File
LANL
TA 53

Los Alamos National Laboratory

ENVIRONMENTAL RESTORATION

University of California
Environmental Restoration, MS M992
Los Alamos, New Mexico 87545
505-665-4557
FAX 505-665-4747

U. S. Department of Energy
Los Alamos Area Office, MS A316
Los Alamos, New Mexico 87544
505-665-7203
FAX 505-665-4504

Date: AUG 3 1 1995
Refer to: EM/ER:95-458

Ms. Barbara Hoditschek, Program Manager
Hazardous & Radioactive Materials Bureau
New Mexico Environment Department
P. O. Box 26110
2044 Galisteo Street, Bldg. A
Santa Fe, NM 87502

RECEIVED

SUBJECT: RESPONSE TO TECHNICAL AREA 53, NE & NW SURFACE IMPOUNDMENTS (FORMER OPERABLE UNIT 1100)

Dear Barbara:

Enclosed is our response to the New Mexico Environment Department's notice of deficiency issued on the subject closure plan.

Please call Gene Gould at 667-0402 or Everett Trollinger at 667-5801, if you have any questions.

Sincerely,

Jorg Jansen, Program Manager
Environmental Restoration

Sincerely,

Theodore J. Taylor, Program Manager
Los Alamos Area Office

Environmental Restoration

4312

TC

JJ/TT/bp

Enclosure: Response to TA-53, NE & NW Surface Impoundments

Cy (w/enc.):
B. Driscoll, EPA, Region 6
D. Griswold, ERD, AL, MS A906
G. Gould, ESA-DE, MS G787
E. Merrill, EM-453, DOE-HQ
T. Taylor, LAAO, MS A316
E. Trollinger, LAAO, MS A316
N. Weber, Bureau Chief, NMED-AIP
J. White, ESH-19, MS K498
S. Yanicak, NMED-AIP
EM/ER File, MS M992
RPF, MS M707

Cy (w/o enc.):
T. Baca, EM, MS J591
G. Rael, ERD, AL, MS A906
W. Spurgeon, EM-453, DOE-HQ
J. Vozella, LAAO, MS A316

CERTIFICATION

I certify under penalty of law that these documents and all attachments were prepared under my direction or supervision in accordance with a system designed to assure that qualified personnel properly gather and evaluate the information submitted. Based on my inquiry of the person or persons who manage the system, or those persons directly responsible for gathering the information, the information submitted is, to the best of my knowledge and belief, true, accurate, and complete. I am aware that there are significant penalties for submitting false information, including the possibility of fine and imprisonment for knowing violation.

Document Title:

Response To Technical Area 53, NE & NW Surface Impoundments (Former Operable Unit 1100)

Name: _____ Date: 8-29-95

Jorg Jansen, Project Manager
Environmental Restoration Project
Los Alamos National Laboratory

or

Tom Baca, Program Director
Environmental Management
Los Alamos National Laboratory

Name: _____ Date: 8/31/95

Joseph Vozella,
Acting Assistant Area Manager of
Environment Projects
Environment, Safety, and Health Branch
DOE-Los Alamos Area Office

or

Theodore J. Taylor
Program Manager
Environment Restoration Program
DOE-Los Alamos Area Office

RESPONSE TO COMMENTS
ON LANL's RESPONSE TO NOD
for
TA-53 MIXED WASTE SURFACE IMPOUNDMENTS

ENCLOSURE A

A-1) LANL Comment 1, Discussion, p.1. ***...In addition, the reference for the new text on p.2-14 will be added to p.7-4.***

This reference should also be noted in the new text for p. 2-14.

Response:

The text will be modified as indicated by the comment.

A-2) *LANL Comment 2, Proposed Text Changes, p. 4, ¶1.* ***The approach for determining whether the closure performance standard has been met is illustrated in Figure 5-1. This approach is based on two methodologies: comparison with screening action levels (SALs) developed using the methodology outlined in the Installation Work Plan (LANL 1993) and development of a baseline risk assessment using methodology identified by the EPA (EPA 1989) . . .***

In addition, information on radioactive constituents detected should be used in preparing the health-based cancer risk assessment. This information is necessary to provide a better estimate of the true cancer risk at these lagoons, and thus to manage the hazardous constituents in an appropriate manner.

Response:

LANL agrees to provide information on potential dose from radionuclides on site for comparison with DOE guidelines for industrial exposure. By voluntarily submitting this data LANL neither admits nor concedes that NMED has authority to regulate, under RCRA or the New Mexico Hazardous Waste Act, source, special nuclear or by-product material as defined in the Atomic Energy Act. Potential dose from radionuclides and potential cancer risk and health hazard from chemicals on site will be tabulated separately. The radionuclide dose and the chemical risk and hazard will be calculated based on a continued industrial land use scenario. See response to comment #34 (Enclosure B).

