.

If the question regarding a very large CSM-equivalent study area is, “Are all EUs in the study area less than a target risk-based concentration?”or“What are the estimated risks (a 95% UCL is needed for the risk assessment EPC)?”, then an alternative approach with less uncertainty is to expand special coverage by increasing the number of SUs (in this case, small EUs) in the sampled subset. Section 3.2.8.2 details the process for an EU or study area so large that it cannot be sampled as a single unit. The very large CSM-equivalent EU is completely divided into many (more than 100) SUs of equal size, or the same strategy can be used to divide a large study area or property into multiple equal-sized SUs (smaller receptor-specific EUs) that are randomly sampled using ISM. The statistical calculations, which are independent of spatial area, establish a sample size of 59 SUs (smaller receptor-specific EUs), which is sufficient for 95% confidence that at least 95% of the EUs in the very large study area are less than or equal to a threshold (such as a risk-based screening level or cleanup goal) if none of the EUs are greater than the threshold and for calculating a 95% UCL. For a 95% UCL, three replicates are collected from a percentage of or all of the smaller EUs to determine the pooled variance for use in calculating the 95% UCL. Section 3.2.8.2 notes key factors that can influence the extrapolation uncertainty and likelihood of making a decision error:

- the variance of the increments (CV of the underlying distribution)
- the percentage of the large study area or large EU area sampled
- the likely magnitude of the average 95% UCL (across all sampled subset EUs) relative to a compliance level (ratio of average 95% UCL divided by compliance level)

Simulation studies suggest sampling 30% or more of the study area to achieve less than 5% false compliance (false negative) decisions (Goodrum et al. 2018).

In all cases when extrapolations are made from sampled areas and used as surrogates for unsampled areas, it is important to use the uncertainty section of the risk assessment to discuss uncertainty in the EPC and risk estimates. Statements on uncertainty from the extrapolation methods used for unsampled areas can and should be supported with a discussion on the demonstrated low variability as measured by the replicates’ RSD and the verification of CMS-equivalent DUs in the risk assessment uncertainty analysis. Greater uncertainty in the representativeness of the subset of SUs sampled to the entire very large EU is tolerated when either the COPC concentrations are far below the risk-based screening levels (or ALs), or the calculated risks and hazards are far below the regulatory thresholds (typically 1 x 10^{-6} risk and hazard index of 1).

8.4 Considerations for Use of Background ISM Data in Risk Assessment

Background concentrations are often used in risk assessment to help refine the COPC list so that chemicals related to site releases can be more easily identified. Risk assessments may compare ambient background concentrations to site concentrations to eliminate COPCs before or after quantitative risk assessment is completed, depending on the stakeholders. ISM can be used for assessing ambient background soil concentrations of both native metals and ubiquitous anthropogenic chemicals such as dioxins and PAHs. While a background threshold value derived from discrete data can sometimes be used to determine if a chemical has been released to the environment and should be included in the quantitative risk assessment, it should not be used to eliminate a chemical based on an ISM site concentration. The comparison of ISM site data to discrete background benchmarks should be done with an understanding of the potential error in the mean based on the ISM result. Details about errors and other important concepts comparing ISM results with a single value are covered in in Sections 6.2,

Section 3.2.5, and Section 3.3.3. Also, as stated previously in this guidance document, while statistics exist to compare ISM samples to other types of samples, these methods are complex and require the assistance of a qualified statistician. Therefore, this section will describe the benefits of using ISM background results for comparison with ISM site results, how to properly plan for including a background ISM comparison in a risk assessment, and briefly describe comparison methods. Sections 3.1.6.2, Section 3.3.4, and Section 6.3 cover these topics in more detail.

