United States
Environmental Protection
Agency

Office of
Emergency and
Remedial Response

EPA 9355.0-91
April 2005
PEER REVIEW
DRAFT

EPA

GUIDANCE ON SURFACE SOIL CLEANUP AT HAZARDOUS WASTE SITES:

IMPLEMENTING CLEANUP LEVELS

GUIDANCE ON SURFACE SOIL CLEANUP AT HAZARDOUS WASTE SITES: IMPLEMENTING CLEANUP LEVELS

PEER REVIEW DRAFT

Prepared for:

Office of Emergency and Remedial Response
U.S. Environmental Protection Agency

Prepared by:

Industrial Economics, Incorporated
2067 Massachusetts Avenue
Cambridge, MA 02140

Dr. Shahrokh Rouhani
New Fields, Incorporated
1349 W. Peachtree Street
Atlanta, GA 30309

April 2005

DISCLAIMER

This document presents current OSWER technical options regarding implementation of cleanup levels for surface soil at hazardous waste sites. EPA and state personnel may use and accept other technically sound approaches, either on their own initiative, or at the suggestion of potentially responsible parties or other interested parties. Therefore, interested parties are free to raise questions and objections about the substance of this document and the appropriateness of the application of the options to particular situations. EPA will, and states should, consider whether or not the approaches in this document are appropriate in each situation. This document does not impose any requirements or obligations on EPA, states, or other federal agencies, or the regulated community. The sources of authority and requirements in this matter are the relevant statutes and regulations (e.g., the Comprehensive Environmental Response, Compensation and Liability Act, and the Resource Conservation and Recovery Act). EPA welcomes public comments on this document at any time and may consider such comments in future revisions of this document.

TABLE OF CONTENTS

1.0 INTRODUCTION

Achieving EPA's goal of protecting human health and the environment typically requires that successful soil cleanup at hazardous waste sites eliminate unacceptably high risks associated with potential exposure to contaminated soils. It is important to achieve this goal in a cost-effective manner. This document provides guidance on approaches to surface soil cleanup that address these considerations and can help risk managers find the most cost-effective way to reduce human health risk.[1] Risk assessors and statisticians should also find this guidance useful in providing information needed to support risk management decisions. Such decisions are not purely statistical; while they require statistical inputs, they are ultimately management decisions.

This guidance does not address development of remediation goals, but rather focuses on how cleanup goals should be implemented at a site in delineating what soil areas to remediate and the concentration above which soil should be removed to achieve a protective cleanup. This guidance does not address the related determination of exposure point concentrations used for risk assessment. This topic is covered in: *Supplemental Guidance to RAGS: Calculating the Concentration Term;* and *Calculating Upper Confidence Limits for Exposure Point Concentrations at Hazardous Waste Sites.*[2] The guidance also does not address site assessment sampling design. Nor does it address post-cleanup confirmatory sampling to be done regardless of how the cleanup goals are implemented. This topic is covered in: *Methods for Evaluating the Attainment of Cleanup Standards, Volume 1.*[3] The *Multi-Agency Radiation Site Survey and Site Investigation Manual* (MARSSIM) provides guidance on survey design for confirmatory sampling for radionuclides,[4] and is based on an area average approach and could be adapted for other contaminants.

A vital concept in this document is the difference between the implementation of a cleanup level as a not-to-exceed level or as an area average. The not-to-exceed option typically entails treating or removing all soil with contaminant concentrations exceeding the cleanup level. The area average option typically involves treating or removing soils with the highest contaminant concentrations such that the average (usually the upper confidence limit of the average) concentration remaining onsite after remediation is at or below the cleanup level. A key factor driving the choice between these options is the basis for the cleanup level. The method used in

---

[1] The guidance is not applicable to sites where ecological risks are driving cleanup decisions, or where contaminants have the potential to leach to groundwater and where such leaching is the basis for cleanup.

[2] U.S. EPA Office of Solid Waste and Emergency Response, *A Supplemental Guidance to RAGS: Calculating the Concentration Term*, Publication 9285.7-081, 1992; and U.S. EPA Office of Emergency and Remedial Response, *Calculating Upper Confidence Limits for Exposure Point Concentrations at Hazardous Waste Sites*, OSWER 9285.6-10, December 2002.

[3] U.S. EPA Office of Solid Waste and Emergency Response, *Methods for Evaluating the Attainment of Cleanup Standards, Volume 1,* Publication 230/02-89-042, February 1989.

[4] U.S. EPA, U.S. DOE, U.S. NRC, *Multi-Agency Radiation Survey and Site Investigation Manual* (MARSSIM), EPA 402-R-97-016, Rev. 1, August 2000.

implementing the cleanup level should be compatible with the method used in establishing the cleanup level. Another important factor is the nature and extent of site assessment data. Therefore, consideration of the approach to implementing cleanup levels should be interwoven with decisions about sampling and risk assessment and should be addressed early in remedial action scoping and delineation sampling for excavation.

This guidance first discusses general concepts important to understand when making decisions about how to implement surface soil cleanup levels at hazardous waste sites. Then, the document presents in detail two recommended options for implementing cleanup levels: not-to-exceed and area average. To help risk managers decide whether to implement cleanup levels as not-to-exceed levels or as area averages, this part of the guidance discusses these options with respect to their advantages, disadvantages, and appropriate use. The remainder of the guidance focuses on potential methods for implementing cleanup levels as area averages and presents three different statistical methods for calculating remediation action levels that ensure that post-remediation area average contaminant concentrations achieve cleanup levels. This emphasis on the area average approach is not intended to suggest that the area average is EPA's preferred approach, but rather to fully explain this new approach as an option for consideration. The purpose of this document is to present the options and discuss factors to consider in the decision to use a not-to-exceed approach or an area average approach. The document does *not* establish EPA policy on how to implement cleanup levels; such decisions will depend on many site-specific factors, some of which are discussed in this document.

2.0 GENERAL CONCEPTS

Exposure unit. The application of cleanup levels at a site usually requires understanding the behavior of the receptors and how these receptors are exposed to contaminants across the geographic area of the site. A key concept is the exposure unit (EU). The exposure unit generally is the geographic area within which a receptor comes in contact with a contaminated medium during the exposure duration.[5] The exposure unit should be defined based on the receptor, exposure medium, and nature of the receptor's contact with the medium.

If the receptor is a resident exposed to soils in his/her yard, the area within which exposure occurs will likely be relatively small; exposure units for residential receptors are typically less than one acre in size. EPA's Soil Screening Guidance assumes a 0.5 acre source area for residential exposure.[6] Other receptors, such as industrial workers and recreational adults may be exposed to contaminants across much larger areas at a site, and a much larger EU may be appropriate. Definition of the exposure unit is critical to the success of a cleanup strategy that involves applying cleanup levels as area averages, since the averages should be calculated for the exposure unit and should accurately reflect the average exposure of receptors in that area.

Exposure point concentration. The exposure point concentration (EPC) normally is an essential element in evaluating exposure and risk at hazardous waste sites, both pre- and post-remediation. The EPC is defined in EPA's *Risk Assessment Guidance for Superfund: Volume III - Part A* as "the average chemical concentration to which receptors are exposed within an exposure unit."[7] The spatial and temporal variability of contamination and the receptor's behavior and activity within the exposure unit can all influence the EPC. For "reasonable maximum exposure" (RME), the Risk Assessment Guidance for Superfund (RAGS) recommends using the average value with a specified level of confidence to represent "a reasonable estimate of the concentration likely to be contacted over time."[8] This average value generally is based on the assumption that contact is spatially random. For risk assessments based on point estimates of the EPC, the EPC is usually calculated as the 95 percent upper confidence limit (UCL) for the mean. For lead, risk generally is estimated using the IEUBK Model which uses an arithmetic mean EPC.

---

[5] U.S. EPA Office of Solid Waste and Emergency Response, *Risk Assessment Guidance for Superfund: Volume III - Part A, Process for Conducting Probabilistic Risk Assessment*, EPA 540-R-02-002, December 2001.

[6] U.S. EPA Office of Emergency and Remedial Response. *Soil Screening Guidance: Technical Background Document.* EPA 540/F-95/128, July 1996.

[7] U.S. EPA Office of Solid Waste and Emergency Response, *Risk Assessment Guidance for Superfund: Volume III - Part A, Process for Conducting Probabilistic Risk Assessment*, EPA 540-R-02-002, December 2001.

[8] U.S. EPA Office of Emergency and Remedial Response, *Risk Assessment Guidance for Superfund, Volume I - Human Health Evaluation Manual (Part A).* Interim Final. EPA 1540/1-89/002, 1989.

Guidance on the calculation of UCLs for EPCs at Superfund sites is provided in: *Supplemental Guidance to RAGS: Calculating the Concentration Term,* and *Calculating Upper Confidence Limits for Exposure Point Concentrations at Hazardous Waste Sites.*[9]

Preliminary remediation goal. Chemical-specific preliminary remediation goals (PRGs) are concentration goals for individual chemicals for specific medium and land use combinations. There are two general sources of chemical-specific PRGs: concentrations based on ARARs, and concentrations based on risk assessment.[10]

Cleanup level. Cleanup levels generally are based on PRGs and are refined by considering the cost and implementability of remedial alternatives, including the technical feasibility of achieving the risk-based PRG, and other criteria outlined in the National Contingency Plan (NCP). The cleanup level generally is a chemical-specific concentration chosen by the risk manager as appropriate for likely future land use based on the PRG and other practical considerations.[11] In some cases, the cleanup level is the same as the PRG. Cleanup levels are documented in the Record of Decision (ROD). Decisions about whether to implement the cleanup level as a not-to-exceed level or as an area average depend to some extent on the degree of uncertainty in the protectiveness of the cleanup. This degree of uncertainty can be determined by many factors, including but not limited to the effectiveness and adequacy of site sampling, the exposure assumptions in the risk assessment, and the toxicity of the chemicals of concern.

Remediation action level. The remediation action level (RAL) is a concept that goes hand-in-hand with the application of the cleanup level as an area average. The RAL in most cases is the maximum concentration that may be left in place within an exposure unit such that the average concentration (or 95% UCL of the average) within the EU is at or below the cleanup level.[12] The RAL is generally determined statistically. Post-remediation sampling typically is needed to ensure that the appropriate cleanup level has been met as an area average.

Nature and extent of contamination. Information about contaminant concentrations, contaminant toxicity, and the spatial and temporal variability of contamination in most cases is critical in determining how to implement cleanup levels at hazardous waste sites. Therefore adequate characterization of contamination at the site is essential. Adequate characterization depends on the spatial structure of contamination and on the sampling design. The number, location, volume, and depth of samples relative to the contamination affect the adequacy of sampling. As illustrated in

---

[9] U.S. EPA Office of Solid Waste and Emergency Response, *A Supplemental Guidance to RAGS: Calculating the Concentration Term,* Publication 9285.7-081, 1992; and U.S. EPA Office of Emergency and Remedial Response, *Calculating Upper Confidence Limits for Exposure Point Concentrations at Hazardous Waste Sites,* OSWER 9285.6-10, December 2002.

[10] U.S. EPA Office of Solid Waste and Emergency Response, *Risk Assessment Guidance for Superfund: Volume III - Part A, Process for Conducting Probabilistic Risk Assessment,* EPA 540-R-02-002, December 2001.

[11] U.S. EPA Office of Solid Waste and Emergency Response, *Risk Assessment Guidance for Superfund: Volume III - Part A, Process for Conducting Probabilistic Risk Assessment,* EPA 540-R-02-002, December 2001.

[12] U.S. EPA Office of Solid Waste and Emergency Response, *Risk Assessment Guidance for Superfund: Volume III - Part A, Process for Conducting Probabilistic Risk Assessment,* EPA 540-R-02-002, December 2001.

Exhibit 1, extrapolating the results of a small number of samples to a large area can be misleading unless the contaminant distribution across the large area is uniform. Clearly, for areas with heterogeneous distribution of contamination (e.g., scattered or dumped), the more extensive the sampling data the more representative they will be of the exposure concentration. There are also situations where the depth of the sample relative to the vertical distribution of contamination at the site can be important in accurate representation of the exposure concentration. This is illustrated in Exhibit 2. For sampling data to accurately represent the exposure concentration, they should generally be representative of the contaminant populations at the same scales as the remediation decisions and the exposures on which those decisions are based. Uncertainty associated with sampling error can be very large, particularly at sites where there is significant spatial heterogeneity in contaminant concentrations.[13] Where there is spatial heterogeneity, sampling variability can lead to erroneous cleanup decisions. Clearly, the more extensive the sampling data, the more accurately they will represent the exposure concentration. Thorough characterization of site contamination typically includes understanding of the following:

- ***Spatial variability.*** Spatial variability can be determined based on historical and physical information on contaminant release, fate, and transport, and on spatial plots of site data. Spatially correlated contamination typically exists when sampling data or other information reveal a structured pattern of highly contaminated areas surrounded by gradually decreasing concentrations. Within an exposure unit, if the data show that high contaminant concentrations occur randomly among samples with low concentrations, then the contamination is not normally considered to be spatially correlated.

