5-503.45

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
REGION 5
77 WEST JACKSON BOULEVARD
CHICAGO, IL 60604-3590

Attachment in Files

APR 06 1999

REPLY TO THE ATTENTION OF: SRF-5J

Mr. Johnny W. Reising
United States Department of Energy
Fernald Area Office
P.O. Box 398705
Cincinnati, Ohio 45239-8705

Subject: Technical Review Comments on Revised Real-Time In Situ Gamma Spectrometry Reports and Documentation to Support the Use of High-Purity Germanium Detectors to Perform Final Soil Certification for Primary Radionuclides

2154

Dear Mr. Reising:

The United States Environmental Protection Agency (U.S. EPA) has reviewed the above-referenced reports and documentation as part of its oversight activities for the Fernald Environmental Management Project (FEMP). The reports and documentation, which are dated January 1999, were prepared by Fluor Daniel Fernald for the U.S. Department of Energy and consist of the following:

- "Comparability of In Situ Gamma Spectrometry and Laboratory Data, Revision 1,"
- "Comparability of In Situ Gamma Spectrometry and Laboratory Data and Decisions for Certification Units, Revision 0,"
- "Radiation Tracking System (RTRAK) Applicability Study, Revision 2"
- Draft Data Validation Checklist for Validating High-Purity Germanium Detector Measurements to Analytical Support Level D
- Updated Section 2.5, (Revision B), for incorporation into User's Manual, entitled "Certification"
- Updated Section 3.7, (Revision B), for incorporation into User's Manual, entitled "Certification Measurements"

The reports and documentation provide the results of recent studies on the real-time gamma spectrometry instruments and proposals for their future use at the site. U.S. EPA reviewed the reports and documentation for technical adequacy and consistency with other FEMP documents and procedures.

The FEMP Real Time Program has proven to be time and cost effective in its current application at the site. Expedited remedial actions, hot spot delineation and pre-certification have all been areas where the real time program has been beneficial. However, U.S. DOE's latest submittals focus on using real time in situ spectrometry for final certification, which raises many issues.

U.S. EPA's review found numerous deficiencies in the application of in situ spectrometry to final certification. Many of the deficiencies appear unresolvable due to the inherent technical limitations of in situ spectrometry and environmental factors at Fernald. Major deficiencies with the application of in situ spectrometry to final certification include:

- Challenges presented by natural radon emanation and problems with definitive radium quantitation prohibit the use of in situ spectrometry as a certification tool for radium isotopes.

Recycled/Recyclable • Printed with Vegetable Oil Based Inks on 50% Recycled Paper (20% Postconsumer)

- Certification decisions for hazardous contaminants will be based on laboratory analytical results. Since certification decisions for radium 226 and some other isotopes (for example, technetium 99) would also have to be based on laboratory analytical results, the overall effectiveness and cost savings resulting from certification usage of in situ spectrometry is not apparent.
- The precision of uranium and radium measurements near their final remediation levels (FRLs) has not been effectively demonstrated.
- Accuracy cannot be effectively defined for in situ HPGe detector measurements, because no matrix spike measurements are performed.

Based on these and other deficiencies, U.S. EPA does not recommend use or further development of in situ gamma spectrometry for final certification of primary radionuclides. U.S. EPA does recommend using the existing Agency-approved excavation and certification strategy while continuing the development and application of in situ spectrometry where it can serve as an effective screening tool.

U.S. EPA's general and specific review comments are enclosed. Please contact me at (312) 886-4591 if you have any questions.

Sincerely,

E Jablonowski

Gene Jablonowski
Remedial Project Manager
Federal Facilities Section
SFD Remedial Response Branch #2

Enclosures (2)

cc: Tom Schneider, OEPA-SWDO
Bill Murphie, U.S. DOE-HDQ
John Bradburne, FERMCO
Terry Hagen, FERMCO
Tom Walsh, FERMCO

ENCLOSURE 1

TECHNICAL REVIEW COMMENTS ON
REVISED REAL-TIME IN SITU GAMMA SPECTROMETRY REPORTS AND
DOCUMENTATION TO SUPPORT THE USE OF HIGH-PURITY GERMANIUM
DETECTORS TO PERFORM FINAL SOIL CERTIFICATION
FOR PRIMARY RADIONUCLIDES

FERNALD ENVIRONMENTAL MANAGEMENT PROJECT

(17 Pages)

Attachment to: M-00893

TECHNICAL REVIEW COMMENTS ON REVISED REAL-TIME IN SITU GAMMA SPECTROMETRY REPORTS AND DOCUMENTATION TO SUPPORT THE USE OF HIGH-PURITY GERMANIUM DETECTORS TO PERFORM FINAL SOIL CERTIFICATION FOR PRIMARY RADIONUCLIDES

TECHNICAL REVIEW COMMENTS ON "COMPARABILITY OF IN SITU GAMMA SPECTROMETRY AND LABORATORY DATA, REVISION 1"