A-3) *LANL Comment 2, Proposed Text Changes, p. 4, 53.* ***. . .Background will be defined as the 95% upper tolerance level (UTL) calculated from concentrations of inorganic constituents measured in soil and tuff similar to that present near the TA-53 lagoons. If existing data are not sufficient to provide a statistically meaningful UTL, then additional samples will be collected near the lagoons in locations believed to be unaffected by releases from Laboratory facilities . . .***

Whether or not existing data are sufficient must be decided now. See Comment A-11 and Comment #2, page 4, in Enclosure B.

Response:

The surface fill and stratigraphy underlying the TA-53 lagoons are comprised of Unit 2B of the Bandelier Tuff. Existing data for this unit of the Bandelier Tuff are sufficient to provide a statistically meaningful UTL for inorganic and radionuclide constituents. Background will be defined as the 95% upper tolerance limit (UTL) on the 99th percentile calculated from concentrations of measured inorganic chemicals. These data have been collected, analyzed, and will be reported in a LANL ER Project Earth Sciences Council document to be issued in September, 1995 (LA-UR-95-1217, 1995).

A-4) *LANL Comment 2, Proposed Text Changes, p. 4, §4.* ***If the maximum concentration in soils, subsoil, and tuff is less than the SAL, then the clean closure performance standard will be met for that constituent . . .***

The SAL is not a clean-up level. As discussed during the meeting on November 9, 1994, the risk assessment will take into account the cumulative impact of all hazardous waste constituents. This statement should be changed accordingly.

Response:

LANL agrees that the SAL is not a clean-up level. However, the SAL comparison is used to evaluate whether further action or analysis (e.g., risk assessment) at a site is warranted. At the request of NMED, the **bold** text (above) will be revised to state the following: "If only one constituent is detected and the maximum concentration in soil, sub-surface soil, and tuff is less than the SAL, then the clean closure performance standard will be met." Multiple detected constituents are addressed in the following sentence (p. 4 of the December 16, 1994 LANL Response to NOD, which will not be changed. Text addressing the constituents evaluated in a baseline risk assessment is presented in the first sentence of the first full paragraph on page 5 of the December 19, 1994 LANL Response to NOD.

A-5) Comment 2, Proposed Text Changes, p. 4, ¶4. ***...If multiple constituents are detected and the total sum of the SAL comparison ration (maximum concentration/SAL) is less than 1 (LANL 1993b), then clean closure will be met (see Section 3.1 for a detailed discussion of multiple constituent comparisons)...***

The document containing this Section 3.1 should be identified and noted in the References. LANL 1993b should also be identified in the References.

Response:

The text will be modified to clarify that the reference is to Section 3.1 of the Closure Plan. The LANL 1993b citation is currently listed in the references.

A-6) *LANL Comment 2, Proposed Text Changes, p. 5, §1. . . .* ***If many TICs are present, or the TIC concentrations appear high, an effort will be taken to positively identify and reliably measure the concentrations (EPA 1990) . . .***

"many" and "high " are imprecise terms. This protocol should be more specific.

Response:

EPA Risk Assessment Guidance for Superfund, Volume 1, 1989 (p. 5-18) states: "If **many** TICs are present relative to the [target analyte list] TAL and [target chemical list] TCL compounds identified, or if TIC concentrations appear **high**. . . then further evaluation of TICs is necessary." Therefore, LANL proposes to collect and analyze additional samples with the full Appendix IX calibration in an effort to accurately identify and quantify possible toxic chemicals under the following conditions: 1) if a TIC is detected in approximately 10% or more of the samples, and is a chemical of potential concern (i.e., an Appendix IX constituent), or 2) if the estimated concentration of an Appendix IX constituent is higher than the screening action level.

A-7) LANL Comment 5, Discussion, p. 8. ***...J-flagged concentrations will not be eliminated from consideration in the baseline risk assessment based on comparisons with screening action levels. The text in Section 3 and Appendix K (see response to Comment 2) has been modified in accordance with the comment and discussions conducted with NMED on 11/5/94.***

This change has not been made in Section 3. Section 3 should be revised accordingly.

Response:

The indicated change will be made in Section 3.

A-8) LANL Comment 6, Discussion, p. 9, ¶2. ***...In addition, the phrase "three times"... will be deleted in order to be consistent with actual field practice.***

Actual field practice should be defined.

