ORISE
OAK RIDGE INSTITUTE FOR SCIENCE AND EDUCATION

April 29, 2010

Bruce Watson, CHP
U.S. Nuclear Regulatory Commission
Two White Flint North
11545 Rockville Pike
Mail Stop: T-8F37
Rockville, MD 20852-2738

SUBJECT: FINAL—INDEPENDENT TECHNICAL EVALUATION AND RECOMMENDATIONS FOR THE APPLICATION OF COMPOSITE SOIL SAMPLING IN DEMONSTRATING COMPLIANCE WITH RELEASE CRITERIA WHEN IMPLEMENTING THE MARSSIM GUIDANCE (RFTA NO. 10-005) DCN: 2012-TR-01-0

Dear Mr. Watson:

Enclosed is the technical evaluation of the composite sampling approaches proposed for implementation during the final status surveys at two sites:

1. The U.S. Department of Energy's West Valley Demonstration Project located in West Valley, NY and
2. The U.S. Army Corps of Engineers' Shallow Land Disposal Area (SLDA) Site located in Parks Township, PA.

Comments provided on the draft report have been incorporated into the final evaluation. You may contact me at 865.576.5073 or Erika Bailey at 865.576.6659 if we may provide additional information.

Sincerely,

Timothy J. Vitkus, CHP
Survey Projects Manager
Independent Environmental
Assessment and Verification

TJV:jc

Enclosure

c: T. Patterson, NRC/NMSS/TWFN 8A23
T. Carter, NRC/FSME/DWMEP T-8F5
File/2012
E. Bailey, ORISE
S. Roberts, ORISE

Distribution approval and concurrence:	Initials
Technical Review	EMB

P.O. Box 117
Oak Ridge, TN 37831

Voice: 865.576.5073 Fax: 865.241.3497 E-mail: Tim.Vitkus@orau.org

FINAL
INDEPENDENT TECHNICAL EVALUATION AND RECOMMENDATIONS FOR THE APPLICATION OF COMPOSITE SOIL SAMPLING IN DEMONSTRATING COMPLIANCE WITH RELEASE CRITERIA WHEN IMPLEMENTING THE MARSSIM GUIDANCE

Phase 1 Final Status Survey Plan, West Valley Demonstration Project* (WVDP), West Vally, New York *and;
Shallow Land Disposal Area* (SLDA*) Site Final Status Survey Plan

At the request of the U.S. Nuclear Regulatory Commission (NRC), the Oak Ridge Institute for Science and Education (ORISE) performed a technical review of two final status surveys (FSS) plans submitted to the NRC. These plans contained composite sampling approaches proposed for implementation during the FSS at two different sites. The approaches and applicable sites are described in the following two documents—referred to in this report as the *Plans* (WVDP and/or SLDA):

1. Argonne National Laboratory. *Phase 1 Final Status Survey Plan for the West Valley Demonstration Project in West Valley, New York*. Argonne, IL; December 16, 2009
2. U.S. Army Corp of Engineers. *Shallow Land Disposal Area (SLDA) Site Final Status Survey Plan, Parks Township, Armstrong County, Pennsylvania*. December 2009.

The review of the composite sampling and analysis approaches that are being proposed to demonstrate compliance with both the derived concentration guideline levels ($DCGL_W$s) and the $DCGL_{EMC}$s can meet several of the data quality objectives discussed in the *Plans*. However, the ORISE review identified several technical deficiencies that were not specifically addressed in the *Plans*. The *Plans* also do not provide adequate detail or otherwise do not include the site-specific information that is necessary to independently evaluate multiple areas of concern.

The discussion of additional considerations will be provided below in the form of examples as to how the composite sampling approach could be applied to a final status radiological survey. The examples will be for 1) a standard MARSSIM final status survey that involves a gamma-emitting contaminant and the relationship between scan sensitivity and sample spacing in Class 1 survey units and 2) a non gamma-emitting contaminant (Sr/Y-90 as is the case for the example used in the WVDP *Plan*). Many of the aspects that are discussed for this Sr/Y-90 example are also of concern in the SLDA *Plan*.

Composite Sampling Overview

Overall, the application of a composite sampling approach will meet a number of the objectives detailed in the *Plans*. Evaluation of the approach was based on guidance and recommendations detailed in the following references:

1. U. S. Environmental Protection Agency. *Guidance on Choosing a Sampling Design for Environmental Data Collection for Use in Developing a Quality Assurance Project Plan, EPA QA/G-5S*. Washington, DC; December 2002.

2. U. S. Environmental Protection Agency. *EPA Observational Economy Series; Volume 1: Composite Sampling, EPA-230-R-95-005*. Washington, DC; August 1995.
3. U.S. Nuclear Regulatory Commission (NRC). *Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM), NUREG-1575; Revision 1*. Washington, DC; August 2000.
4. ASTM International. *Standard Guide for Field Subsampling for Environmental Waste Management Activities, Designation: D 6051-96*. West Conshohocken, PA. October 1, 2006.
5. R. O. Gilbert. *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold, 1987.

This examination first evaluated the conditions under which a composite sampling approach is appropriate and would be considered advantageous, and concurrently evaluated the disadvantages associated with composite sampling. These are summarized in Table 1.

Table 1:
Composite Sampling Overview

Advantages	Disadvantages
Reduces analytical costs.	Should not be used for establishing surrogate ratios.
Provides better estimate of mean concentration in the study area.	Lost information is a concern when testing to determine if a substance exceeds a threshold (dilution).
Identifying units that have the highest contaminant levels.	Cannot be used when action levels (DCGLs) are near analytical detection limits.
	Lost information when temporal or spatial variability is a concern.
	Cannot be used when integrity of individual sample values change, such as loss of volatile contaminants, due to the physical compositing mechanism.
Uses and Considerations for Applying Composite Sampling	
Useful when the size of the pattern or feature of interest is smaller than the spacing between sampling locations.	
User must account for potential introduction of large additional errors due to heterogeneous nature of the contaminant in the matrix, or the matrix itself.	
Aliquots used to form the composite must be of equivalent weight/volume and the individual aliquots and the composite itself must be well homogenized.	
Must account for the dilution factor when evaluating the result against a threshold, most commonly a hotspot or legal action threshold.	
In most cases, the user must maintain the ability for re-testing of individual samples that made the composite to retrieve potentially lost information.	