GENERAL COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Section #: Not applicable (NA) Page #: NA Line #: NA
Original General Comment #: 1
Comment: The comparability study involves a number of comparisons of laboratory methods, especially for analysis of total uranium and radium 226. Work has been started on a "Multi-Agency Radiation Laboratory Protocols (MARLAP) Manual" involving the Department of Energy (DOE), U.S. Environmental Protection Agency (U.S. EPA), and five other federal agencies. The method comparison data and other information generated while resolving data discrepancies should be provided to the MARLAP group to support development of the MARLAP Manual.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 2
Comment: In situ gamma spectrometry using a high-purity germanium (HPGe) detector has not yet been accepted as a source of definitive data, especially for certifying that soil in an area meets final remediation levels (FRL) for primary radionuclides. As discussed in original specific comment 4 below, the radium 226 HPGe results appear to have a definite low bias, which is unacceptable for making final certification decisions. In addition, as discussed in comments on the draft data validation checklist for validating HPGe detector measurements to analytical support level (ASL) D, the current calibration method that uses multiple readings of a point source that are mathematically manipulated is quite different from the analytical method that measures an overall, relatively uniform source. Therefore, the use of data from the HPGe detector for final certification purposes is premature.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 3
Comment: The comparability study demonstrates reasonable agreement between HPGe detector measurements and laboratory results for uranium and thorium 232. However, some concerns still exist regarding HPGe detector measurements of radium 226. In this comparability study, a new correction algorithm is provided that differs from the one proposed in the user's manual for the real-time instruments. In fact, the new algorithm results in corrected radium 226 concentrations that are markedly lower than those using the user's

manual method. In addition, little information exists to justify the use of the revised correction algorithm. The text states that Section 8.0 of the comparability study presents a validation of the correction algorithm. However, Section 8.0 does not provide adequate information to validate the correction algorithm. To begin with, none of the validation results apply to radium 226 concentrations near the FRL of 1.7 picocuries per gram (pCi/g). Table 4 in Appendix J shows area-average laboratory analysis results from 0.70 to 0.99 pCi/g and in situ HPGe results from 0.58 to 0.87 pCi/g. Furthermore, for the majority, if not all, of the data in the validation study, comparability between data points can be demonstrated without the use of a correction algorithm. Therefore, the new correction algorithm is not appropriately tested and its use is not validated. The text should be revised to further address these issues.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 4
Comment: The comparability study raises a general concern regarding HPGe measurements and data quality. During physical sampling activities, strict laboratory quality control (QC) must be adhered to. One element associated with laboratory QC involves measurement of spiked samples and their duplicates. In this case, the measured value can be directly compared to a known value. However, this is not the case with HPGe measurements. In situ gamma spectrometry can only be used to approximate the true value. Therefore, a flaw exists with regard to HPGe QC criteria. According to the text in Section 9.0, ASL D requires HPGe values to be within 20 percent of the laboratory values. However, because the laboratory values are allowed to be within 20 percent of the true (known) values, HPGe results are allowed to potentially be within 40 percent of the known values. Because a 40 percent allowance is unacceptable, one of the following two solutions should be considered:(1) lowering the QC acceptance criterion for laboratory samples and HPGe measurements (combined) so that the total propagated difference is less than 20 percent or (2) ensure that the sum of the average relative percent differences (RPD) for physical samples and matrix spikes, and for HPGe measurements and physical samples, (the total RPD) is less than 20 percent.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 5
Comment: In various places, the text states that the total propagated uncertainty (TPU) of laboratory data is higher than the TPU for HPGe measurements. The text states that this results from the fact that available HPGe data only account for counting errors and that the HPGe system will register many more counts. However, because the HPGe system relies on comparison to laboratory data to approximate the "true" value, errors associated with laboratory measurement should be included when calculating HPGe measurement errors. Unless complete TPUs for the HPGe system are provided, discussion on measurement error between the two methods is

useless. If complete TPUs for the HPGe system are not provided, only the counting error associated with laboratory measurements should be included.

SPECIFIC COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 2.8 Page #: 2-11 Line #: 4
Original Specific Comment #: 1
Comment: The text states that some "statistically significant differences" exist between concentrations measured by different detectors. Table 2-5 shows some pairs of data sets as differing at the 95 percent confidence level. The text should be revised to identify the statistical test used to support the conclusion that some pairs of data differ significantly.

Commenting Organization: U.S. EPA Commentor: Saric
Figures #: 3-2 and 3-3 Page #: NA Line #: NA
Original Specific Comment #: 2
Comment: Figures 3-2 and 3-3 show additional study locations. To increase the useability of these figures, Figures 3-2 and 3-3 should be revised to include the symbols for low, medium, and high concentration areas that are used in Figure 3-1. Also, Figure 3-1, which shows the entire facility, should be revised to indicate the general locations of the detailed maps in Figures 3-2 and 3-3. These modifications would help the reader relate the study areas to the overall pattern of contamination within the facility.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 5.3 Page #: 5-2 Line #: 10
Original Specific Comment #: 3
Comment: Section 5.3 discusses total uranium analytical results, and the accompanying tables and appendixes provide data for areas 0 through 10. However, the text of Section 3.2.1 discusses collection of data from areas 12 through 19 to supplement data already available from areas 1 through 10. Either the tables in Section 5 and associated appendixes should be revised to include data for areas 12 through 19, or the omission of these data should be justified in the text.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 7.4 Page #: 7-5 Line #: 5
Original Specific Comment #: 4
Comment: This section discusses the data in Tables 7-3 and 7-4 and compares results from laboratory analyses and HPGe detector measurements for radium 226. However, the text does not note that the HPGe detector results are lower than the results from both laboratory methods for 6 of the 10 areas in Table 7-3 and 12 of the 16 areas in Table 7-4. If all three methods give equal results, then the HPGe detector should yield the lowest results one-third of the time. Using the binomial distribution, the probability of the actual results occurring by chance alone is 0.0197 for Table 7-3 data and 0.0001 for Table 7-4 data. These probabilities indicate a systematic difference between the HPGe and laboratory results for radium 226, especially at the 1-meter height (Table 7-4). The

