

# Advanced MARSSIM Topics HPS Annual Meeting Providence, RI

Eric W. Abelquist; eric.abelquist@orau.org Oak Ridge Associated Universities



### Introduction

- Advanced MARSSIM Topics
  - Building surface survey design for multiple contaminants; DQA using unity rule
  - Soil surface survey design for multiple contaminants in Class 1—scan MDC implications
  - Double Sampling if survey unit fails statistical test



### Final Status Survey

- Objective: Demonstrate that residual radioactivity in each survey unit satisfies release criteria
- Builds on data from HSA and survey results from scoping and characterization
- Survey design includes 1) scans to identify hot spots and 2) random (statistical) samples for determining average contamination levels in survey unit



### Final Status Survey (cont.)

- Null hypothesis (H<sub>0</sub>): Residual radioactivity
   exceeds the release criteria
  - H<sub>0</sub> is treated like a baseline condition, assumed to be true in the absence of strong evidence to the contrary
- Decision errors occur when H<sub>0</sub> is rejected when it is true (Type I), or when H<sub>0</sub> is accepted when it is false (Type II)



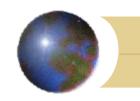
### Final Status Survey (cont.)

- Two statistical tests are used to plan and evaluate final status survey data
  - Wilcoxon Rank Sum (two-sample test)
  - Sign Test (one-sample test)
    - 1) When the contaminant is **not** in background, or is present at small fraction of DCGL
    - 2) When surface activity assessment performed with average background subtracted from each measurement

### Sign Test Example for Surface Activity Assessments

- Concrete floor potentially contaminated with Am-241, Co-60, Cs-137 and SrY-90
- Class 2 survey unit area is 320 m²
- DCGLs: 130 dpm/100 cm<sup>2</sup> Am-241

   11,400 dpm/100 cm<sup>2</sup> Co-60
   34,400 dpm/100 cm<sup>2</sup> SrY-90
   44,000 dpm/100 cm<sup>2</sup> Cs-137
- Separate alpha and beta measurements with gas proportional detectors



### Gross Activity DCGLs

- Gross alpha DCGL same as Am-241
- Gross beta DCGL must consider relationship of Co-60, SrY-90 and Cs-137
- Results of characterization provides relative ratios: 0.3 Co-60, 0.2 SrY-90 and 0.5 Cs-137

$$DCGL = \frac{1}{0.3/11,400 + 0.2/34,400 + 0.5/44,000} = 23,000 \, dpm/100 \, cm^2$$



- Background levels are assessed for each surface material encountered in the survey unit (natural radioactivity from Th, U, K-40)
- Group surface types with like background levels (drywall, steel, wood)— as opposed to individual backgrounds for each
- Obtain 10 to 20 measurements of background across the surface material



### Survey Instrument DQOs

- Gas proportional detector used for surface activity measurements
- Static MDC (in dpm/100 cm<sup>2</sup>) should be less than 50% DCGL

$$\frac{MDC = 3 + 4.65\sqrt{C_B}}{\varepsilon_i \ \varepsilon_s T(P.A./100)}$$

### Survey Instrument DQOs (cont.)

- Assume  $\varepsilon_i$  was determined for each beta contaminant: 0.41, 0.59, and 0.46 for Co-60, SrY-90, and Cs-137, respectively
- Determine weighted total efficiency:

rad fr	action	<u> </u>	<u> </u>	weighted $\varepsilon_{\text{tot}}$
Co-60	0.3	0.41	0.25	0.031
Sr/Y-90	0.2	0.59	0.5	0.059
Cs-137	0.5	0.46	0.5	<u>0.115</u>
	Total efficiency			0.205

### Survey Instrument DQOs (cont.)

- Assume average background counts on concrete floor is 360 cpm
- Calculate beta MDC for a 1-min count:

$$MDC = \frac{3+4.65\sqrt{360}}{(0.205)126/100} = 350 \ dpm/100 \ cm \ 2$$

Is this MDC less than 50% of DCGL<sub>w</sub>?
 (yes, gross DCGL was 23,000 dpm/100 cm²)