Discussion of Review Results

The reviewer evaluated the technical basis for composite sampling in the *Plans* based on Table 1 information. It is the reviewer's interpretation of the technical basis appended in the WVDP *Plan* that the primary focus was to provide a means for reducing the probability of Type II error, or

failing to reject the null hypothesis (H_0) which states that the survey unit exceeds the release criteria. Overall, it is the reviewer's opinion most tenets provided in the technical basis are sound in the proposed approach. However, the neither *Plan* clearly accounts for ensuring that a Type I error does not occur, that is, rejecting the H_0 when it is true. This opinion is based on several factors. These are:

1. The documents lack specific information as to how the presence of multiple radionuclides will be accounted for in the final status survey plan and reporting, other than committing to apply the sum-of-ratios (SOR) calculation to the analytical results. The *Plans* must include SOR considerations in up front planning. This information should include:
    a. Added detail as to how the presence of multiple radionuclides of interest (ROIs) will impact the necessary number of samples to account for hot spots in Class 1 survey units, when the actual scan minimum detectable concentration (MDC) is greater than required scan MDC for many of the radionuclides. The WVDP *Plan* mistakenly compares the Table 5 scan MDC to the CG_{EMC}s listed in Table 1. The CG_{EMC}s listed in Table 1 are for a 1 m^2 area. Therefore, unless sample spacing will be no more than that, many of the listed MDCs are not adequate.
    b. Furthermore, the *Plan* does not account for the impact of the proposed composite sampling approach on any of the ROIs except for Sr-90. Of particular significance are those ROIs with low DCGLs such as I-129 and Np-237. The combination of the analytical limit on detection together with the low DCGL will undoubtedly result in lost information as to when an action level has been exceeded.
    c. Lastly, the WVDP *Plan* has not addressed the possible reduction in the dose criteria that should be included in adjusted DCGLs and therefore action levels to account for those radionuclides contributing less than 10% of the dose.
2. The *Plans* do not adequately address accounting for residual hot spots. One of the primary deficiencies related to hot spot considerations is that discussed above in item 1a. There are a number of other considerations that are unaccounted for. One of which is the threshold investigation level. It is the reviewer's opinion that the investigation thresholds discussed in the WVDP *Plan* will not identify potential residual contamination, not only from the unacceptable scan MDCs but also could result from the composite sample strategy. Both the overall composite sample approach and the investigation thresholds are the primary focus in the remainder of this report.
3. The decision logic in both *Plans* is not appropriate for investigations for Class 2 and 3 survey units. That is, the WVDP *Plan* discusses comparing composite results to the CG_W times an appropriate area factor. Class 2 and 3 results must be compared to a fraction of CG_W. This is necessary to prevent missing contamination in a Class 2 or 3 area due to the dilution of the individual samples and therefore overlook misclassification of a survey unit. An almost identical oversight was also identified in the SLDA *Plan*. That plan also proposes to use composite sampling in Class 2 and 3 areas. Composite sampling is not recommended in Class 2 and 3 survey units because of the lost information is a concern when testing to determine if a substance exceeds a threshold (dilution). If classified correctly, samples from Class 2 or 3 survey units should never exceed the $DCGL_W$. By compositing, information will be lost as to whether the $DCGL_W$ has been exceeded. With the SLDA $DCGL_W$ for Th-232 of 1.4 pCi/g—without even accounting for reduction in this level once the SOR is applied— accounting for a reduced threshold will require reanalysis of individual aliquots from virtually every sample location as will be seen in the discussions that follow.

The composite sample approach discussed in the *Plans* is highly suspect in its ability to adequately assess the final status radiological conditions. This is a direct result of site conditions and objectives being subject to 3 of the 5 composite sample disadvantages listed in this report's Table 1 overview.

This report provides further details and possible solutions, in the form of examples, to more clearly demonstrate where inherent issues may arise with the *Plans'* proposed composite sampling. The parameters/assumptions that are to be used in the example are provided in Table 2. Because several sections of the plan are non-specific and therefore lack all necessary information for direct evaluation, the table cites those parameters that were either taken directly from the WVDP *Plan* or calculated based on information provided in the WVDP *Plan*, or indicates that the parameter is for example only to allow for completion of the evaluation.

Table 2:
Evaluation Parameters/Assumptions

Parameter	Source
Minimum number of systematic samples: 8	Section 4.7, page 27
Aliquots per composite location: 5	Section 4.3, page 22
Subsurface soil $DCGL_W$ Sr-90: 130 pCi/g	Table 3, page 21
Sr-90 Area Factor 250 m^2 = 1.1	Not available in the WVDP *Plan*. Extrapolated from Table 3 values
Sr-90 Area Factor 92 m^2 = 2.6	Table 3, page 21
Sr-90 Area Factor 50 m^2 = 31.6	Not available in the WVDP *Plan*. Interpolated from Table 3 values
Sr-90 Area Factor 20 m^2 = 45	Not available in the WVDP *Plan*. Interpolated from Table 3 values
Sr-90 Area Factor 1 m^2 = 56	Table 3, page 21
Sr-90 scan MDC: --	Table 5, page 58
Am-241 scan MDC: 30 pCi/g	Table 5, page 58
Subsurface soil $DCGL_W$ Am-241: 15 pCi/g	Example only
Am-241 Area Factor for 250 m^2 = 2.5	Example only
Subsurface soil $DCGL_{EMC}$ for 250 m^2 Am-241: 37.5 pCi/g	Example only
Am-241 Area Factor for 50 m^2 = 5	Example only
Subsurface soil $DCGL_{EMC}$ for 50 m^2 Am-241: 75 pCi/g	Example only
Modeled survey unit size: 2000 m^2	Table 3, page 22
Modeled hot spot sizes: ~50 and 20 m^2	50 m^2 based on proposed composite plan of 8 systematic locations, 5 aliquots, distributed over 2000 m^2. 20 m^2 is an example reduction in hot spot size.
Survey unit classification: Class 1, 2 and 3	Example and WVDP *Plan* discussions

Survey Planning Example Evaluation

Example 1

First a review will be presented of a "typical" MARSSIM survey plan where hot spots must be considered in the sample spacing which will then be compared with the proposed composite approach. In this example, assume Am-241 is the contaminant of concern in a Class 1 survey unit and using the DQO process, the plan results in eight systematic soil samples required for the statistical evaluation. However, the required and actual scan MDC must be evaluated to account for potential hot spots. Sample spacing is 250 m^2. From the table of parameters, the actual scan MDC is less than the required scan MDC. Therefore the regulatory agency can be assured that this sample design plan is adequate to demonstrate compliance for both the $DCGL_W$, hot spots >250 m^2 in area concentrations between 15 and 30 pCi/g, and any residual hot spots <250 m^2 that are at concentrations >30 pCi/g. Systematic soil sampling will identify the lower activity, large area hot spots and higher activity, smaller area hot spots of concern will be identified through surface scanning and/or sampling.