high coefficient of variation for the radium 226 results, which results from many of the data points being near the minimum detectable activity, indicates that parametric tests for bias among the analytical methods have very low statistical power. Therefore, nonparametric tests, such as the binomial test discussed above, are more appropriate for comparisons of laboratory results and HPGe detector measurements. The text should be revised to discuss the observed low bias in the HPGe detector results for radium 226, including how the low bias is related to the correction algorithm.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 7.4.2 Page #: 7-7 Line #: 3
Original Specific Comment #: 5
Comment: Based on the data provided in Tables 7-5 and 7-6, it appears that HPGe and laboratory data for radium 226 are less comparable for HPGe measurements collected at the 1-meter detector height. Therefore, it is not clear why the 1-meter HPGe detector height was selected for acquiring additional data points. The text should be revised to clarify this issue.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 7.6.2 Page #: 7-10 Line #: 9
Original Specific Comment #: 6
Comment: The correction algorithm discussed in Section 7.6.2 is based on all available radium data. The correction algorithm should be validated by collecting new HPGe data, calculating HPGe results using this algorithm, and comparing the results to analytical results from physical samples.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 7.6.2 Page #: 7-10 Line #: 14
Original Specific Comment #: 7
Comment: The text states that the two-method average of laboratory gamma and alpha analytical results was used because these averages most likely represent the true radium concentrations. However, it is not clear what this statement is based on. In addition, the laboratory data should be supported by QC measurements such as matrix spike and matrix spike duplicate results. The method of measurement that yields results closest to the spiked sample concentrations is usually more appropriate. If multiple spikes are used that yield mixed results between the two methods, then an average of the results from the two methods may be more appropriately used. However, such QC information is not provided in the text. The text should be revised to include additional justification for selection of the two-method average.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 7.6.2 Page #: 7-11 Line #: 8
Original Specific Comment #: 8
Comment: The text states that "3.35 - 2.0 = 1.28." The text should be revised to correct this arithmetic or transcription error. The correction factors should then be checked to verify that no such

errors exist in the data sets used for calculations and in the calculations themselves.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 9.4 Page #: 9-4 Line #: 18
Original Specific Comment #: 9
Comment: The text presents an equation for calculating the minimum detectable concentration of radionuclides. The text should be revised to cite the source for this equation and to justify its applicability to HPGe measurements.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 9.6 Page #: 9-6 Line #: 22
Original Specific Comment #: 10
Comment: This section concludes that in situ HPGe measurements produce definitive (ASL D) data. However, although the calibration technique produces apparently acceptable results in the studies detailed here, it is quite different from the measurement technique. Therefore, the calibration technique may not be sufficiently robust to produce acceptable results under other circumstances such as heterogeneous distribution of radionuclides. In addition, radium 226 HPGe results are biased low, which is not acceptable for certification purposes. The text should be revised to address these issues.

Commenting Organization: U.S. EPA Commentor: Saric
Appendixes #: A, B, and C Page #: NA Line #: NA
Original Specific Comment #: 11
Comment: These three appendixes contain many pages of data. However, the pages of these appendixes are not numbered. The appendixes should be revised to include page numbers as in Appendix D.

Commenting Organization: U.S. EPA Commentor: Saric
Appendix #: G Page #: G-3 Line #: 19
Original Specific Comment #: 12
Comment: The text states that Attachment 3 includes an example report generated by the GammaVision software. However, the appendix contains only two attachments, neither of which is a report generated by the software. Either the GammaVision report should be added or the text should be modified.

Commenting Organization: U.S. EPA Commentor: Saric
Appendix #: G Page #: G-4 Line #: 19
Original Specific Comment #: 13
Comment: The equation in Line 19 defines the five-point weighted average used to smooth the data. However, the equation does not include a term for the original (central) data point. The equation should be revised to include a term for the original data point. Also, the coefficient in the denominator is 16, but the sum of the coefficients in the numerator is only 12. The equation should be corrected so that the sums of the coefficients are equal in the numerator and denominator.

Commenting Organization: U.S. EPA Commentor: Saric
Appendix #: J Page #: Table 3 Line #:NA

Finally, if consecutive duplicate measurement criteria (that is a 20 percent RPD) were instituted, four of the six certification units would have estimated data for radium 226. This issue should be discussed further in Real-Time Workgroup meetings and the text revised accordingly.