- ***Temporal variability.*** Temporal variability refers to contaminant conditions that may vary over time due to factors such as migration or natural degredation of contaminants in the environment. While the time frame for sampling should provide data that are representative of the time frame of potential exposure, site characterization should note any uncertainties due to mismatches between short timeframes for sampling and longer time frames for exposures.

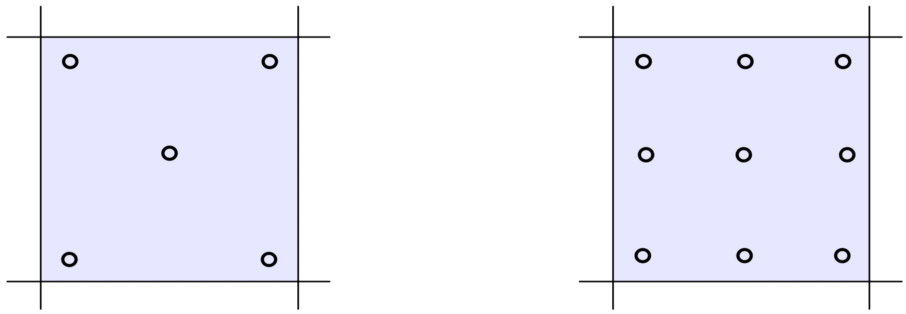
[13] *Technical and Regulatory Guidance for the Triad Approach: A New Paradigm for Environmental Project Management,* prepared by The Interstate Technology & Regulatory Council Sampling, Characterization and Monitoring Team, December 2003, provides more in-depth discussion of this issue.

Exhibit 1

SAMPLING UNCERTAINTY - LATERAL HETEROGENEITY AND SAMPLE SIZE

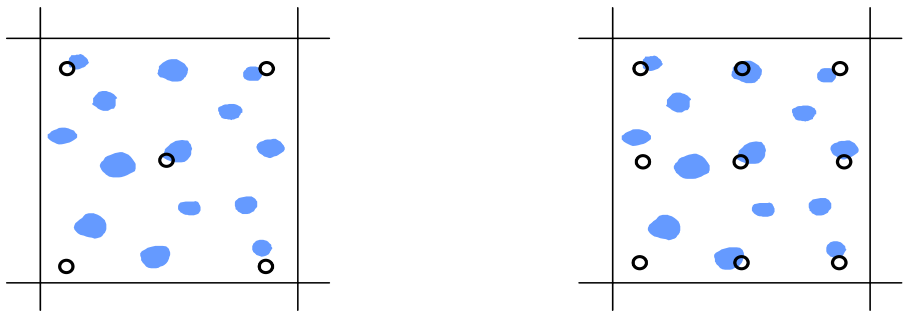
Uniform Concentration Distribution (area average = 20 ppm)

Small Sample (Sample = 20 ppm) | Large Sample (Sample = 20 ppm)



Scattered High Concentrations (area average = 20 ppm)

Small Sample (Sample = 10 ppm) | Large Sample (Sample = 20 ppm)



Dumped High Concentrations (area average = 20 ppm)

Small Sample (Sample = 10 ppm) | Large Sample (Sample = 20 ppm)

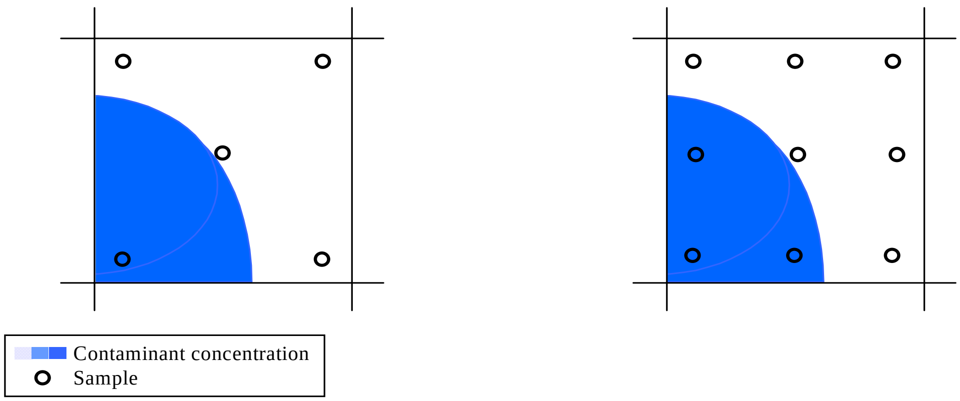
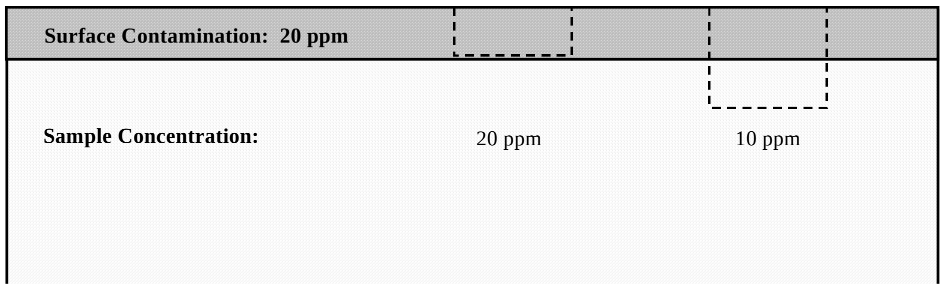


Exhibit 2

SAMPLING UNCERTAINTY - VERTICAL HETEROGENEITY AND SAMPLE DEPTH



- ***Acute and chronic toxicity.*** Contaminants present at hazardous waste sites may pose human health risks from short-term exposures, as well as from long-term exposures. Therefore contaminants need to be evaluated for their acute and chronic toxicity, and the toxicity generally should be matched to the exposure duration and frequency.

Nature of exposure. Information on the spatial and temporal variability of contamination and on the behavior and activities of the receptor guides assumptions about exposure and is important in determining how to implement cleanup levels at hazardous waste sites.[14] Specific characteristics of exposure may include:

- ***Random and non-random exposure.*** Understanding how a receptor moves across an exposure unit and contacts contaminated soils is a key consideration in determining whether it is appropriate to apply a cleanup level as an area average. At most sites, it is reasonable to assume that random exposure occurs over the long-term. Short-term exposures, however, may be non-random. For example, a resident may move randomly across his/her property spending equal amounts of time in all areas over the long-term period of residence, but intense short-term exposure may occur as a result of a construction project, such as building a shed. Another example of intense short-term exposure would be a child playing in a sand box. Both types of exposure, together with chronic and acute toxicity of contaminants, should generally be addressed in developing a strategy for surface soil cleanup.

[14] EPA's *Supplemental Guidance for Developing Soil Screening Levels for Superfund Sites*, OSWER 9355.4-24, December 2002, describes various mechanisms of human exposure to surface soil.

- ***Size of exposure unit.*** Averaging site data for evaluating exposure and risk normally requires that the size of the exposure unit be appropriate for the receptor being considered. If the exposure unit is larger than the area throughout which the receptor is exposed, then the average for the area will not accurately reflect actual exposure. For example, an adult resident may have a larger area of exposure than a child. If the adult-sized exposure unit is used for the child, then a cleanup strategy based on an area average for the adult-sized exposure unit may not be protective for the child. Similarly, a worker at an outdoor commercial facility (e.g., garden center) may be exposed to a larger area than a customer.

- ***Exposure duration.*** Similar to the size of the exposure unit, the exposure duration should be consistent with the time scale of receptor activities. In addition, the exposure duration determines which toxicity criteria are used. Appropriate acute or subchronic toxicity values should be used for short-term exposures, and chronic toxicity values should be used for long-term exposures.

Conceptual site model (CSM). The application of cleanup levels at a site typically requires knowledge of the distribution of contaminants in soil and how that distribution is related to the sample data and to risks from potential exposure. The CSM is a comprehensive representation of the site that documents these current site conditions and relationships. It generally characterizes the distribution and variability of contaminant concentrations across the site based on knowledge of contaminant, release, fate, and transport mechanisms, and on sampling data. As shown in Exhibit 1, this information is critical to developing a sampling plan that generates data that accurately represent exposure concentrations. At sites with spatially correlated contamination (e.g., dumped high concentrations), statistical sampling is more powerful when data from different contaminant populations are separated. The CSM also identifies all potential exposure pathways, migration routes, and potential receptors.[15] The CSM is initially developed from existing site data. It is a key component of the Remedial Investigation/Feasibility Study (RI/FS) and Data Quality Objectives (DQO) process, and it should be continually revised as new site investigations produce updated or more accurate information.

[15] The process for developing a conceptual site model is described in: U.S. EPA Office of Solid Waste and Emergency Response, *Supplemental Guidance for Developing Soil Screening Levels for Superfund Sites,* OSWER 9355.4-24, December 2002, p. 4-7.

3.0 RECOMMENDED OPTIONS FOR IMPLEMENTING CLEANUP LEVELS

This section presents two recommended options for risk managers to implement cleanup levels: (1) as not-to-exceed levels, or (2) as area averages. Implementing cleanup levels as area averages instead of not-to-exceed levels represents a less stringent and less costly option. Implementing cleanup levels as not-to-exceed levels may not be necessary in situations where receptor exposure activity is random but may still be the preferred option based on consideration of the nine criteria for remedy selection, and more specifically compliance with ARARs, the degree of uncertainty in the risk estimates, the adequacy of site characterization, the level of community acceptance, and other criteria discussed in this section.[16]

In determining what soils to remediate and to what levels, risk managers at hazardous waste sites may need to decide between remediating all soils to concentrations at or below the cleanup levels specified in the ROD, or leaving in place some soils with concentrations above the cleanup levels while ensuring that the estimated post-remediation EPC for a given exposure unit is below the cleanup level. These options can be characterized as implementing the cleanup level as a "not-to-exceed" value or as an "area average."

Implementing the cleanup level as a not-to-exceed value normally means that soil removal or treatment will continue until the analysis of soil samples indicates that all soil with contaminant concentrations exceeding the cleanup level has been removed or treated. As shown in Exhibit 3, after remediation is complete, the highest concentration within the EU should be at the cleanup level, and the average or UCL concentration (i.e., post-remediation EPC) within the EU should be lower than the cleanup level. If the risk-based PRG is chosen by the risk manager as the cleanup level, then applying it as a not-to-exceed level should result in a post-remediation EPC that is below this protective level.

The area average approach involves removing or treating the areas of the EU with the highest contaminant concentrations until the average or UCL concentration (i.e. post-remediation EPC) is at or below the cleanup level.[17] (See Exhibit 3.) This approach normally requires establishing a cleanup level that is the desired post-remediation EPC, and making a statistical determination of a "remediation action level (RAL)," the level to which all contaminant concentrations in soil within an EU are reduced to ensure that the estimated post-remediation EPC for the EU is at or below the cleanup level. The RAL is itself a maximum concentration, or not-to-exceed level, for the purposes of site remediation. Remediating or removing all soil with contaminant concentrations above the

---

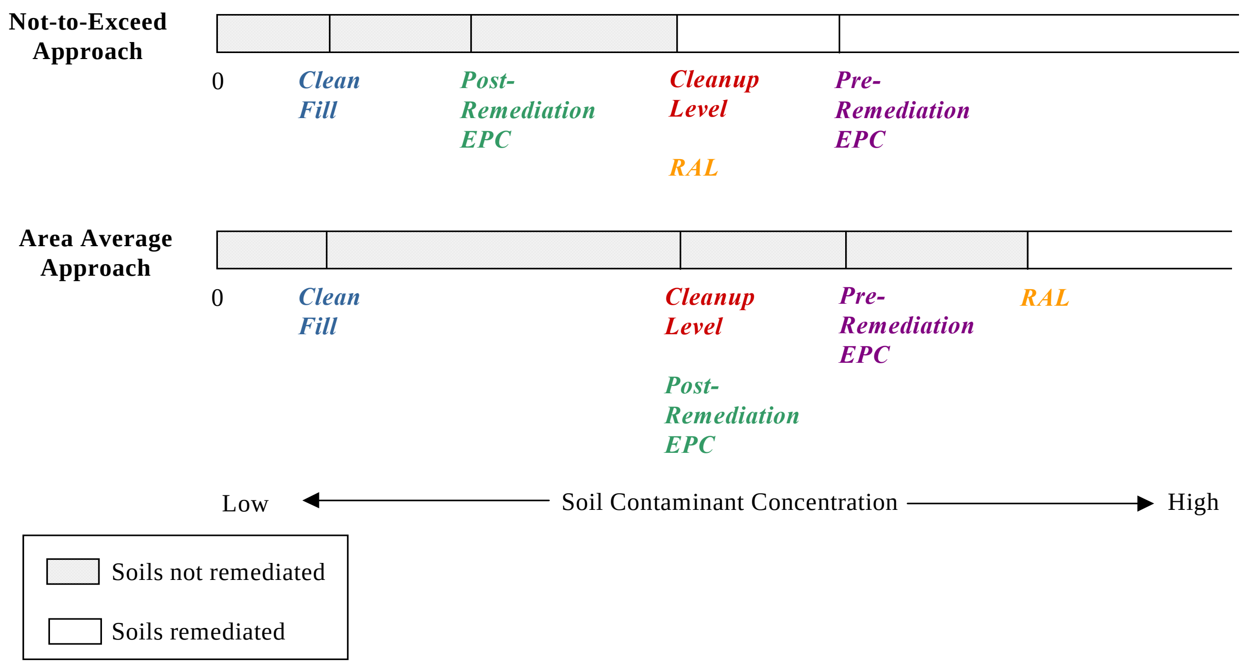
[16] See *Guide to Selecting Superfund Remedial Actions*, OSWER Publication 9355.0-27FS, April 1990, for description of nine criteria. Compliance with ARARs is one of the nine criteria. Generally, it is not appropriate to use the area average approach if an ARAR is either designated as a not-to-exceed level or was developed based on factors other than average risk.

