RAC RESPONSES TO PEER REVIEWER COMMENTS

Task 6: Sampling Protocols

Radionuclide Soil Action Level Oversight Panel

July 1999

"Setting the standard in environmental health"

RAC

Risk Assessment Corporation
417 Till Road, Neeses, SC 29107
Phone 803.536.4883 Fax 803.534.1995

RESPONSES TO REVIEWER COMMENTS ON RSAL REPORT FOR TASK 6: *SAMPLING PROTOCOLS*

This document repeats the reviewer's comment and then follows it with an indented response. In some cases, the comments do not question the Task 6 report but make a general statement that does not require a response. The Peer Review Team and RSALOP (or Panel) comments are addressed in the following discussion.

RESPONSES TO PEER REVIEW TEAM COMMENTS

Reviewer A

Review Summary

The content of the above named report came as a surprise to me. I assumed based on the title that the objective of the report was to develop a **sampling protocol** or *multiple* **protocols** (because of the plurality in the title). Finding that was not the case was a disappointment. Possibly this situation was a result of RSALOP or RFCAB not properly briefing the referees as to the purpose of this particular task. To be more informative about the contents of the report, I recommend RAC retitle this report: "Considerations for Developing a Soil Sampling Protocol for RFETS."

> *RAC* will consider changing the title for the Task 6 report. During recent discussions with the Panel, it was agreed that the protocol would be directed toward the final status survey. The report will include a discussion on sampling in support of remedial action; however, a sampling protocol for remedial action support will not be developed. The report will clarify that the protocol is directed toward the final status survey.

Once having realized that a protocol was not developed and presented in the document, I expected a description of *cutting-edge* scientific considerations. What I found was largely textbook rhetoric and a summary of DOE and EPA methods. If such a presentation was indeed the product expected by RSALOP, then RAC cannot be faulted for providing it. However, the membership of RSALOP includes engineers, full professors and several members with many years of engineering and environmental science experience. Thus, I cannot help but wonder if a review at a depth no greater than that contained in documents already available to RSALOP, is really useful to that committee.

> *RAC* believes that the information presented in the Task 6 report, including "textbook rhetoric and a summary of DOE and EPA methods" is essential to provide an understanding of the underlying concepts of soil sampling. Although some Panel members have experience in various aspects of sampling, not every panel member is familiar with these concepts. We believe that some elementary discussion and presentation of current regulatory guidance is beneficial to the Panel. As far as "*cutting-edge* scientific considerations," recent discussions with the Panel have resulted in areas of

the sampling protocol that will be investigated in more detail. Important aspects of the sampling protocol, which the Panel has elected to investigate, include radiation detection instrumentation for use at the RFETS and further investigation into hot-spot detection and methods to compare hot-spot activities against the action levels.

Here I would like to note that none of the really difficult issues of developing a soil sampling protocol for RFETS were addressed or even mentioned in the RAC report. These issues include, but may not be limited to, the following:

(1) Detection of hot particles and small areas of contamination (those on the order of a few square meters or less or a few kilograms or less),

RAC and the Panel agreed during the last monthly meeting (June 1999) that the sampling protocol will not address the problem of hot particles. We agree that this is an important current area of research for plutonium contamination. The Panel has elected to research the hot-spot detection methods in further detail at a workshop.

(2) Spatial correlations among measurements which effects the sample size estimates to reach any level of statistical confidence sought,

Spatial correlations among measurements is a very important concept. This comment was also provided by Panel member Victor Holm. The sampling protocol will be directed toward the final status survey. After remediation of a contaminated area, the spatial correlations will probably not exist, thus, removing this consideration from use in the sampling protocol. However, it may apply to areas that have not been remediated (characterization data indicated the area was below the action levels) and for which a final status survey will be conducted. Ideally, the characterization survey and sampling is conducted to a level of quality that the data can also be used for the final comparison to the action levels. This is an area that the Panel may wish to further investigate to account for final status surveys of unremediated areas.

(3) Lack of correspondence between the surrogate measurement technique (*in situ* measurement of the gamma-ray from ^{241}Am) which views a relatively large sample mass (probably several thousand kg) and small areas (or volumes) of high contamination

Radiation detection instrumentation and the detection of small areas of contamination is an area the Panel has decided to investigate in more detail. A workshop will be held to learn more about methods available for detecting plutonium in soil.

(4) Lack of distinction between uncertainty and variability in measurements, averages, predicted values by surrogates, etc.

We will review the report and incorporate the suggested changes in the use of such terms as uncertainty and variability, averages, and predicted values by surrogates.

(5) Methods, sample size, etc., for detecting excess uranium contamination in the presence of natural uranium.

The milestone release date of the Task 6 report was scheduled before the calculation of the action levels for each radionuclide. This has limited the consideration of several important details of the sampling protocol. Because the action levels for the uranium isotopes are not known, it is difficult to determine if the contribution from natural uranium will be important (i.e., if the natural levels are significantly less than the action levels, differentiation may not be required). There are methods using ratios of the uranium isotopes to determine if the uranium is naturally occurring or contamination. The final date for the Task 6 report has been delayed until after the action levels are calculated to allow investigation into these types of topics.

Another omission that I think should be corrected is that the report should provide a definition of a "sampling protocol." The closest to this I found was the list of 10 elements on p. iii. It would be equally useful to include a short discussion of the how a "sampling protocol" fits into the overall goals of the remediation activities ongoing at RFETS.

The report will be revised to incorporate a definition of "sampling protocol" and a discussion will be added to describe how a "sampling protocol" fits into the overall goals of the final status survey. However, as noted in previous comment responses, a "sampling protocol" will not be developed for remediation activities. Remediation activities would require the development of several protocols based on the particular processes that resulted in the contamination event. For instance, spills would require different considerations than windblown contamination, which also would also be different from leaking tanks.

My detailed comments, of which there are many, are found in the following section.

Detailed Comments

p. iii and p. i The report states "The report provides recommendations to the RSALOP for consideration in developing a sampling protocol for the RFETS in support of the effort to conduct an independent assessment and calculate radionuclide soil action levels for Rocky Flats."

I would normally expect that a sampling protocol would be used to determine an average contamination, spatial patterns or trends, compliance with remediation standards, or other such specifics. Is this not the case? The sentence above implies the protocol is limited in use for an independent assessment and to calculate radionuclide soil action levels. If so, then the referees are

again without the proper information as to the purpose of the programs. However, if the statement is wrong, as it appears to me, then the sentence should be significantly revised.

RAC will revise the statement to better clarify the purpose of the sampling protocol in the Task 6 report.

The opening sentence on p. iii implies that RSALOP will develop the sampling protocol, but on p. 12, it states "Therefore, including an Oversight Panel member representative on the RFETS planning team is an important consideration to ensure an understanding of the final sampling protocol." The latter statement implies that an RFETS planning team – which may or may not include a RSALOP member – will be developing the protocol. The sentence on p. iii and p. 12 do not agree as to *who* has the responsibility of developing the protocol. This should be better explained.

RAC will incorporate this recommendation into the report.

p. 1 The terms "sampling units" and "survey unit sizes" are used in the 4th bullet without definition. [I later found a definition for "survey unit" on p. 14. It should be presented earlier if the term is used there.]

RAC will incorporate this recommendation into the report.

p. 1, 5th bullet states: "The variance in contamination measurements must be estimated before the number of samples can be determined..." It should state "...before the number of samples [required for xxxx or to reach zz precision of the mean, etc.] can be determined...".

RAC will incorporate this recommendation into the report.

This bullet also mentions estimating the variance by "professional judgement." I have seen such statements in the literature, but I wonder if RAC is really advocating that. If not, it should be stated with much more care.

RAC will modify this statement in the report.

p. 2, The report states "Radiation detection equipment can be used to *scan* [my italics] the entire survey area...". The authors have not defined what they mean by "scan" and their lack of precision in defining that concept indicates to me a lack of understanding of the technique. Does this mean a moving detector, e.g. vehicle mounted or an aerial survey (e.g. using aircraft) or does this mean using *in situ* spectrometry to monitor the entire "survey unit". *In situ* spectrometry is normally a static procedure, so I don't understand the word *scan*. If the potentially contaminated area has an area that is a large fraction of, or multiples of, areas equal to one or more square kms,

then I don't believe it is feasible to survey (or *scan* in the words of the authors) the entire area by *in situ* methods. Large areas would require vehicle or aircraft mounted detectors. I remain suspect of the authors understanding of these measurement techniques. Certainly the text requires clarification if it is to be useful and not misleading to a lay audience.

> *RAC* will define the word "scan" in the report. The word "scan" in terms of radiation field survey techniques for the detection of radionuclides in soil involves moving a detector across the soil surface. The FIDLER, a large area (12.5 cm. diameter by 0.16 cm. thick) NaI(Tl) scintillation crystal, detects the 17 keV and 60 keV X and γ rays from plutonium and ^{241}Am in the soil. The instrument and detector are portable and can be used to "scan" the soil surface. As far as in situ techniques, researchers at the U.S. Department of Energy's Environmental Measurements Laboratory are currently working on the problem of the identification of hot-spots using in situ (static) gamma-ray spectrometry techniques in support of the MARSSIM methodology. However, at the present time, modeling results for the action levels for each radionuclide are not available, thus, preventing a recommendation on instrumentation based on the detection capabilities of different instruments versus the action levels. The Panel has decided to delay the release of the final Task 6 report until after the action levels are calculated. In addition, the Panel is currently preparing a workshop in which experts in radiation detection methods for field applications will discuss potential methods for use at Rocky Flats.

This section also states "Therefore, a correlation between the plutonium contamination and the americium contamination must be established to ensure that areas of elevated americium contamination detected also provide coverage for other radionuclides." I cannot understand this. A quantitative relationship (not a correlation) is used to predict a radionuclide of interest (e.g., Pu) which is difficult to measure from a surrogate radionuclide which is easier (or less expensive) to measure (e.g., Am). How does establishing that relationship ensure proper coverage for other radionuclides that likely have different chemical properties and different degrees of migration?