SPECIFIC COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 2.1 Page #: 2-1 Line #: 27
Original Specific Comment #: 1
Comment: The correction factor for radium 226 discussed in Section 2.1 is different from the correction factor provided in the user's manual for the real-time instruments. According to the creators of the original correction factor, the algorithm is appropriate for correcting HPGe results and additional verification should be performed. However, a new correction factor has now been developed. Past data that used or was used to develop the original correction factor should be assessed to ascertain if comparability is preserved, and independent testing of this new correction factor should be conducted for verification purposes. The text should be revised accordingly.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 2.2.1 Page #: 2-3 Line #: 10 through 18
Original Specific Comment #: 2
Comment: The text presents three cases for passing or failing certification. However, the three cases are not complete (do not include all possibilities), and some cases are logically implied but impossible, such as the average, but not the UCL on the average, exceeding the FRL. These cases should be revised to reflect the following: (1) the UCL is less than the FRL, and the maximum result is less than the hot-spot criterion; (2) the UCL is greater than the FRL, and maximum results are irrelevant; and (3) the maximum result is greater than the hot-spot criterion, and UCL results are irrelevant.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 3.1 Page #: 3-1 Line #: 16
Original Specific Comment #: 3
Comment: The text states that nine certification unit failures of FRL for thorium 232 result from artifacts of the laboratory analytical procedure. However, the text does not discuss how many other thorium analytical results are unreliable and should also be disregarded. In the absence of more specific data, it must be assumed that all associated thorium 232 results collected in Area 1, Phase I (A1P1), are doubtful. Therefore, very few reliable data sets remain for that radionuclide. The text should be revised to address this issue.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 3.2 Page #: 3-3 Line #: 15
Original Specific Comment #: 4
Comment: The text states that none of the certification units in A1P1 failed for thorium 232 based on alpha spectrometry data. However,

these data do not appear in the report. These data should be included for independent calculation purposes.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 3.4 Page #: 3-5 Line #: 3
Original Specific Comment #: 5
Comment: The conclusions presented on Page 3-5 are unsupported. For uranium, most of the individual results are less than 20 parts per million (ppm). Furthermore, all of the uranium data, except for a few points in certification units O-20, P19-23, and P18-40 of A1P1 and in certification unit 3 of Area 8, Phase I (A8P1), were less than 40 ppm, which is less than half of the FRL. Only a single data point exceeded the FRL. Because essentially all results are far removed from the FRL, it is not surprising that decisions made based on different analytical methods are identical. As noted in original specific comment 3 above, most laboratory thorium 232 results are suspect. Therefore, few fully viable data sets for that radionuclide are available for comparison. Finally, as noted in the review comments for the "Comparability of In Situ Gamma Spectrometry and Laboratory Data, Revision 1," report, the developed (radon-corrected) HPGe results for radium 226 are biased low relative to laboratory results. Such a bias, in conjunction with the tendency for HPGe results to have lower variability, would tend to result in the passing of certification units that would fail on the basis of laboratory results. The text should be revised to address these issues.

TECHNICAL REVIEW COMMENTS ON
"RADIATION TRACKING SYSTEM (RTRAK) APPLICABILITY STUDY,
REVISION 2"

GENERAL COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #:1
Comment: The calibration of the RTRAK based on HPGe measurements raises several concerns. First, efficiency calibrations were made by developing calibration algorithms to relate static RTRAK measurements to HPGe measurements. Calibration verification with the RTRAK system was made by taking repeated static measurements in the Uranium in Soils Identification Demonstration (USID) area and Drum Bailing Area (DBA) and comparing those measurements to collocated HPGe measurements. Although the algorithm is reasonable for low activities in the USID area, calibration could not be demonstrated in the DBA for elevated activity. In fact, all of the RTRAK results are biased low compared to the collocated HPGe measurements. Because the bias could approach 20 to 50 percent of the actual HPGe results, areas where measurements are two to three times the FRL warrant concern. The primary problem with the calibration algorithm is that the majority of data points correspond to data points for areas indicative of background levels. These low levels, therefore, govern the calibration curve. Additional calibration data should be collected in areas of higher activity.

Second, calibration verification should be conducted and verification demonstrated for areas that have higher activities. Because the RTRAK system is the first step in determining general areas of elevated contamination, as well as areas that have hot spots (two times the FRL, three times the FRL, and so on), calibration and subsequent verification should focus on these high-activity areas. These issues should be addressed and the text revised accordingly.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 2
Comment: Calibration of the radiation scanning system (RSS) is similar to that for RTRAK, where a calibration algorithm was developed based on collocated HPGe and static RSS measurements. However, in this case, calibration verification is demonstrated by comparing dynamic RTRAK and RSS survey results. Similarly, this verification was conducted in the USID (background) area and DBA. Although the USID area showed comparatively similar RTRAK and RSS results, the DBA demonstrated effects resulting from heterogeneities. However, because calibration of the RTRAK system is questionable, comparison of RTRAK to RSS results is also questionable. In this case, RSS calibration verification should have been conducted using static measurements collocated against the HPGe system. These issues should be addressed and the text revised accordingly.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 3
Comment: Although theoretical detection levels corresponding to area and radionuclide activities have been developed, it has not been demonstrated that these levels are achievable in the field. Based on the data provided in Table 4-19 for the RTRAK system and Tables C-1 through C-3 for the RSS, it is not clear if these instruments can adequately identify contamination hot spots. This issue should be further addressed and the text revised accordingly.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 4
Comment: It is not clear why a 31-centimeter (cm) detector height was used for the HPGe system for calibration purposes. The text states that the 31-cm height produced the best results for calibration purposes. However, because the HPGe detector field of view is almost four times larger than that of the static RTRAK or RSS field of view, it is not clear how the RTRAK and RSS instruments could be used to demonstrate hot spot detection. The text should provide a more adequate explanation for use of the 31-cm HPGe detector height because the RTRAK and RSS systems may be typically used in heterogeneous areas. Otherwise, it appears that the RTRAK and RSS systems should be calibrated against the 15-cm HPGe detector height. This issue should be further considered and the text revised accordingly.