8.4.1 Benefits of using ISM site and ISM background datasets in risk assessment

If ISM background data are properly collected, they can be used to compare and determine which chemicals detected in ISM site samples are likely present in similar or greater concentrations than background conditions. This information can be used before the quantitative risk assessment to eliminate COPCs or after the risk estimates are completed to aid in determining which chemicals are COCs. If the background comparison is completed after risk estimates are calculated, the comparison of COCs detected in ISM background to ISM site concentrations can provide information about how much of the total site risk for exposure pathways and scenarios is attributable to background conditions. Since both background and site data provide mean concentrations, they are relatively easy to compare and use to make decisions about COPCS or COCs. In addition, planning to collect an adequate number and matching soil types for ISM background and site soils will enable the best comparison results. The statistical comparison methods are also fairly straightforward to use.

8.4.2 Planning for ISM background data collection and use

Background data from an appropriate reference area are often used to evaluate site data for environmental projects. Statistical methods typically applied to compare discrete site sample concentrations to discrete background concentrations are applicable to compare an ISM site dataset to an ISM background dataset. Section 3.1.6.2 presents aspects to consider in planning for selecting and defining ISM background SUs and DUs. Figure 3-10 in Section 3.1.6.2 provides a depiction of the way background DUs could be configured when background areas the same size as site DUs are not contiguous and/or not square/rectangle grids. Ideally, ISM background samples should be comparable to site data in the following ways:

- same sample range of depths (for example, surface soil defined as 0 to 1 ft bgs)
- same soil type (such as sand or loam)
- same increment density (for example, 30 increments per ½ acre)
- same number of increments and replicates in the DUs
- same field methods
- same analytical methods

Assumptions should be clearly stated in the planning stage and revaluated after the data are available. Power is the probability of detecting a difference between background and site concentrations, given that a difference truly exists. A power analysis based on the expected variation and desired CI should also be completed in the planning stage, so the number of increments collected in each SU or EU is more likely to meet the acceptable decision errors (see Sections 3.3.3.1 and Section 3.3.4.1). See Section 3.3.4 for more details on planning for background comparisons.

8.4.3 ISM background comparison methods

Although USEPA guidance on hypothesis testing (USEPA 2002d, b, 2009) was developed with discrete sampling in mind, these methods are also applicable to ISM data comparisons. Two fairly easy statistical methods useful for comparison of ISM background and site data are means versus upper tails (see Section 3.3.4). Evaluating assumptions and completing power analyses before sampling and after data are available are also recommended.

Section 6.3 describes how to evaluate some of the assumptions underlying certain statistical tests, such as independence of the samples and normality of the distribution of the data. ISM sampling designs commonly do not have a sufficient number of samples to reject normality regardless of the true distributional form. However, as noted in the 95% UCL part of Section 3.2.4 and Section 6.3.2, an assumption of normality is generally appropriate for ISM data due to the CLT.

If background and site data are both generated using ISM, comparisons of means can be made using hypothesis testing. Section 3.3.4.1 provides guidance about the use of one- or two-sided hypothesis tests and what to do if the distributions of one or both ISM background or site datasets do not appear normally distributed. ANOVA can also be used to compare the variability of the means between both datasets if the variances are equal. ANOVA is useful for comparing multiple site DU means to background DU means, but it only identifies if there is a statistical difference between the groups, not the groups that differed, necessitating follow-up by individual comparisons for each site DU to the background DU. Furthermore, for 95% confidence ($\alpha = 0.05$), a minimum of five ISM replicates are typically needed from each DU. While ANOVA is a viable option for statistically comparing ISM site data to ISM background data, it is best conducted by a statistician or other

professional well versed in environmental statistics.

Section 3.3.4.2 discusses how to compare the upper tails of the two distributions from the ISM background dataset and ISM site dataset. Statistical power to detect differences will be low if there are a limited number of replicates in the ISM dataset. Commonly, at least $n = 8$ observations per group is desired before using hypothesis tests to compare upper tails (such as the Quantile test).

Decision errors and power analyses are discussed in Section 3.3.3.1 and Section 3.3.4.1 Acceptable decision error probabilities should be determined in the planning phase and reevaluated when the data are available. Power analyses should be completed in the planning stage to determine the number of increments to collect in each SU but are also useful to interpret the false exceedance of background error in a one-sample hypothesis test.