> *RAC* will modify this statement and clarify that the surrogate method being discussed involves using americium as a surrogate to the estimation of plutonium contamination.

p. 7. 1st paragraph numbered item 2). It should read "deposited inventories..."

> This editorial comment will be incorporated into the report.

On p.10, I found what I assume is the part of the "bottom line" of this report, that being that RAC approves of the RFETS soil sampling programs: "Overall, the RFETS sampling program meets the industry standard and ensure the collection of quality data." I note this here because when I reached the end of p. 35 and then suddenly found the REFERENCE LIST, I felt there to be an absence of any concluding remarks. The sentence I refer to seems to be of an importance to warrant inclusion in a closing section. I found no other such definitive statements in the report,

but possibly there were some. If there are other such definitive findings, the report would do well to highlight them some way or include them in a "Concluding Remarks" section.

RAC will provide a conclusion section in the report with all "definitive" statements and recommendations. *RAC* will also clarify the statement that "Overall, the RFETS sampling program meets the industry standards and ensures the collection of quality data." This statement referred to the overall documented program in terms of quality assurance programs, DQO methodology, procedure development and documentation, and reporting. As stated in the Task 6 draft report, *RAC* does not advance the soil field sampling methods used at Rocky Flats, such as the CDH sampling methods or the Rocky Flats sampling method for comparison to the action levels. In addition, we have pointed out in the report that composite sampling, currently being conducted at Rocky Flats, is not acceptable for comparison to the action levels.

p. 10, The use of an excessively long string of modifiers, e.g. "specific radionuclide soil action level sampling protocol" makes interpretation laborious. I suggest improving the presentation style here (and elsewhere), e.g., "a sampling protocol to determine the soil action levels for specific radionuclides."

We will revise the report to incorporate the above suggestion.

p. 12, The report states "...radionuclide concentrations in soil are not deterministic, rather they are uncertain..." This is stated very imprecisely. In actuality, radionuclide concentrations (which are always an average over whatever volume of sample is measured) can be determined extremely accurately. If the authors means that concentrations as determined by *in situ* methods (which indeed are less accurate), concentrations determined by surrogate measurements, or the average concentration value, are uncertain, then they should say so.

RAC will revise this statement to clarify that the soil sample data will provide an estimate of the population distribution for each radionuclide. These distributions can then be used in the stochastic assessment of the sum-of-ratio calculations.

This section goes on to say the "[uncertain] concentrations can be represented by "a probability distribution." I don't believe individual measurements of concentrations are very certain (see above). Do you mean the distribution can represent the population of concentration values or do you mean the uncertainty distribution of the mean? Lest the authors think that I nit-pick here, let me assure you that differentiating the concepts of variability and uncertainty is what defines state-of-the-art assessment techniques and such a level of rigor is absent throughout the entire document.

See above comment response.

p. 13, The last sentence states "...will not need the same level of investigation to achieve the soil action levels." I don't see that a soil action level is something that can be "achieved" by an investigation; an SAL is a limiting value of contamination. Consequently, the meaning of this sentence escapes me entirely.

RAC will revise the statement.

p. 17, Equation 3-2, as written, subtracts a unitless quantity (in parentheses) from the plutonium concentration over the 0-3 cm depth range. It seems to be written improperly.

To reduce confusion, we will put additional parentheses in the equation to further outline the mathematical operational progression. In addition, the sign for multiplication in the equation will be changed from a (•) to a (×).

p. 17, The first sentence of the last paragraph should state: "The **required** sampling depth is also dependent on...".

RAC will incorporate this into the report.

p. 18, The 4th paragraph refers to "The recommended profile stages are 0-3 cm intervals, as conducted by Colorado..." Does this mean the profile *increments* (a more conventional term than *stages*) are recommended by RAC or by Colorado State University?

The paragraph will be revised to state that *RAC* is recommending the profile "increments" that are currently used by CSU in their research studies of soil at Rocky Flats for use in soil sampling for comparison to the action levels.

p. 20, The first paragraph of Section 3.6 states "...systematic grid evaluates whether the residual radioactivity in an area exceeds the soil action levels for contamination conditions that are approximately uniform across the survey unit." I assume this means that systematic grid sampling evaluates the residual radioactivity **when** the contamination is approximately uniform. A crucial word is left out or some other point is being made that I have not grasped.

RAC will revise the paragraph to clarify the meaning.

The same paragraph talks about "surface scanning" which (again) is not a proper name for any field measurement technique I have ever heard of. I can only guess the term refers to *in situ* spectrometry, however such measurements are generally static and have nothing to do with scanning – unless a moving detector is used, such as a vehicle or aircraft mounted unit – but there have been no references to such devices and they would not likely have the sensitivity required.

The wording "scanning" again appears on p.21, 2nd paragraph but no additional explanation is provided.

> As discussed in a previous response to a similar concern involving the use of the term "scan," *RAC* will define the word "scan" in the report. The word "scan" in terms of radiation field survey techniques for the detection of radionuclides in soil involves moving a detector across the soil surface. In addition, the Panel is currently preparing a workshop in which experts in radiation detection methods for field applications will discuss potential methods for use at Rocky Flats.

I could not grasp the entire discussion on Scan MDC presented on p. 21. I don't believe this is due to my lack of understanding of the field as I have considerable experience with measurements of radioactivity in the field and of detection limits. I don't see any citations on this material. Is it original to RAC or does it have another origin? In any case, it needs a considerable rewrite. In particular, I found this sentence not to be understandable: "One method of determining values for comparing capabilities of the detection equipment involves modifying the soil action levels using a correction factor that accounts for the difference in area and the resulting change in dose."

> *RAC* will revise the discussion on page 21 of the Task 6 report. Additional work is currently being conducted on the detection of hot spots (i.e., small areas of elevated activity) that will be incorporated into the final report.

Detection limits are sufficient for comparing sensitivity of instruments or particular measurement techniques; in fact, that is one of their purposes. Thus, I am at a loss in understanding this section. For that reason, I did not review the rest of this section up to Sec. 3.7. The rest of the material in this section also probably needs improvement but I did not attempt reviewing it as noted.

> See previous response from prior comment.

p. 23, The r^2 value for the empirical relationship given as eq. 3-10 is quite high (0.87) for field studies, indicating a correlation of 0.93! What suggests to the author that "it appears that additional studies would be needed to provide an accurate correlation that would apply to all areas surrounding the 903 pad?

> The basis for the statement that "additional studies should be conducted" involved statements by the author showing comparisons to other research studies and differences in the onsite and offsite ratios between soil concentrations of ^{241}Am and $^{239,240}Pu$. We will add additional discussion in this area of the report to clarify the reviewer's concerns.

p. 23, Neither the Mann-Whitney test (based on the rank-sum) or the Kolmogorov-Smirnov test (for comparing distributions or goodness-of-fit) are parametric tests as the authors claim. These are both non-parametric tests. This level of inaccuracy in this report is perplexing.

The Mann-Whitney and Kolmogorov-Smirnov tests are nonparametric tests as pointed out by the reviewer. This was an unfortunate error and it will be corrected.

p. 24, 2nd paragraph. The crucial word **"not"** is left out of the sentence: "When the underlying distribution is not symmetric, the mean is **not** equal to the median."

The actual error in this statement results from having the word "not" before symmetric. The sentence will be revised to state that "When the underlying distribution is symmetric, the mean is equal to the median." The above sentence then follows into the next sentence which correctly states that "When the underlying distribution is not symmetric, these tests are still true tests of the median but only approximate tests of the mean."

p. 24, 4th paragraph. Incorrectly failing to release an area which in actuality satisfies the release criteria is a 'false-positive' result because the null hypothesis (of no significant difference between the measurements and the release criteria) is incorrectly rejected. The text states this is a Type II error but in actuality this is a Type I error. This misapplication is carried into Section 3.8.2. [I discuss this topic in more detail later in this review.]

See comment response for Section 3.8.2 below.

p. 24, last sentence. I think the authors do a disservice to RSALOP by the vagueness and lack of useful information which is exemplified in the last sentence: "...if the background concentrations of the radionuclides present at the RFETS are only a small fraction of the soil action levels, there is no need to consider background." Data on background levels of uranium, plutonium, and americium is available in the open literature. Why not include at least mean values of these nuclides for the Rocky Mountain States and thereby give some useful information? In that case, the sentence could be rephrased, "...if the soil action levels are many times the average background of uranium, plutonium, and americium (give the values here, from my knowledge, I estimate 35 Bq/kg, ~0.5 Bq/kg, <0.1 Bq/kg, respectively), then there is no need to consider background."

RAC will include a section on radionuclide background concentrations for areas near the Rocky Flats site.

p. 25, One of the single most important elements of a "sampling protocol" is that of determining the required number of samples to meet a particular level of precision, yet, I found section 3.8.2 to be the weakest section of the report and to also have a number of inaccuracies. The following comments apply to section 3.8.2.

Section 3.8.2.1 First sentence claims that the required number of samples depends on the "anticipated variability" of the measurements. This is most assuredly wrong. Possibly the authors

mean that an initial estimate of the required number of samples depends on the "anticipated variability", however, such estimates would have to be revised as better information is collected.

RAC will revise the sentence to indicate that "an initial estimate" and not "anticipated variability" is required to calculate the number of samples.

Section 3.8.2.2 It is stated: "When testing the mean, the recommended null hypothesis (H_0) is that the residual radioactivity in the survey unit exceeds the action levels." A null hypothesis by convention states that the statistic (such a sample mean) that is being compared to another sample mean or to fixed value (such as a RSAL) is the result of random sampling and that any difference is due to chance. Thus, the recommended null hypothesis should be that there is **no** difference in the residual radioactivity measured in the survey and the action level – exactly the opposite of what is written in the text. Note that what I am describing is a conventional two-sided hypothesis test. A single-sided test, closer to what is described in the text, can also be computed which tests H_0 against an alternative hypothesis, where the alternative is stated as "the measurements tend to exceed the RSAL." However, the null hypothesis still proposes the **null** condition, i.e., there is no significant difference and that any observed difference is due to chance.