[17] Because the area average approach assumes random exposure within the exposure unit, all areas (including clean areas) within the exposure unit should be included in the calculation. As discussed in Section 4.0, there are different methods for spatially correlated and non-spatially correlated data to address hot spots (i.e., areas with particularly high contaminant concentrations).

RAL should enable risk managers to ensure that the estimated post-remediation EPC achieves the cleanup level.

Exhibit 3

RECOMMENDED OPTIONS FOR IMPLEMENTING CLEANUP LEVELS



Key issues in applying the area average approach include defining the EU over which to average concentrations, and determining what locations to remediate in order to reduce the EU EPC to a level at or below the cleanup level. In addition, it is important to consider the possibility of acute or subchronic health effects from short-term exposure to contaminants at concentrations above the cleanup level. Because soils with contaminant concentrations exceeding the cleanup level will be left onsite, it is important to ensure that those concentrations are not so high that they pose acute or subchronic health risks if exposure to them occurs. Therefore, if this approach is used, the RPM should conduct a separate assessment of potential acute effects to determine the contaminant concentration at which acute effects are likely to occur. The RAL should be below that concentration to ensure protection against acute effects. If acute toxicity data are insufficient to either determine whether the RAL is protective for acute effects or to establish an alternative protective level, then the area average approach should not be used.

In general, the method used to determine the RAL should be compatible with the method used to determine the EPC. To the extent that the RAL is designed to ensure that the average post-remediation concentration (i.e., post-remediation EPC) is at or below the cleanup level, calculation of the post-remediation EPC should be consistent with calculation of the pre-remediation

EPC, in terms of the averaging basis, EU size, and contaminant distribution. However, in some instances the methods may be different. For example, the distribution of the contaminant concentration may be different pre- and post-remediation, or the EU size may be different if the land use is changing.

When deciding to apply the cleanup level as a not-to-exceed level or as an area average, a variety of factors come into play, including:

- ***Exposure.*** Using the area average approach is based on the assumption of random exposure. If exposure is not random across the EU, but rather receptors spend more time in areas of high concentration, then remediating soils such that the average post-remediation concentration achieves the cleanup level may not be protective of the receptor with non-random exposure. The exposure assumptions made in determining which approach to use should be consistent with the exposure assumptions used in the risk assessment and the determination of the EPC.

- ***Size of exposure unit and adequacy of sampling.*** If the size of the exposure unit is different pre- and post-remediation due to change in the land use, the sampling may not be appropriate to support implementation of either approach. For example, if the sampling was designed for large, non-residential, pre-remediation exposure units, and future residential land use requires smaller exposure units, the sampling may not be extensive enough to adequately characterize exposures within the smaller unit. In this case, it may be necessary to do more extensive sampling so that the sample accurately represents the post-remediation exposure.

- ***Toxicity.*** If the cleanup level is based on acute exposure, it should be implemented as a not-to-exceed level, because any short-term exposure exceeding the cleanup level could cause adverse health effects. Even if the cleanup level is based on chronic exposure, care should be taken in implementing it as an area average based on an RAL, because it is important to ensure that the RAL (which may be significantly higher than the cleanup level) is protective of acute effects. For example, a single instance of a child ingesting a handful of soil containing malathion could cause acutely toxic effects in the child. Therefore the RAL must be low enough to be protective against acute effects of single and short-term exposures, as well as chronic effects of long-term exposure. At present EPA does not have acute toxicity criteria, therefore consultation with a toxicologist may be necessary to determine if the RAL is sufficiently protective for acute effects. Risk managers can seek assistance in identifying appropriate toxicity values from EPA's regional toxicologists and risk assessors. In addition, the Agency for Toxic Substances and Disease Registry (ATSDR) publishes Minimal Risk Levels (MRLs) that may, in some cases, be suitable for use as acute toxicity values.[18] In the absence of sufficient acute toxicity data, the not-to-exceed approach should be used if acute exposures are likely.

- ***Confidence in the protectiveness of the cleanup level.*** Uncertainty about the protectiveness of the cleanup level may indicate that it is most appropriately implemented as a not-to-exceed level. If the cleanup level is the risk-based PRG calculated as the 95 percent UCL of the average post-remediation contaminant concentration, there is less than a five percent chance that average exposure at that level will pose significant risk. Therefore in instances where there is adequate data coverage and exposure units are well defined an area average approach may be appropriate. However, if the cleanup level is less "conservative" because it is not based on an average post-remediation EPC, or is significantly higher than the PRG because of practical or technical considerations (e.g., cost, implementability), then it may not be protective if implemented as an area average.

- ***ARARs.*** Federal and state requirements for cleanup must also be considered to ensure compliance. If the cleanup level is an ARAR that the state or tribe designates as a not-to-exceed level, or an ARAR that was developed based on factors other than average risk, then it may not be appropriate to implement the cleanup level as an area average.

- ***Quality and quantity of site characterization data.*** Confidence in the degree to which contaminant concentration data accurately represent soils that the receptor contacts over time may influence the decision to use the area average approach. The discussion on page 5 and a report prepared by The Interstate Technology & Regulatory Council Sampling, Characterization and Monitoring Team[19] emphasize the importance of adequate sampling. Site managers should also refer to *Data Quality Objectives Process for Hazardous Waste Site Investigations* for guidance on what constitutes high quality site characterization data.[20] In addition the *Guidance for Data Usability in Risk Assessment (Part A and Part B)* may be helpful. Site characterization data should meet the Data Quality Objectives described in

---

[18] ATSDR MRLs were developed in response to a CERCLA mandate and represent the highest exposure levels that would not lead to the development of non-cancer health effects in humans based on acute (1-14 days), subchronic (15-364 days), and chronic (365 days and longer) exposures via oral and inhalation pathways. MRLs are based on non-cancer health effects only. MRLs are available from ATSDR's website, http://atsdr1.atsdr.cdc.gov:8080/mrls.html.

[19] Technical and Regulatory Guidance for the Triad Approach: A New Paradigm for Environmental Project Management, prepared by The Interstate Technology & Regulatory Council Sampling, Characterization and Monitoring Team, December 2003.

[20] U.S. EPA, *Data Quality Objectives Process for Hazardous Waste Site Investigations,* EPA QA/G-4HW, EPA/600/R-00/007, January 2000; U.S. EPA, *Guidance for Data Usability in Risk Assessment (Part A),* EPA 9285.7-09A, April 1992; and U.S. EPA, *Guidance for Data Usability in Risk Assessment (Part B),* EPA 9285.7-09B, May 1992.

the guidance. However, if site characterization or sampling data are insufficient to provide confidence in the use of the area average method, then the cleanup level should be implemented as a not-to-exceed level because it generally provides more certainty about the protectiveness of the cleanup. The area average approach is specifically intended for situations where adequate site characterization data are available. Applications of area average methods to sites with limited or incomplete data are inappropriate. However, if the quality of site characterization data is the only factor limiting the use of the area average approach, it may be more cost-effective to spend more on sampling to improve the quality of the data before deciding to implement the cleanup level as a not-to-exceed level where the area average approach could save on remediation costs.[21]

- ***Cost-effectiveness.*** In considering implementation of the cleanup level as an area average, it is important to identify the associated costs and cost savings. The area average approach is likely to be more cost-effective for removal remedies than for in situ treatment remedies. The marginal cost of additional excavation and disposal of soils with concentrations between the cleanup level and RAL usually will be relatively high compared to the marginal cost of additional treatment to reduce soil concentrations to the cleanup level. In addition, the area average approach may require more extensive site characterization (as described above) and statistical analysis. These additional costs may be offset by remediation cost savings due to the less extensive cleanup required. Risk managers should consider these trade-offs to determine the most cost-effective approach.

- ***Community acceptance.*** Community acceptance is an important consideration in using the area average approach. Because an area average cleanup will leave in place some soils with contaminant concentrations that are above the cleanup levels set forth in the ROD, the community may not be confident in the remedy's protectiveness. The community will likely be more comfortable knowing that all soils achieve the ROD-specified cleanup level.

- ***Statistical expertise.*** The area average approach requires statistical expertise. As discussed in Section 4.0, various methods requiring varying degrees of expertise are available. The iterative truncation method requires a basic understanding of area averaging, and most risk managers should be able to implement it without specialized expertise. The confidence response goal method is somewhat more sophisticated, requiring some statistical training to understand, but it is formulaic and can be automated. The geostatistical method requires highly specialized statistical expertise and usually will have

[21] Determining the point at which increased sampling offsets remediation cost savings will depend largely on the sampling costs and marginal remediation costs which are very site-specific.

significant resource implications. The availability of resources or statistical support should be considered in deciding on an appropriate approach and in project scoping.

Taking these and other appropriate factors into consideration, the risk manager should decide whether it is appropriate to implement the cleanup level as a not-to-exceed level or as an area average. In instances where there is adequate data coverage and the exposure units are well defined, an area average approach may be appropriate. Specific situations in which this approach would not be appropriate are when:

- Exposure within the EU is non-random;
- The cleanup level is based on acute exposure;
- The cleanup level is not derived from a UCL of average post-remediation concentrations (e.g., it considers factors other than risk);
- The quality of site characterization data is not optimal, and it is not worth investing in additional sampling;
- Given the site conditions (complexity, size, characterization, contaminant distribution), it is not cost-effective to do the necessary sampling and statistical analysis;
- The community most likely will not accept leaving soil onsite with contaminant concentrations that exceed the cleanup level.

4.0 METHODS FOR IMPLEMENTING CLEANUP LEVELS AS AREA AVERAGES

Site remediation typically involves making cleanup decisions for large volumes of contaminated soil based on data from many, very small soil samples. Such decisions are made with the help of statistical methods. Simple, classical statistical approaches assume that there is no spatial correlation between contaminant concentrations within an EU, and they rely on the mean and standard deviation of the sample data to accurately represent the exposure concentration in the EU. More sophisticated, geostatistical approaches are designed to estimate the spatial distribution of contaminant concentrations in the EU based on sample data, and in this way explicitly address differences in scale between sample data and cleanup decisions. Both simple and geostatistical methods generally are available for calculating RALs when implementing cleanup levels as area averages. There is a growing recognition that geostatistical methods are generally more appropriate because contamination at most sites has spatially correlated distribution patterns and more than one contaminant population. Geostatistical methods are generally more rigorous and account for the relationship between contaminant concentrations and the size, location, and geography of the EU. Simpler, classical statistical approaches may be applicable in some situations where contaminant concentrations are not spatially structured (i.e., have a uniform distribution) or where areas of an exposure unit can be delineated into subsections within which contamination is likely to be homogeneous. The appropriate method depends in large part on the distribution of contaminants at the site.

Non-spatial techniques can be used when there is no spatial correlation between contaminant concentrations at the exposure unit level. A good example is a landfill site where waste was disposed across a large area, a process resulting in small randomly-located spots of high contaminant concentrations interspersed within areas of lower concentrations. On the other hand, spatial techniques, or geostatistics, normally are used when there is a spatial correlation between contaminant concentrations. For example, a site where the major source of contamination is a unlined liquid waste storage lagoon should have very high contaminant concentrations in and around the lagoon and lower concentrations at increasing distances from the lagoon. The contaminant distribution pattern at a site can be determined based on historical and physical information and spatial plots of site data. If the data show that contaminant concentrations are homogeneous across the exposure unit, or high contaminant concentrations occur randomly among samples with low concentrations, then statistical determination of RALs can be performed using simple, non-spatial techniques. If, on the other hand, sampling data reveal a structured pattern with highly contaminated areas surrounded by gradually decreasing concentrations, then statistical determination of RALs using spatial techniques may be more appropriate. If the exposure pattern is a mixture of isolated hot spots and areas with lower random concentrations, then spatial techniques or a combination of spatial and non-spatial techniques should be used.

