DEPARTMENT OF THE ARMY
SENECA ARMY DEPOT ACTIVITY
5786 STATE RTE 96, P.O. BOX 9
ROMULUS, NEW YORK 14541-0009

February 28, 2005

REPLY TO
ATTENTION OF
Caretaker Office

J-7

'05 MAR -3 A11:04

RECEIVED
REGION 1

Mr. James Kottan
U.S. Nuclear Regulatory Commission
Region 1
Division of Nuclear Materials Safety
Nuclear Materials Safety Branch 2
475 Allendale Road
King of Prussia, PA 19406-1415

SUG-1275
04008526

SUBJECT: Response to Request for Additional Information Concerning NRC License Termination Report for Seneca Army Depot Activity (Control Number 135163) - phone conversation from January 27, 2005

Dear Mr. Kottan,

The United States Army is pleased to submit the additional information requested regarding the License Termination Report for Seneca Army Depot Activity (SEDA) in Romulus, New York. The NRC, in a phone call on January 27, 2005 clarified their request for additional information pertaining to retrospective power curves.

In comments provided on August 9, 2004, the NRC made the request to: "Please discuss the statistical methods you used for determining compliance to the DCGLs relative to the null hypothesis recommended in MARSSIM and presented in Table 5-4 of your LTP. Also please provide the retrospective power curves." The Army responded to the comment in a letter dated September 2, 2004 explaining the statistical methods used; however, retrospective power curves were not provided at that time.

As requested, the retrospective power curves are being provided. The CD provided with this letter contains the following:

- The file *Summary Tables.pdf.* This file summarizes for each of the survey units included in the license termination the results of the WRS test, the Quantile test, the background median plus Lower Bound of the Grey Region (LBGR), and the Power test.
- A folder *Retrospective Power Curves.* This folder contains both the alpha and beta radiation retrospective power curves for each of the survey units included in the license termination.
- A folder *Supporting Information*, which contains the following folders:
    - *Kruskal-Wallis Test Data,* which contains the tables and calculations used to perform the Kruskal-Wallis (K-W) Tests on the survey units; and
    - *WRS_Quantile_Power Calculations*, which contains the tables and calculations used to perform each the WRS test, the Quantile test, and the Power test for the survey units. This information was used to create the retrospective power curves.

Also included with this letter is Attachment A, which provides a discussion of each of the tests used in the retrospective power curve development.

135163

NMSS/RGN1 MATERIALS-002

The goal of the License Termination Report for SEDA, which follows the *Multi-Agency Radiation Survey and Site Investigation Manual* (MARSSIM; NRC, 2000) and other applicable guidance, is to demonstrate that the license termination requirements for NRC license SUC-1275 (NRC Docket No. 040-08526) have been met and to remove SEDA from Licenses SUC-1380, 45-16023-01NA, SUB-834, BML 12-00722-07, and STC-133.

We appreciate the opportunity to provide you with this additional information for a report that is of great importance to the United States Army. Should you have any questions regarding the document, please do not hesitate to contact me (607) 869-1235.

Sincerely,

Stephen M. Absolom
Installation Manager

Attachment A
Discussion of Retrospective Power Curves
NRC License Termination Report for Seneca Army Depot Activity
(Control Number 135163)

Retrospective power curves for the statistical analyses involving the alpha and beta radiation field measurements were generated for the *NRC License Termination Report for Seneca Army Depot Activity*. The following describes the methodology and the assumptions used, and a brief summary of each step in the process. Files used to perform the calculations that are provided in the attached CD have been referenced in italics.