RAC disagrees with this comment because the reviewer is referring to the null hypothesis as being the state where there is no difference between the sample data results and the action levels. In hypothesis testing we begin by making a tentative assumption about a population parameter. This tentative assumption is called the null hypothesis and is denoted by H_o. We then define another hypothesis, called the alternative hypothesis, which is the opposite of what is stated in the null hypothesis. This alternative hypothesis is denoted by H_a. The hypothesis-testing procedure involves using data from a sample to test the two competing statements indicated by H_o and H_a.

Let μ_o denote the specific numerical value of the population mean being considered in the null and alternative hypotheses. In general, a hypothesis test concerning the values of the population mean μ must take one of the following three forms:

$H_o: \mu \geq RSAL$	$H_o: \mu \leq RSAL$	$H_o: \mu = RSAL$
$H_a: \mu < RSAL$	$H_a: \mu > RSAL$	$H_a: \mu \neq RSAL$

In many situations, the choice of H_o and H_a is not obvious; in such cases, judgment on the part of the data user is needed to select the proper form of H_o and H_a. However, as the above example forms show, the equality part of the expression (either $\geq$, $\leq$, $=$) always appears in the null hypothesis. The statement of the null hypothesis in the report did not include the equality portion of the statement in the null hypothesis; this may be what confused the reviewer. In selecting the proper form of H_o and H_a, keep in mind that the alternative hypothesis is what the sampling study is attempting to establish. Thus, asking whether the data user is looking for evidence to support $\mu < RSAL$, $\mu > RSAL$, or $\mu \neq RSAL$ will help determine H_a.

We will revise the null hypothesis to state that:

Ho: $\mu \geq$ RSAL and Ha: $\mu <$ RSAL

Note: This is an example case where we are testing the mean of the radionuclide soil concentrations against the action levels. The Panel may determine that another population parameter will be applied to the action levels such as the median, 95 percentile, etc.

The null hypothesis as stated above assumes that the survey unit is contaminated above (or equal) to the action level; thus, the statistical analysis must prove that this assumption is wrong. If the statistical analysis provides evidence that the null hypothesis is not true, then we accept the alternative hypothesis that the survey unit soil concentrations are less than the action levels.

The statement of the null hypothesis above is designed to be protective of human health and the environment as well as consistent with current methods used for demonstrating compliance with regulations. EPA, NRC, and DOE have adopted the use of this null hypothesis in MARSSIM for statistical analysis of contaminated sites. The EPA has also provided null hypotheses stated in this manner in their soil cleanup standards document (U.S. Environmental Protection Agency, 1989, *Methods for Evaluating the Attainment of Cleanup Standards - Volume 1: Soils and Solid Media*, EPA 230/02-89-042, Office of Policy, Planning, and Evaluation, Washington, D.C.). It is acknowledged that site contamination conditions, such as lack of measurement techniques with appropriate detection sensitivities, may preclude the use of the null hypothesis that the survey unit is assumed to be contaminated. Another problem arises when there is a high variability in background concentrations of the radionuclides. If the background concentrations are near the proposed action levels, it becomes difficult to distinguish between background and contamination concentrations. The U. S. Nuclear Regulatory Commission uses the null hypothesis that the survey unit concentration is indistinguishable from the concentration in the background area, when background variability is an issue (C.V. Gogolak, A.M. Huffert, and G.E. Powers, 1995, *A Nonparametric Statistical Methodology for the Design and Analysis of Final Status Decommissioning Surveys,* NUREG-1505, Draft Report for Comment, U.S. Nuclear Regulatory Commission, Division of Regulatory Applications, Office of Nuclear Regulatory Research, Washington, D.C.). Statistical tests are then performed to demonstrate that the survey unit is indistinguishable from background.

The Panel is currently planning a workshop on radiation detection instruments for use at the Rocky Flats site. In addition, the release of the Task 6 report has been delayed until the action levels are determined so that additional considerations in this area may be investigated.

Section 3.8.2.3 This section incorrectly states "A false positive error would result in the release of a survey unit containing residual radioactivity..." Because of the incorrect statement of the null hypothesis (see above), the examples given for both the Type I and Type II errors are incorrect. Given that the null hypothesis proposes **no** d difference except that due to chance, a Type I error (false positive) would result in **cleanup** (not release) of a survey unit **that does not** contain residual radioactivity in excess of the RSAL. The next paragraph gives an incorrect example of a

Type II error. It states "A false negative error would result in either unnecessary cost because of remediation .. that are below the release criterion. Rather, a Type II error (false negative) would result in the **release** (not cleanup) of an area that is **above** (not below) the release criterion.

As stated in the above comment response, the null hypothesis is not incorrect. Therefore, the Type I and Type II errors are stated correctly.

The following table provides an explanation of false positive and false negative conclusions based on our proposed null and alternative hypotheses:

Decision based on the sample data is	The true condition is	
	Clean	Contaminated
Clean	Correct decision	False positive (Type I error) released as clean when it is really contaminated
Contaminated	False negative (Type II error) unnecessary remediation	Correct decision

In addition to the inaccuracies discussed above, this section is exceedingly weak because it gives no guidance on calculating the required number of samples. It is hard to believe such a section would be presented without some of the conventional mathematical formulae available from the literature for estimating sample sizes (there are several formulae available depending on the type of distribution type that is likely).

We will provide guidance and conventional mathematical formulae on calculating the required number of samples in the revised report.

Furthermore, a major consideration was not discussed or even mentioned which is important to determining the number of samples required, that being, the phenomena (sometimes considered a problem) of spatial correlation. Because concentrations of radioactive contamination (or natural elements) in the environment represent single values sampled from a continuum over a two-dimensional space, each value is related to nearby values (i.e., they are correlated) as a result of the events or process which deposited the radioactivity. Spatial correlation has important consequences to estimating the local (or regional) spatial variance and for correctly stating confidence limits. Failure to realize the presence of positive correlation in measurement data (e.g., environmental measurements) leads to a confidence interval that is too narrow. Two narrow of confidence limits results in a false sense of security that the required precision has been achieved. For the purposes of comparing measurements with guidelines, standards or historical values, recognizing and determining the spatial correlation of measurements may be important. Gilbert (1987, Chapters 4,8,16, 17) and Cressie (1991) should be consulted for methods to determine the existence of, and the degree of, spatial correlation. Cressie (1991) shows how to compute the number of samples (n) required to reach the same precision, if spatial correlation exists, as compared to the same number of independent samples.

See comment response given previously on the subject of spatial correlation.

References:

Gilbert, R. O. Statistical Methods for Environmental Pollution Monitoring. Van Nostrand Reinnhold Co. 1987.

Cressie, N. Statistics for Spatial Data. New York, John Wiley & Sons, Inc. 1991.

p. 26, First bullet item: The BEIR V report (incidentally, the reference should not be to BEIR which is an acronym for the title, but to the National Research Council that published the report) does not state "public health risk is modeled as a linear function of dose" but rather the report derives risk estimates which are linear functions of dose for a variety of solid cancers. Note also that leukemia risk is a linear-quadratic function of dose.

RAC will incorporate these suggestions into the report.

p. 26, First bullet item states "When the risk is linear, much higher values..." but should say "When the risk is linear and without a threshold, much higher values...".

RAC will incorporate this suggestion into the report.

p. 26, Second bullet item states: "The soil radionuclide action levels are not free of error." The words "not free of error" should be replaced "cannot be determined with absolute precision." Error and imprecision due to uncertainty do not necessarily have the same meaning.

RAC will incorporate this suggestion into the report.

A "Concluding Remarks" section at the very end would be useful because the report ends abruptly without tying together the variety of topics covered.

RAC will develop a conclusion section for the report.

Reviewer B

Introductory note: for convenience, overall comments are presented first, and more detailed comments are presented on a page-by-page basis. Purely editorial comments are introduced by the word "Editorial". From my perspective, RAC need not respond in writing to any of the comments and suggestions labeled "Editorial".

Overall Comments

The draft is thorough and quite credible as to the substantive topics dealt with. It will, however, be hard going for the non-specialist, but I do not see a good solution to that problem. As noted in my comments on the Draft Report on Task 2, I recommend the preparation of a separate, free-standing brief summary that presents the essence of the draft, especially its recommendations and conclusions, in a form more understandable by non-specialists. This will be a tough challenge for this particular part of the project.

> *RAC* will include a brief summary of the report that is written to help nontechnical people to understand our objectives and recommendations.

A particularly strong aspect of the Draft was the identification of the specific recommendations from the body of the Draft in the Executive Summary. This feature should be considered for all subsequent drafts as well. In fact, when reviewing the actual Draft, it was sometimes hard to identify the recommendations on the first reading, since I was focussing on the detailed flow of the analysis. Therefore I suggest consideration be given to highlighting the recommendations in the text in some way, perhaps with italics, or special sub-headings.

> *RAC* will attempt to provide a better presentation of the recommendations throughout the report.

On their merits, I support all of the recommendations provided by RAC. However, the basis for certain of RAC's more important (and likely more controversial) recommendations needs to be beefed up. In my specific comments, I have indicated where this is needed.

The Executive Summary did not adequately incorporate RAC's conclusions from Section 2.6. I recommend that this be done in the Executive Summary, but only after significant improvements are made to Section 2.6 (see my detailed comments later on Section 2.6, page 10 of the draft).

> See comment response on Section 2.6 below.

Also, in both the body of the draft and the Executive Summary, it would be good to specify exactly to whom the recommendations are really directed (and thus from whom a response is expected).

RAC will identify the intended responsible parties in the revised report.

The end of the Draft left me hanging. The three sections are: Introduction, Review, and Recommendations. There is no final section headed Conclusions. Perhaps it is not necessary, and the Executive Summary fully serves that purpose. I suggest, however, that the authors consider adding a brief conclusion section.