### Survey Instrument DQOs (cont.)

- $\bullet$  Assume  $\varepsilon_i$  for alpha measurements calibrated to Th-230 is 0.44, and background is 2 cpm
- Alpha MDC is calculated:

$$MDC = 3 + 4.65\sqrt{2} = 69 \ dpm / 100 \ cm \ 2$$
  
 $(0.44)(0.25)126 / 100$ 

Alpha MDC is just slightly greater than 50% of DCGL for 1-min count



### Sign Test Example—DQO Inputs

- Unity rule is used for survey design
- Characterization data used for planning:

Survey Unit (in cpm) Reference Area

Gross  $\alpha$  8  $\pm$  5 (1 $\sigma$ )

 $2.0 \pm 0.4 \ (1\sigma)$ 

Gross  $\beta$  1544  $\pm$  562 (1 $\sigma$ )

 $360 \pm 45 \ (1\sigma)$ 

Type I and II decision errors set at 0.05

# Convert gross activity DCGLs to cpm using efficiencies

Gross alpha DCGL:

 $(130 \text{ dpm}/100 \text{ cm}^2)(0.44)(0.25)(126/100) = 18 \text{ cpm}$ 

Gross beta DCGL:

 $(23,000 \text{ dpm}/100 \text{ cm}^2)(0.205)(126/100) = 5940 \text{ cpm}$ 

Since Unity Rule is used: DCGL = 1



### Calculate the relative shift— $\Delta / \sigma$

LBGR is set at expected concentration:

$$(8-2)/18 + (1544 - 360)/5940 = 0.53$$

 Variability for measurements should consider that Sign test involves subtracting mean background from gross measurement:

$$\sigma$$
 total =  $\sqrt{\sigma_s^2 + \sigma_r^2}$ 



### Measurement Variability

Gross alpha variability

$$\sigma total_{\alpha} = \sqrt{5^2 + 0.4^2} = 5 cpm$$

Gross beta variability

$$\sigma total_{\beta} = \sqrt{562^2 + 45^2} = 564 cpm$$



### Calculate the relative shift— $\Delta/\sigma$

Normalized standard deviation is:

$$\sigma^2 = \left(\frac{5}{18}\right)^2 + \left(\frac{564}{5940}\right)^2 = 0.086; \quad \sigma = 0.29$$

- $\bullet \Delta/\sigma = (1 0.53)/0.29 = 1.60$
- MARSSIM Table 5.5 provides N = 17

### Survey Results for Concrete Floor

Gross Alpha	Gross Alpha	Gross Beta	Gross Beta	Weighted Sum
(cpm)	(dpm/100 cm2)	(cpm)	(dpm/100 cm2)	(sum of fractions)
1	-7	839	1854	0.03
15	94	2209	7158	1.03
2	0	380	77	0.00
1	-7	991	2443	0.05
12	72	540	697	0.59
6	29	2702	9067	0.62
23	152	1856	5792	1.42
10	58	788	1657	0.52
3	7	2400	7898	0.40
11	65	439	306	0.51
5	22	902	2098	0.26
8	43	390	116	0.34
2	0	912	2137	0.09
13	79	450	348	0.63
6	29	1604	4816	0.43
4	14	12223	45927	2.11
3	7	671	1204	0.11
Average	Average	Average	Average	Average
7	39	1782	5506	0.54

### Data Quality Assessment

- Average gross alpha = 39 dpm/100 cm<sup>2</sup>
   (DCGL = 130 dpm/100 cm<sup>2</sup>)
- Average gross beta = 5506 dpm/100 cm<sup>2</sup>
   (DCGL = 23,000 dpm/100 cm<sup>2</sup>)
- Average sum of fractions is 0.54
- Does survey unit pass?
  - How many samples are less than DCGL of 1? S+ is 14
  - Critical value is 12; since S+ > C.V.; survey unit passes
- DCGL<sub>EMC</sub> would be evaluated for each hot spot identified

# Sign Test Example – Co-60 and Am-241 in Soil (Class 1)

- Site contaminants: Co-60 and Am-241
- Class 1 survey unit area is 1,500 m²
- Two strategies for determining need for additional samples based on scan MDCs
- RESRAD version 5.95 used to get DCGLs:
  - 3.4 pCi/g for Co-60; 11.8 pCi/g for Am-241



### Sign Test – DQO Inputs

- Unity rule is used for survey design
- Characterization data used for planning:
  <u>Survey Unit</u>