1. **Kruskal-Wallis Test:** Per Section 13.3 of NUREG-1505 (NRC, 1998), the Kruskal-Wallis (K-W) test was performed on the reference area data using the different types of materials present or different measurement locations to determine if there was sufficient variability in background for the Scenario B null hypothesis. The K-W test was performed for the following background datasets:

- **Building 722** *(NRC Building 722 K-W Test.xls)*: Data from Building 722 were grouped by type of surface or material measured for each direct measurement. The phoswich data had nine types of surfaces, while the floor monitor data had three types. Two datasets (beta phoswich and beta floor monitor) showed significant variability at a Type I error (α) of 0.05. One dataset (alpha phoswich) showed significant variability at an α of 0.2. One dataset (alpha floor monitor) did not demonstrate significant variability (i.e., the calculated value of K was less than all critical values [K_c] listed in Table 13.1 of NUREG-1505). However, to maintain consistency for all survey unit datasets using this reference area, Scenario B was used for the alpha floor monitor measurements despite the K-W test result.

- **2002 Igloos** *(NRC Igloos 2002 K-W Test.xls)*: Since the measurement surfaces within each background igloo were the same (i.e., concrete), the data were grouped by individual background igloo (i.e., Igloo A1107, B0806, C0912, D0405, E0403). Both the alpha phoswich and beta phoswich datasets demonstrated significant variability at an α of 0.05.

- **Building 2078** *(NRC Building 2078 K-W Test.xls)*: Locations for background measurements taken from Building 2078 were not available; as a result, to group measurements for the K-W test, the data for both floor monitor and hand-held gas proportional measurements were visually inspected and were grouped on the basis of "high" or "low" measurements. For the floor monitor data, there was a break between beta measurements of 945 and 1065 counts per minute (cpm) when ranked in order of magnitude. As a result, locations with beta measurements of 945 cpm or lower were grouped as "low" and locations with beta measurements of 1065 cpm or higher were grouped as "high". For the hand-held data, there was a break between ranked beta measurements of 192 and 233 cpm, and the data were grouped accordingly. As a result of this data grouping,

the alpha hand-held, beta hand-held, and beta floor monitor data demonstrated significant variability at an α of 0.05. The alpha floor monitor data showed significant variability at an α of 0.2.

2. **Calculation of Lower Boundary of the Gray Region (LBGR):** The calculation for the LGBR (i.e., level that is distinguishable from background) for Scenario B was performed per Section 13.4 of NUREG-1505. The recommended default value ($3w$) was calculated and used for all background datasets, except for the Building 722 alpha floor monitor. Because the K-W test did not demonstrate significant variability for the Building 722 alpha floor monitor dataset, a value of zero was used as the LBGR for the background tests with that dataset. The LBGR calculation spreadsheets are included on the CD in the K-W spreadsheets noted above in bullet 1.

3. **Wilcoxon Rank Sum (WRS) Test Comparing Survey Unit Data and Background:** The WRS test was performed on the survey unit and reference area data using Scenario B, per Section 6 of NUREG-1505. The null hypothesis used for this test was that the difference between the survey unit median and the background median is less than the LBGR (i.e., the survey unit is indistinguishable from background). The critical value for the WRS test was calculated using the equation in Table A.4 of NUREG-1505. To determine the effect of tied ranks on the critical value, an example calculation for the critical value accounting for ties was performed for the 306 Room 10 alpha phoswich and alpha floor monitor datasets (*306R10 Power.xls*). For the phoswich, the difference between the initial critical value (1332.86) and the critical value accounting for ties (1332.74) was not significant. Likewise, the difference for the floor monitor was also not significant (514.13 initial, 514.06 with ties). It was concluded that ties would not significantly affect the critical value, and ties were not considered for the other tests.

Five alpha floor monitor datasets were found to exceed background based on the WRS test: Building 306 Room 10, Building 306 Room 11, Building 306 Room 13, Building 2073 Room 3, and Building 2084 Room 3. The WRS tests for each survey unit are included in the ****Power.xls* files (e.g., *306R10 Power.xls*, *306R11 Power.xls*, etc.). The WRS test results are summarized in the *2002 Buildings*, *Building 612*, and *Igloos* spreadsheets in the *Summary Tables.pdf* file.