We will provide a conclusion section in the revised report.

The list of acronyms is a useful feature.

This draft had a few typos, etc., which I have tried to identify is as much detail as possible. Many of them were probably due to excessive reliance on the "spell check" feature of the word processing program used.

Detailed Comments

Page iii, first paragraph. Probably editorial, but perhaps not. Is the Oversight Panel *itself* going to develop and carry out its own sampling protocol at Rocky Flats? Without going back to the basic documents, I thought the purpose of this Task was to help the Panel provide its views on any protocols to be used by the actual Rocky Flats team (DOE and contractors). The wording in the first paragraph seems to state that the Panel will be developing and conducting its own protocol.

RAC will clarify the purpose of the sampling protocol in the revised report.

Page iii, second paragraph. The draft states that "several areas...were considered acceptable." This implies that other areas were not acceptable (though I did not find such a conclusion in the body of the draft). If some areas are acceptable, and some are not, I recommend that both groups be listed in the Executive Summary, or that at least a reference be given to a new section in the draft (see General Comment above) that identifies both groups.

RAC will incorporate these suggestions into the report.

Page iii, #3. Editorial. Isn't the recommendation specifically for the MARSSIM "classification scheme", or would another system do?

We will clarify the MARSSIM classification scheme in the revised report.

Page 1, first paragraph. See earlier comment on Page iii, first paragraph.

See prior comment response for page iii, first paragraph.

Page 1, first bullet and elsewhere. Editorial. I believe that the generally accepted phrase is "DQO process" (not "DQO*s* process"), and that "DQOs" is used when specifically discussing the objectives, not the process leading to them. If I am correct, this can be fixed by a careful search and replace review of the entire draft.

We will incorporate these changes into the revised report.

Page 1, second bullet. Editorial. I suggest ending the sentence as follows: "...in inappropriately releasing contaminated sites for restricted or unrestricted use."

RAC will consider this editorial comment for incorporation into the report.

Page 1, third bullet. Editorial. Shouldn't "comparing" replace "evaluating"?

RAC will change "evaluating" to "comparing."

Page 2, second paragraph. Editorial. The DQOs/DQO matter mentioned just above. I will not identify any additional examples in these comments, though there are many.

RAC will ensure appropriate usage of DQO and DQOs in the revised report.

Page 3, first paragraph. While the summary of this section (page 10) gives the conclusion, I suggest telegraphing the conclusion right at the start of the section.

RAC will incorporate this suggestion into the revised report. *RAC* will also clarify the statement that "Overall, the RFETS sampling program meets the industry standards and ensures the collection of quality data." This statement referred to the overall documented program in terms of quality assurance programs, DQO methodology, procedure development and documentation, and reporting. As stated in the Task 6 draft report, *RAC* does not advance the soil field sampling methods used at Rocky Flats, such as the CDH sampling methods or the Rocky Flats sampling method for comparison to the action levels. In addition, we have pointed out in the report that composite sampling, currently being conducted at Rocky Flats, is not acceptable for comparison to the action levels.

Page 3, second paragraph. Editorial. While SOP is found in the acronym list, I don't think it has been used (and identified) earlier in the body of the report.

We will ensure that acronyms are spelled out when they first appear in the report.

Page 4, first full paragraph. Editorial. Does "operable unit" need to be defined?

We will define operable unit in the report.

Page 5, Table 2-2. Editorial. RFI and RI are not defined in this draft. For this purpose, perhaps a footnote to the table would meet the need, since I don't think the acronyms are used later.

We will define RFI and RI in the report.

Page 5, heading 2.2.2. Editorial. Does EMD need to be defined?

We will define EMD in the report.

Page 7, end of first partial paragraph. Editorial. "Very" should probably be "verify", and HPGe, while in the acronym list, isn't spelled out in the text until page 9 (this last is a very minor point, but does show how carefully I read the draft).

RAC will incorporate these suggestions into the report.

Page 10, Summary section. The entire Section 2 was really more of a description rather than a review, until we reach 2.6, which is the heart of the matter. I strongly urge that 2.6 be expanded to include more clearly RAC's views. The RFETS program is "considered adequate" (last paragraph). Some skeptics might wonder if RAC left out the words "barely" or "marginally". Others, with a different attitude, might think the word "fully" was omitted. RAC should be as clear as possible, to minimize the chance of such interpretations by others. In the same vein, earlier in 2.6, if RAC believes that the RFETS team has used the right references, developed solid guideline documents, etc., it should say so clearly and unequivocally. If the RFETS team did only a so-so job, RAC should say that instead. Finally, I urge RAC to offer its own views on whether the four methods are all necessary. (I have my own opinion, but that is not relevant. RAC's opinion is, however, highly relevant to the Oversight Panel.) Simply noting that each method is used for different purposes is not quite enough. Are all the purposes of equal importance? Are some methods more "standard" than others, while the Rocky Flats situation also requires the addition of non-standard methods? Does the existence of four methods, while perhaps complicating the level of detail, etc., provide for a fuller understanding of site conditions, risks, etc.? In short, I recommend that this section be expanded to provide a full presentation of RAC's conclusions, and that appropriate features of this new section be incorporated into the Executive Summary.

RAC will provide additional text to clarify our view of the RFETS sampling program to ensure that no interpretation of our findings is necessary.

Page 12, first paragraph. Elsewhere, the DQO process is flat out said to be iterative (which I think is correct), while here it is characterized as "likely" to be iterative. RAC should decide which it believes is correct, and only use that descriptor. (I believe this is the only place where "likely" is used.)

RAC will remove the word "likely" from the sentence.

Page 12, third paragraph. Editorial. "Consideration" should be replaced by "recommendation", and the sentence re-worded appropriately.

We will replace "consideration" with "recommendation."

Page 12, fourth paragraph. Editorial. "Active" should be replaced by "action".

We will change "active" to "action."

Page 13, third paragraph. Editorial. Remove the comma after "thus".

RAC will incorporate this change into the report.

Page 13, fourth paragraph. Editorial. "Probably" should probably be "probability". "CDPF" should be "CPDF".

RAC will incorporate this change into the report.

Page 13, fifth paragraph. I don't think "population" by itself is quite the right word, and the sentence may need to be expanded to be clear. Perhaps it should be "...that the estimated overall measured group of samples meets..." In general, this paragraph should be re-written to read more clearly than it does now.

RAC will rewrite this paragraph.

Page 13, last partial paragraph. This sentence needs re-writing. *Investigations* in and of themselves do not achieve soil action levels, only cleanup activities do that.

RAC will rewrite this sentence.

Page 15, end of section 3.3.2. If it is possible, I recommend RAC add a short paragraph explaining the implications if this recommendation is adopted by the Oversight Panel. This is one of only two specific recommendation that, in my opinion, calls for such elaboration.

RAC will elaborate on this recommendation.

Page 17, second (complete) paragraph. Would it be helpful to some audiences to briefly describe what the background is due to?

RAC will describe the background reference in the Litaor (1999) study.

Page 17, third paragraph. Editorial. "Intensive" should be replaced by "extensive".

We will change "intensive" to "extensive."

Page 17, fourth paragraph. I recommend that the last sentence ("...profile is likely to be...") be elaborated on. In particular, in what way(s) is the profile likely to be vastly different?

We will elaborate on the reason that the radionuclide profile distribution would be disturbed by remedial activities.

Page 18, first partial paragraph. Editorial. I suggest the following change: "In contrast, other pathways, such as plant uptake...."

RAC will incorporate this change into the report.

Page 18, end of Section 3.4. If RAC really believes the 3 cm interval is the way to go, the case should be made as strongly as possible in this section, if for no other reason than it apparently will mark a major change (and may be strongly resisted by some). In addition, this specific recommendation should be carried forward as stated to the Executive Summary, not generalized as it is now.

RAC will provide additional discussion on the incremental sampling recommendation.

Comment on all of Section 3.5 and Section 3.6. This is about as clear an explanation of these topics as I have ever seen. In particular, I suggest the use of additional "for examples", as done on the bottom of page 21, to give some reality for non-specialist readers.

RAC will consider incorporating additional examples in the report for this purpose.

Page 20, last paragraph. Editorial. I suggest a comma after "(e.g., standard deviation)".

We will incorporate this suggestion into the report.

Section 3.7. Without having taken the time to do the math, I suspect there is not a really large difference between Eq. 3-9 and Eq. 3-10, even though one covers one isotope and the other, two isotopes. RAC's observations on this would be a helpful addition to this section. More important, if RAC really believes surrogates would help (it is included in the recommendations in the Executive Summary), a stronger case should be made in this section, and perhaps an actual suggested conceptual approach, with an estimate at least of how quickly it could be done if not the probable cost.

Equations (3-9) and (3-10) are similar and an important omission occurred in Equation (3-9). Although Shierman (1994) provides the equation as ^{239}Pu, discrimination between ^{239}Pu and ^{240}Pu was not possible using alpha spectroscopy. Therefore, Equation (3-9) will be revised to indicate that it is also using ^{241}Am to estimate ^{239,240}Pu. *RAC* will also expand the discussion on the use of surrogates at Rocky Flats.

Page 24, second paragraph. My rusty recall of statistics tells me that when the distribution *is* symmetric, the mean is equal to the median. Thus I recommend that "not" be removed from the sentence ending "the mean is equal to the median". If my recall is wrong, ignore this suggestion.

The statement will be corrected as indicated in the response to the same comment from Reviewer A.

Page 24, fourth paragraph. The Type II error concept is cited here, but not defined until page 25. The answer may simply be to eliminate the parenthetical in this paragraph.

RAC will remove the reference to Type II in this sentence.

Section 3.9, page 26. In my view, this is one of the most important recommendations in the Draft, and one that I particularly strongly support. It behooves RAC to elaborate somewhat on the general implications (cost, whether it would lead to significant schedule delays, etc., etc.)