Response: -

The text will be modified to indicate that equipment is rinsed once.

A-9) LANL Comment 11, Proposed Test Changes, p. 13. ***In addition, if hazardous constituents are detected and confirmed before closure is certified, appropriate action will be taken to address the presence of these constituents.***

"appropriate action" is vague. Possible appropriate actions, e.g., vadose zone investigation/characterization /remediation should be specified.

Response:

The text will be modified to indicate the range of possible responses, varying from further vadose zone characterization to remediation if that might be required. The text will indicate that specific actions that might be taken will depend on the nature and extent of the detected hazardous constituents, and cannot be specified in advance.

A-10) LANL Comment 12, Discussion, p.14, ¶1 ...***(LANL, 1994a; EPA, 1989)***...

EPA 1989 and LANL 1994a should be identified in the References

Response:

The citations for LANL 1994(a) and EPA 1989 are listed in the references for Appendix K. These will be added to the references for the body of the text.

A-11) *LANL Comment 18, Discussion, p.19, ¶1.* ***LANL proposes to identify a subset of the Laboratory background data that applies to soils and tuff similar to that present in the vicinity of the TA-53 lagoons. These data will be used to establish background concentrations for comparison with measured metals or radioactive constituents in the soils and tuff underlying the lagoons, If existing applicable data are not sufficient to provide a meaningful statistical analysis of the background concentration, additional samples will be collected at locations near the lagoons that are believed to be unaffected by releases from the TA-53 facilities.***

This needs to be more specific. Data sufficiency, statistical methodology, and, if necessary, a sampling and analysis plan to determined background need to be decided now and included in the closure plan. If existing data are to be used, they must be presented in the closure plan, along with a justification for their use, before the closure plan can be approved. See HRMB Comment No. A-3.

Response:

See response to comment A-3.

A-12) LANL Comment 21, Proposed Text Changes, p. 21, ¶3. ...***the liner will be decontaminated by steam cleaning followed by rinsing with clean water.*** ~~*The wash and rinse water will be dispensed of to the impoundments.*~~ ***The liner will be field screened...***

The rinse water also must be tested. Provision for proper disposal of the rinse water, if it is found to be hazardous, should be noted in closure plan.

Response:

The rinse water is a decontamination waste, and will be handled as indicated for such wastes. This will involve sampling and proper disposal as described in Section 5.3.2 of the closure plan.

A-13) LANL Comment 22, Discussion, p.22. ***LANL agrees with the comment and will submit one Final Closure Report instead of a series of reports. Section 6.0 and Appendix K will be completely revised and all descriptions and references to the series of reports will be revised.***

The report referred to in this comment is an interim report, and should contain the information listed in NMED comments (Comment No. 22) dated October 27, 1994. It should be submitted after characterization sampling has been completed and before implementation of closure starts. NMED approval of risk management and consequent implementation activities, based on the sampling results and risk assessment, will be required. If a sampling and analysis plan for delineation of hot spots, removal of contaminated material, and confirmatory sampling is necessary, NMED must approve prior to implementation.

Response:

LANL will conform with the procedures stipulated in the comment, and will modify the text accordingly.

A-14) LANL Comment 22, Proposed Test Changes, P. 22, ¶1. ***...Application of the above process to demonstrate clean closure will be documented in a ~~series of reports~~ the Final Closure Report to be submitted to NMED, as described in Section ~~5.2.2~~ 6.0.***

See HRMB Comment No. A-13 above.

Response:

See response to Comment #13 above.

A-15) Comment 30, Discussion, p. 32. ***See response to Comments 2 and 18.***

There is nothing in the responses to Comments 2 and 18 that indicate that changes will be made as requested to p. K-3.

Response:

The response to Comment 2, referred to in the response to Comment 30, indicates that changes will be made to both Section 3 and Appendix K. Proposed changes to the text for Appendix K were included in the response to Comment #30. See also the response to Comment No. A-3 in this document.

ENCLOSURE B

1) Responses to technical review of the NOD comments for the TA-53 surface impoundments closure plan dated December 16, 1994:

#2 *(page 3) A minimum of 10 samples per exposure unit is necessary to give a 95% UCL of the mean. Please provide plots of observations versus concentration to indicate distribution and show how the 95% UCL of the mean was calculated. If LANL is unable to demonstrate this, LANL must either take more samples per exposure unit or use the highest measured value for the concentration term in order to calculate the reasonable maximum exposure. If LANL use[s] the highest measured value and cannot demonstrate a 95% UCL, they must state this. Reference: EPA publication 9285.7-081, May 1992 "Supplemental Guidance to RAGS: Calculating the Concentration Term."*