Example 2

The second example will use the same general scenario as Example 1, with the following exception: a less sensitive instrument with a scan MDC of 40 pCi/g will be used for scanning, yet available funding requires that sample sizes remain essentially equivalent to the original plan. Therefore, composite sampling and retesting are factored into the plan. Because the sample design of 8 systematic samples is no longer adequate to address hot spots of concern, the sample density must be increased. Using the approach provided in the WVDP *Plan*, 8 composite samples will be acquired with each composite consisting of 5 aliquots. Therefore, each aliquot for the composite sample represents an area of 50 m^2—assuming equidistant spacing—and the composite in total represents the original 250 m^2 area. Because the composite better represents the total area than a discrete sample, the hot spot of concern is now ≤50 m^2 with Am-241 concentrations >75 pCi/g. Again large area, low activity hot spots are addressed via sampling, with some caveats, to be further defined, and small, higher activity hot spots identified through scanning. This same approach can be continued as a function of decreasing scan sensitivity by increasing the required number of systematic composite samples, whereby each aliquot will represent a smaller and smaller area. Individual samples that comprised the composite sample must be retained for re-testing of the individual aliquots when a threshold concentration is exceeded. This analysis of discrete samples will enable determining whether a $DCGL_{EMC}$ has been exceeded and better locate the area of concern for further investigation.

From Table 1, two of the composite sampling overview advantages were beneficially used. Specifically, these are reduced analytical cost and a better estimate of the mean. However, there remains a concern with exceeding a $DCGL_{EMC}$ threshold. Because of the dilution factor inherent to composite sampling, the investigation threshold must be reduced accordingly. The survey plan would then specify that if this pre-determined composite sample concentration threshold were to be

exceeded, the individual samples making up the composite would each be analyzed. This step is necessary to evaluate what information has been lost because of compositing that would have been retained by collecting and analyzing individual samples.

An example of re-testing individual aliquots when a threshold is exceeded is provided in the following iteration:

1. Determine the most conservative *a priori* investigative threshold.
2. This threshold is a function of the number of aliquots in each composite, i.e. a dilution factor.
3. For the example above, the composite sample analysis will readily provide sufficient information as to whether the $DCGL_{EMC}$ for the 250 m^2 area represented by the sample has been met. That is, an analytical result less than 37.5 pCi/g.
4. The Am-241 concentration in a composite sample is 16 pCi/g.
5. This value is less than the scan MDC but greater than the $DCGL_W$. There were no anomalies identified during gamma scanning.
6. The concentration is less than the allowable $DCGL_{EMC}$ for both a 50 m^2 area and the 250 m^2 area.
7. The *a priori* threshold determined should be the $DCGL_{EMC}$ for 50 m^2 divided by the number of aliquots in the composite or 75pCi/g/5 = 15 pCi/g. This assumes a contaminant of concern is not present in background or at very low levels. Therefore the activity in the composite could all be contributed by a single aliquot or contributed from several to all of the aliquots. This is the primary information lost when compositing.
8. Therefore, the sample described potentially exceeds the hot spot limit and further analysis is necessary.

NOTE: The above example is one of the primary issues identified in the WVDP *Plan* based on the review of Section 7.5 and 7.6. Also of significant note, is the discussion in Section 7.5 where the document states, "The generic process for demonstrating CG_{EMC} compliance is the same for Class 1, Class 2, and Class 3 units." Compliance with the CG_{EMC} should never be a consideration in Class 2 and Class 3 units; else the units have been misclassified. Furthermore, the investigative threshold of composite samples discussed above must be further reduced in Class 2 and 3 units to ensure there are no locations above the $DCGL_W$. In the above example the threshold would be 15 pCi/g/5 or 3 pCi/g above which there could be contribution to the total activity from a hot spot that has been diluted, thereby requiring re-testing of individual aliquots. Similarly, in the SLDA *Plan*, the Th-232 investigation level would be 0.3 pCi/g above background.

<u>Example 3</u>

The third and final example is presented to illustrate the deficiencies in the proposed composite approach for Sr-90 detailed in Section A.3.1 of the WVDP *Plan*. This example validates a number of the conclusions reached in the WVDP *Plan* for correctly identifying a hot spot of

specified size at a given proportion. First, a virtual evaluation of the initial sampling plan using 8 discrete samples distributed in a random-start/systematic manner to detect a hot spot of 50 m^2 in a 2000 m^2 area was examined using Visual Sample Plan (VSP) v. 5.9. There is no scan MDC for Sr-90 and a surrogate relationship cannot be established. The probability of at least one sample being randomly placed within the hot spot boundary was estimated based on 10 applications of the VSP random-start/systematic sample placement. The results are shown in Table 3.

Table 3:
Illustration of Random Sampling Location Probability to Identify a 50 m^2 Hot Spot