Non-spatial techniques typically allow rapid calculation of RALs by assuming that contaminant concentrations are devoid of any spatial structure or correlation, and that the sampling is unbiased and accurately represents exposure concentrations. Non-spatial techniques are based on the mean and standard deviation of the sample contaminant concentration data and how those metrics change as soils with high contaminant concentrations are replaced with clean fill during remediation. Calculated RALs will differ for each EU depending on the mean, standard deviation, and distribution of contaminant concentration data. Even if two EUs have the same average concentration, the RALs will differ if the variances and distributions are different. Non-spatial statistical computations yield exposure concentrations that are independent of the size, location or geometry of the EU. If the sample soil concentrations display spatial structures or correlations, or if the samples do not accurately represent exposure or are collected in a biased way (e.g., over-sampling of areas thought to have high concentrations), then application of non-spatial statistical techniques should result in unreliable RALs. RPMs should avoid this pitfall and ensure successful implementation of these approaches by using thorough site characterization based on sufficient sample sizes and unbiased sampling.

Spatial techniques, on the other hand, use available contaminant concentration data to model the spatial correlation underlying the measured values, thus directly accounting for differences in the different scales of sample data and EUs in the cleanup decision process. EU grids are placed over a map of the area of interest and, using an averaging geostatistical technique called block kriging, average EU concentrations are determined for pre- and post-remediation conditions. Currently, block kriging is the only commonly used procedure that is capable of estimating area averages based on correlated point measurements. An alternative geographic approach is currently being developed to accommodate spatially correlated contamination and match the sample representativeness to the scale of the remediation decision. The approach involves dividing EUs into discrete remedial subsections that reflect different contaminant populations and within which contamination is homogeneous.[22] This general approach has been used at at least one site.[23] It requires extensive sampling and Monte Carlo simulation to adequately represent each remediation subsection, and its advantages over geostatistics are not well established. As is the case for non-spatial techniques, when using geostatistical techniques, RPMs should use good sample data and site characterization in order to apply the techniques successfully.[24] An added consideration, however, can be the relatively high analytic costs of applying geostatistical techniques. They usually require highly specialized statistical expertise, and are time- and resource-intensive. If this approach is a likely option, time and resource issues should be addressed in the project scoping.

---

[22] Evan Englund, U.S. EPA, National Exposure Research Laboratory, March 2005, personal communication.

[23] Ryti explored the relationship of the statistical properties of exposure concentrations and the EU geometry as part of the remedial design of the Piazza Road dioxin site (an EPA Superfund site in Missouri). Ryti divided EUs into remedial subunits, referred to as "cells." Using a Monte Carlo simulation approach, soil contaminant concentrations for each cell were estimated. During each simulation round, the cell with highest simulated concentration was selected. If the EU containing the selected cell had an average concentration in excess of the cleanup level, the cell was targeted for cleanup. This process was repeated until the cell with the next highest simulated concentration had a value less than the cleanup level. The spatial statistical characteristics of correlated/structured data demonstrated by Ryti are formalized in geostatistical block kriging techniques. Ryti, Randall T., *Superfund Soil Cleanup: Developing the Piazza Road Remedial Design*, Journal of Air and Waste Management, 43:197-202, February 1993.

[24] Composite samples cannot be used for these methods because they do not represent contaminant concentrations at specific locations.

In this section, we present three statistical methods that have been suggested for use in implementing cleanup levels as area averages -- iterative truncation, confidence response goal (CRG), and geostatistics. They reflect the spectrum of simple to complex. The two at the simple end of the spectrum, the iterative truncation and CRG methods, are non-spatial and have limited applicability. The most complex, geostatistical method, is spatial and more widely applicable because most sites exhibit spatial structure to contamination. All three methods focus on calculating RALs that should ensure that the average soil contaminant concentration within an EU after remediation is at or below the cleanup level. Exhibit 4 summarizes the pros and cons of the three approaches which are discussed in detail below.

Iterative Truncation Method

As discussed for purposes of this guidance, the iterative truncation method is a simplistic approach to calculating the RAL for surface soil cleanup. It is based on the identification and removal of soils with high contaminant concentrations to lower estimated post-remediation EPCs to levels at or below the cleanup levels. Iterative truncation is used for non-spatial data, it assumes that each sample is an uncorrelated, unbiased representation of a remediation area within the site or EU. As indicated, iterative truncation involves removing (truncating) high values in the sample concentration measurements and calculating a hypothetical post-remediation EPC. For this reason, it is inappropriate to use composite samples. Each iteration entails replacing the next highest value with the concentration of clean fill.[25] With each iteration, a new EPC is calculated and compared to the cleanup level. At the end of the process when the estimated post-remediation EPC is at or below the cleanup level, an RAL that will achieve that EPC is specified.

In practical terms, the steps involved in iterative truncation are as follows:

- Order the sampling data from lowest to highest concentration.
- Starting with the highest concentration, remove a sample and replace it with the post-remediation concentration (e.g., concentration found in clean fill).
- Recalculate the post-remediation EPC for the new data set and compare the resulting EPC to the cleanup level.
- If the EPC is higher than the cleanup level, repeat the process iteratively until the EPC is less than or equal to the cleanup level.

[25] To ensure that the concentration in clean fill is achieved during remediation, the concentration for clean fill should either be based on data from likely fill material or it should be tested during construction to ensure that it does not exceed the estimated value.

Exhibit 4

SUGGESTED STATISTICAL METHODS FOR IMPLEMENTING CLEANUP LEVELS AS AREA AVERAGES

NON-SPATIAL

Iterative Truncation Method

Pros:
- Simple; no statistical expertise needed.

Cons:
- Very sensitive to highest contaminant concentrations in the sample; if the highest sample concentrations are not representative of the highest concentrations at the site, the resulting RAL may not be protective.

Cautions:
- Inappropriate for use with composite data.
- Inappropriate for use with spatially correlated data.
- If sampling data are biased such that higher concentration areas are over-sampled, the resulting RAL will be unnecessarily low.

Confidence Response Goal Method

Pros:
- Less sensitive than iterative truncation to the representativeness of the highest sample concentrations.
- Accounts for different statistical distributions of contaminant concentration data.

Cons:
- May entail some statistical expertise.
- Difficult to communicate results to public, due to mathematical complexity.

Cautions:
- Inappropriate for use with composite data.
- Inappropriate for use with spatially correlated data.
- If sampling data are biased such that high concentration areas are over-sampled, the resulting RAL will be unnecessarily low.

SPATIAL

Geostatistical Method

Pros:
- Can be used with spatially correlated data.
- Can be used with biased sample data (e.g., over-sampling of hot spots).

Cons:
- May entail geostatistical expertise and specialized software.
- More costly and time consuming than non-spatial methods.

Cautions:
- Consider the value of the information gained from geostatistical approach to ensure that the anticipated benefits justify the costs.

- When the calculated post-remediation EPC in the data set is less than or equal to the cleanup level, the highest sample concentration remaining in the data set is designated the RAL.

During each iteration, removal and replacement of selected concentrations impacts the distribution of the remaining data. Such changes may require reevaluation of the distribution of the resulting data set at each iteration. These iterative reevaluations and EPC calculations can be performed efficiently by utilizing ProUCL.[26]

Schulz and Griffin applied the iterative truncation method to actual site data for a residential exposure unit at western mining site contaminated with arsenic. Applying iterative truncation to the data set of 36 soil samples resulted in an RAL which indicated remediation of three areas. If the cleanup goal were applied as a not-to-exceed level, 16 of the 36 areas would require remediation.[27]

Generally, the iterative truncation method will fail to produce an adequate RAL for cleanup if site characterization is incomplete. Because this method is solely reliant on the sampling data, it is sensitive to the highest contaminant concentrations in the sample. If the highest sample concentrations are not representative of the highest concentrations in the EU and there are actually areas with higher concentrations, then the resulting RAL may not be protective. To examine the sensitivity of the method to representative sampling, Schulz and Griffin applied the iterative truncation method to sampling data from two sites. They determined the RAL using the sampling data and then increased the two highest sample concentrations by 50 percent and recalculated the RAL. The recalculated RAL was sufficiently lower than that based on the actual sampling data, as to require remediation at an additional location.[28] Obviously, the more complete and representative the sampling data, particularly at the high concentrations, the more reliable this method.

Some sampling plans may be biased towards areas of known or suspected contamination. Applying the iterative truncation approach to such data sets will result in unnecessarily low RALs. The iterative truncation method uses the sample data to calculate the post-remediation EPC, as high concentration samples are removed. Because the low concentration areas are not adequately represented in biased data, applying the method would result in unrealistically high EPCs, as samples are iteratively removed. Therefore to get the average at or below the cleanup level would require an unnecessarily low RAL, and the advantages of area averaging would not be realized.

---

[26] Singh, A., A.K. Singh, R.W. Maichle, *ProUCL Version 3.0, User Guide,* EPA, April 2004. (http://www.epa.gov/nerlesd1/tsc/images/proucl3apr04.pdf)

[27] Schulz, T.W. and S. Griffin, Practical Methods for Meeting Remediation Goals at Hazardous Waste Sites, *Risk Analysis,* Vol. 21, No. 1, 2001.

[28] IBID.

The iterative truncation method is also sensitive to the variability of sample concentrations. If the variability is high and the number of samples small, the RAL resulting from this approach may be over-protective. Since the RAL is a sample concentration (i.e, the highest sample concentration next to the one that gives an EPC less than or equal to the cleanup level when replaced with the concentration of clean fill), gaps in the data may force the RAL lower than necessary to achieve the cleanup level. To use this method with confidence, it is important to have good site characterization based on extensive, unbiased, and representative sampling, and the resulting data should adequately represent random, long-term exposure to receptors.

Due to these limitations, it is important to consider the following criteria in determining whether iterative truncation is an appropriate method to use.

- ***Sample size is sufficient.*** Small sample sizes translate to large uncertainty in estimates of post-remediation EPCs and resulting RALs.

- ***Sampling design yields a representative distribution of measurements within the EU.*** Simple random sampling may fail to represent a patchy distribution of contaminants. Similarly, over-sampling high concentration areas may fail to represent random movement of receptors.

- ***Assumptions about post-remediation distribution of concentrations are reasonable.*** If these assumptions are shown to be incorrect by post-remediation sampling, the process for developing RALs may need to be repeated and additional remediation may be required.

Confidence Response Goal Method

Bowers, Shifrin and Murphy developed a method for calculating what they term the "confidence response goal" (CRG) which under this guidance is the same for all practical purposes as an RAL, that is, a not-to-exceed level for EU remediation that will ensure that the area average for the EU is at or below the cleanup level.[29] The basic premise of this method is that the CRG can be expressed as a function of the mean and standard deviation of contaminant concentration, and the cleanup level. The average post-remediation concentration is determined from the pre-remediation distribution that is truncated at the CRG, and a second superimposed distribution that represents the concentration of contaminant in clean fill. The average concentration of the post-remediation distribution is a weighted average of the portion of the pre-remediation distribution with concentrations below the CRG and the concentration of the clean fill which replaces all pre-remediation concentrations that exceed the CRG. The original method was based on the assumption that the sample data are derived from a spatially uncorrelated lognormal distribution and accurately represent the distribution of contamination in the EU. Subsequent investigations by

[29] Bowers, T.S., N.S. Shifrin, and B.L. Murphy, Statistical Approaches to Meeting Soil Cleanup Goals, *Environmental Science and Technology,* 30, 1437-1444, 1996.

Singh and Singh indicated that although CRG method is a reasonable procedure, its application based on the assumption of lognormality of contaminant concentrations can result in RALs that are not adequately protective. Specifically, Singh and Singh simulated sampling from a lognormal distribution of contamination and found that the process yields very high RALs such that no cleanup is recommended, even when the true average concentration exceeds the cleanup level.[23]

Schultz *et al.* demonstrated that the CRG method yields reliable results if certain other types of distributions are used. For this purpose, they developed equations for computing the CRG (c^*) for normal, exponential, and uniform continuous distributions. The equations for normal and exponential distributions are presented below.[24]

Normal distribution:

$$CUG = c_o + (\mu - c_o)\Phi(\frac{c^* - \mu}{\sigma}) - \sigma\phi(\frac{c^* - \mu}{\sigma})$$

where,

CUG	=	cleanup goal (same as the cleanup level)
c_o	=	contaminant concentration in clean fill
c^*	=	CRG (same as RAL)
μ	=	mean contaminant concentration
σ	=	standard deviation of contaminant concentrations
$\Phi(.)$	=	standard normal cumulative density function
$\phi(.)$	=	standard normal density function

Exponential distribution:

$$CUG = c_o e^{-\frac{c^*}{\theta}} - c^* e^{-\frac{c^*}{\theta}} - \theta e^{-\frac{c^*}{\theta}} + \theta$$

where,

CUG	=	cleanup goal (same as the cleanup level)
c_o	=	contaminant concentration in clean fill
c^*	=	CRG (same as RAL)
θ	=	mean contaminant concentration used in the exponential model

[23] Singh, A.K., and A. Singh, Review of the Confidence Removal Goal Approach for Site Cleanup, Technology Support Center, U.S. EPA, Las Vegas, Nevada, 1996.

[24] Schultz, B., A.K. Singh, A. Singh, An Evaluation of the Confidence Response Goal Approach for Making Remediation Decisions at Superfund Sites, *Environmetrics,* 13 725-732, 2002.