Co-60  $1.1 \pm 0.4 (1\sigma)$ 

Am-241  $3.8 \pm 0.8$  (1 $\sigma$ )

Type I and II decision errors set at 0.05



### Calculate the relative shift— $\Delta / \sigma$

■ LBGR is set at expected concentration:

$$1.1/3.4 + 3.8/11.8 = 0.65$$

Standard deviations from survey unit are normalized according to MARSSIM eqn I-17:

$$\sigma^2 = \left(\frac{\sigma_{Co-60}}{DCGL_{Co-60}}\right)^2 + \left(\frac{\sigma_{Am-241}}{DCGL_{Am-241}}\right)^2$$

## Calculate the relative shift— $\Delta/\sigma$ (cont.)

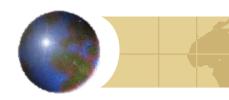
Normalized standard deviation is:

$$\sigma^2 = \left(\frac{0.4}{3.4}\right)^2 + \left(\frac{0.8}{11.8}\right)^2 = 0.0184; \quad \sigma = 0.14$$

- $\bullet \Delta/\sigma = (1 0.65)/0.14 = 2.5$
- MARSSIM Table 5.5 provides N = 15

# Determining Need for Additional Samples in Class 1 Survey Unit

- In Class 1 areas, determine if sample size is sufficient for hot spots that may be present
- Based on n, average area bounded by sample points (a') represents largest hot spot that could exist, and not be sampled
- Requires comparison of actual scan MDCs for radionuclides to required scan MDCs



### Hot Spot Considerations (cont.)

- If Actual Scan MDC < Required Scan MDC—then initial sample size sufficient</p>
- If Actual Scan MDC > Required Scan MDC—then calculate Area Factor that corresponds to actual Scan MDC:

 $AreaFactor = \frac{Scan MDC(actual)}{DCGL_w}$ 

## Assess data needs for Elevated Measurement Comparison test

- 1.25" x 1.5" Nal used for scans; scan MDC for Co-60 is 5.8 pCi/g and for Am-241 is 45 pCi/g
- Area bounded by systematic samples,
   a', is 1500/15 = 100 m<sup>2</sup>
- Area factors associated with a' are:
  - 1.22 for Co-60 and 1.34 for Am-241

## Area Factors for Co-60 and Am-241 (from RESRAD)

Area (m²)	<u>Co-60</u>	<u>Am-241</u>
3000	1	1
1000	1.02	1.00
300	1.10	1.06
100	1.22	1.34
30	1.58	2.46
10	2.33	4.39
3	4.95	9.50
1	11.6	159

June 26, 2006 27

### Assess data needs for EMC

Determine required scan MDC:

(DCGL<sub>W</sub> Area Factor)

Co-60: (3.4 pCi/g)(1.22) = 4.15 pCi/g

Am-241: (11.8 pCi/g)(1.34) = 15.8 pCi/g

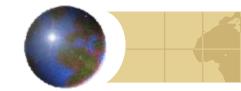
 Actual scan MDCs for both radionuclides are greater than the required scan MDCs additional samples are needed for EMC

## Strategies for Determining Number of Additional Samples

- Determine which of the two radionuclides is the "driver" for the additional samples
  - Calculate ratio of actual-to-required scan MDC
  - Radionuclide with largest ratio is the driver
- Determine reasonable ratio (or range of ratios) between the radionuclides
  - Scan MDC, DCGL, and area factors are determined for specific radionuclide mixture

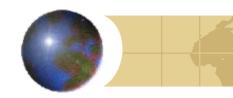
# Determine "driver" for the additional samples

- Actual-to-required scan MDCs:
  - $^{20}$  Co-60: 5.8/4.15 = 1.4
  - Am-241: 45/15.8 = 2.8
- Therefore, Am-241 is the driver
- AF = actual scan MDC/DCGL,
   AF = 45/11.8 = 3.81
- Interpolate to get area that corresponds to this AF: a' is 13.1 m<sup>2</sup>; new sample size = 1500/13.1 = 115



## Determine reasonable ratio (or range of ratios) between the radionuclides

- Requires determination of the scan MDC for a specific mixture of Co-60 and Am-241 ...and modeling to determine DCGL<sub>W</sub> and area factors for this specific mixture (simply stated, this approach is more work!)
- Let's briefly review scan MDC calculations