Response:

LANL agrees to use either the maximum chemical concentration or the 95% UCL of the mean when calculating dose, potential health hazard, or cancer risk. The 95% UCL of the mean will be based on a minimum of 10 samples. A Shapiro-Wilks statistic will be used to indicate whether the data are distributed normal or log-normal. Formulas for the computation of the 95% UCL of the mean will be provided for normal or log-normal distributed data. If the data do not meet the criteria for either a normal or log-normal distribution, a non-parametric procedure will be used to derive the 95% UCL of the mean. If the 95% UCL of the mean is greater than the highest measured concentration, then the highest measured value will be used as the concentration term (p. 3, EPA 1992, "Supplemental Guidance to RAGS: Calculating the Concentration Term).

(Page 4) LANL must assure that determination of background concentrations for naturally occurring metals is conducted by using samples from same strata and soil types found at the site. LANL may compare their site specific values to the facility-wide background study to assure that the results are within facility-wide range. Also, LANL must provide a site-specific sampling and analysis plan to determine background concentrations for naturally occurring metals. This plan may incorporate relevant facility-wide background study results.

Response:

See response to comment A-3 (Enclosure A).

(page 5) If LANL cannot adequately indicate 95% UCL of the mean then the highest concentrations should be used including those detected at greater than or equal to five times the limit of detection (see comment #2 (page 3) above).

Response:

See response to comment #2 (page 3) above.

(pages 4 & 5) If LANL uses the Tolerance Interval for determining the background upper (tolerance) critical limit, and the Confidence Interval to average the verification sample results from a closure/remediation activity, the resulting comparison would involve comparing totally different single parameters (comparing a mean to a maximum with an unknown sample distribution). This is not how the statistical methods were intended to be used. LANL must propose a consistent and acceptable method for comparing background results with sample results.

Response:

LANL will compare the maximum sample concentrations to the background upper tolerance limits. This comparison will be conducted as prescribed in Statistical Comparisons to Background, Part I (Los Alamos National Laboratory Environmental Restoration Project, Assessments Council, March 1995, LA-UR-95-1217) (attached).

(page 6) ***"The additive effects of multiple constituents is similarly evaluated by adding the ratios of the SAL comparison values (maximum concentration/SAL) for each constituent with similar toxic end point (e.g., cancer, kidney effects, liver effects, etc.)"*** *Calculating risk assessment based on toxic endpoint organ is not acceptable. LANL must quantify exposure and calculate toxicity assessment as outlined in the Risk Assessment Guidance for Superfund (December 1989).*

Response:

The text to which NMED refers (i.e., page 6 in **bold**, above) addresses the screening approach used to evaluate whether further action or analysis (e.g., risk assessment) at a site is warranted. However, NMED's comment refers to risk assessment, which is addressed in Appendix K of the Interim Status Closure Plan.

However, LANL proposes to change the approach for the screening assessment to follow the current guidelines proposed by LANL's Decision Support Council for the Environmental Restoration Project (i.e., chemicals will be grouped and evaluated separately according to carcinogenicity, noncarcinogenic health effects, and radionuclides for the multiple constituent analysis). This is similar to RAGS approach for calculating risk (i.e., risk related to chemical carcinogens, chemical noncarcinogens [Exhibit 8-1, p. 8-3, RAGS 1989] and radionuclides [final paragraph, Section 10.7.3, p. 10-33, RAGS 1989] should be tabulated separately).

According to EPA guidance for risk assessment (RAGS 1989), the process of calculating potential health hazard related to noncarcinogens is an iterative process that allows for considering effect and mechanism of action (i.e., toxic end point) (p. 8-14 RAGS 1989): " . . . the assumption of dose additivity is most properly applied to compounds that induce the same effect by the same mechanism of action. Consequently, application of the hazard index equation to a number of compounds that are not expected to induce the same type of effects or that do not act by the same mechanism, although appropriate as a screening-level approach [in the risk assessment], could overestimate the potential for effects. . . If the [hazard index] HI is greater than unity as a consequence of summing several hazard quotients of similar value, it would be appropriate to segregate the compounds by effect and by mechanism of action and to derive separate hazard indices for each group."

Therefore, LANL proposes to follow RAGS and will consider effect and mechanism of action if the HI is greater than unity as a consequence of summing hazard quotients for compounds that are not expected to induce the same type of effects.