SPECIFIC COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 2.1 Page #: 2-1 Line #: 7
Original Specific Comment #: 1
Comment: The text states that Figure 2-1 shows the location of the USID and South Field study areas. In addition, line 19 on page 2-4 states that Figure 2-1 shows the USID study area. However, Figure 2-1 shows the "Southfield Area" but not the USID area. The figure should be corrected to show the USID area.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 4.1.1 Page #: 4-17 Lines #: 1 through 4
Original Specific Comment #: 2
Comment: These lines present equations 6 through 8. Each of these equations includes the term "t^2" in the denominator of the right-hand side. The text should be revised to define this term.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 4.5 Page #: 4-31 Line #: 19
Original Specific Comment #: 3
Comment: This section compares RTRAK measurements to HPGe measurements. RTRAK results should also be compared directly to the corresponding laboratory analytical results. This direct comparison would minimize the propagation of error in the measurements and calculations. This same comment applies to Section C.3, which describes the RSS.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 4.5.2 Page #: 4-36 Line #: 17
Original Specific Comment #: 4
Comment: The text states that Table 4-22 compares certain HPGe results with corresponding RTRAK results. However, Table 4-22 is omitted from both this section and the table of contents. Either the text should be revised to include Table 4-22, or the citation in the text should be corrected.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 5.2.1 Page #: 5-4 Line #: 27
Original Specific Comment #: 5
Comment: The text states that some results are summarized in Table 5-3. This citation should be corrected to read "Table 5-2."

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 6.0 Page #: 6-4 Lines #: 4 and 5
Original Specific Comment #: 6
Comment: The bold-faced text of this conclusion is incomplete; therefore, the text is unclear. The text should be revised to include the missing words for clarification.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: A.3.1 Page #: A-3 Line #: 25
Original Specific Comment #: 7
Comment: The text states that pulse height is proportional to gamma ray energy. Actually, pulse height in a multichannel analyzer is proportional to gamma ray flux (number of gamma rays), but the channel on which the pulse appears is proportional to the individual gamma ray energy (that is, its frequency or inverse wavelength). The text should be revised accordingly.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: A.4.1 Page #: A-8 Line #: 21
Original Specific Comment #: 8
Comment: The text describes intermittent limitations on the quality of global positioning system (GPS) data. It would be useful to note that quality is generally a function of the number of satellites with an unobstructed line of sight to the receiving antenna. This section should be revised to note that the number of satellites from which data are actually received, as well as other parameters such as discontinuities in apparent position, are used for QC purposes in assessing real-time data.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: C.1 Page #: C-1 Line #: 21
Original Specific Comment #: 9
Comment: The text states that the RSS is essentially identical in operation to the RTRAK, with a gain of 5.84 kiloelectronvolts (keV) per channel. However, line 17 on Page A-4 states that the RTRAK has a gain of 5.85 keV per channel. Because input data (channel number and peak energy) are the same for the RSS and the RTRAK, the same calculation result, including rounding off, should be presented for both systems. The text should be revised accordingly.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: C.4 Page #: C-13 Line #: 19
Original Specific Comment #: 10
Comment: The text concludes that strong statistical evidence indicates that the RTRAK and RSS measurements are equivalent. Because of the irregularities of the results from the USID area, which are discussed in Section C.4, it is more accurate to state that the two systems produce equivalent results at levels that reasonably exceed the minimum detection concentration. The text should be revised accordingly.

TECHNICAL REVIEW COMMENTS ON DRAFT DATA VALIDATION CHECKLIST FOR VALIDATING HIGH-PURITY GERMANIUM DETECTOR MEASUREMENTS TO ANALYTICAL SUPPORT LEVEL D

GENERAL COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 1
Comment: As a result of Real-Time Workgroup meetings, it was generally understood that the ratio of low- to high-energy photon fluence would be evaluated during the data validation process. This step would provide a means for identifying a high-activity source located some distance underneath the soil surface. For this reason, the data validation checklist should be revised to include an element to account for data related to the ratio of low- to high-energy photon fluence.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 2
Comment: The attached "In-Situ HPGe Gamma Spectrometry Detector Annual Calibration Checklist" is generally deficient for supporting production of definitive data because it relies on multiple readings of point sources, followed by mathematical manipulation to generate a result. This sort of calibration is quite different from the use of the instrument to measure activity in a defined volume of space. This issue should be discussed during Real-Time Workgroup meetings.

SPECIFIC COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 1.1 Page #: 1 Line #: NA
Original Specific Comment #: 1
Comment: The data validator is instructed to determine the completeness of the data package. The checklist should either be revised to cite completeness criteria or provide a reference for these required elements.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 2.9 Page #: 4 Line #: NA
Original Specific Comment #: 2
Comment: The first phrase of this sentence is unclear. The phrase should be revised to read, "Were the daily checks (efficiency and resolution) performed with...." or whatever is meant.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 5.2 Page #: 6 Line #: NA
Original Specific Comment #: 3
Comment: The action items referred to here should be "...according to 5.3 or 5.5" and "...according to 5.4 and 5.6," respectively. This modification would provide explicit references to the low-activity procedures for evaluating duplicate field results.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 5.4 Page #: 7 Line #: NA
Original Specific Comment #: 4
Comment: It is not clear why an allowance of 35 percent RPD is instituted for nonconsecutive duplicate results. Because the HPGe system averages radionuclide concentrations over relatively large areas, minor misalignments in repositioning the detector should not result in major disagreements in duplicate results. A major misalignment would be prevented by proper use of the GPS. Furthermore, certification measurements are intended to be taken in areas considered to demonstrate homogeneous contamination. For precertification HPGe measurements, a 35 percent RPD may be feasible; however, for certification, a 20 to 25 percent RPD would be more appropriate. This issue should be discussed during Real-Time Workgroup meetings and the text revised accordingly.