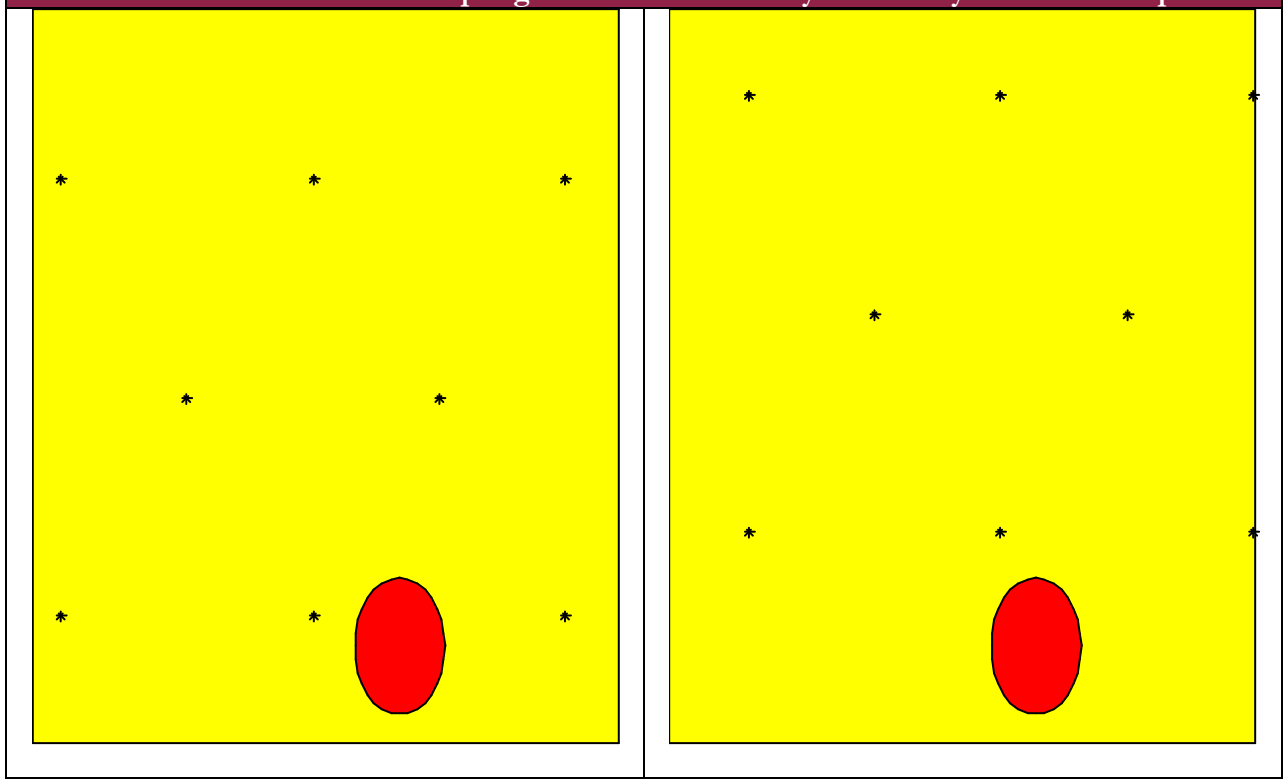
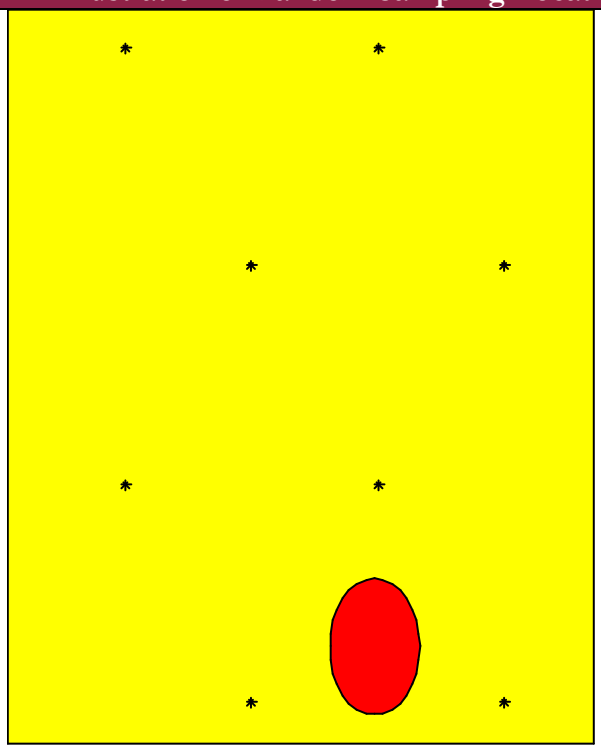
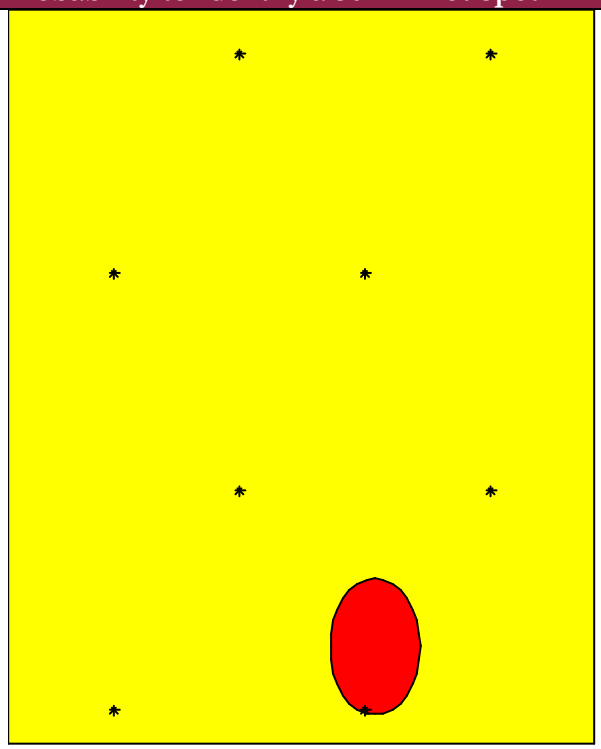
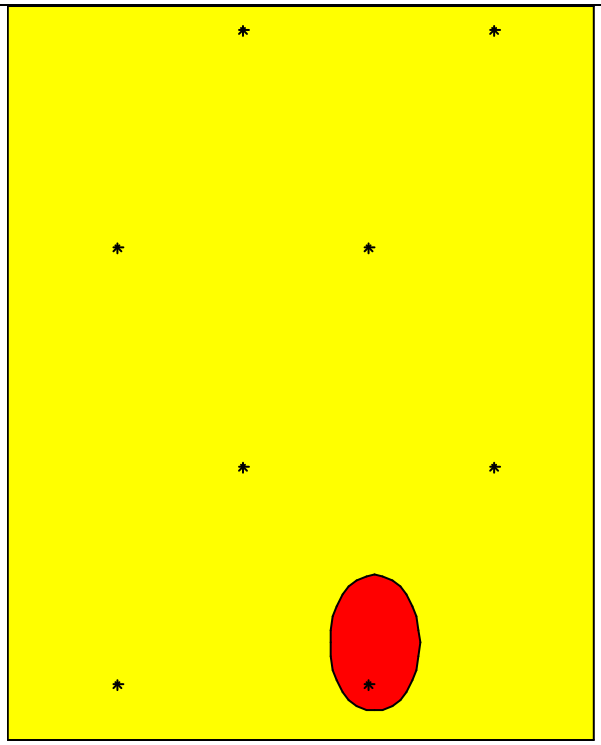


Table 3:
Illustration of Random Sampling Location Probability to Identify a 50 m^2 Hot Spot