These equations can be used for determining the CRG (c*). Site assessors can readily solve these equations in a spreadsheet calculation or compiler language. Schultz *et al.* also developed a CRG equation for the uniform continuous distribution. A similar CRG equation can be developed for the gamma distribution which has proved to be a useful model for skewed data sets.[25]

While the potential remediation cost savings associated with the use of the CRG method are compelling, risk managers should be aware of the need for statistical expertise in applying the method correctly and the difficulty in communicating the results to the public and gaining community acceptance. Communities typically are concerned that the RAL be protective, and may find the explanation of the CRG method difficult to understand. The mathematical complexity of the CRG method may be a detriment.

The difficulty of community acceptance is illustrated by a site that was planned to be redeveloped as an urban park. In this case, there were zoning restrictions against industrial and commercial use, and the property was zoned for residential use. However, the future use of the proposed park was not specified (e.g., passive green space, a botanical garden, athletic fields). The RAL determined using the CRG method suggested considerable savings in cleanup costs compared to implementing the cleanup level as a not-to-exceed value. Citizens favoring the park's use as athletic fields turned out in force at the proposed plan public meeting and voiced their objections to the remedial action level calculated using the CRG method. These citizens were opposed to any implementation of the cleanup level as an area average because receptors could potentially have intense direct contact with the soil over a small exposure unit (e.g., football field). The citizens were very distrustful of the scientists' explanation of the highly statistical CRG method. The lack of community acceptance of the CRG approach forced the remedial project manager to implement the PRG as a not-to-exceed level.

This example illustrates the importance of making sure that the implementation of the cleanup level is appropriate to the potential exposures, and that the community is educated from the outset regarding plans for developing cleanup levels and implementing them as area averages. EPA's guide to Superfund community involvement provides suggestions for presenting technical material to communities.[26]

Geostatistical Methods

While non-spatial methods assume that the contaminant concentrations within an EU are uncorrelated, at many sites contaminant concentrations reveal clear spatial patterns, where highly impacted zones are surrounded by marginally impacted areas with gradually decreasing contaminant concentrations. Geostatistical techniques are statistical procedures designed to process spatially

[25] Singh, A., A.K. Singh, and R.J. Iaci, *Estimation of the Exposure Point Concentration Term Using a Gamma Distribution*, Technology Support Center Issue, EPA/600/R-02/084, EPA Technology Support Center for Monitoring and Site Characterization, Las Vegas, Nevada, October 2002.

[26] U.S. EPA, *Superfund Community Involvement Handbook*, EPA 540-K-01-003, April 2002. http://www.epa.gov/Superfund/action/communityinvolvement.htm.

correlated data. The presence of spatially correlated data is quite common at hazardous waste sites because of structured patterns in the distribution of contamination. For example, sites impacted by migration of contaminants from a concentrated localized source, such as an unlined lagoon for liquid waste storage, might exhibit spatial patterns of contamination. Contaminant concentrations in and around the lagoon may be higher than those at greater distances from the lagoon. Spatial correlation also arises because sampling data are sometimes collected in a biased fashion and may be clustered around hot spots of contamination. In such areas, soil samples can be influenced by the same phenomena and are therefore not independent. This type of spatially correlated data is suitable for geostatistical analyses. While geostatistics can accommodate biased data, excessive bias can provide misleading results. For example, geostatistical interpolation would over-estimate the extent of the impacted area in a situation where a highly sampled hot spot is surrounded by unsampled non-impacted areas.

Exhibit 5 displays an example of a structured pattern of contamination. At this site, soil concentrations display a well defined spatial distribution consistent with the release pattern of the original site contaminants. Under such conditions, cleanup computations can be performed using spatial geostatistical techniques that take into account the size and location of the targeted EU.

The computational steps involved in determining RALs for soil cleanup using geostatistical techniques are outlined below. These steps, as depicted in Exhibit 6, are computationally intensive and require the use of specialized software. (Appendix A provides an overview of available software.)

- ***Step 1: Iso-concentration Mapping.*** This step involves an initial analysis of available contaminant concentration data to determine whether a transformation of data is warranted, whether segregation of site data into separate populations should be pursued, or whether anomalous values should be eliminated.[27] After any necessary manipulation of the data, the resulting data set is subject to variogram analysis.[28] This is a process for computing and modeling the spatial correlation underlying the measured values. A function is fit to the sample variograms to determine the variogram model. This involves examining all pairs of measured values. For each pair, the difference between the measured values, the angular direction between them, and their separation distance are computed. Pairs along a selected direction with similar separation distances are grouped, and for each group, the sample variogram is computed as one-half the mean of the pairs' squared differences. In order to ensure the reliability of the selected variogram model, the

[27] Use of log transformed data in block kriging normally would yield results that do not correspond to the block averages. Therefore, log transformation of data is not recommended unless the geostatistical software used is capable of producing the correct back-transformed block averages.

[28] Variogram analysis is an assessment of spatial correlation. Englund and Sparks (1988) define the variogram as a plot of the variance (one-half the mean squared difference) of paired sample measurements as a function of the distances (and optionally of the direction) between samples.

computation process should be conducted consistent with the applicable ASTM standard guides listed in Appendix A. Iso-concentrations maps based on the variogram model are then generated to depict the extent of contamination.[29]

- ***Step 2: EPC Computation.*** At this stage, a grid of EUs is superimposed on the site map and average contaminant concentrations (EPCs) are estimated for each EU, as depicted in Exhibit 6. The geostatistical EPC computational process is called block kriging and is described in detail in Appendix B. For sites with well defined hot spots, care must be taken to position the EU grids such that individual EUs cover hot spots without undue mixing with surrounding, less-impacted areas. Those EUs where EPCs are higher than cleanup levels are referred to as the exceeding EUs.

- ***Step 3: RAL Determination.*** In this step, the iso-concentration map is used to identify zones that must be remediated in order to reduce the EPCs of the exceeding EUs to levels less than or equal to cleanup levels. This is an iterative process. During each iteration, points with estimated values in excess of a given cutoff concentration are replaced with concentrations equal to post-remedial conditions. Upon replacement, EPCs of exceeding EUs are re-computed. If there are still exceeding EUs, the cutoff concentration is decreased iteratively until full compliance is attained. The final cutoff concentration is defined as the RAL.

- ***Step 4: Cleanup Extent.*** Once the RAL is determined, the original iso-concentration map is used to define zones with concentrations in excess of the RAL. The contoured zone is the area that requires remediation. In many instances, existing data can also be used as part of the post-remediation confirmatory analyses.

It is important to note that geostatistical techniques are not a substitute for collecting sample data; the reliability of the results depends on adequate sampling data. The adequacy of the data can be evaluated by examining the kriging standard deviations for each estimate. Insufficient sampling is particularly evident when the kriged values are low and their standards deviations are high. In such cases, additional sampling generally will be necessary.

[29] In case of a "not-to-exceed" cleanup criterion, the iso-concentration map can be used to delineate the extent of the area requiring remediation.

Exhibit 5

EXAMPLE OF STRUCTURED SOIL CONTAMINATION

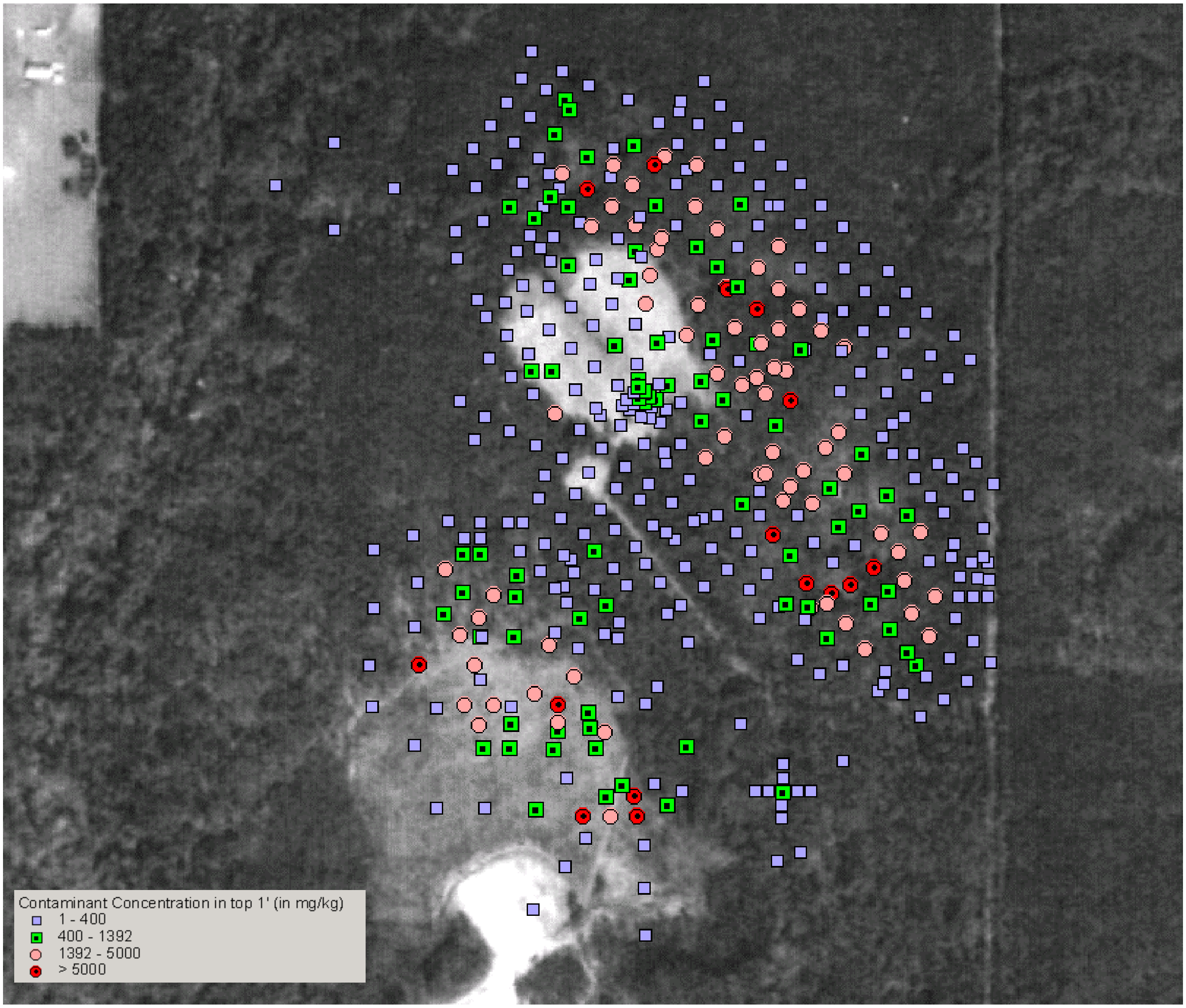


Exhibit 6

SAMPLE COMPUTATIONAL STEPS FOR CLEANUP OF STRUCTURED SOIL CONTAMINATION

Step 1. ISO-concentration Mapping

- Exploratory Data Analysis
- Variogram Analysis
- Point Kriging/Contouring

Step 2. EPC Computation

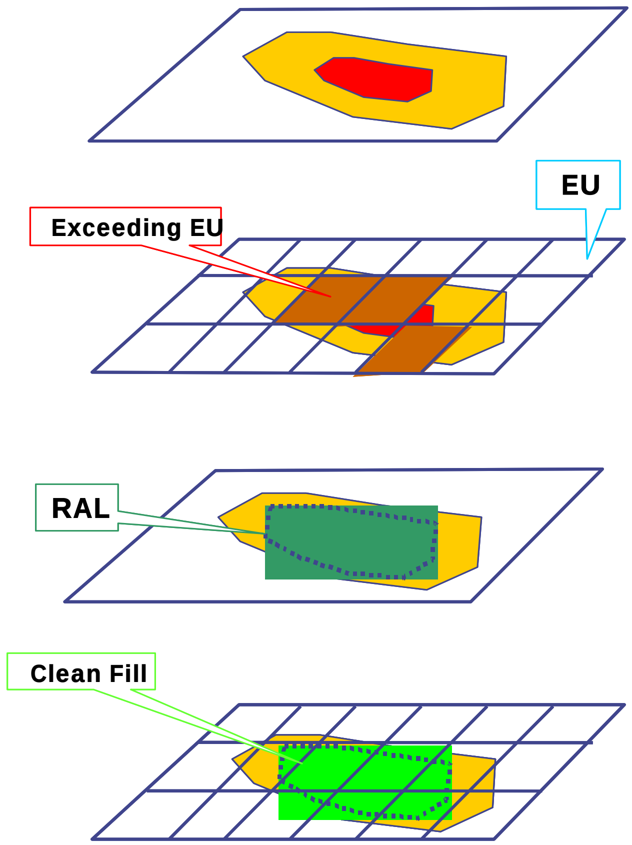
- Define Exposure Unit Grid
- Block Kriging to Compute EPC for each EU block
- Identify Exceeding EU blocks

Step 3. Determining RAL

- RAL = concentration that, if removed, would bring all EU blocks into compliance

Step 4. EPC Confirmation

- Confirmatory sampling and/or EPC Recomputation
- All EU blocks in compliance



Despite the broad applicability of geostatistical techniques, their computationally-intensive, time-consuming nature may limit their practical use. In general, geostatistical techniques are not recommended for cases where the following conditions exist:

- Site historical and physical information, or pre-remediation plots of contaminant concentrations indicate the presence of a random, non-spatial contamination pattern;
- Use of non-spatial statistical methods that ignore the spatial correlation and biased nature of data would not lead to overly-conservative, cost-prohibitive cleanups;
- Adequate technical/computational resources are not available; or
- The anticipated benefits from geostatistical analysis do not justify the costs.