4. **Quantile Test:** The Quantile test was performed per Section 7 of NUREG-1505 to detect differences in only a fraction of the survey unit data versus the reference area data. Per NUREG-1505, it is required for the Scenario B null hypothesis that a survey unit passing the WRS test must also pass the Quantile test. Values of k, r, and α for the test were determined from Table A.7b from NUREG-1505. For numbers of survey or reference area measurements that did not exactly match those listed in Table A.7b, the closest values were used. If k of the r largest ranks were from the survey unit, the null hypothesis was rejected. Per EPA 230-R-94-004 (EPA, 1994), if the r-th largest measurement was among a group of tied (i.e., equal-in-value) measurements, r was increased to include the tied measurements. The value of k was increased by the same number of measurements. Two datasets (Building 306 Room 10 alpha floor monitor and Igloo C0401 alpha phoswich) failed the Quantile test with background. The Quantile tests for each survey unit are included in the ****Power.xls* files. The Quantile test results are summarized in the *2002 Buildings*, *Building 612*, and *Igloos* spreadsheets in the *Summary Tables.pdf* file.

5. **WRS and Quantile Tests Comparing Survey Unit Data and the $DCGL_W$:** This second WRS test was performed for datasets that failed either the initial WRS test or Quantile test with background. The null hypothesis for this test is that the difference between the survey unit median and the background median is less than the LBGR plus the $DCGL_W$ (i.e., the test is rerun after adding the $DCGL_W$ in cpm to the LBGR). The six datasets that failed either the initial WRS test or Quantile test all passed the WRS test with the $DCGL_W$, indicating that each met the release criterion. In addition, all six datasets passed the Quantile test. The detailed WRS and Quantile tests for each survey unit and the $DCGL_W$ are included in the ****Power.xls* files. The WRS and Quantile tests for each survey unit and the $DCGL_W$ are summarized in the *DCGL Comparison* spreadsheet in the *Summary Tables.pdf* file.

6. **Generation of Retrospective Power Curves:** Retrospective power curves were generated using the methods described in Section 10.5 of NUREG-1505. The larger of the standard deviations from the survey unit measurements and background measurements (shaded in yellow on the spreadsheets) was used in the calculation. The power (i.e., probability of survey unit failing) was determined at the survey unit median equal to the background median measurement plus the LBGR (i.e., the distinguishable level above background). For the datasets that underwent a comparison with the $DCGL_W$, additional power curves were generated and the power was determined at the survey unit median equal to the background median plus the LBGR plus the $DCGL_W$. The desired power for the statistical tests was 0.95. The power calculations for the comparison of survey unit data with background are presented in detail in the ****Power.xls* files and are summarized in the *2002 Buildings*, *Building 612*, and *Igloos* spreadsheets in the *Summary Tables.pdf* file. The power calculations for the comparison of the survey unit data with the $DCGL_W$ are also presented in detail in the ****Power.xls* files and are summarized in the *DCGL Comparison* spreadsheet in the *Summary Tables.pdf* file. The power calculation results are discussed in further detail below.

- **2002 Buildings:** Of the 33 alpha phoswich datasets, 5 datasets had a calculated power less than 0.95 (ranging from 0.74 to 0.93). Four of the 33 beta phoswich datasets also had a calculated power less than 0.95 (ranging from 0.87 to 0.93). Of the 24 alpha floor monitor datasets, 8 datasets had a calculated power less than 0.95 (ranging from 0.37 to 0.91). Three of the 24 beta floor monitor datasets had a calculated power less than 0.95 (ranging from 0.66 to 0.92).

  All but one of the 2002 Buildings datasets with a calculated power less than 0.95 were collected from Building 5. The alpha and beta phoswich datasets from Building 5 that resulted in a calculated power of less than 0.95 consisted of 5 to 8 measurements. The alpha floor monitor datasets from Building 5 that resulted in a calculated power of less than 0.95 consisted of 2 to 14 measurements. The alpha floor monitor datasets from Building 5 Room 2 and Building 306 Room 13 had standard deviations greater than the background standard deviation, which contributed to the reduced power. The beta floor monitor datasets from Building 5 that resulted in a calculated power of less than 0.95 consisted of 2 to 3 measurements.