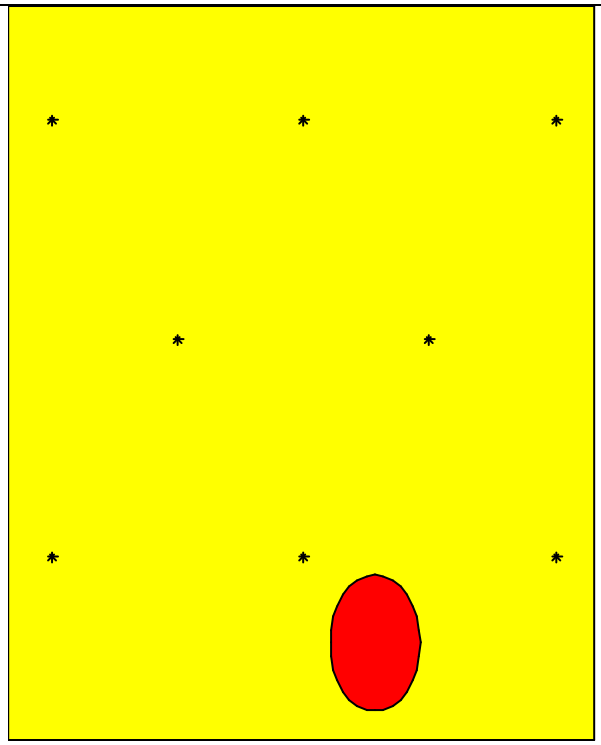
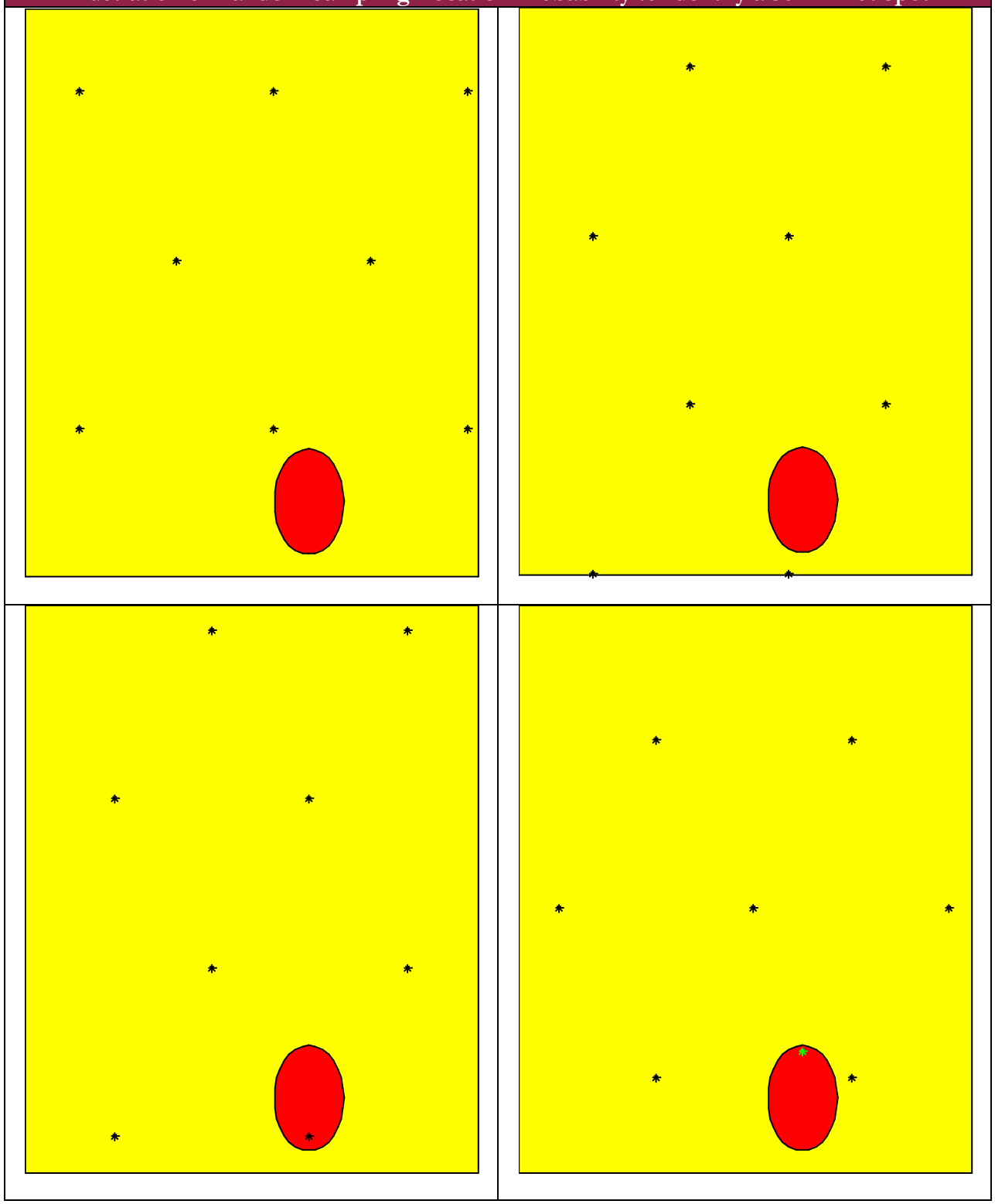


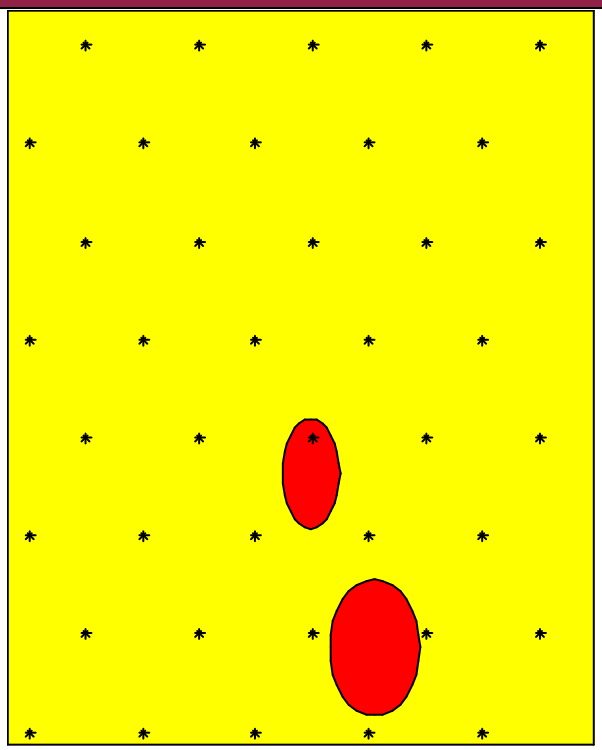
Table 3:
Illustration of Random Sampling Location Probability to Identify a 50 m^2 Hot Spot



This example provides an estimated probability where the hot spot is identified in 4 out of 10 tries. A simplified binomial distribution can be established where either the hot spot is successfully sampled or not sampled. The probability of success that a sample location will be within the hot spot is approximately 0.2 as each sample location represents a 250 m^2 area and the hot spot can occupy up to 50 m^2 of that area. The cumulative probability of successfully locating the hot spot more than 40% of attempts is 0.03. This examination clearly illustrates that additional samples in the form of discrete or composite samples are necessary to increase the probability of successfully identifying a hot spot of the example dimension.