Example Application

To illustrate the potential application of geostatistical methods to determine and apply RALs, we present a case study of a site with a structured pattern of soil contamination (previously shown in Exhibit 5).

- ***Step 1: Iso-concentration Mapping.*** Exhibit 7 displays the cumulative probability distribution of soil concentrations of the site contaminant of concern. This distribution exhibits multiple population characteristics due to the presence of a cluster of elevated samples. The variograms for the complete data set, as well as for the segregated data set without clustered measurements above 5,000 ppm, are shown in Exhibit 8. As shown in Exhibit 8A, the inclusion of the elevated measurements masks the true underlying correlation of the site data. Therefore, the elevated cluster is segregated in the variogram analysis in order to clearly decipher and model the underlying spatial correlation (Exhibit 8B). The variogram model is used to estimate contaminant concentrations at locations throughout the EU. These estimated concentrations are computed as weighted averages of nearby measured values. This estimation process is directly dependent on the variogram model, which is used to attain minimum variance and eliminate bias. The resulting values, together with data on elevated cluster, are used to generate the iso-concentration map, as displayed in Exhibit 9.

Exhibit 7

SAMPLE PROBABILITY PLOT OF SOIL CONTAMINANT OF CONCERN

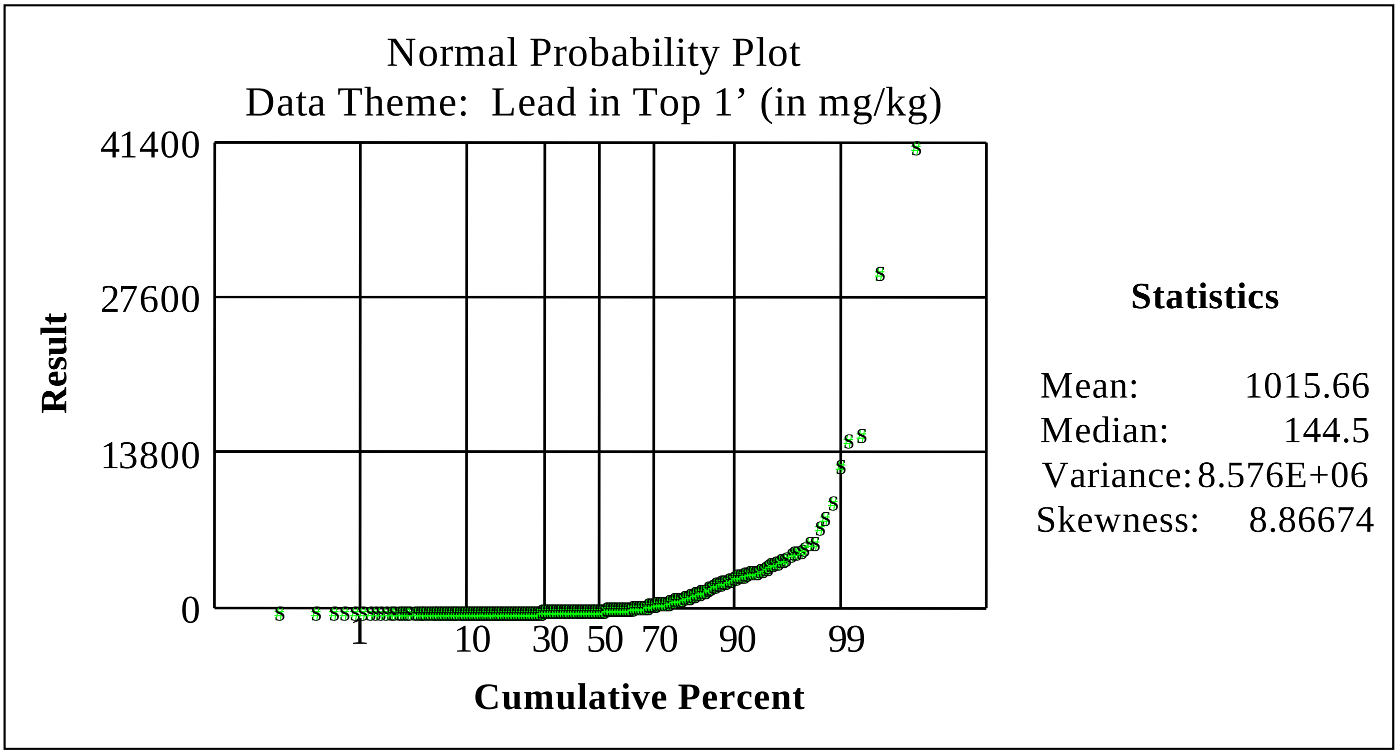
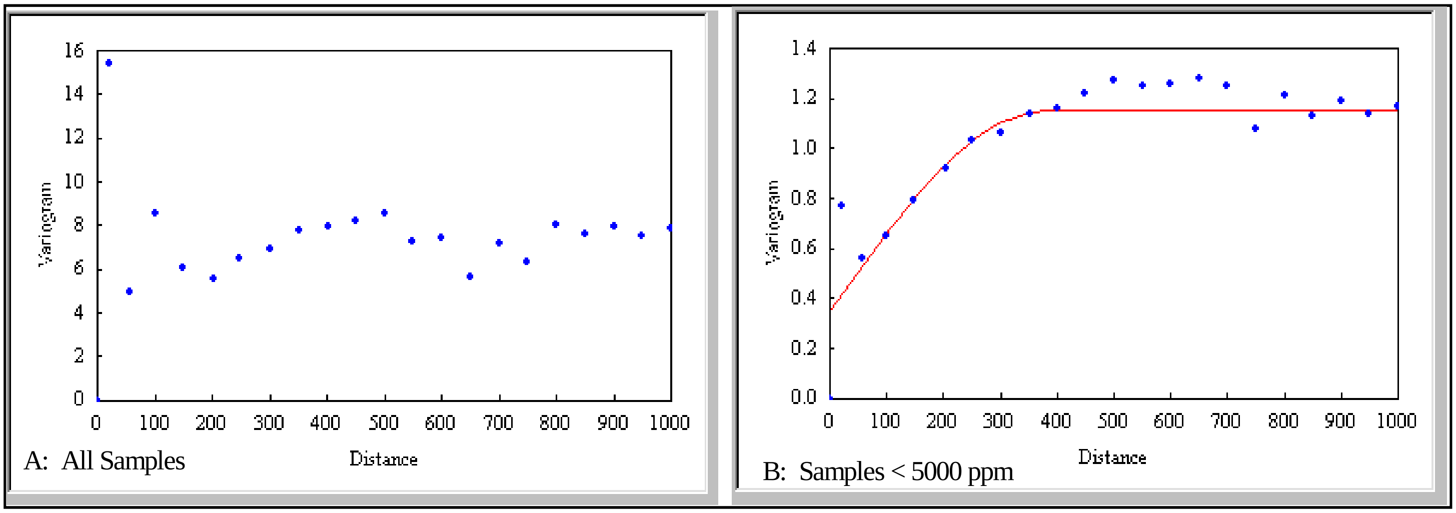


Exhibit 8

SAMPLE VARIOGRAMS OF SOIL CONTAMINANT OF CONCERN



- ***Step 2: EPC Computation.*** Using the available soil contaminant concentration data, the selected model variogram, and the cleanup level, exceeding EU cells are identified and highlighted in Exhibit 10. This example assumes a residential scenario, where the EPC is computed as average contaminant concentrations over half acre EUs. For purposes of this example, the cleanup level is 400 ppm. Exhibit 8 displays the EU grid, showing exceeding EUs with bold (red) outlines.

- ***Step 3: RAL Determination.*** The RAL is determined iteratively. In these iterations, the post-remedial concentrations are defined as having a 5 ppm concentration with a 1 ppm standard deviation, which is assumed to be the background concentration of local soil for the investigated contaminant of concern. The RAL is calculated to be 750 ppm.

- ***Step 4: Cleanup Extent.*** Upon determination of the RAL, the extent of zones that require remediation for each scenario is depicted using the iso-concentration map of the contaminant of concern (Exhibit 11).