A multiple lines of evidence approach with qualitative comparisons of site and background ISM data in conjunction with statistical comparisons of means and/or upper tails can provide more confidence in a decision regarding how to evaluate a native metal or ubiquitous anthropogenic chemical as a COPC or COC in a risk assessment. An example of a qualitative approach is comparison of RSDs among replicates from a site ISM DU to the background ISM DU. Because the ISM results are means, the variability (RSD) in the ISM replicate results from a contaminated site DU may be higher than that in the ISM background replicates. Discrete data from a contaminated area typically have wider variability in their distribution than discrete background data, but ISM datasets have less variability than discrete because ISM data distributions are from mean concentrations (ITRC 2012). Although ISM field replicates measure sampling precision, comparison of RSDs from ISM site to ISM background data may indicate a contaminant release has occurred on the site but cannot be used to support the hypothesis that site concentrations are similar to background concentrations. Including graphical analyses and figures will aid in showing important differences between ISM site and ISM background distributions. Simple graphical analysis can also provide useful information and serve as a semi-quantitative means of comparison. Background comparisons accompanied by visual presentations of site versus background data such as scatter plots, histograms, or box plots provide more information about how the data compare.

8.5 Use of ISM for Post-Remediation Risk-Based Confirmation Sampling

The application of ISM is most straightforward at the post-remedial confirmation sampling stage of an environmental investigation. The objective of confirmation sampling is commonly to compare average soil concentrations to remedial action criteria within defined DUs that are remediation units. ISM can be used for confirmation sampling conducted to evaluate if a remedial action meets risk-based benchmarks because a properly designed ISM sampling plan can provide a more robust estimate of mean residual contaminant concentrations in a DU than a sampling plan reliant on discrete samples.

In confirmation sampling, the usual objective is to determine whether there is sufficient evidence to conclude that the true DU mean concentration is less than the remedial goal associated with the remediation. Generally, this involves comparison of a 95% UCL to the remediation goal. Section 3.2.5 discusses the topic of statistical hypothesis tests to support this objective and why the use of a 95% UCL is a simple and appropriate way to implement steps 5 and 6 of USEPA's DQO process as it is equivalent to conducting a hypothesis test with a 5% error rate.

The most common example for the application of ISM in post-remediation confirmation sampling design is excavation of contaminated soil from a DU, which allows ready access to the assumed outer margins of the DUs. The sidewalls and floors of excavations can be treated as one or more DUs and can be sampled for confirmation of adequate soil removal. If the spatial boundaries of the excavation DUs are different than the site-specific EU for risk-based cleanup goals, then either combining excavation DUs into EUs or subdividing DUs into SUs with spatial scales equivalent to EU spatial boundaries is appropriate. Use of ISM for the collection of confirmation samples is illustrated in Example Set 1 in Section 3.1.6.

ISM is applicable to confirmation sampling if the criterion for successful excavation is achieving a mean concentration in soil DUs (sized appropriately for the project objectives, for example, risk-based screening levels or remediation goals) below the ALs at the excavation boundaries. Small volumes of contaminated soil within otherwise clean excavation sidewalls or floors do not necessarily pose a significant risk to human health and the environment. The specification of an appropriate DU area in relation to confirmation sampling should be addressed during the planning process.

8.6 Risk Communication Suggestions for Explaining ISM

The goal of risk communication is for all stakeholders to have a common understanding of the results, processes, and assumptions used in risk assessment as well as how the risk assessment is used in risk management. Use of ISM for risk assessment input may necessitate communicating with stakeholders and others the differences between discrete and ISM sampling, as well as how the proper use of ISM enhances the defensibility of risk assessment results in comparison to

reliance on discrete sampling. Section 9 of RISK-3 (ITRC 2015) provides a general overview of risk communication, including issues related to risk perception and strategies for communicating technical information to diverse groups of stakeholders. This section focuses on the key differences between ISM and traditional discrete and composite samples and provides suggestions on addressing common misconceptions about ISM. For further information on risk communication, see ITRC's Risk Communication guidance (ITRC 2020).