Table 4 illustrates the same 10 applications of a random-start/systematic sampling strategy where the number of sample points is increased to 40 locations. Also added to further illustrate the discussion that will follow is the inclusion of a second, smaller hot spot of 20 m^2.

Table 4:
Combined Illustration of Random Sampling Location Probability to Identify either a 50 m^2 or 20 m^2 Hot Spot



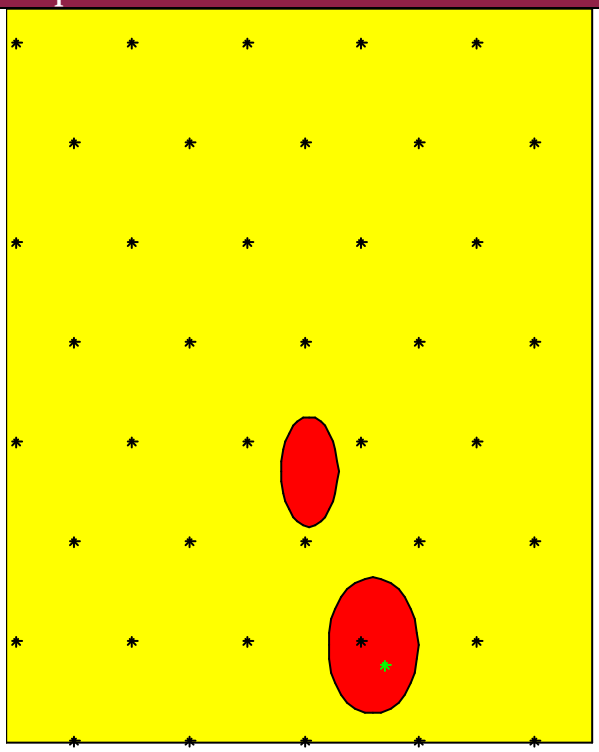


Table 4:
Combined Illustration of Random Sampling Location Probability to Identify either a 50 m^2 or 20 m^2 Hot Spot