#4 *(page 7) PCBs are identified as hazardous constituents in 40 CFR 261, Appendix VIII. PCBs are also hazardous constituents discussed in Subpart S guidance, in which a conservative health-based action level is provided as an example. Although TSCA has set "cleanup standard" for PCBs based upon land usage, HRMB is concerned about the "protectiveness of health and the environment" because the example Subpart S health-based action level is more conservative than the most conservative TSCA PCB standard. For screening purposes, HRMB recommends calculating the screening action level as described in Subpart S guidance using the most recent toxicological data for PCBs.*

Response:

For screening purposes, LANL proposes to follow the current guidelines proposed by the LANL's Decision Support Team for the Environmental Restoration Project. These guidelines represent the most current accepted practices for evaluating sites at LANL and have been and are being presented to EPA for their approval as they are developed and revised. Specifically, EPA Region 9 Preliminary Remedial Goals (PRGs) have been adopted by LANL as screening action levels for identifying chemicals of concern (COCs). Region 9 PRGs are calculated from a target cancer risk of 10^{-6} and a hazard index of 1 under a residential exposure scenario and include all direct pathways of exposure (i.e., incidental soil ingestion, dermal exposure, and inhalation of volatiles and particulates). Region 9 PRGs are considered to be more representative of potential site exposure than either Subpart S health-based action levels, which consider only one pathway of exposure (i.e., incidental ingestion of soil), and SALs which include only two pathways of exposure at most (i.e., inhalation of

volatiles and ingestion). Region 6 EPA has been consulted and agrees with this approach. Thus, LANL proposes to use the Region 9 PRGs as screening levels for all chemicals, including PCBs. For PCBs, the numerical value of the Region 9 PRG is equal to the screening level calculated as described in the Subpart 5 guidance.

#11 (page 13) This response does not but should indicate that monitoring will continue until closure is certified.

Response:

The text will be modified to indicate that monitoring will continue until closure is certified.

#13 *(page 17) Are 3 samples enough? See response to comment #2 (page 3) above.*

Response:

As indicated in the LANL response to Comment #2, Enclosure B, a minimum of ten samples will be used to calculate the 95% UCL of the mean.

#18 *(page 19) See comments for item number 2 (page 4) above. There should be a sampling and analysis plan to determine background levels for naturally occurring metals.*

Response:

See response to comment A-3 (Enclosure A).

#30 *(page 32) Are three samples per exposure unit enough to produce a curve to determine the 95% UCL? See response to comment number 2 (page 3) above.*

Response:

LANL agrees to use a minimum of 10 samples to calculate the 95% UCL of the mean. See response to comment number 2 (page 3) above.

#32 *(page 33) Again, are three samples per exposure unit enough to produce a curve to determine the 95% UCL? See response to comment number 2 (page 3) above.*

Response:

LANL agrees to use a minimum of 10 samples to calculate the 95% UCL of the mean. See response to comment number 2 (page 3) above.

#34 *The guidance default value for Exposure Frequency (EF) is 365 days not 274 for residential land use. LANL must perform a baseline risk assessment using the conservative residential risk scenario for comparison purposes. Additionally, future land use is a major consideration. Therefore, LANL should utilize a residential land use scenario, a hazard index of 1 or less, 10^{-6} or less increase in cancer risk over background. Risk assessment calculations based on other assumptions may be presented in addition to the most conservative scenario.*

Furthermore, because of this site's long history, the nature of historic activities, and the lack of complete knowledge of process, it is important to characterize all risks, including that associated with radioactive constituents, to human health and the environment. If there are radioactive constituents present, then by their very nature they are hazardous to a person's health. Because health risk is being evaluated here, it is important to look at the health risk posed by the combination of all contaminants of concern, including radioactive isotopes. Therefore, LANL should include radioactive isotope sampling and radioactive concentration terms in the risk assessment. NMED understands that the radioactive waste, if necessary, will be remediated under a different authority.

Please refer to the EPA Risk Assessment Guidance for Superfund Volume I Human Health Evaluation Manual (Part A) Interim Final, Section 10.7.3 Combining Radionuclide and Chemical Cancer Risks, and other EPA Risk Assessment Guidance documents, for precautions to be taken when combining radiological and chemical risk assessments.