Exhibit 9

SAMPLE ISO-CONCENTRATION MAP OF SOIL CONTAMINANT OF CONCERN

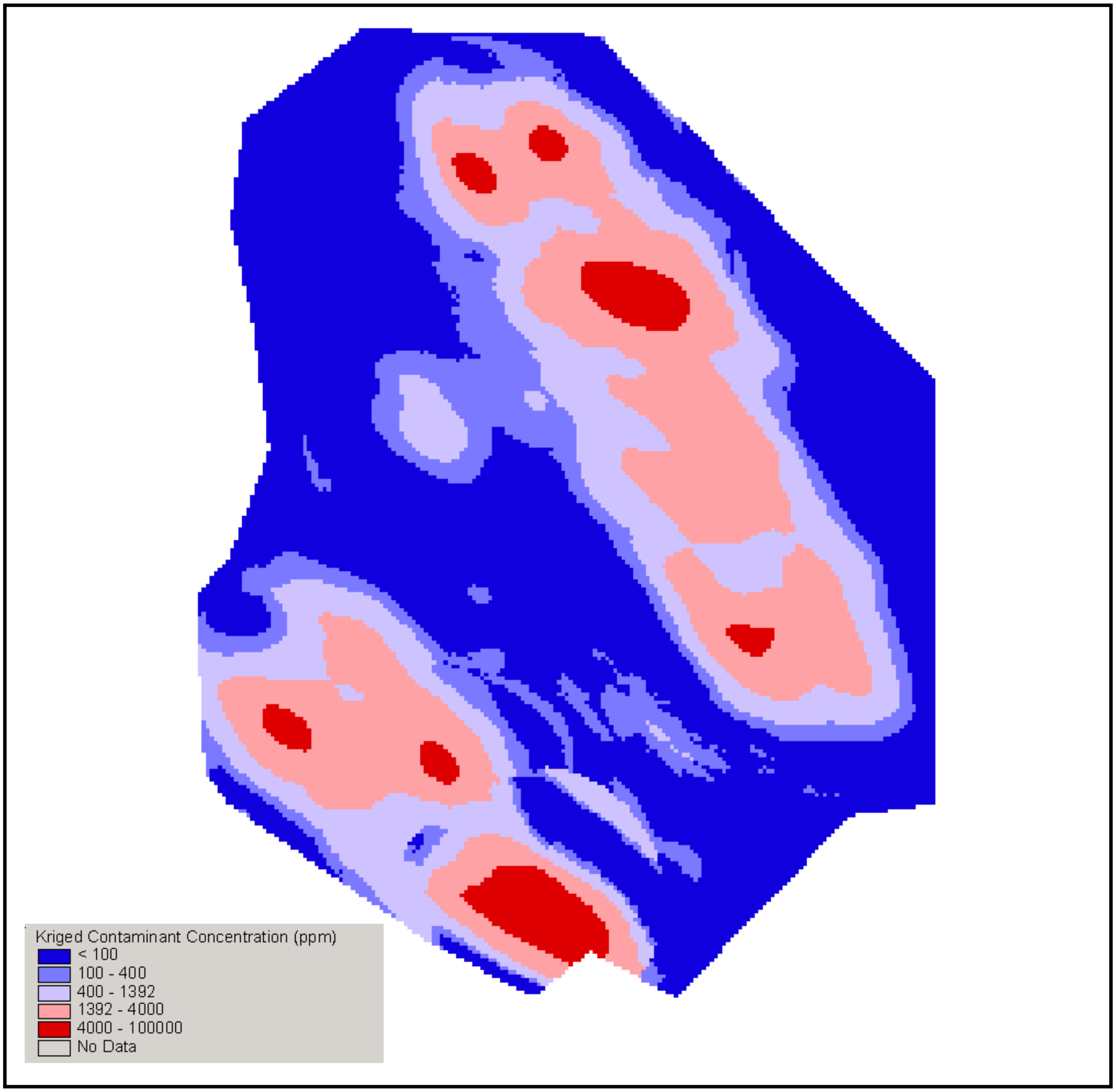


Exhibit 10

SAMPLE EXPOSURE UNIT GRID

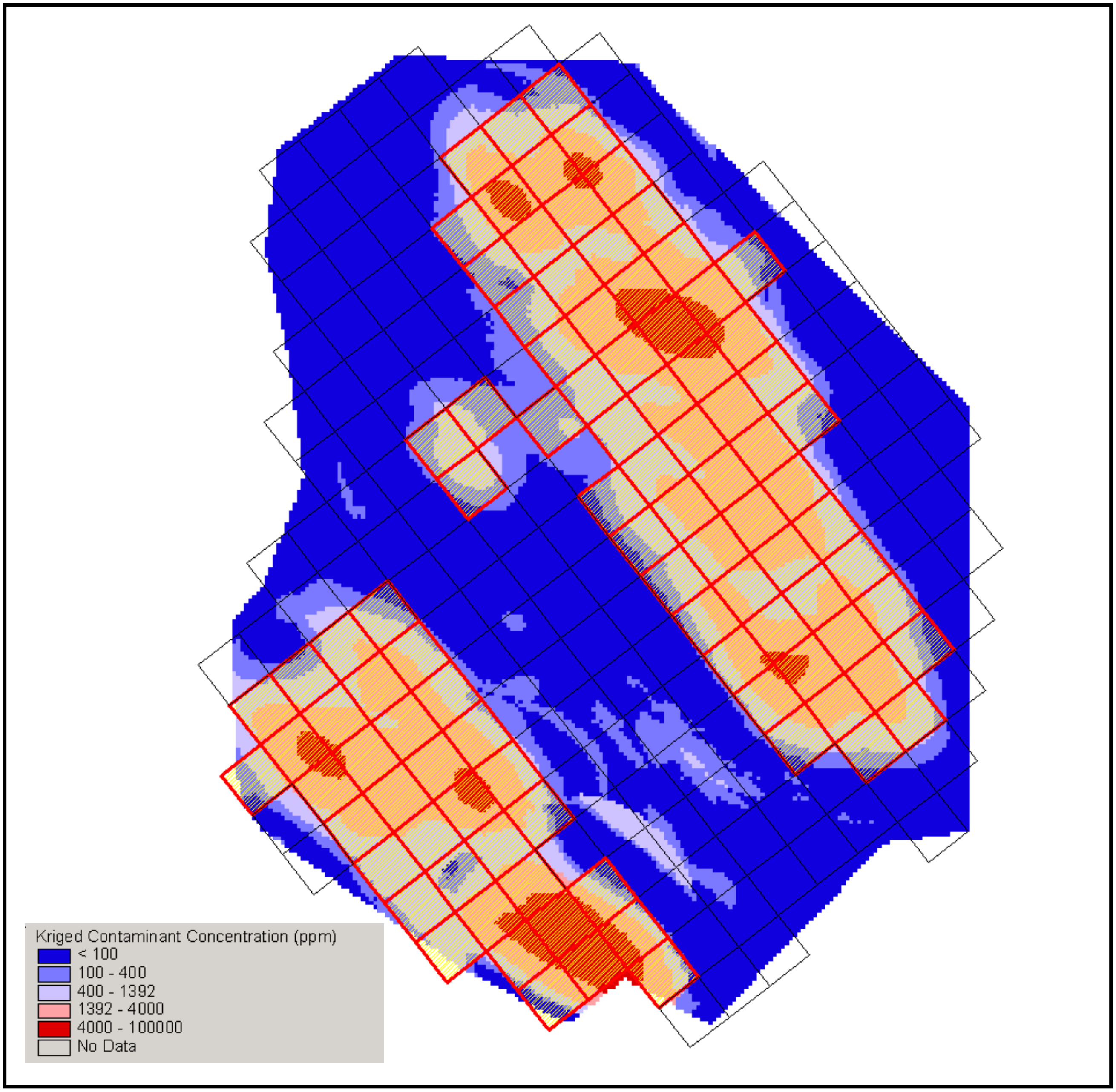
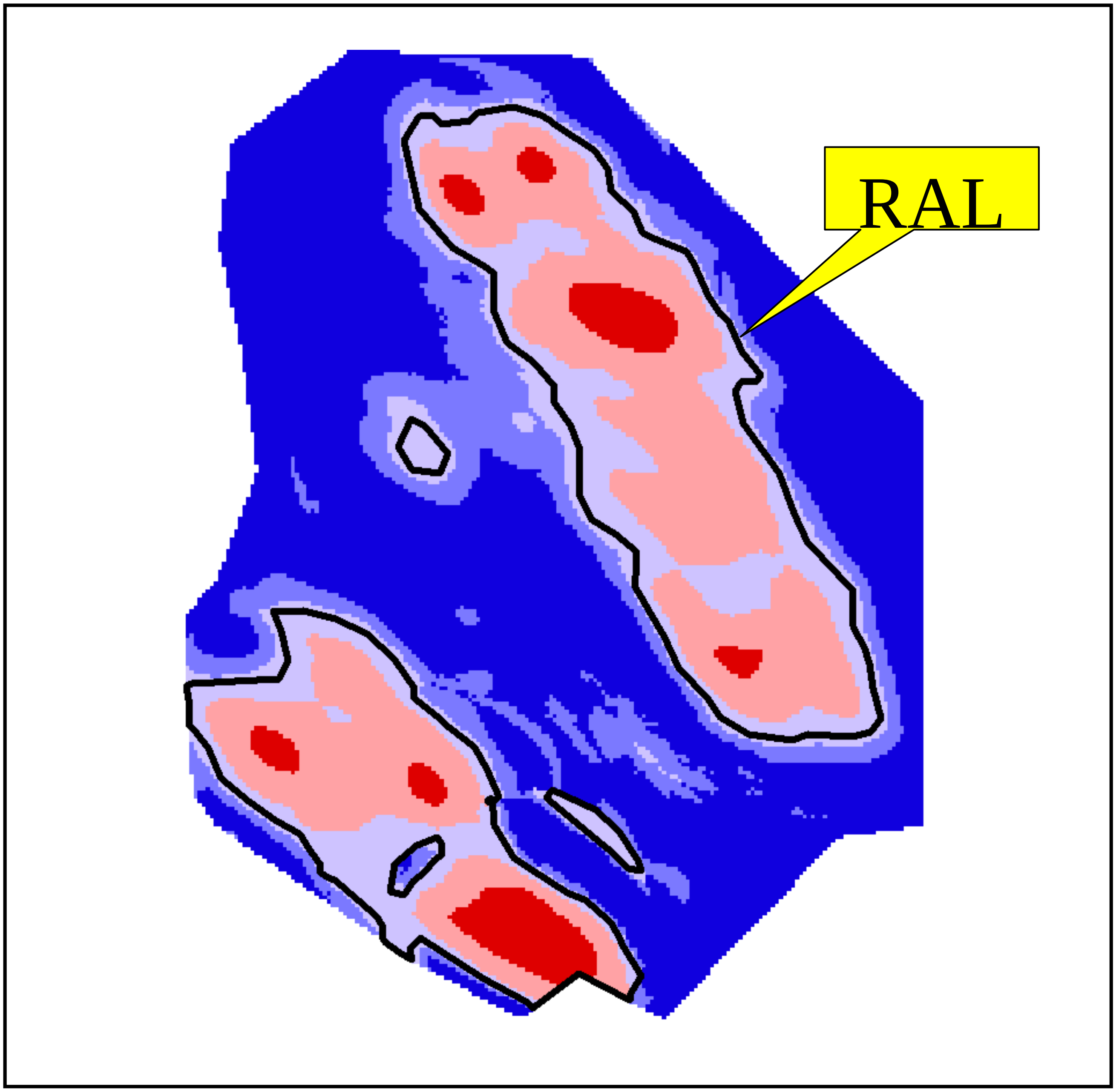


Exhibit 11

SAMPLE RAL DELINEATION FOR SOIL CONTAMINANT OF CONCERN



5.0 APPENDIX A: GEOSTATISTICAL REFERENCE SOURCES AND SOFTWARE

For detailed explanation of geostatistical techniques and applications, readers are referred to Matheron (1971); Journel and Huijbregts (1978), Isaaks and Srivastava (1989); and ASCE (1990a and 1990b). The American Society of Testing and Material (ASTM) provides a series of standard guides to perform geostatistical analyses. These guides are:

- ASTM, Standard Guide for Content of Geostatistical Site Investigations, D5549-94, 1994.
- ASTM, Standard Guide for Analysis of Spatial Variations in Geostatistical Site Investigations, D5522-96, 1996.
- ASTM, Standard Guide for Selection of Kriging Methods in Geostatistical Site Investigations, D5523-96, 1996.
- ASTM, Standard Guide for Selection for Simulation Approaches in Geostatistical Site Investigations, D5524-96, 1994.

As noted, the use of geostatistics requires specially designed software. EPA has taken the lead in promotion of geostatistics by producing the first public-domain software package, known as GEO-EAS (Geostatistical Environment Assessment Software), developed by Englund and Sparks (Englund and Sparks 1988). This package was followed by another EPA package, known as GEOPACK, developed by Yates and Yates (EPA 1990). The successful results of application of GEO-EAS prompted the EPA to recommend its use in spatial environmental data analysis (EPA 1989 and 1990). EPA Region V FIELDS (FIeld EnvironmentaL Decision Support) Team has also developed software for processing and analysis of spatial data that can be used in conjunction with geostatistical software.[30]

Many public domain and commercial software packages are also available. Among them is GSLIB (Deutsch and Journel 1992) which is public domain software containing a comprehensive library of geostatistical Fortran codes. Commercial software, such ISATIS (Geovariances 2000) and ArcGIS Geostatistical Analyst (ESRI 2002), provides comprehensive tools to perform geostatistical analysis.

[30] Http://www.epa.gov/region5fields/htm/software.htm.

6.0 APPENDIX B: SAMPLE GEOSTATISTICAL PROCEDURES FOR COMPUTING EPC

[Consider adding this to the UCL Guidance or perhaps more appropriately to RAGS Appendix C, Using Geostatistics to Represent the Concentration Term for Probabilistic Risk Assessment (EPC Guidance).]

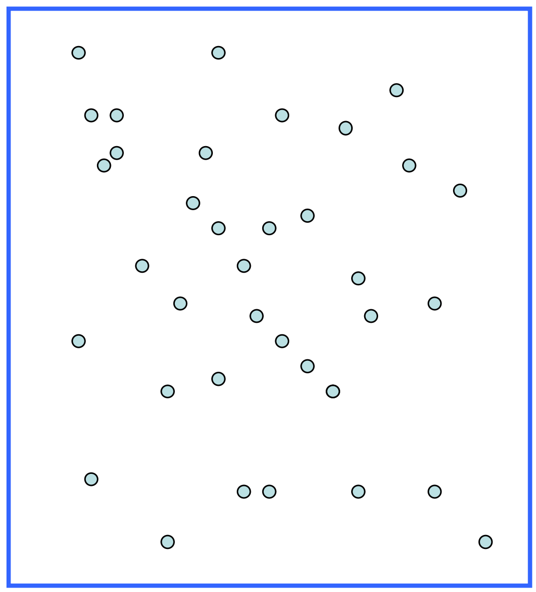
Exposure point concentration (EPC) is one of the key variables in estimating exposure in risk calculations. For purposes of this guidance, the EPC is not a point value but rather an average value for an exposure unit (EU). In cases where data are collected randomly and are not spatially correlated, the EPC is usually estimated as the upper confidence level (UCL) of the arithmetic mean concentration. However, due to the typical bias of field data toward impacted zones, the collected data can be biased, clustered, and correlated. Under such conditions, the use of non-spatial statistical techniques could lead to overly-conservative (i.e., elevated) EPCs. In such cases, geostatistics can be used to compute the EPC in order to avoid overly-conservative results.[31]

Approach

Computation of the EPC for a specified EU can be conducted by using the geostatistical estimation process known as block kriging.[32] One of the available approaches for this process is demonstrated below.

Step 1. Determine the spatial correlation between any two points (i.e., the variogram γ i, j) based on the available data.

Step 2. Define the extent of the Exposure Unit (EU) and identify its nearby measured values (z_1, z_2 , ..., z_n).



Available Site Data

[31] Excessive bias resulting from extensive hot spot sampling and no sampling in surrounding unimpacted areas, however, can provide misleading results using geostatics. In such cases, geostatistical interpolation would over-estimate the impacted areas.

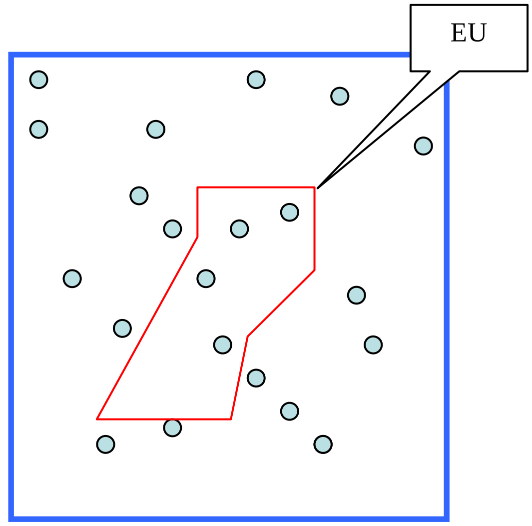
[32] Englund and Sparks (1988) defines kriging as a weighted-moving-average interpolation method where the set of weights assigned to samples minimizes the estimation variance, which is computed as a function of the variogram model, the locations of the samples relative to each other, and the point or block being estimated. Each estimated value is accompanied by its estimation (kriging) standard deviation.