TECHNICAL REVIEW COMMENTS ON UPDATED SECTION 2.5, (REVISION B), FOR INCORPORATION INTO USER'S MANUAL, ENTITLED "CERTIFICATION"

GENERAL COMMENT

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 1
Comment: Because of the extent of comments and unresolved issues associated with HPGe comparability study documents, incorporation of updated Section 2.5 entitled "Certification" into the user's manual is not acceptable at this time. Issues related to HPGe comparability study documents should be discussed during Real-Time Workgroup meetings.

SPECIFIC COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Figure #: 2.5-1 Page #: NA Line #: NA
Original Specific Comment #: 1
Comment: The figure discusses surveying sampling locations in the field to locate and finalize the given areas. However, it is not exactly clear what the text means by "survey." The text should be revised to clarify what type of survey will be conducted.

Commenting Organization: U.S. EPA Commentor: Saric
Figure #: 2.5-2 Page #: NA Line #: NA
Original Specific Comment #: 2
Comment: According to this figure, if a certification unit fails because of a large variation in the data, additional physical samples or HPGe measurements will be performed. However, if the unit still fails due to variability during further physical sampling or measurements, it is not clear what action would be taken. If the unit fails certification based on additional sampling data, either the certification unit should be further remediated or other DOE management decisions should be explored. The text should be revised accordingly.

TECHNICAL REVIEW COMMENTS ON UPDATED SECTION 3.7, (REVISION B), FOR INCORPORATION INTO USER'S MANUAL, ENTITLED "CERTIFICATION MEASUREMENTS"

GENERAL COMMENT

Commenting Organization: U.S. EPA Commentor: Saric
Section #: NA Page #: NA Line #: NA
Original General Comment #: 1
Comment: Because of the extent of comments and unresolved issues associated with the HPGe comparability study documents, incorporation of updated Section 3.7 entitled "Certification Measurements" into the user's manual is not acceptable at this time. Issues related to HPGe comparability study documents should be discussed during Real-Time Workgroup meetings.

SPECIFIC COMMENTS

Commenting Organization: U.S. EPA Commentor: Saric
Figure #: 2.5-2 Page #: NA Line #: NA
Original Specific Comment #: 1
Comment: According the this figure, if a certification unit fails because of a large variation in the data, additional samples or HPGe measurements will be taken. However, if the unit still fails due to variability upon further sampling or measurements, it is not clear what action would be taken. If the unit fails certification based on additional sampling data, either the unit should be further remediated or other DOE management decisions should be explored. The text should be revised accordingly.

Commenting Organization: U.S. EPA Commentor: Saric
Section #: 3.7.5 Page #: 3.7-3 Line #: 23
Original Specific Comment #: 2
Comment: The guidance bullet states that if any HPGe detector result is a possible data outlier, it will be remeasured and an average of the two measurements (two out of two or the closer of two out of three) will be used for certification. No comparable remeasurement procedure exists for laboratory measurements; laboratory measurements are either rejected or considered valid and used as generated. The use of selected composite in situ measurements will tend to bias the overall statistics toward the mean. This will, in turn, decrease the probability that a unit will be rejected for certification. The use of such data manipulation must be thoroughly justified. One alternative procedure would be to remeasure a certification unit using a new, randomly generated set of 16 HPGe measurement locations, when an apparent data outlier is detected in the original measurement set. The text should be revised accordingly.

Commenting Organization: U.S. EPA Commentor: Saric
Section #:3.7.5 Page #: 3.7-3 Line #: 24
Original Specific Comment #: 3
Comment: The intent of the second guidance bullet is unclear. It appears as though the specific criteria for HPGe measurements may lead to unnecessary field time for areas that would be considered clean.

For example, applying the criteria established in the guidance bullet for the data associated with certification results of A8P1 in the January 1999 comparability study would result in at least 14 additional measurements. Both HPGe measurements and laboratory analysis indicate that this unit meets FRLs. Furthermore, if these criteria are to be instituted, an RPD higher than 20 percent should be allowed for additional measurements. The text in this bullet should be re-evaluated and revised as necessary.