- **Building 612:** Of the 28 alpha hand-held gas proportional datasets, 20 datasets had a calculated power less than 0.95 (ranging from 0.46 to 0.92). All 28 of the beta hand-held datasets had a calculated power of 1.0. Of the 23 alpha floor monitor datasets, 2 datasets had a calculated power less than 0.95 (ranging from 0.50 to 0.91). All 23 of the beta floor monitor datasets had a calculated power of 1.0.

  The alpha hand-held datasets from Building 612 with a calculated power of less than 0.95 consisted of 9 to 47 measurements. Based on that wide range of survey unit measurements, the fixed number of background measurements (32) may be more responsible for the reduced power. The two alpha floor monitor datasets from Building 612 that resulted in a calculated power of less than 0.95 consisted of 2 and 4 measurements.

- **Igloos**: All 120 of the alpha phoswich datasets had a calculated power less than 0.95 (ranging from 0.072 to 0.30). All 120 of the beta phoswich datasets had a calculated power of 1.0.

  Since each igloo had the same number of alpha measurements (30), the change in calculated power appears to be primarily based on the change in standard deviation of the survey unit data. As discussed in Section 3.3.3 of the *NRC License Termination Report for Seneca Army Depot Activity* , elevated alpha measurements were consistently taken at the vent screen on the upper rear wall of each igloo – these elevated measurements were observed in both the background and affected igloos, and were attributed to the presence of radon progeny. The effect of these vent measurements can be seen in the file *A0201 K-W Test and Power – rev bkgd.xls*. As an exercise, the measurement from the vent location was removed from each background igloo dataset, the K-W test was re-run, and the LBGR was re-calculated. As a result of removing these vent measurements, the LBGR was reduced from 13.3 to 10.7. More importantly, the standard deviation of the background data decreased from 12 cpm to 4.1 cpm. Correspondingly, the calculated power for that dataset increased, from 0.30 to 0.92.

- **DCGL Comparison:** Of the six datasets that were compared with the $DCGL_W$, two datasets (Building 306 Room 13 alpha floor monitor and Igloo C0401 alpha phoswich) had a calculated power at the background median plus LBGR plus $DCGL_W$ of less than 0.95 (0.86 and 0.16, respectively). The calculated power for these datasets in the $DCGL_W$ comparison is the same as the calculated power for these datasets in the background comparison.

The information presented above follows the methodology discussed via telephone conversation with NRC personnel on January 27, 2005. The procedures followed from NUREG-1505 were to demonstrate sufficient variability in background (i.e., the Kruskal-Wallis test), calculate the LBGR, perform the WRS and Quantile Tests for Scenario B, and generate retrospective power curves. Of the 462 alpha and beta datasets evaluated, 164 (i.e., 35 percent) had a calculated power of less than 0.95, while 298 (i.e., 65 percent) datasets had a calculated power of 0.95 or better. All survey units had at least one dataset with a calculated power of 0.95 or greater. While additional measurements at some survey units may have resulted in an increased statistical power, it is very unlikely that a different

outcome to the tests (i.e., the survey unit fails rather than passes) would have resulted based on those additional measurements.

Cited References:

EPA, 1994. *Statistical Methods for Evaluating the Attainment of Cleanup Standards; Volume 3: Reference-Based Standards for Soils and Solid Media*, EPA 230-R-94-004, U.S. Environmental Protection Agency, June.

NRC, 1998. *A Nonparametric Statistical Methodology for the Design and Analysis of Final Status Decommissioning Surveys*, NUREG-1505, U.S. Nuclear Regulatory Commission.

General References:

Abelquist, 2001. *Decommissioning Health Physics: A Handbook for MARSSIM Users*, Institute of Physics Publishing, Philadelphia, PA.

NRC, 2000. *Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)*, NUREG-1575, Revision 1, U.S. Nuclear Regulatory Commission, August.