RAC will expand the discussion on Section 3.9. We are not in a position to estimate costs and schedule delays that may be incurred by Rocky Flats because of the implementation of this recommendation. However, Rocky Flats is currently using independent contractors to perform independent confirmatory investigations on the buildings at Rocky Flats. Therefore, it would seem reasonable that similar activities should be conducted for the action levels in soil.

Reviewer C

General comments

It was hard to immediately determine the overall goal of the report. I eventually concluded that the goal was to determine a sampling protocol for determining the spatial extent of cleanup needed for areas which do not pass the soil action level. Presumably, this would require additional sampling in the future. Because of the tremendous amount of soil sampling that has already been done, it was not immediately apparent to me that more sampling will be required in the future. My recommendation is to add a short paragraph at the beginning of the EXECUTIVE SUMMARY and at the beginning of the INTRODUCTION, explaining what is anticipated in the future with regard to determining a course of action based on the eventual soil action level. Then explain that more soil sampling may be required, thus a sampling protocol is needed.

> During recent discussions with the Panel, it was agreed that the protocol would be directed toward the final status survey. Therefore, after remediation of an area that exceeds the action levels is completed, additional soil sampling will be required. *RAC* will add a discussion on this topic in the EXECUTIVE SUMMARY and at the beginning of the INTRODUCTION. Ideally, areas that have not required remediation (characterization data has provided evidence that the area was below the action levels), the characterization survey and sampling is conducted to a level of quality that the data can also be used for the final comparison to the action levels. This is an area that the Panel may wish to be further investigate to account for final status surveys of unremediated areas.

Much of the material in section 2, especially parts 2.1, 2.2, 2.3 and 2.4 just give reference to other documents, without providing any real insight to the reader. It forces the reader to refer to many other documents to understand how RFETS does QA, QC, SOP's, Analytical Requirements, and Validation Requirements. If the philosophy on all these issues would be difficult to summarize, then it should be made clear that this report is making no such attempt. Based on my experience, all the documentation in the world attempting to assure that good, sound data is obtained can easily fail to do so. Many times, so much effort is placed on the "paper trail", that the actual care taken in the conduct of the work can become secondary. It really boils down to having good, competent & careful people do the sampling, sample preparation, and laboratory analysis-under the guidance of very experience leaders.

> *RAC* will provide a statement in the report that indicates the purpose of the RFETS sampling program review. It would be quite difficult to provide a summary on the philosophy for all aspects of the REFETS program. We agree that having procedures and programs in documented form does not provide assurance that field or analytical errors will not occur. We will stress this concept in the report.

The section on "Data Quality Objectives" is important, but it could benefit from some specific "what ifs". For example, what about the need to consider unforeseen circumstances such as those

affecting soil stability (e.g. off-road vehicles, fire, flood, high wind, etc.). How will these sorts of things affect the DQOs?

> We will expand the discussion on DQOs to provide examples of the factors that need to be considered during the process.

Most of Section 3 was quite well done, and it reflected a good understanding of the historical literature and most of the more subtle issues concerning Pu & Am behavior in soils.

> Although no response is required for this comment, it is helpful to *RAC* when reviewers provide comments on sections of the report that they feel were done well. These sections can then be compared to those sections that reviewers have identified for revision and helps us understand what level of discussion is expected.

Specific Comments

Page 1, 4th bullet: Water as well as wind can factor into contaminant dispersion. Material can, in some circumstances, accumulate in depressions, etc., thus the statement about concentrations decreasing with distance, although generally true, is not always so. I think there are exceptions to the statement about larger survey units having lower mean concentrations. It depends on scale of sampling as well as heterogeneity of contamination and random chance.

> *RAC* agrees with the reviewer that there are exceptions to the statements made in this bulleted item. We will revise this paragraph to indicate these are general statements and that exceptions are possible. That is why we stress in the first sentence that sampling units must be defined based upon knowledge of the site.

Page 2, last sentence of bullet paragraph: The relationship of Pu and Am has been quite well established already. This is acknowledged later in the report.

> We will change this sentence to indicate that relationships between americium and plutonium have been established at the site based on several studies.

Page 2, first sentence of next to last paragraph: RAC reviewed more than just the RFETS sampling program, as is apparent in section 3.

> *RAC* will modify the introduction to clarify the extent of the review.

Page 3, third paragraph: Cite literature for the CDPHE and RFETS methods. I don't find in the reference list. This is quite important for this document.

We will provide references for the CDPHE and RFETS methods.

Page 4, Table 2-1: a 1 in 20 field duplicate is not adequate for the RFETS.

Table 2-1 provides a review of the current RFETS quality control sample collection frequency. Table 3-2 provides a method to determine an adequate number of samples based on the desired confidence level. We will provide additional discussion on the RFETS review in Section 2 to clarify those areas of the program that *RAC* deems adequate.

Page 7, last line of first paragraph: I don't think HPGe surveys can be "verified" with soil data, because, not only does depth distribution affect the detector response, but also rocks, moisture, and micro-topography.

We agree that HPGe surveys are affected by depth distribution of the radionuclides and soil properties. In general, the depth distribution needs to be investigated to ensure that the correct parameters are used in the conversion from instrument count rate to soil concentration. As an example (from NUREG-1506, "Measurement Methods for Radiological Surveys in Support of New Decommissioning Criteria", U.S. Nuclear Regulatory Commission, Draft Report for Comment, July 1995), for undisturbed soils a negative exponential profile with depth has frequently been found to be an adequate model for deposited radionuclides, that is:

$$S = S_o \exp^{[(-\alpha/\rho)\rho z]}$$

where S is the activity per unit volume of soil (pCi cm^{-3}) at depth z (cm), S_o is the activity per unit volume at the soil surface (pCi cm^{-3}), α is the reciprocal of the relaxation length of the exponential distribution (cm^{-1}), and ρ is the soil density (g cm^{-3}). This expresses the profile in terms of the soil mass per unit area, pz (g cm^{-2}), with the degree of penetration into the soil represented by the depth parameter α/ρ (cm2 g^{-1}). This type of profile has the maximum concentration at the soil surface (S_o) and decreases with depth. If the value of α/ρ approaches infinity, the source distribution approaches a plane atop the ground, and if α/ρ equals 0, the source distribution is uniform with depth. As an example, assume a soil density of 1.5 g cm^{-3} and an α/ρ value of 0.2 cm2 g^{-1} (which is a typical value for an aged fallout deposit), the corresponding relaxation depth for the exponential profile would be 3.33 cm, meaning that the concentration would be reduced to 1/e, or 37%, of the surface value at this depth. For in situ measurements, the value of α/ρ can be determined from the analysis of soil samples from different depth increments. The fraction of the total activity below a given depth (log value) can be plotted versus the mass depth, ρz. The slope of the line is then the value of α/ρ.

RAC will change the statement from "verify the HPGe surveys" to "verify the radionuclide depth distribution for use in the HPGe survey conversion factors."

Page 8, line 1 (and elsewhere): What is the "undesirable top layer"? Be more specific and scientific.

> *RAC* will describe the meaning of "undesirable top layer" in terms of the types of materials typically removed before sampling at Rocky Flats.

Page 15, third paragraph, line 5: Does Pu activity in soil refer to concentration? It is not clear.

> The "plutonium activity in soil" reported by Litaor et al. (1995) was referring to concentration (Bq/kg). We will change "activity in soil" to "soil concentration."

Page 19, third paragraph: I can't agree with statement about a site needing to be homogeneous to use random sampling. Better consult a statistician.

> This statement will be revised. The site need not be homogeneous to perform random sampling. However, random sampling is not recommended because of the unpredictable spacing between sample locations. Therefore, sample locations may be clustered in certain areas, with other areas having fewer sample locations.

Pages 22, 23, equations 3-9 and 3-10. I recommend comparing these to see if they are really different.

> Equations (3-9) and (3-10) are similar and an important omission occurred in Equation (3-9). Although Shierman (1994) provides the equation as ^{239}Pu, discrimination between ^{239}Pu and ^{240}Pu was not possible using alpha spectroscopy. Therefore, Equation (3-9) will be revised to indicate that it is also using ^{241}Am to estimate ^{239,240}Pu. *RAC* will also expand the discussion on the use of surrogates at Rocky Flats.

Page 29, second bullet: An analytical replicate of RFETS soil will not serve purpose due to micro heterogeneity. In other words, the soil cannot be adequately mixed. This can be overcome with many replicates and large sample volumes.

> We note on page 29 that analytical replicates can only be used by the analyst as an internal control tool and not as an unbiased estimate of analytical precision. Therefore, field replicates are better indicators of the total error for sampling. We agree that many replicates and large sample volumes will provide a better estimate of error than will a small number of replicates.

Reviewer D

General Comments

The report seems adequate in general, but RAC needs to clarify some issues. General comments are listed below and specific page-by-page comments are listed in a separate section.

1) Soil sample depth. In DOE/EPA/CDPHE (1996), surface soils are defined to be from the ground surface down to 15 cm. Sub-surface soils are defined to be from 15 cm to the groundwater table. Since groundwater and surface water pathways were dismissed in DOE/EPA/CDPHE (1996), only surface soil radionuclide soil action levels (RSALs) were developed. RAC has stated that they, too, will ignore the groundwater pathway (p.27 Task 2) "subject to reinterpretation based on new findings..." RAC should explicitly note in the introduction to this report that the RSALS that they plan to develop are for surface soils only. As is, this information is available only by noting that none of the sampling methodologies are applicable to depths greater than 15 cm. If the groundwater pathway becomes more important for any reason, it may be necessary to characterize deeper soils and also to make a decision as to whether to apply surface soil RSALs to subsurface soils or to develop a separate set of subsurface RSALs.

RAC will incorporate this recommendation into the revised report.