ENCLOSURE 2

ADDITIONAL TECHNICAL REVIEW COMMENTS ON
REVISED REAL-TIME IN SITU GAMMA SPECTROMETRY REPORTS AND
DOCUMENTATION TO SUPPORT THE USE OF HIGH-PURITY GERMANIUM
DETECTORS TO PERFORM FINAL SOIL CERTIFICATION
FOR PRIMARY RADIONUCLIDES

FERNALD ENVIRONMENTAL MANAGEMENT PROJECT

(6 Pages)

Attachment to: M-00893

TECHNICAL REVIEW COMMENTS ON "COMPARABILITY OF IN SITU GAMMA SPECTROMETRY AND LABORATORY DATA AND DECISIONS FOR CERTIFICATION UNITS, REVISION 0"

GENERAL COMMENTS

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: NA (general comments on statistical methods)
Original General Comment #: 1
Comment: There are many methods available for comparing two sets of data. This report compares the results of *in-situ* gamma spectrometry measurements with laboratory data by:

1. using the average absolute percent difference between the certification unit (CU) means for each contaminant derived from each set of data, and
2. comparing the results of the compliance determination that would be made based on the upper confidence level (UCL) for the CU mean indicated by each set of measurements.

Results of the first comparison demonstrate that the two sets of measurements are almost always within 35% RPD. Exceptions are identified that indicate large values for the RPD occur only when the results are near the detection limit of the methods.

Results of the second comparison show that almost all CUs demonstrate compliance with the required remediation levels regardless of the data set used. Exceptions generally occur when unusually high laboratory measurements are observed without corroboration from the *in-situ* data. The *in-situ* method integrates over a much wider area than is covered by the collocated soil sample. The laboratory analysis of the soil samples is only expected to produce higher results than the *in-situ* method when small, isolated areas of elevated activity occur within the *in-situ* field of view and are included as part of the soil sample.

MARSSIM and EPA QA/G-9 (*Guidance for Data Quality Assessment*) provide guidance on evaluating data. Part of this evaluation is selecting the appropriate statistical tests. The basis for selecting these tests should be discussed, and the discussion should describe how these tests meet the objectives for the survey.

Both methods consider the mean values obtained in each CU. However, page 2-5 lists the mean, the 95% UCL, and the maximum concentration as considerations for compliance demonstration. There are additional comparisons that should be considered based on individual data points for the co-located data including correlation, regression, and graphical methods.

Correlation and regression analysis can identify the degree of association between co-located data. Outliers that affect the results may be clearly identified and subject to further investigation. Graphical analysis would better summarize the large volume of data presented in this report. Scatter plots

comparing the results of the two analytical methods would demonstrate the degree of correlation, or lack thereof. A fundamental tool for comparison of alternative decision-making methods is the operating characteristic curve (OCC). An empirical estimate of the curve can be constructed by plotting the cumulative rejection percentage on the vertical axis and the contamination level on the horizontal axis. A good decision tool will produce an OCC that remains relatively flat at concentration far below the remediation level, rises sharply as the remediation level is approached, and approaches 100% at contamination levels slightly above the remediation level. OCCs may be constructed for data sets containing either co-located or non-co-located data points, and may include 1 or more CUs providing a broader range for the comparison. The empirical OCC permits a rapid visual determination of three important features of each measurement method:

1. ability to maintain a low false positive rate (Type II error) at contamination levels far below the remediation level,
2. ability to maintain a low false negative rate (Type I error) at contamination levels above the remediation level, and
3. a sharp rise in the rejection rate as the remediation level is approached

An important consideration is the application of the results of the comparison to other areas at FEMP. The current analysis is not constructed to facilitate this consideration. The current analysis reports only the average mis-classification rates averaged over all contamination levels, and the results are inconclusive. The mis-classification rates must be reported as a function of the contaminant concentration level to permit extrapolation to future surveys. If the areas and CUs where these methods might be applied in the future contain concentration higher than those encountered in this study, the current analysis is inadequate to predict the performance of the two measurement methods for making decisions regarding compliance in these areas. Constructing an OCC will help improve the comparison and permit extrapolation of the results to other areas.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: NA
Original General Comment #: 2
Comment: The majority of the time the decision made based on *in situ* gamma spectrometry results and laboratory data on whether or not the release criterion has been exceeded is exactly the same. In the few instances where there is a disagreement, the average is below the final remediation level and the difference in the uncertainty causes the disagreement. The comparison based on the 95th upper confidence level is consistent with Superfund guidance. However, this comparison only addresses the Type I decision error where a contaminated site may be incorrectly released. Using hypothesis testing (as recommended in MARSSIM, Superfund's Soil Screening Guidance, and EPA's QAD guidance documents) both Type I and Type II decision errors can be addressed. Evaluating the Type II decision error can help determine if the reason for these differences is based on an insufficient number of measurements

(not enough power) or if the level of contamination is a problem. Also, the overall survey design can be evaluated to determine if it can be used effectively during subsequent decommissioning activities at the site.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: NA
Original General Comment #: 3
Comment: There is no discussion of the development of data quality objectives (DQOs) for this project and how they might affect the interpretation of results. The development of the final remediation levels is not discussed. Have the certification unit boundaries been compared to the assumptions used to develop the final remediation levels to ensure they are compatible? What other assumptions were used to develop the survey design? The level of detail provided with this report is insufficient to evaluate the overall demonstration of compliance. Only the statistical test results are provided, and only those results have been reviewed.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: Section 3
Original General Comment #: 4
Comment: Section 3 discusses the comparability of the results. The methods used to compare the data include the assumption the data are normally distributed. There is no discussion of whether or not this assumption has been verified. The use of nonparametric statistics is recommended in MARSSIM and Superfund's Soil Screening Guidance to avoid verifying assumptions of normality. Graphically representing the results of a preliminary data review (as recommended in MARSSIM and EPA QA/G-9) would improve the readability of this section. The use of posting plots, histograms, and power curves could be used to visually interpret the results, compare the two methods, and evaluate the effectiveness of the survey design.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: Not applicable (NA)
Original General Comment #: 5
Comment: The references (page R-1) are not all referred to in the text. For those that are, it is difficult to determine which reference is being cited. The text makes extensive references to other reports with the assumption that the reader has access to these reports. For example, page 2-1, 4th paragraph, last sentence; page 2-4, beginning sentences of Section 2.2.2; page 2-5, 1st sentence following the definition of relative percent deviation (RPD, also referred to as percent relative deviation).