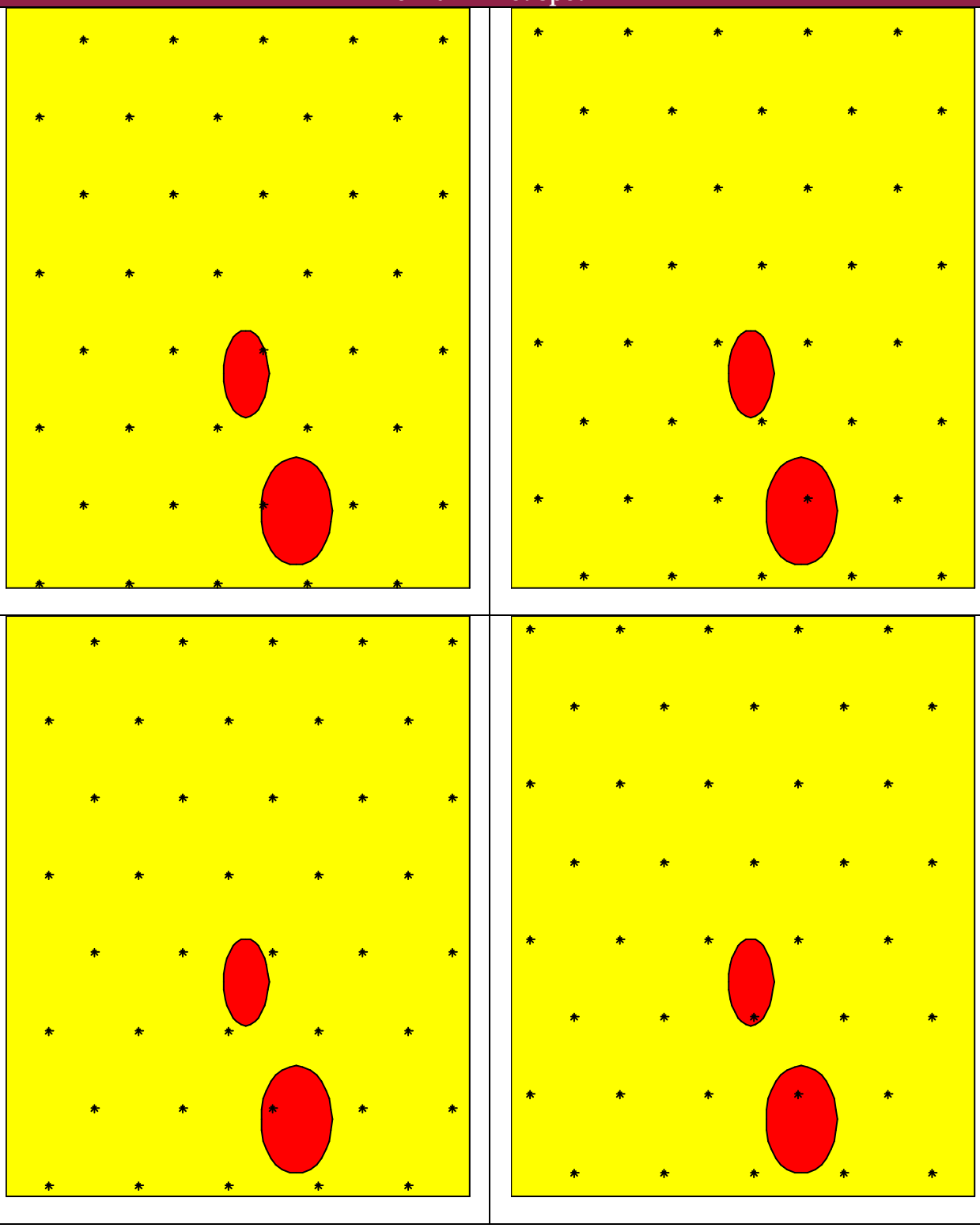
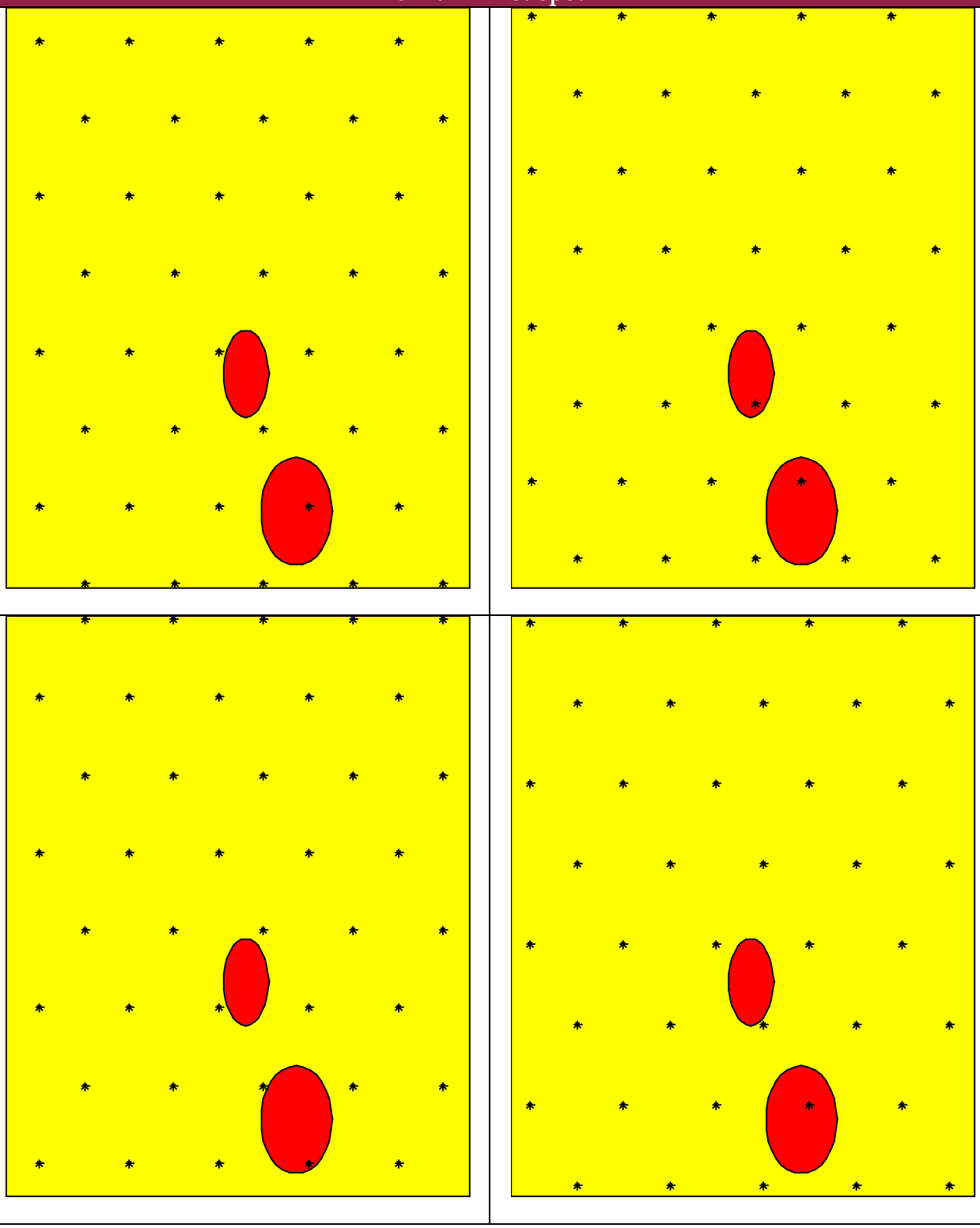


Table 4:
Combined Illustration of Random Sampling Location Probability to Identify either a 50 m^2 or 20 m^2 Hot Spot



For this illustrated increased sample density, the larger hot spot is sampled in virtually every case, thereby providing a high confidence level that this condition is identified. However, when reducing the hot spot area of concern by a factor of 2, yet increasing sample density by a factor of 5, the probability of identifying the smaller hot spot is again less than 0.5.

In a "typical" final status radiological survey, the reduced probability of identifying hot spots is accounted for through the iteration of adjusting sample spacing to satisfy required scan MDC sensitivity. The approaches discussed in the *Plans*, where the contaminant of concern is a hard-to-detect radionuclide or have low DCGLs relative to the investigation level and laboratory MDC, does not address this issue adequately. Recommendations for consideration are provided below.