2) Terminology. If the RSALs are to be determined stochastically, I think that some new terminology is required for clarity. As is, I have difficulty determining whether RAC is talking about a point value RSAL or an RSAL distribution since the acronym RSAL is used for both (e.g., see comment 12.). RAC should put some thought into new nomenclature for this and future reports. It seems to me that the term RSAL is best reserved for the point value RSAL. The RSAL distribution, as I understand it, is really the distribution of contaminant concentrations in soil that result in the target dose at a site (I'll call this the target dose concentration-- C_{DT} --here), given uncertainties in at least some of the input parameters to the concentration-dose model. The point value RSAL is a specific value along the C_{DT} distribution (e.g., the x^{th} percentile). It is inappropriate and confusing to call a specific value of the C_{DT} distribution an RSAL and also to call the C_{DT} distribution an RSAL as well.

RAC will consider nomenclature to remove the confusion between action levels that are point values and those that would be distributions.

3) Use of the soil sampling data. A discussion of the how the stochastic RSALs will be compared to measured radionuclide concentrations for sites with single and multiple radionuclides should be more explicit. As is, such a discussion buried in Section 3.2. I think that a brief introductory section detailing the use soil sampling data for comparison to RSAL (point values or distributions) is necessary (i.e., move section 3.2 and revise some). After all, the whole purpose of the report is to recommend soil sampling procedures for the purpose of gathering contaminant concentration data for comparison to the newly developed stochastic RSALs.

RAC will incorporate this recommendation into the report.

Specific page-by-page comments

1. p. vii:
The acronyms EMD, RFP, and RSAL are used in the text and not defined. Define here.

We will define the acronyms as suggested by the reviewer.

2. p.1. 3rd bullet 1st sentence: The verb 'evaluating' doesn't seem right. Shouldn't it be 'comparing'?

RAC will change "evaluating" to "comparing."

3. p.1. 5th bullet: Aren't there adequate prior sampling studies for most areas?

Data from prior sampling studies may be used to estimate the variance of the population for the purpose of calculating the number of required samples. However, several of these studies have generally been performed for characterization purposes. Therefore, the variance in the measured contaminant distributions before remedial action will not be the same after remediation.

4. p.2, 1st paragraph: It is stated that the "equipment must be capable of detecting contamination at the level desired. This requires knowledge of the detection equipment, including detection efficiencies for the contamination of concern." It seems to me that, since the charge for the current task was to recommend specific soil sampling prodders, it would have been advantageous to perform Task 5 (Independent Calculation) before the current task. Then, RAC could apply its knowledge of the detection equipment to make recommendations as to the appropriate detection equipment to use to evaluate contaminant concentrations at the RSALs.

The final Task 6 report has been delayed until the completion of Task 5. The Panel is currently preparing a workshop in which experts in radiation detection methods for field applications will discuss potential methods for use at Rocky Flats.

5. p.5 Section 2.3 , 1st paragraph: The use of "RFETS" as a subject in this sentence should be corrected. The entity(ies) that developed the document should be identified. The same comment applies to p.7 2nd paragraph and p. 13, 4th paragraph.

RAC will remove "RFETS" from these statements and replace with the appropriate entity.

6. p.7 first paragraph: There appear to be several typos in the offset quote. Is RFP right? If so, define in the acronym list. In the last sentence, should 'very' be 'verify'?

"RFP" will be replaced with "RFETS," and "very" should have been "verified". Note that previous review comments have questioned the use of the term "verify" and the meaning of this statement will be changed accordingly.

7. p.7 third para and throughout document. Be consistent in use of metric system. It's irritating to switch back and forth between cm and inches. I recommend sticking with the metric system.

The revised report will use metric units as the standard with English units in parentheses.

8. p. 12 last paragraph, 2nd to last sentence. A word is missing after 'population'--shouldn't it be 'population statistics'?

RAC will correct this sentence.

9. p.13, first paragraph. It would help to finish the last sentence with something like ", which represents uncertainty in the SR value given the measured concentrations."

RAC will incorporate this suggestion into the report.

10. p. 13. It seems to me that the first alternative for setting RSALS (developing a distribution of SR values) is useful for determining whether measured contaminant concentrations at a site are acceptable, but is of limited use in setting soil remediation goals (i.e., once it is determined that measured contaminant concentrations are unacceptable at a site and a remediation effort is required, what contaminant concentrations will be acceptable to leave behind.) Of course, you could measure contaminant concentrations after the remediation is complete , develop a post-remediation distribution of SR values and require further remediation if the results are unacceptable. However, it is useful to have remediation goals specified as contaminant concentrations in soil that, if attained, will allow the site to be declared 'clean'. As such, I believe that the latter method (i.e., choosing a point value RSAL for each radionuclide) is more useful for setting soil remediation goals.

RAC points out that 'RFETS and the Oversight Panel need to consider which method to apply to determine RSALS'. It seems to me that this should be an issue explicitly discussed prior to Task 5. This, and the specification of acceptable probabilities are key issues and should not be buried.

A similar comment was received from Panel member Victor Holm. *RAC* will work with the Panel to provide a recommendation of the process to be used for the sampling protocol.

11. p. 18, 3rd full para. Is the 'Rocky Flats profile method' the same as the 'Vertical Soil Profile Method' or is it the 'Rocky Flats Method'. RAC should be consistent.

We will revise this paragraph to ensure consistency in the use of terminology.

12. p.24. 3rd and 4th paras. Here, it seems to me RAC is assuming a point value RSAL. Specify.

The nonparametric tests are based on a point value RSAL. As stated in previous responses, *RAC* will work with the Panel to provide a recommendation of the process to be used for the sampling protocol.

13. p. 24, last para. Here, mention is made of the possible use background concentrations. However, I noticed that Table 2-1, p.4 implies that background concentrations will not be considered. If the implication in Table 2-1 is true, then that should be reflected here as well.

Table 2-1 provides the current quality control sample collection frequency used at RFETS. The discussion in this table applies to field blanks and not background measurements. Blanks are typically used to make sure that contamination does not exist in the laboratory that may affect sample results.

14. p.25, 2nd para. RAC states that " The statistical tests are further defined by the method chosen for comparison to soil action levels". RAC should be more explicit about the differences in the statistical tests required by each method.

We will provide additional discussion on the statistical tests and their use in comparison to the action levels.

References

US DOE, US EPA, CDPHE (1996) Action Levels for Radionuclides in Soils for the Rocky Flats Cleanup Agreement Final (October 31, 1996).

Reviewer E

General comments

The basic objective of the report is not apparent. The opening paragraph states that it is to provide recommendations for sampling protocols for Rocky Flats. To a large extent, the report does not achieve that objective. It is not clear whether it simply intends to present a general introduction to elements that should be considered for soil sampling protocols, or, instead, a considered evaluation that is specific to the Rocky Flats situation. The report vacillates between these extremes, and fails to do either satisfactorily. Since the former purpose already is exhaustively addressed in existing reports, most notably the MARSSIM manual developed jointly by EPA, DOE, DOD, and NRC, it appears to this reviewer that the most useful objective would be to focus specifically on Rocky Flats, and to deal with general considerations only to the extent needed to provide background for recommendations for this site. Unfortunately, the report stops short of specific recommendations for Rocky Flats in many places. Examples include:

1) An important consideration is that more than one radionuclide will have to be considered in designing the sampling protocols (Section 3.2). The report fails to explore the implications of this for the specific radionuclides actually present at Rocky Flats, except indirectly in Sections 3.4 and 3.7. What is the variability in radionuclide ratios at Rocky Flats? How important are they? To what extent are the answers pathway dependent? And what are the implications for sampling protocols?

> The milestone for the draft Task 6 report was due before Task 5 (independent calculations); therefore, information on dominant pathways for each radionuclide were not available. The milestone for the final Task 6 report has been delayed until after completion of Task 5. This will allow *RAC* and the Panel to investigate additional areas of the sampling protocol in relation to the calculated action levels.

2) A fundamental choice for setting the number of samples required is the decision between use of parametric or non-parametric statistical analyses and, for the latter, between use of the Wilcoxon Rank Sum and the Sign tests. The report should make and support recommendations that are specific to the Rocky Flats site.

> *RAC* will provide specific recommendations in the revised report for the Rocky Flats site.

3) What areas at Rocky Flats fit, in the view of the authors, into MARSSIM's Classes 1, 2, and 3?

> Defining which areas of Rocky Flats fall within the Class 1, 2, or 3 categories is not within *RAC's* scope of work on the sampling protocols. However, the MARSSIM classification scheme begins with the assumption that all impacted areas being evaluated for release have a potential for radioactive contamination above the action levels. This means that all areas are initially considered to be Class 1 areas unless some basis for reclassification as Class 2, Class 3, or nonimpacted is provided. The basis for changing

the classification of an area depends on data obtained during previous investigations include characterization surveys. Determining the classification of areas at Rocky Flats again requires knowledge of the action levels, which will be determined in Task 5. For instance, all areas with contamination in excess of the action levels before remediation are Class 1 areas. Class 2 areas have contamination that are not expected to exceed the action levels. Class 3 areas are not expected to contain any residual radioactivity or are expected to contain levels of residual radioactivity at a small fraction of the action levels

4) What specific kinds of survey units should be considered in protocols for Rocky Flats, based on levels and homogeneity of contamination, relevant pathways, and other site conditions? For what depths should sampling protocols be considered, based on pathways that may be present at Rocky Flats (e.g., inhalation, ground shine, ingestion, and ground water)?

RAC will provide some guidelines on the factors to be considered in defining survey units. However, it would be futile to attempt to provide a basis for defining survey units that would encompass all possible variations at Rocky Flats. This is the purpose of the DQO process—to ensure that all factors relevant to a study are considered before sampling is initiated.

In terms of sampling depth, *RAC* is conducting the action level study for the surface soil (i.e., 15 cm). Subsurface pathways such as groundwater are not being considered at this time.

5) What is the recommendation regarding measures to assure completeness of successful data collection and analysis?

Recommendations on completeness are discussed in Section 3.10.1.6 of the Task 6 report.

6) What are the recommendations regarding data validation?