### Scan MDC Determination

- Minimum detectable count rate based on signal detection theory & human factors(NUREG –1507)
- Relate MDCR in cpm to minimum detectable exposure rate (MDER) based on NaI characteristics
- Microshield™ to model specific radionuclide(s) concentration and geometry—yields exposure rate
- Scan MDC = MDER/CF, where CF is the exposure rate to concentration correction factor

### Scan MDCs For 1.25"x1.5" NaI Detector

<b>Radionuclide</b>	Scan MDC (pCi/g)
Cs-137	10.4
Am-241	44.6
Th-232	2.8
Co-60	5.8
Processed Uranium	115
Enriched Uranium (3%)	137
Ra-226	4.5

# Determine Additional Sample Needs for 30% Am-241 and 70% Co-60

- Based on process knowledge and characterization, 95% confidence level on fractional amount of Am-241: 0.3 to 0.8
- Determine additional sample needs for lower bound ratio: 30% Am and 70% Co
- Calculate scan MDC—expect it to be between 5.8 pCi/g (Co-60) and 45 pCi/g (Am-241): scan MDC = 7.8 pCi/g

# Determine Additional Sample Needs for 30% Am and 70% Co (cont.)

- $\bullet$  DCGL<sub>W</sub> for 30% Am/70% Co = 4.8 pCi/g
- AFs also generated for this mixture; the AF for 100 m² (initial a') is 1.22
- Required scan MDC: (4.8)(1.22) = 5.9 pCi/g—need more samples
- New AF = 7.8/4.8 = 1.62; new a' is 29 m<sup>2</sup>; new sample size = 1500/29 = 52 samples

# Determine Sample Needs for Various Am-241 to Co-60 Ratios

 Scan MDC for multiple radionuclides involves Microshield™ modeling or by observation

$$scanMDC = \frac{1}{\frac{f_1}{scan MDC_1} + \frac{f_2}{scan MDC_2}}$$

 For mix of 80% Am/20% Co: scan MDC using Microshield™ was 19.1; using eqn it's 19.2 pCi/g

## Additional Sample Needs as a Function of Am/Co Mixture

<u>Mixture</u>	<u>DCGL</u>	Scan MDC	Add'l Samples
0%Am	3.4	5.8	62
30%Am	4.8	7.8	52
50%Am	6.6	10.3	42
65%Am	9.3	13.4	29
80%Am	14.7	19.2	0
90%Am	13.1	26.9	35
100%Am	11.8	45	115

### Conclusion: Strategies for

### Determining Number of Add'l Samples

- Using Am-241 as the "driver" requires 115 additional samples...no effort to get scan MDC for multiple radionuclides
- If range of ratios between Am and Co can be justified (e.g., 30 to 80% Am-241), the conservative additional sample size is 52...but approach requires effort to get scan MDC, DCGL, and Area Factors

## If Scan MDC Is NOT Sufficient – Reduce Scan MDC By:

- Slowing scan speed to increase observation interval; however, practical limit of several seconds on observation interval (can't keep on scanning slower)
- Use more sensitive instrument (increase efficiency)
- Accept more false positives, which requires training technicians to pause and flag spots more frequently



## What If No Scan Capability At All?

- Radionuclides include alpha and beta emitters (H-3, C-14) and low energy gamma and x-ray emitters (e.g., Fe-55)
- Perform systematic sampling in survey unit & analyze samples, and assess with posting plot
- Perform second stage sampling based on results of first sampling stage
  - At locations where samples exceed DCGL<sub>w</sub>
  - Where posting plot indicates potential for contamination



- Consider soil compositing
- Consider revising the DCGL<sub>W</sub> via dose modeling – use realistic scenarios (environmental pathways) and sitespecific parameters

## What if you fail the statistical test? Double Sampling...perhaps

- If the mean > DCGL<sub>W</sub>, then you fail w/o even performing the statistical test
- Re-do the final status survey—always the case when additional remediation is necessary
- If the mean < DCGL<sub>W</sub>, can you collect additional samples and perform the statistical test again?