SPECIFIC COMMENTS

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 2.0
Original Specific Comment #: 1
Comment: The concept of the relative standard deviation (RSD, or coefficient of variation) should be defined in this section. The RSD is used as part of the data evaluation in Section 3.0.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Figures #: 2.0 Page #: 2-5
Original Specific Comment #: 2
Comment: Page 2 through 5 notes that a greater number of measurements or CUs will increase the representativeness of each data set and reduce variability. However, it is not clear if the contaminants are assumed to be homogeneously distributed or if the presence of small areas of elevated activity is suspected. If small areas of elevated activity are present, increasing the number of measurements may result in an increase in variability.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 2.1
Original Specific Comment #: 3
Comment: The selection of measurement methods and sample locations does not address whether the areas and soils reflect different types of materials (e.g., native soils and imported backfill). The text should describe surface conditions such as ruts, furrows, undulations, etc., which might impact the *in-situ* measurements by modifying the detector's field of view. There is no discussion of soil properties that might lead to elevated radon accumulations at or near the surface. The characterization of soil properties should discussed in more detail along with the potential impact on the analytical results and how this might affect the variability in the results.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 2.2.1
Original Specific Comment #: 4
Comment: The goal stated in the first sentence does not mention small areas of elevated activity as a potential concern. However, the third bullet includes the maximum concentration as a consideration for demonstrating compliance, which implies that small areas of elevated activity might be a concern. MARSSIM guidance describes an integrated survey design that combines concerns of homogeneous and heterogeneous contaminant concentrations into a single survey design. Also, the use of the term "hot spot" is not recommended because it implies there is automatically a health risk when this may or may not be the case.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 2.2.2
Original Specific Comment #: 5
Comment: A brief summary of the comparability discussion between HPGe and laboratory data would be helpful for readers who may not be familiar with the December 1998 Comparability Study.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 2.3
Original Specific Comment #: 6
Comment: The role of "hot spot criteria" in comparability should be discussed. This is particularly important for the *in-situ* measurements since they represent the "average" activity over a large area. A CU that demonstrates compliance using the *in-situ* results may fail based on the "hot spot" criteria. MARSSIM provides guidance on evaluating the dose or risk from a small area

of elevated activity based on the size of the area and the radionuclide-specific relationship between area and dose or risk.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 3
Original Specific Comment #: 7
Comment: The data listed in Table 3-5 and Table 3-6 have not been corrected for the different numbers of measurements from *in-situ* and laboratory analyses. For example, there are 457 laboratory results and 399 *in-situ* results for the A1P1 32-cm and 100-cm data sets. The analysis based on the RPD does not take the different number of results into account. In assessing the variability of the data, the standard error (standard deviation divided by the square root of the number of measurements) might be considered which takes into account the different number of data points.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 3.1
Original Specific Comment #: 8
Comment: The discussion addressing comparability of total uranium should clearly state the assumption about the isotopic distribution. This is especially important for a site like Fernald where different mixtures of uranium were used.

The UCL for nine CUs were artifacts associated with the laboratory procedure, which is described in a separate document. The discussion should include a brief explanation of the important details and assess their significance to the analysis.

The discussion of anomalous conditions leading to high near-surface radon accumulations should include the rationale for identifying these results as anomalous instead of accurate representations of the ^{226}Ra concentrations. Given the significance of this finding and the implications in correcting for radon equilibrium, there is a need to present more information on the nature or condition of the soil that could have produced these results. There should be an assessment of how prevalent this condition might be at Fernald (e.g., is this an isolated incident) and how this information can be used to identify similar conditions at other locations.

Commenting Organization: U.S. EPA Commentor: Jablonowski
Section #: 3.2
Original Specific Comment #: 9
Comment: The basis for selecting the 20% RPD and 35% RPD action levels for comparability should be identified and explained. EPA QA/G-5 (*EPA Guidance for Quality Assurance Project Plans*) includes a discussion of evaluating comparability as a survey objective. For this analysis selecting a single value rather than a range of values may provide a criterion that is more technically defensible, or at least easier to defend.

Also, a discussion concerning changes made in the way laboratory results were calculated and why this resulted in lower ^{232}Th concentrations should be included.

Commenting Organization: U.S. EPA Commentor: Jablonowski

Section #: 3.3
Original Specific Comment #: 10
Comment: The percent relative standard deviation (RSD) reported in Table 3-6 do not use the same evaluation criteria used for RPD. The discussion of representativeness should identify a criterion for deciding whether or not the results are acceptable. EPA QA/G-5 (*EPA Guidance for Quality Assurance Project Plans*) includes a discussion of evaluating representativeness as a survey objective.

This section should include a discussion about the interpretation of the results and their applicability to other areas of the FEMP. See the general statistics comments for additional comments.