8.6.1 ISM reduces uncertainty in the mean

Uncertainty in the risk estimates from the use of ISM sampling and analysis protocols derives from the reduced uncertainty in the average concentrations of contaminants in soil or sediment. Scientific studies demonstrate lower variability between ISM replicates than discrete samples. Section 2.2.2 and Section 2.5.2 discuss the soil science and sampling theory that contributes to the larger variation between discrete samples' mean concentrations and includes a study by (Becker 2005) where lead is the COPC. Studies by Brewer et al. (Brewer, Peard, and Heskett 2016) (Brewer, Peard, and Heskett 2017) examined the effects on variability between discrete samples and ISM replicates within a DU for three different COPCs – arsenic, lead, and PCBs. Arsenic was used as a termiticide and preservative for building products at a former facility that manufactured ceiling and wallboard material treated with arsenic at Site A. Lead in incinerator ash at a former municipal incinerator was investigated at Site B. PCBs at a former radio broadcasting station was the focus for Site C. The ISM design of three replicates from each DU with either 54 increments (arsenic and lead) or 60 increments (PCBs) per DU provided robust evidence of reduced variability between samples as compared to discrete samples. Further sound evidence for reduced variability with ISM was demonstrated upon examination of 20 iterations of random groupings of 10 discrete samples from each of Sites A, B, and C. The largest variation was seen in Site C's PCB random grouping of discrete samples, where the 95% UCL ranged from 9.4 to greater than 1,000,000 mg/kg. The 95% UCL for arsenic discrete datasets of $n = 10$ ranged from 403 to 776 mg/kg from Site A. Lead 95% UCLs from Site B ranged from 201 to 439 mg/kg, straddling the risk-based concentration and preliminary cleanup goal for commercial/industrial receptors. These studies demonstrated that the mean concentrations from discrete samples are more prone to elicit inaccurate risk conclusions resulting in incorrect decision-making for protection of human health and the environment.

From a statistical standpoint, analysis of multiple ISM replicates collected with the same sampling protocol (sampling method and number of increments in an EU) provides a direct measure of the variance in the mean. ISM variances in the mean concentration are typically much lower than that for discrete sample data. A figure may be useful for conveying the population differences in variances between discrete and ISM datasets, and one option is a modification of what was originally Figure 4-3 in ITRC ISM-1 (ITRC 2012). Lower variability, as is characteristic with ISM, yields a 95% UCL with a smaller CI width, and as discussed in Section 8.3.3, a small mean-to-95%-UCL width provides more confidence in the estimate of the true EU mean. In comparison to the results from several discrete sample results, several ISM replicate results typically have a tighter upper CI. The lower variance among the ISM replicates results in less uncertainty in the EPC and in the associated risk estimate.

8.6.2 ISM provides a more representative EPC

ISM protocols typically provide much greater spatial coverage than discrete sampling. Because the use of ISM samples provides better spatial coverage than randomly placed discrete samples, EPCs based on a properly designed ISM sampling plan (one with an adequate number of increments and replicates in a DU) are more representative of a receptor's potential exposure over a long-term period than random discrete samples. The EPC is only one factor in the equations for estimating risks that are upper bound estimates used to meet the goal of ensuring that risks are not underestimated. The higher confidence in the EPC reduces the uncertainty in the risk estimates and strengthens the scientific defensibility of risk-based decision-making. A common concern with the use of ISM is that there could be areas of elevated soil concentrations within a DU, and that ISM will "average away" these concentrations. ISM does provide an estimate of the average concentration within a DU but no information on possible spatial patterns of contamination therein unless the DU is subdivided into mutually exclusive SUs. However, it must also be acknowledged that there is a much higher probability of hitting a small area of elevated soil concentrations with three 50-increment ISM sample replicates than with, for example, 10 or 15 discrete samples collected in the same area. The trade-off is between diluting the concentrations from the elevated area by including the area within the larger DU mean – and potentially missing the area altogether when using discrete samples. Concerns with identifying subareas of much higher soil concentrations are particularly valid when the CSM suggests that such areas could exist, and as noted in Section 8.2.3.1, sampling designs can be developed to address these concerns.

Click Here to download the entire document.