Step 3. Consider the EU concentration as a linear combination of its surrounding point values:

$$Z_{EU} = \sum_{i=1}^{n} W_i Z_i$$

where:

Z_{EU}	=	concentration over EU
Z_i	=	random variable representing concentration at point i
W_i	=	unknown weight of the Z_i



Nearby Measured Values

Step 4. To minimize the estimation variance of Z_{EU} (i.e., the most accurate linear estimate of Z_{EU}), subject to the unbiasedness criterion (i.e., no systematic error), solve the kriging system below to determine the unknown estimation weights (w_1, w_2,…,w_n).

$$\begin{bmatrix} \gamma_{1,1} & \gamma_{1,2} & . & . & \gamma_{1,n} & 1 \\ \gamma_{2,1} & \gamma_{2,2} & . & . & \gamma_{2,n} & 1 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \gamma_{n,1} & \gamma_{n,2} & . & . & \gamma_{n,n} & 1 \\ 1 & 1 & . & . & 1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ . \\ . \\ w_n \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma_{1,EU} \\ \gamma_{2,EU} \\ . \\ . \\ \gamma_{n,EU} \\ 1 \end{bmatrix}$$

where:

$\gamma_{i,j}$	=	variogram between measured values z_i and z_j
$\gamma_{i,EU}$	=	variogram between the measured value z_i and the block representing the exposure unit EU
w_i	=	estimation weight of the measured value z_i
μ	=	Lagrange multiplier

Step 5. Compute the block-average concentration over the EU (Z^*_{EU}) and its measure of accuracy (σ_{EU}):

$$Z^*_{EU} = \sum_{i=1}^{n} w_i z_i$$

$$\sigma_{EU} = 2\sum_{i=1}^{n} w_i \gamma_{i,EU} - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \gamma_{i,j} + \gamma_{EU,EU}$$

where:

Z^*_{EU}	=	estimated average concentration over EU
σ_{EU}	=	kriging standard deviation of Z^*_{EU}

Step 6. Compute EPC as the Upper Confidence Limit (UCL_{EU}) of average concentration over EU (95% UCL shown below)[33]:

$$EPC = UCL_{EU} = Z^*_{EU} + 1.64\sigma_{EU}$$

The results of block kriging (i.e., the block-average concentration and its standard deviation) can be used to calculate the EPC as the upper confidence limit of the block average concentration based on a pre-defined tolerance level (e.g., 95%). This is a computationally intensive process that requires the use of specialized software, such as EPA's GEO-EAS program (Englund and Sparks 1988). Detailed descriptions of block kriging are provided in textbooks, such as Journel and Huijbregts (1978) and Isaaks and Srivastava (1989). Standard Guide for Selection of Kriging Methods in Geostatistical Site Investigations, produced by the American Society for Testing and Materials (ASTM), provides further guidance on selecting and conducting the appropriate kriging procedure.

Example Application

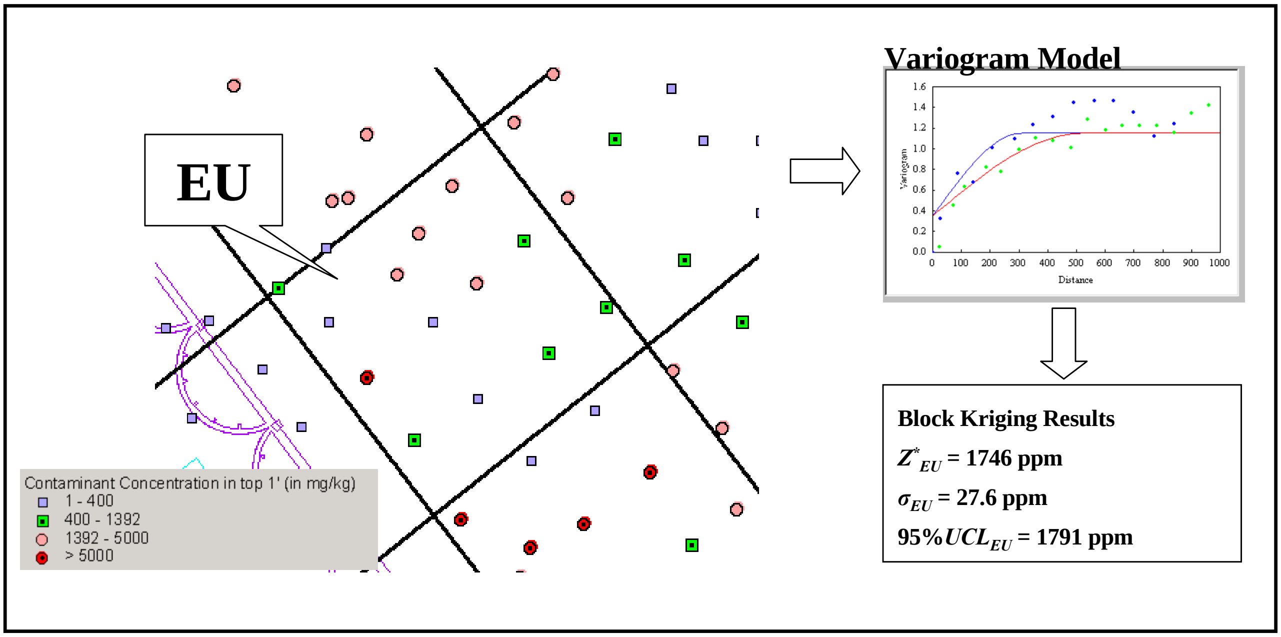
Exhibit B-1 illustrates a case study in which the EPC over a block, representing a 2.4-acre EU is computed. This EU represents the exposure domain of an ecological receptor to soil concentrations of a contaminant of concern. This data set displays a high degree of spatial correlation due to the nature of past releases at the site. In the block kriging process, measured concentrations of the investigated contaminant within and outside of the EU are used. This is accomplished by computing estimation weights that incorporate the location of sampling points relative to each other, as well as relative to the EU. In general, data points that are isolated and close

[33] In block kriging, the area mean is calculated as the average of all point values that constitute the block (e.g., EU). Based on the Central Limit Theory, such an average has a strong tendency toward normal distribution, regardless of the distribution of the underlying point values. Therefore, the UCL is calculated in accordance with the normality assumption of the area mean value, using the block average and its corresponding estimation standard deviation.

to the EU receive higher estimation weights, while clustered data far from the EU are assigned lower weights.

Exhibit B-1

EXAMPLE OF COMPUTING EPC USING BLOCK KRIGING



Key Potential Method Advantages

Computation of the EPC through block kriging may offer the following advantages.

- The computational process can accommodate biased, clustered and correlated data sets. In fact, block kriging is the only commonly used process that can estimate area averages based on correlated point measurements.
- The correlation of data automatically governs the estimation process (i.e., estimation is performed based on the modeled variogram).
- The location, shape and size of the EU are directly incorporated in the estimation process.
- The resulting EPCs are specifically computed for each EU.

Generally, the advantages of block kriging are mainly achieved through computationally-intensive and time-consuming processes that require specialized software and training. However, if results of simpler methods are deemed acceptable (i.e., they do not produce unreasonably high estimates of EPCs for individual EUs at a site), then more complicated geostatistical analyses can be forgone in favor of non-spatial statistical computations.

7.0 BIBLIOGRAPHY

ASCE Task Committee on Geostatistical Techniques in Geohydrology, S. Rouhani, Chairman and Principal Author, Review of Geostatistics in Geohydrology, I. Basic Concepts, *ASCE Journal of Hydraulic Engineering*, 116(5), 612-632, May 1990a.

ASCE Task Committee on Geostatistical Techniques in Geohydrology, S. Rouhani, Chairman and Principal Author, Review of Geostatistics in Geohydrology, II. Applications, *ASCE Journal of Hydraulic Engineering*, 116(5), 633-658, May 1990b.

ASTM, *Standard Guide for Analysis of Spatial Variations in Geostatistical Site Investigations*, D5522-96, 1996.

ASTM, *Standard Guide for Content of Geostatistical Site Investigations,* D5549-94, 1994.

ASTM, *Standard Guide for Selection for Simulation Approaches in Geostatistical Site Investigations,* D5524-96, 1994.

ASTM, *Standard Guide for Selection of Kriging Methods in Geostatistical Site Investigations*, D5523-96, 1996.

Bowers, T.S., Shifrin, N.S., and Murphy, B.L., Statistical Approaches to Meeting Soil Cleanup Goals, *Environmental Science & Technology*, 30, 1437-1444, 1996.

Cohen, J.T., Bowers, T.S., Lampson, D.W., Lampson, M.A., and Merrill, D.E., Quantification of Exposure Area Cleanup Thresholds When Contaminant Levels are Uncertain, *American Statistical Association 1997 Proceedings of the Section on Statistics and the Environment*, Alexandria, VA, 1997.

Deutsch, C. V., and A. G. Journel, *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, 1992.

Englund, E., and A. Sparks, *GEO-EAS (Geostatistical Environmental Assessment Software) User's Guide*, EPA 600/4-88/033, ENMSL, Environmental Protection Agency, Las Vegas, 1988.

ESRI, *ArcGIS Geostatistical Analyst*, http://www.esri.com/software/arcgis/arcgisxtensions/geostatistical/index.html, 2002.

Geovariances, *ISATIS Reference Guide*, www.geovariances.fr, 2000.

Isaaks, E.H., and R.M. Srivastava, *Applied Geostatistics*, Oxford University Press, 561 p., 1989.

Journel, A.G., and C.J. Huijbregts, *Mining Geostatistics*, Academic Press, London, 600 p., 1978.

Matheron, G., The Theory of Regionalized Variables and Its Applications, *Les Cahiers du Centre de Morphologie Mathmatique de Fontainebleau*, No. 5, 211 p., 1971.

Ryti, Randall T., *Superfund Soil Cleanup: Developing the Piazza Road Remedial Design,* Journal of Air and Waste Management, 43: 197-202, February 1993.

Singh, A., A.K. Singh, and R.J. Iaci, *Estimation of the Exposure Point Concentration Term Using a Gamma Distribution,* Technology Support Center Issue, EPA/600/R-02/084, EPA Technology Support Center for Monitoring and Site Characterization, Las Vegas, Nevada, October 2002.

Singh, A., A.K. Singh, R.W. Maichle, *ProUCL Version 3.0, User Guide*, EPA, April 2004.

Singh, A.K., A. Singh, *Review of the Confidence Removal Goal Approach for Site Cleanup,* Technology Support Center, U.S. EPA, Las Vegas, Nevada, 1996.

Schultz, B., A.K. Singh, A. Singh, An Evaluation of the Confidence Response Goal Approach for Making Remediation Decisions at Superfund Sites, *Environmetrics,* 13 725-732, 2002.

Schulz, Terry W., and Susan Griffin, Practical Methods for Meeting Remediation Goals at Hazardous Waste Sites, *Risk Analysis,* Vol. 21, No. 1, 2001.

U.S. Environmental Protection Agency, *Calculating Upper Confidence Limits for Exposure Point Concentrations at Hazardous Waste Sites,* OSWER 9285.6-10, December 2002.

U.S. Environmental Protection Agency, *Supplemental Guidance for Developing Soil Screening Levels for Superfund Sites,* OSWER 9355.4-24, December 2002.

U.S. Environmental Protection Agency, *Superfund Community Involvement Handbook*, EPA 540-K-01-003, April 2002.

U.S. Environmental Protection Agency, *Risk Assessment Guidance for Superfund: Volume III - Part A, Process for Conducting Probabilistic Risk Assessment,* EPA 540-R-02-002, December 2001.

U.S. Environmental Protection Agency, *Multi-Agency Radiation Site Survey and Investigation Manual* (MARSSIM), EPA 402-R-97-016, Rev. 1, August 2000.

U.S. Environmental Protection Agency, *Data Quality Objectives Process for Hazardous Waste Site Investigations,* EPA/600/R-00/007, January 2000.

U.S. Environmental Protection Agency, *Soil Screening Guidance: Technical Background Document,* EPA 540/F-95/128, July 1996.

U.S. Environmental Protection Agency, *Supplemental Guidance to RAGS: Calculating the Concentration Term,* EPA 9285.7-08, May 1992.

U.S. Environmental Protection Agency, *Guidance for Data Usability in Risk Assessment (Part A)*, EPA/9285.7-09A, April 1992.

U.S. Environmental Protection Agency, *Guidance for Data Usability in Risk Assessment (Part B)*, EPA/9285.7-09B, May 1992.

U.S. Environmental Protection Agency, *Risk Assessment Guidance for Superfund: Volume 1: Human Health Evaluation Manual Part B, Development of Risk-Based Preliminary Remediation Goals,* Washington, D.C., December 1, 1991.

U.S. Environmental Protection Agency, *Guide to Selecting Superfund Remedial Actions*, OSWER 9355.0-27FS, April 1990.

U.S. Environmental Protection Agency, *Risk Assessment Guidance for Superfund*, *Volume 1 - Human Health Evaluation Manual (Part A).* Interim Final. EPA 1540/1-89/002, 1989.

U.S. Environmental Protection Agency, *Methods for Evaluating the Attainment of Cleanup Standards, Volume 1: Soils and Soil Media*, EPA230/02-89-042, 1989.