[Answer: Yes, provided that it was agreed upon with the regulator during the DQO process]

# Two-Stage or Double Sampling for Final Status Surveys

- Two-stage sampling: survey design is specifically intended to be conducted in two stages
- Double sampling: when the survey design is one-stage design, but an allowance is made for a second set of samples to be taken
- Type I error increases with double sampling, but by no more than a factor of two

## Draft NRC NUREG-1757, vol. 2, Appendix C

- "...double sampling should not be used as a substitute for adequate planning. If it is to be allowed, this should be agreed upon with NRC staff as part of the DQO process." (p. C-6)
- Note double sampling allows for one more set of additional samples...i.e., no triple sampling
- Additional samples should be collected randomly, and added to initial sample data set

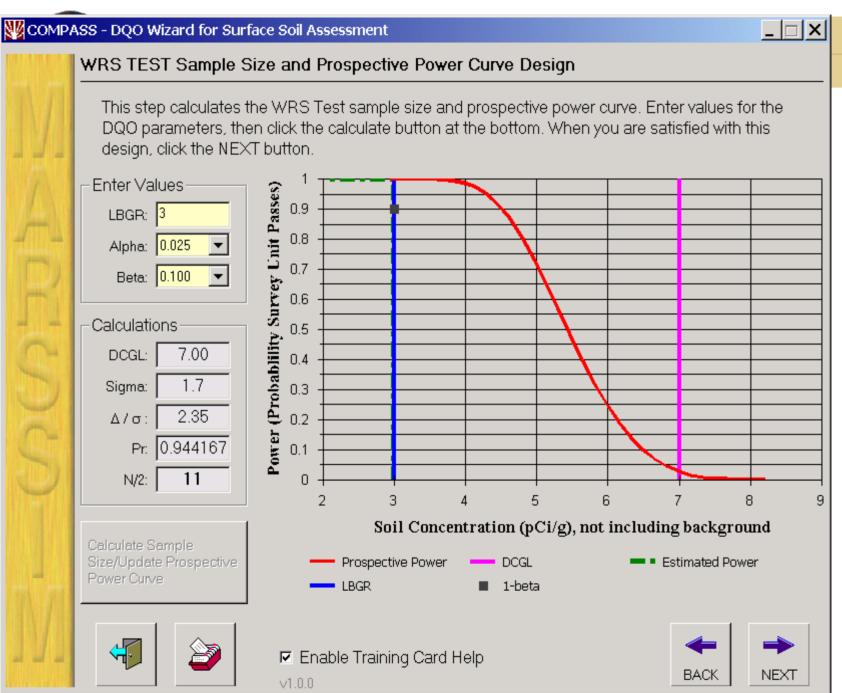


### Double Sampling Notes

- Double sampling should never be necessary for Class 2 or 3 survey units (data should all be less than DCGL<sub>W</sub> in these SUs)
- Statistical failures can occur as a result of poor characterization efforts; bad estimates of mean and std dev in the SU—so while the mean may be less than DCGL<sub>W</sub>, the sample size is too small to reject the H<sub>0</sub> for the actual std dev in the SU (assess by retrospective power curve)

## Double Sampling Example

- Survey design using WRS test for Th-232,
   DCGL<sub>W</sub> = 7 pCi/g
- Limited characterization results in 4.1 pCi/g mean and 1.7 pCi/g std dev in survey unit; bkg had Th-232 concentration of 1.1 pCi/g (net 3 pCi/g in survey unit—set LBGR at 3 pCi/g)
- Relative shift:  $\Delta/\sigma = (7-3)/1.7 = 2.35$ ; Type I error = 0.025; Type II error = 0.1
- $\bullet$  N/2 = 11 samples



June 20, 2000 47



## Double Sampling Example

- Survey design was implemented and 11 samples collected in SU and bkg, but actual standard deviation turns out to be 3.06 pCi/g (1.7 pCi/g was planned)
- The retrospective power curve illustrates the impact of this significant underestimate of standard deviation



#### **Basic Statistical Quantities Summary**

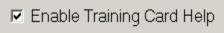
Summarizes the basic statistical quantities such as the mean, median, minimum value, maximum value, and standard deviation for the systematic sample, reference area samples if applicable and the estimated values provided from the DQO process.