Response:

It is LANL's current policy that the TA-53 surface impoundments will remain under continued industrial land use. The TA-53 surface impoundments are located within an industrial setting at the eastern edge of TA-53. The Los Alamos Meson Physics Facility (LAMPF) is located within the western edge of TA-53, and is currently active in providing basic research involving subatomic particles, isotope production, radiochemistry, research, solid-state physics research, and

accelerator technology development. TA-72 is immediately east of TA-53 and is the location of the active small-arms firing range used by the Laboratory's security guards. TA-20, immediately south of TA-53, was the location of several firing sites where experiments involving the use of explosives were conducted. Eighty-three PRSs, including landfills, firing sites, waste and product storage areas, underground storage tanks, septic systems, outfalls, and surface impoundments are located within this area which comprises OU 1100. These PRSs are being addressed under the Hazardous and Solid Waste Amendment (HSWA) Corrective Action Program. It is LANL's current policy that the lands comprising OU 1100, including the TA-53 surface impoundments, will remain under continued laboratory administrative control into the foreseeable future. The specification of continued industrial use for the TA-53 lagoons is in accordance with EPA guidance (p. 9, RAGS, Volume I, Part B, Development of Risk-based Preliminary Remediation Goals, 1991) which states that "Sites that are surrounded by operating industrial facilities can be assumed to remain industrial areas unless there is an indication that this is not appropriate."

Further, LANL and Region 6 EPA have concurred on cleanup standards based on continued industrial use for ongoing accelerated cleanups at other locations on LANL property. The designation of continued industrial use for the TA-53 lagoons is consistent with this approach.

Therefore, LANL proposes to calculate risk based on a continued industrial land use scenario considering each surface impoundment as an industrial exposure unit. Potential dose from radionuclides and potential cancer risk and health hazard from chemicals on site will be calculated using industrial exposure scenario input parameters, including exposure frequency.

In addition, LANL proposes to tabulate risk related to chemical exposure and dose related to radionuclide exposure separately. The closing paragraph in Section 10.7.3 (p 19-33, RAGS 1989) states that because of differences between the manner in which radionuclide and chemical toxicity values and exposures are calculated that " . . . the two sets of risk estimates should be tabulated separately in the final baseline risk assessment."

2) Response to Technical Area (TA 53, NE & NW Surface Impoundments (Former Operable Unit 1100), dated March 10, 1995 in which LANL responded to the Teleconference Comments of January 24, 1995 with NMED:

LANL's Comment #1 Proposed Text Changes, page 2, third complete paragraph. "Available information indicated that there is also no current risk associated with infiltration to groundwater..." Please remove this entire paragraph or provide enough hydrological data to support this assumption.

Response:

The text will be deleted as requested in the comment.

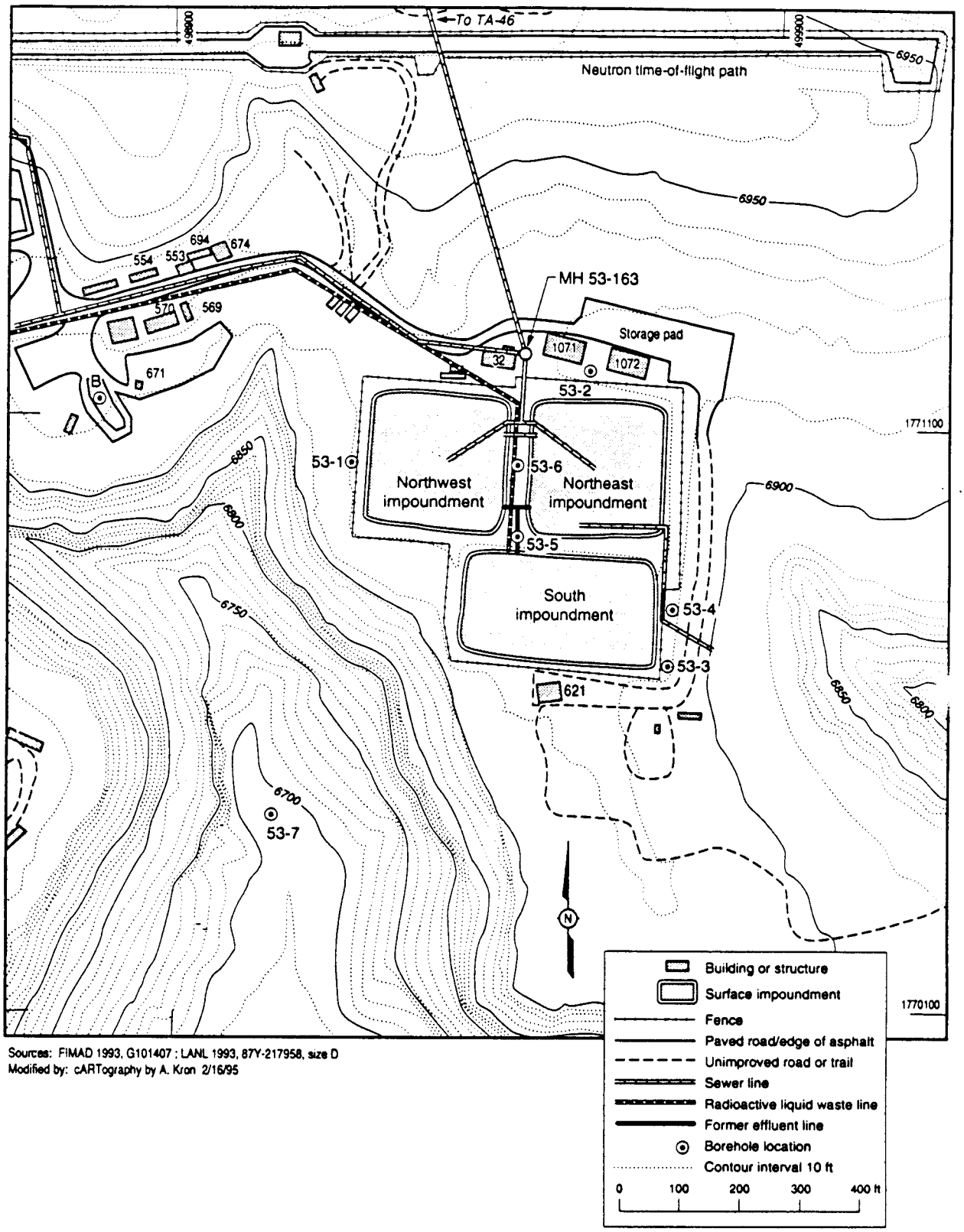
LANL's Comment #2, page 3, general comment. *Notes from the telephone conference reflect that NMED suggested that LANL completely remove the abandoned piping. LANL's response describes removing only the contaminated portion of the piping. It is technically acceptable to leave the uncontaminated portion of piping in place if it has been adequately characterized and shown to pose no risk to human health and the environment. This is the strategy to be utilized for the surface impoundments themselves. The sampling and analysis strategy described for the inactive piping appears acceptable. However, a map should be provided showing the proposed sampling locations.*