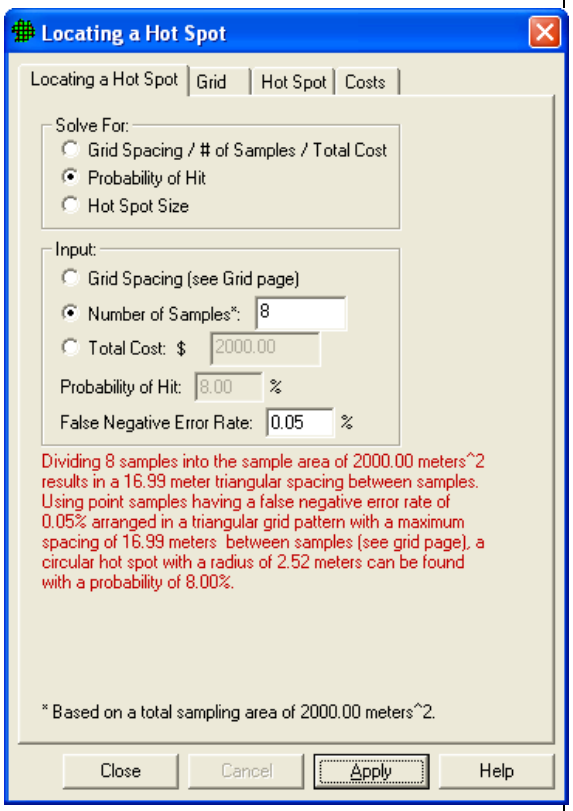
Recommendations

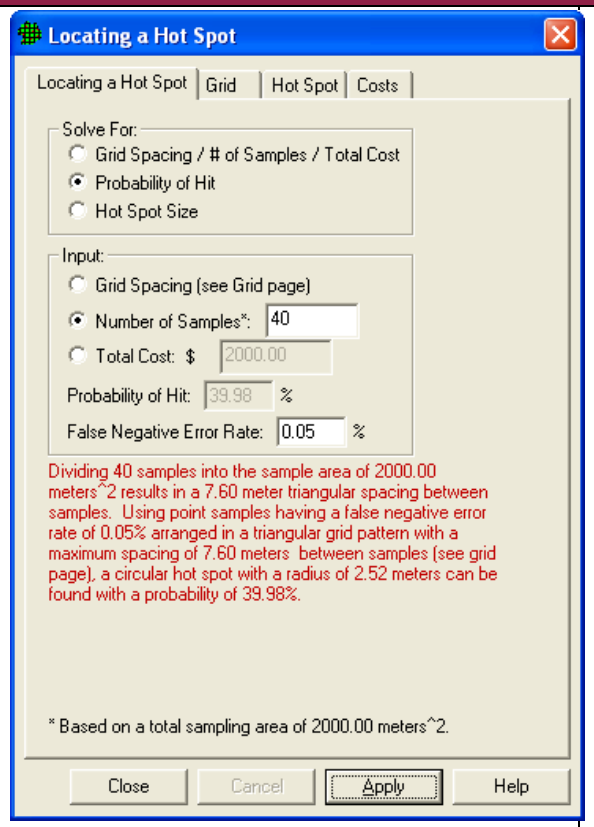
The site-specific conditions involved and discussed in the *Plans* are such that the MARSSIM guidance cannot be followed specifically when the scenario involves a hard-to-detect radionuclide in soil and surrogate relationship cannot be established, or involving very low DCGLs relative to background. For this situation, the reviewer recommends a technical justification that involves the adaptation of a probability-based design for locating hot spots. However, before preparing such a design, an *a priori* hot spot size of concern and the associated $DCGL_{EMC}$ would have to be determined. Once the hot spot size is determined, then either a discrete or composite sampling approach can be applied to provide a high level of probability that the hot spot will be sampled. There remains yet a problem associated with this *a priori* hot spot. The problem is how to select the *a priori* size as there will likely remain stakeholder concerns for yet smaller hot spots that could again be missed as seen in the Table 4 illustration. Therefore, the technical justification is encouraged to also include additional dose modeling details as to the impacts from any other small hot spots that could go undetected and potential contribution to the total dose from all remaining source terms across the site.

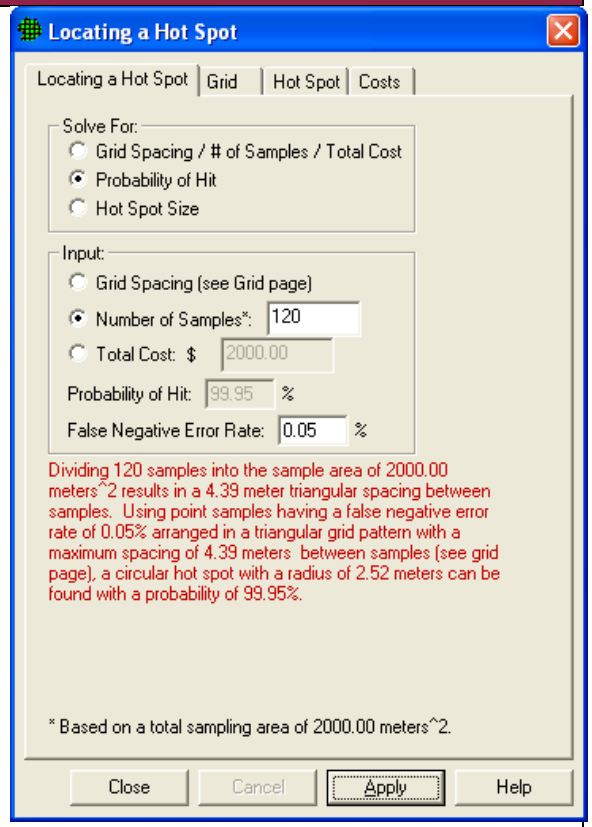
The above recommended solution must stem from an adequate site characterization. With a properly planned characterization, both the maximum contaminant concentrations should be identifiable as should the smallest of the contaminated areas. These data points can then be combined with area factor tables to determine the minimum hot spot area of concern. The site's technical justification should include the highest concentration that could reasonably be encountered. This information could be obtained from characterization surveys or investigations during remedial action support surveys. Also, this concentration may then be used to determine the associated allowable hot spot size that corresponds to that concentration. The spacing of the systematic composite samples could then be derived from that information using a probability-based sampling design.

The recommendation was evaluated using the VSP sampling design function for locating a hot spot. The input parameters were a 20 m^2 hot spot, a 2000 m^2 survey unit, and a false negative rate of 0.05%. The number of samples was input as 8, 40, or then 120. The probability of identifying the hot spot was 8%, 40%, to >99% for the 8, 40, and 120 sample plans, respectively. The screen shots are shown in Table 5.

Table 5:
VSP Screen Shots: Probability of Locating a Hot Spot as a Function of Sample Size







Conclusion

With this information, a robust final status survey soil sampling plan can be developed that will address not only the average residual concentrations across the survey unit but hot spot considerations as well. The hot spot issue can be addressed adequately by first generating sufficient characterization data to model the smallest hot spot of concern. The size of the hot spot will then determine the probability success rate for identifying the hot spot. For situations similar to the example discussed in the recommendation section, where a large number of samples was required, the corresponding increase in analytical costs can be offset with composite sampling. The number of composite samples would be expected to range from 12 to 40 in the example, dependent upon the number of aliquots to form each composite—between 3 and 10. Normally, no more than 10 aliquots should form a composite, which minimizes the range at the aforementioned 12. The number of aliquots must be carefully considered as well, and will be directly related to DCGL, the analytical detection limit, and the investigative threshold. Additional guidance on selecting the optimum number of aliquots for a composites is discussed in the references: *EPA Observational Economy Series; Volume 1: Composite Sampling, EPA-230-R-95-005* and Gilbert's *Statistical Methods for Environmental Pollution Monitoring*.

In summary, the reviewer recommends that the *Plans* be revised, with particular emphasis on the composite sampling approach. This technical justification should provide sufficient detail to allow for an independent evaluation, expand on the impacts of lost information especially as it relates to the investigation thresholds, and further evaluate the methods and procedures necessary to adequately account for multiple radionuclides of interest both in planning and results evaluations.