Statistic	Survey Unit	Background	DQO Results
Sample Number	11	11	N/2=11
Mean (pCi/g)	5.08	0.87	3
Median (pCi/g)	5.20	0.90	N/A
Std Dev (pCi/g)	3.06	0.43	1.7
High Value (pCi/g)	9.20	1.50	N/A
Low Value (pCi/g)	1.70	0.20	N/A

Because the difference between a survey unit measurement and a reference area measurement exceeds the DCGLw AND the difference of the survey unit average and reference area average is less than DCGLw, the WRS Test will be conducted after reviewing the retrospective power curve.







V1.0.0

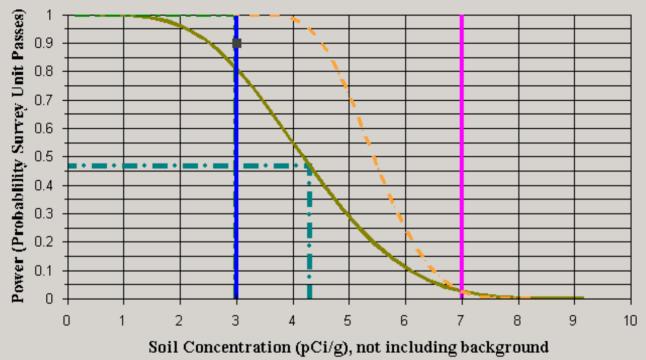






#### Retrospective Power Curve Design

This step compares the retrospective power curve with the prospective power curve for the selected survey unit. The legend below will assist in the interpretation.



- Prospective Power ■ 1-beta ■ • Actual Power

LBGR ■ • Estimated Power

DCGL — Retrospective Power





▼ Enable Training Card Help

V1.0.0







## Double Sampling Example

- The retrospective power curve shows that the actual net concentration is 4.2 pCi/g (not 3 pCi/g), but still less than the DCGL<sub>W</sub>
- The increased std dev (3.06 vs. 1.7) results in the "shallow" retrospective power curve
- The probability of passing at the actual concentration of 4.2 pCi/g is ~ 47%



#### Statistical Test Summary

Summarizes the results of the WRS Test.

The result of the statistical test is the decision to reject or not to reject the null hypothesis (indicated by Pass or Fail, respectively).

Data	Туре	Adjusted Data	Rank	Ref Rank 🛕
0.2	R	7.2	9.5	9.5
0.2	R	7.2	9.5	9.5
0.6	R	7.6	11	11
0.8	R	7.8	12.5	12.5
0.8	R	7.8	12.5	12.5
0.9	R	7.9	14.5	14.5
0.9	R	7.9	14.5	14.5
J	_	<b>-</b> .		

Sum of Ranks: 253

Sum of Ref Ranks: 154

Critical Value: 156

Result: Fail





☑ Enable Training Card Help v1.0.0





## Double Sampling Example

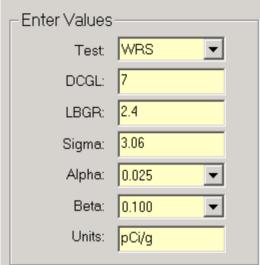
- W<sub>r</sub> = 154; CV = 156—SU just fails (remember SU passes if W<sub>r</sub> > CV)
- Now calculate WRS sample size based on actual std dev and desired power (> 70%) at actual concentration in SU (use COMPASS Practice)
- N/2 = 17 for revised DQOs and actual data
- Already have 11 samples, so need to collect 6 additional random samples in both the SU and reference area
- Regulator has approved Type I errors up to 0.05

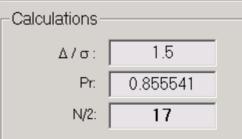




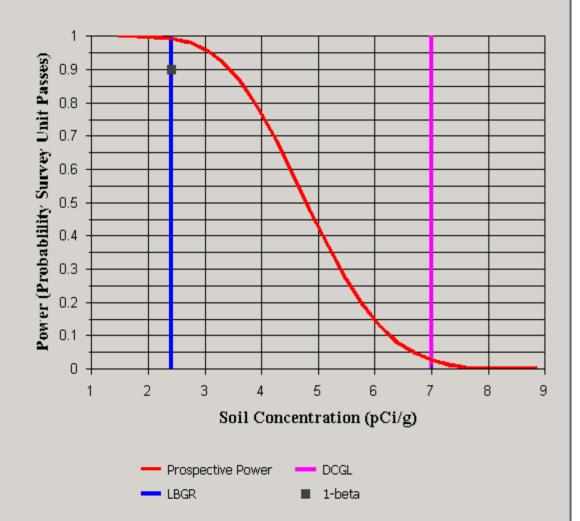
#### Statistical Tests and Prospective Power Practice

Enter the values required below. After these values are entered, you can view the prospective power curve on the right. Click the help button for detailed descriptions of each field.