Data validation is discussed in Section 3.10.2 of the Task 6 report. Rocky Flats uses current standard guidelines for data validation as described in Section 2.4. As stated in Section 3.10.2 of the report, no national standard currently covers data validation of radiochemistry concepts adequately. However, *RAC* does not intend to attempt to develop such a standard under the soil sampling protocol scope of work.

The report continues to reflect, unnecessarily in the view of this reviewer, lengthy discussion about the role of RSALs vs. the regulatory dose limit (e.g., Section 3.2) -- that perhaps is attributable to the discussion in Report #2. While it is quite true that it is necessary to choose a value for the *degree of assurance* that the dose limit will be met (e.g., the 95% confidence level), and that this should invoke projections of the probability distributions of the doses attributable to various soil levels via various pathways, it is also true that at the end of the day the dose limit is

the bottom line that must be met, and that derived secondary quantities like RSALs will have to be chosen, albeit often with conservancy, that will assure that the primary dose limits are met. However, this process is not really relevant to the choice of soil sampling protocols, as discussed in Section 3.2. The relevant point for sampling protocols is whether or not the significant radionuclide ratios are constant in a sampling unit, and therefore do not require independent analysis.

The discussion in Section 3.2 was provided to the Panel to emphasis that several methods are available to compare the soil sample data to the action levels. In Section 3.2 we provided two methods that could be adopted for this study. As stated in the comments from Panel member Victor Holm, *RAC* will need to assist the Panel in making these technical decisions.

Additional page-specific comments

p. iii. para. 2: The "data quality objectives" process is advanced as the reason for recommending use of the RFETS sampling program without defining the process.

We will revise this paragraph.

para. 3: Many of these recommendations (e.g., "select survey units...,employ statistical methods...,establish quality assurance requirements...") are so obvious that they do not warrant mention unless the report makes specific recommendations on how they should be implemented.

RAC will provide additional discussion in the revised report to emphasize the importance of these recommendations in the final sampling protocol.

p. 3. Para. 2. The MARSIM manual was not developed by NRC, as the reference implies. It was developed by an interagency work group chaired by EPA, and included members from EPA, NRC, DOE, and DOD. I recommend the initial textual citation of this key reference should read "...EPA, NRC, DOE, and DOD *Multi-Agency Radiation Survey and Site Investigation Manual* methodologies (MARSSIM 1997).", that subsequent references read "MARSSIM (1997)," and that the citation in the list of references (Section 4) read "MARSSIM. 1997. *Multi-Agency Radiation Survey and Site Investigation Manual.* Final. NUREG-1575, EPA 402-R-97-016. EPA, DOE, DOD, and NRC." (The manual was printed with a NUREG number simply because NRC had funds readily available at the time the work group completed its work, not because it was an NRC work product.)

We will incorporate this comment into the revised report.

Para.3. Last two sentences. These statements appear to be contradicted on p.15.

The contradiction will be corrected.

Para.4. What are "controlled" documents and "procedural" titles?

The sentence will be corrected to remove "controlled" and "procedural."

p. 4 Table 2-1 entries appear without any introduction to the meaning of these terms, which are not explained until p. 29 (using, in some cases, different terminology). In addition, the footnote raises questions that are not assessed, the answers are simply asserted in terms that might apply anywhere (but obviously do not -- so what is so different about Rocky Flats, and does the author agree?).

RAC will add a discussion to explain the meaning of Table 2-1. As stated in previous responses, we will provide a discussion explaining what areas of the Rocky Flats program are adequate and those that are considered inadequate.

p. 4-6 Sections 2.2, 2.3, 2.4 cite lists of references without any evaluative comments.

See previous comment responses on these citations. We will provide additional discussion on those areas of the program that were considered adequate.

p. 9 HPGe surveys are cited w/o explanation.

A discussion of HPGe survey methods will be added to the report.

p. 12 Para. 3, final sentence. I should think that the purpose of including Oversight Panel members on the RFETS planning team would be to insure an *acceptable* protocol, not just to ensure "understanding" it.

RAC will change the meaning of the statement as noted by the reviewer.

Para 5. Pathway dependence is omitted in (3-1).

Equation (3-1) is a generalized equation for the sum-of-ratios calculation. The equation is stating that the sum of the soil concentrations divided by their respective action levels must be equal to or less than 1. We acknowledge that the action levels are derived from the various exposure pathways.

Final para. "A discrete soil action level value will not be provided for each radionuclide..." See general comment above. Surely they will, at the end of the day, and that is what the soil sampling protocols should be based upon. Generation of the distributions mentioned is essential to setting the action levels (along with the primary dose level and the level of confidence that it will be met), but their generation does not mean that discrete levels will not be chosen.

We agree that discrete levels will need to be chosen. This paragraph will be revised.

p. 13 Para. 3. What is a stochastic action level? How does one act on it? Stochastically?

The wording will be revised and changed from "stochastic soil action levels" to "soil action level distributions."

p. 15 Para.1. What does "...typically do not include the area of the survey unit in the calculations" mean?

The paragraph will be revised. The equations to determine the required number of sample points do not include the size of the survey unit. Therefore, the survey unit concept provides a method to ensure adequate coverage of the survey unit with sample locations.

Para. 3. The significance of the statement "However, the study was based upon a "spatial estimation..." is unclear.

The sentence will be deleted.

p. 17 Paras. 1 and 4. The statements "approximately 50% of the total inventory" are in the top 3 cm and "the data indicate that a large majority" is in the 0-3 cm depth profile are in conflict.

The term "large majority" will be revised to reflect that 50% of the total inventory resides in the top 3 cm of the soil.

pp. 18-19 The section on small areas of elevated contaminants needs more work. People are not exposed for a year, let alone a lifetime, to small areas. Averaging takes place, and should be accounted for in the modeling. To much time and expense is wasted of such unimportant fluctuations, which usually have no public health implication. *The need here is for development of a suitable area over which measurements should be averaged, and a protocol for identifying really "hot" spots, which pose a different kind of exposure hazard.*

The Panel has decided to further investigate the implication of hot spots in terms of averaging.

p. 19 Para. 3.5.2. Each sampling strata [stratum?] is more or less [more?] homogeneous than the sample [unit] as a whole.?

The words "or less" will be deleted from the sentence.

p. 20 Is it really useful to set out equations 3-3 and 3-4 to show a 7% difference in the number of samples required?

Equations (3-3) and (3-4) were only provided as examples of how to determine the spacing of a systematic grid pattern for either a triangular or square grid.

p. 21 It is not clear what the text and equations 3-5 to 3-8 contribute that could not be conveyed with a couple of short sentences. The reference to the sum of ratios rule following eq. 3-8 is unclear, since this rule normally applies to different radionuclides and pathways, not to different levels of contamination by the same radionuclide through the same pathway.

The discussion in Section 3.6 will be revised. Currently, *RAC* and the Panel are investigating hot spot identification and averaging issues.

p.24 Last sentence. This should be elaborated for the conditions at Rocky Flats.

Task 6 has been delayed until Task 5 is complete. This will allow for an evaluation of background radionuclide concentrations in comparison to the action levels.

p. 25 The discussion of the null hypothesis needs considerable elaboration (especially the last sentence, and not just for the mean).

See previous comment response to Reviewer A on the null hypothesis.

p. 26 The recommendation in 3.9 needs elaboration, both in terms of justification for the particular recommendation, and in terms of exploration of alternative means for verification of the results of the final status survey. Why are the usual EPA procedures not adequate for Rocky Flats?

RAC is not sure what the meaning of "usual EPA procedures" that the reviewer is mentioning. This section is discussing independent confirmatory investigations. This would require an independent contractor to review and evaluate the results of the study. This may also involve the collection of soil samples by the independent contractor to confirm the results of the study.

p. 32 Section 3.10.1.5. I suspect that the same measurement system for a radionuclide is only required on a pathway specific basis

> We would agree that comparability involves using the same measurement system for a media. For instance, the same measurement system would generally be used for water sample analysis to ensure that all water sample data could be compared.

Eq. 3-11 was particularly helpful so that the concept of % completeness could be taken into hand.

In summary, this is an unsatisfactory and often confusing draft. This reviewer is admittedly not an expert on soil sampling (nor, for that matter, on general sampling) protocols. However, the report should be intelligible to an educated reader, and to this one, in too many places, it is not. In addition to careful technical and usage editing throughout, it needs major rewriting and clarification in many places.

RESPONSES TO PANEL MEMBER COMMENTS

LeRoy Moore

As a lay person who has followed sampling "from a distance," I find this an impressive survey of the many aspects of the complex process of sampling. But the main body of the text seems more like just that -- a survey --rather than the recommendation of protocols for use in remediation to RSAL levels that needs to be made to DOE and its regulators. The RSALOP insisted on having RAC recommend sampling protocols because we discovered that at present protocols do not exist. This suggests that DOE and the regulators could do as they wished.

The Executive Summary seems in this regard an improvement over the actual text it purportedly summarizes, for the ES lists eleven specific recommendations made in the report. I suggest that a conclusion be added to the basic text repeating these eleven recommendations plus any that may, after comments, be added. Also the language within the text needs to specify that a recommendation is being made. As is, it's not always clear that the conclusion of each section culminates in an actual recommendation.

Re. the recommendations as listed in the Executive Summary, # 2 needs to be made as a real recommendation, not an item for consideration. Hence, change 2 to read: "Assess multiple radionuclides."

> *RAC* will add a conclusions section to the report. In addition, we will make sure all recommendations are stated clearly in the text.

3.3 Survey Units: This seems a helpful way to approach what is known re. areas of contamination, not so good for unknown hot spots. Would it be feasible to add doing a survey with a gamma detector to pick up americium deposits, by means of which Pu hot spots could be identified? This would results in identifying new Class 1 survey units. Some of this kind of survey has been done in the past. How complete is it? Does a record exist that is adequate? Or is more of this kind of work needed?