Response:

Proposed borehole locations are indicated in the new Figure 2-3, which was included with LANL's previous responses. A copy of this figure is also attached to the current response.

Final Note

LANL has not developed the specific modifications to the text of the closure plan that would be consistent with the above responses. When concurrence is reached with NMED regarding the modifications proposed in this response, new text, tables, and figures (as appropriate) will be developed that combines all agreed-to modifications.



Sources: FIMAD 1993, G101407 ; LANL 1993, 87Y-217958, size D
Modified by: cARTography by A. Kron 2/16/95

Figure 2-3. Current Configurations of the TA-53 Surface Impoundments

Statistical Comparisons to Background

Part I

Los Alamos National Laboratory

Environmental Restoration Project
Assessments Council

March 28, 1995

INTRODUCTION

The main purpose of this policy paper is to provide guidance to Los Alamos National Laboratory (the Laboratory) Environmental Restoration (ER) Project personnel for the selection of appropriate methods for statistical comparisons between potential release site (PRS) data and naturally occurring concentrations of metals (including radionuclides) in soils. Any background comparison approach consists of two components. The first is the assembly of a defensible set of background data. This document describes comparisons with the Laboratory-wide set of background data collected by Longmire (Longmire at al. 1994, 1142). The second component is the selection of the statistical method(s) used to compare site data with background data. Two statistical methods are presented. The first compares the site concentration data with a statistic representing the upper percentile of background concentrations. The second is a group of methods designed to detect a distributional shift between site data and background data. Although guidelines for the application of these methods are presented in this document, each data analysis report should briefly describe the statistical analysis method chosen and justify its application to the data in question. For PRS data, particular attention should be paid to background comparisons of beryllium and arsenic, because background concentrations of these elements exceed risk-based screening action values. These reports should also justify the use of Laboratory-wide background concentration data or present the rationale for using site-specific background concentration data.

Comparisons of PRS data with background concentrations are needed as part of the Resource Conservation and Recovery Act (RCRA) corrective action process. The guidance provided by this document for statistical comparisons between PRS data and naturally occurring concentrations of metals (including radionuclides) in soils may also be applicable to background comparisons of certain anthropogenic compounds (e.g., radionuclides distributed from nuclear fallout or organic chemicals associated with urban activities).

SUMMARY OF REGULATIONS AND GUIDANCE GOVERNING STATISTICAL COMPARISON TO BACKGROUND

The Environmental Protection Agency (EPA) guidance documents supporting the Comprehensive Environmental Response Compensation and Liability Act (CERCLA) and RCRA programs provide specific information on how to design background studies and how to statistically compare site data with background data.