COMPASS v1.0.0





Summarizes the basic statistical quantities such as the mean, median, minimum value, maximum value, and standard deviation for the systematic sample, reference area samples if applicable and the estimated values provided from the DQO process.

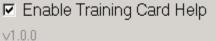
Statistic	Survey Unit	Background	DQO Results
Sample Number	17	17	N/2=11
Mean (pCi/g)	5.15	0.86	3
Median (pCi/g)	5.20	0.90	N/A
Std Dev (pCi/g)	2.54	0.39	1.7
High Value (pCi/g)	9.20	1.50	N/A
Low Value (pCi/g)	1.70	0.20	N/A

Because the difference between a survey unit measurement and a reference area measurement exceeds the DCGLw AND the difference of the survey unit average and reference area average is less than DCGLw, the WRS Test will be conducted after reviewing the retrospective power curve.





Enable Training Card Help



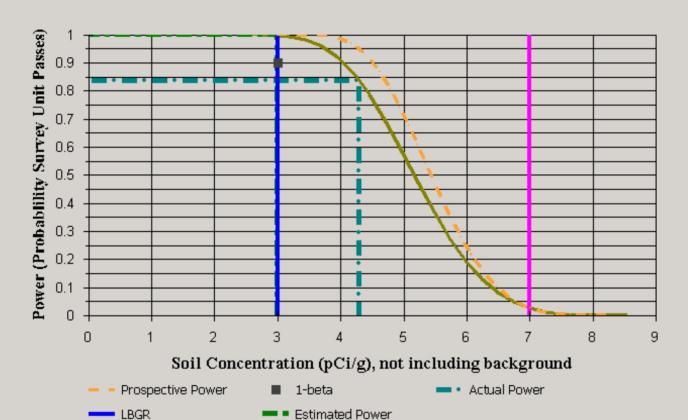






#### Retrospective Power Curve Design

This step compares the retrospective power curve with the prospective power curve for the selected survey unit. The legend below will assist in the interpretation.



Retrospective Power





DOGL

V1.0.0

▼ Enable Training Card Help







#### Statistical Test Summary

Summarizes the results of the WRS Test.

The result of the statistical test is the decision to reject or not to reject the null hypothesis (indicated by Pass or Fail, respectively).

Data	Туре	Adjusted Data	Rank	Ref Rank 🛕
0.2	R	7.2	15.5	15.5
0.2	R	7.2	15.5	15.5
0.4	R	7.4	17	17
0.6	R	7.6	18	18
0.7	R	7.7	19.5	19.5
0.7	R	7.7	19.5	19.5
0.8	R	7.8	21.5	21.5
	_			

Sum of Ranks: 595

Sum of Ref Ranks: 391

Critical Value: 354

Result: Pass





▼ Enable Training Card Help

V1.0.0





## Conclusions from Double Sampling Example

- The survey unit passed the second time around—after 6 more samples were randomly collected and added to data set
- Retrospective power curve is a diagnostic tool to assess survey unit failures
- The overall Type I error (after double sampling) is greater than 0.025, but less than 0.05



## Final Thoughts on Advanced MARSSIM Topics

- MARSSIM lessons learned are being made available—ORS journal (June 2003); DDSC web site (<a href="http://www.orau.gov/ddsc/">http://www.orau.gov/ddsc/</a>); NUREG-1757; COMPASS; NRC's improved D&D web site
- Additional work needed with scan MDCs, subsurface contamination, dose modeling and area factors—share/publish your MARSSIM experiences



### References

- Operational Radiation Safety, June 2003
- "Decommissioning Health Physics: A Handbook for MARSSIM Users." Institute of Physics, Bristol, UK; July 2001
- HP Newsletter "Double Sampling A Statistical Approach for MARSSIM Users" July 2003