> Radiation detection surveys (scanning or in situ) are a necessary component of the sampling protocol as noted in the Task 6 report. Soil sampling alone will be inadequate for identifying hot spots. The future Panel workshop, on radiation detection instrumentation and applications to Rocky Flats, may answer the question of which instruments are useful for this purpose.

3.4 and 3.5 The point is made several times in these two sections that "composite samples are . . . unable to detect individual areas of elevated activity." For purposes of detecting areas needing remediation, composite samples in fact are no help at all. Shouldn't the protocols specify that all suspected hot spots will be sampled separately?

The Panel has decided to investigate hot spots in more detail. The results of these discussions will be incorporated into the revised report. In general, hot spots should be sampled to determine radionuclide soil concentrations and the extent of the hot spot.

3.6 Explain the sentence at the end of the first full paragraph on p. 21: "The area factor is the magnitude by which the concentration within the small area of elevated activity can exceed the soil action level while maintaining compliance with the dose limit." How could this be, if the action level corresponds with the dose limit?

A presentation by *RAC* was provided at the June Panel meeting regarding the area factor and applications to hot spots. Additional revisions will be made to Section 3.6 regarding hot spots based upon ongoing discussions with the Panel.

3.10.1.6 Completeness. If "completeness is not intended to be a measure of representativeness," how can the concept of "representativeness" be incorporated into the protocols? Doesn't it need to be?

Completeness is a term that refers to the percentage of measurements that are determined to be valid. Representativeness reflects the degree to which sample data accurately and precisely represent a characteristic of a population, parameter variations at a sampling point, or an environmental condition. Representativeness is a qualitative parameter that is primarily concerned with the proper design of the sampling program. Representativeness is most appropriately satisfied by being certain that a sufficient number of samples are collected and the sampling locations are carefully positioned at representative locations.

3.10.2 Data Validation. The final paragraph on p. 23 states the "no national standard currently covers data validation of radiochemistry concepts adequately," and that "each DOE site has developed site-specific data validation procedures for radiochemistry data." No reference is made (unless I missed it) to such procedures for Rocky Flats. Do they exist in some documentable form? Are they adequate for the task?

Data validation guidelines have been developed at Rocky Flats. Section 2.4 provides a summary of the guidelines available for various analyses. The quality of data validation is highly dependent on the person performing the validation. This involves having personnel that are familiar with the analytical method being used and the potential laboratory errors that are possible. Guidelines generally provide the basics of data validation; however, personnel with extensive field and laboratory experience provide assurance that data validation is performed correctly.

Victor Holm

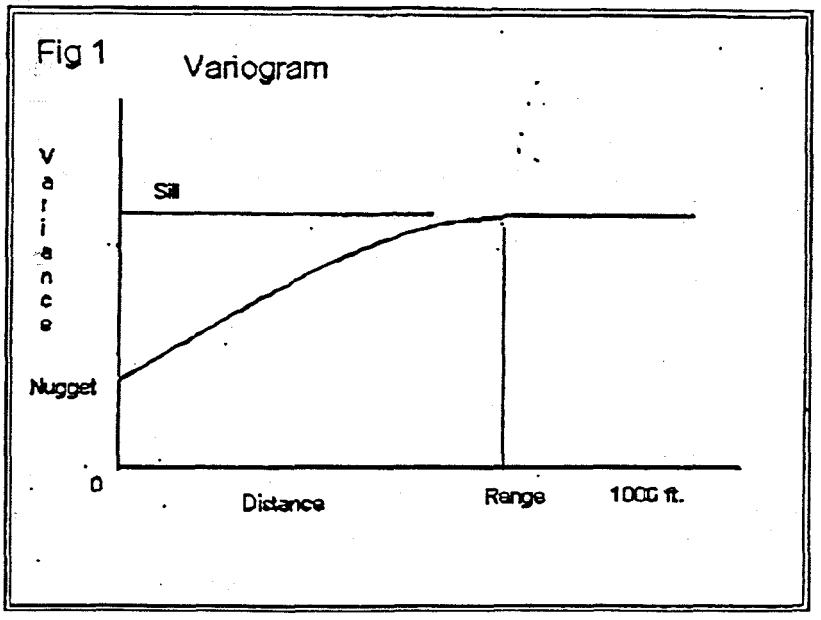
I could find little to fault in your Draft Task 6 Report. Your explanation of the existing soil sampling procedure and the Data Quality Objectives is very clear and for the first time I now understand DQO's. Section 3.2 on Multiple Radionuclide Considerations I found very interesting. You are correct in that ultimately it will be up to the panel to determine how the RSALs will be determined; but, it will up to RAC to help us through what may be a very technical decision. The confusion that some of the peer reviewers had with this concept indicates that it will require substantially more explanation than was possible in this report. What we need to determine is what is the practical effect of the two options.

> *RAC* will work with the Panel to provide information to help with these technical decisions.

I am still very confused about what is being proposed in sections 3.3 thru 3.8. The discussion on follows MARSSIM which I also had real problems with. I do not know if the problem is that I don 't understand what is being proposed or whether I disagree with the survey methods. I was instrumental along with others in including this task in the study because I did not agree with MARSSIM. After reading your report I am more confused than ever. Perhaps the easiest way for me express my doubts is to explain how I would approach the problem.

I have many years experience in open pit mining especially in the area of what is referred to as grade control. In most ore bodies there is a gradual diminishing of grade (metal content) as the boundaries of the ore body are reached. The purpose of grade control is to determine the limit of the economic ore. The grade that divides ore from waste is termed cutoff grade which is similar to the RSAL. One naturally does not wish to send waste to the mill or ore to the waste dump. The boundary between waste and ore is often not simple. Small high grade ore pockets may exist outside the main ore body .If real these small pockets may contain large profits. All of this is very similar to the problem of contamination. The consequences of committing type 1 and type 2 errors may not be same. The result of sending a high grade pocket to the waste dump may be much greater than sending waste to the mill. The mining industry has spent huge resources on these problems. Starting in the eighties much of this technology was transferred to pollution cleanup. The field of statistics involved is called geostatistics. The underlying principles of geostatistics is that samples in environmental systems are not independent and not normally distributed; as a result the use of classical statistics is invalidated. Naturally distributed trace elements whether in an ore body or in a contaminated plume are spatially correlated. That is samples taken near each other will tend to be have similar values. Samples further away will tend to have less similarity . In fact it is possible to plot a graph showing thc relationship between variance and distance. Such a graph is called a variogram (Fig 1). There is usually a variance between samples taken even at the same location: This variance is caused by sampling errors, analytical errors and very small scale variability . This variance is called the nugget. Samples taken further than some distance from other samples show no correlation, in other words, they are independent, this distance is called thc range. The variance at the range is called the sill and is generally about equal to the sample variance. This graph provides the basis for determine several

important parameters. First since it represents a relationship between variance and distance it can be used to select a sample spacing for a given confidence. It can also be used to determine the location where additional sampling will most increase the confidence. It also provides, through use of what is termed the spherical model, the value of the kriging variance. This parameter allows the use of kriging.



Kriging is a method of estimating the amount of contamination in a survey unit. It can be shown to be best linear unbiased estimate. Geostatistics has also developed a nonparametric method of estimation. This method is multiple indicator kriging. The sample values are first dived into classes, then a variogram is developed for each class. Each class is then kriged and then combined into a distribution, The result is a probability distribution of the contamination in each survey unit. This is a very brief synopses of what is a new but, expanding field of statistics.

Using geostatistics and my experience in grade control, I would approach the problem in the following manner. First I would select several different areas that represent the contaminated area and conduct orientation surveys at each site. This would consist of taking many closely spaced samples and using the best analytical methods available (no surrogates for instances). The results of this survey would be used to develop a variogram and the sample spacing for the main survey. To me the main effort should be applied to the characterization and not the final survey. The size of the survey unit should be determined not by the sample spacing or level of contamination; but, rather by the method of remediation to be used. It is analogous to the selective mining unit (smu) which is the smallest area for which a decision is made on whether it is ore or waste. The smu is determined mainly by the size of the mining equipment selected; it is this area that is kriged. In mining it is also normally about equal to the sample spacing. At this point surrogate or other analytical methods may be used, for instance HPG gamma instruments.

As an example lets say that a specially modified CAT 623E elevating scraper is to be used. The cut would then be about ten feet wide. . If the cut was 6" deep it would take about 100 feet to fill the scraper. This would then become thc selective remediation unit (sru). If the contamination in the unit is greater than the RSAL it would be remediated. If the contamination was less then it would be left unremediated. If there is a concern about small hot spots, the probability distribution of the activity within each sru can be used to determine the probability of hot spots existing. Using this method it is possible that a sru with a mean activity below the RSAL will still be remediated if the probability of it containing hot spots exceeds some confidence level.

Since the final status survey has already been done for those areas that were below the level for remediation only an independent confirmatory survey needs to be done. For areas that were remediated a new final status survey is required. Since the activity is no longer spatially correlated geostatistics is no longer required. The methods outlined in your report and MARSSIM could be used. It is very unlikely that any survey unit that was remediated will be found to exceed the action level after remediation.

I realize that remediation strategies are not within the scope of the contract and I do not wish you as the contractor to give a detailed evaluation of the ideas presented here. Rather I wish to stress the intimate connection between the characterization survey and remediation. If this survey is properly performed and the remediation process makes use of appropriate statistical methods the final survey becomes a simple quality assurance check.

> The sampling protocol will be directed toward the final status survey. Because this was not made clear in the beginning of the Task 6 report, we understand the confusion to which the reviewer is referring. After remediation of a contaminated area, the spatial correlations will probably not exist, thus, removing this consideration from use in the sampling protocol. However, it may apply to areas that have not been remediated (characterization data indicated the area was below the action levels) and for which a final status survey will be conducted. Ideally, the characterization survey and sampling is conducted to a level of quality that the data can also be used for the final comparison to the action levels. This is an area that the Panel may wish to be further investigate to account for final status surveys of unremediated areas.