The CERCLA document, *Guidance on Data Useability in Risk Assessment (Part A)* (EPA 1992, 1166), recommends collecting background data prior to collecting site data. If the comparison of background data with site-derived data for a given constituent does not show a difference statistically, that constituent is eliminated from further analyses. The CERCLA guidance also suggests that the number of background samples collected from a site be based on the "minimum detectable difference" procedure (EPA 1989, 0303). Data analysts unfamiliar with this approach should contact the statistical specialists designated by the Assessments Council Chairperson.

Background comparisons for groundwater monitoring data are addressed in the RCRA document, *The RFI Guidance* (EPA 1989, 0088). Methods for comparing data derived from upgradient wells with data from downgradient wells are presented in the RCRA groundwater statistical analysis document (EPA 1989, 1141). These statistical methods are codified in 40 CFR Part 264, *Statistical Methods for Evaluating Ground-Water Monitoring from Hazardous Waste Facilities: Final Rule* Federal Register Tues. Oct. 11, 1988.

Statistical methods used for background comparisons of groundwater can be applied to background comparisons for data from other media as stated in the preface of the RCRA groundwater statistical analysis document (EPA 1989, 1141):

> "This scenario can be applied to other non-RCRA situations involving the same spatial relationships and the same null hypothesis. The explicit null hypothesis for testing contrasts between means, or where appropriate between medians, is that the means between groups (here monitoring wells) are equal (i.e., no release has been detected), or that the group means are below a prespecified action level (e.g., the ground-water protection standard). Statistical methods that can be used to evaluate these conditions are described in Section 5.2 (Analysis of Variance), 5.3 (Tolerance Intervals), and 5.4 (Prediction Intervals)."

The RCRA groundwater monitoring guidance states that the specific approach proposed by the owner/operator should be submitted to EPA for approval, especially where methods other than those presented in the guidance are used. Statistical methods presented below are consistent with those found in the analysis of variance and tolerance interval sections of the RCRA groundwater statistical analysis document (EPA 1989, 1141).

BACKGROUND COMPARISON APPROACHES

Background Comparison Approaches at Other DOE Facilities

Based on information presented at the 1994 Technical Information Exchange Workshop, most DOE facilities have funded a facility-wide background analysis of all potentially impacted media (soil, sediment, surface water, groundwater). Most facilities have information on the soil horizon or other data that support a site-specific comparison with background. No single statistical test for comparing site data with background data is used throughout the DOE facilities studied.

Background comparison approaches at specific facilities include

- Hanford Site (Richland, WA): a background conceptual model has been developed. This model includes metals (including certain radionuclides) and considers transport between soil, sediment, surface water and groundwater.

- Oak Ridge Reservation (Oak Ridge, TN): a statistically significant difference between site data and background data is required before including the constituent in a risk assessment.

- Sandia National Laboratories (Albuquerque, NM): UTLs of metals and radionuclides are calculated based on historical data collected within and outside of PRSs. UTL calculations are made if there is an adequate number of "detects" reported by the analytical laboratory. The UTL is calculated using a method dependent on the distributional properties of the analyte.

Background Comparison Approach at the Laboratory

Use of Laboratory-Wide Background Data

The strategy at the Laboratory has been to collect samples that are representative of Laboratory-wide background variations in metal concentrations in soil and tuff (see Longmire at al. 1994, 1142) for comparison with PRS data (see the Appendix for more detail on the statistical distribution of naturally occurring metals). These Laboratory-wide background data are used as the default background data for making comparisons in the initial RCRA facility investigation (RFI) screening assessment process. At present, 47 soil samples (A, B or C soil horizon) have been analyzed using EPA SW 846 methods, and 50 soil (A, B or C soil horizon) and 38 tuff samples have been analyzed using Instrumental Neutron Activation Analysis (INAA). The INAA data represent total elemental concentrations and are useful for background comparisons with data

from PRSs for which a comparable method of analysis was used. Additional data from soil, sediment and volcanic tuff background analyses will be added to the Laboratory-wide background data base during 1995 and will be made available to ER Project personnel and other interested parties as soon as validated data are received.

Use of Site-Specific Background Data

Variation in certain elements makes comparisons with Laboratory-wide background concentrations less valid than comparisons with site-specific background concentrations. For example, due to natural variability in element concentrations of Bandelier Tuff, samples of background soils collected from technical areas (TAs) at the east and west ends of the Laboratory are likely to show enrichment or depletion in certain elements when compared with Laboratory-wide background values. This variation exists because Bandelier Tuff was derived from a zoned magma chamber in which some elements were concentrated at the top and others were concentrated at the bottom. During the eruption resulting in deposition of the Bandelier Tuff, magmas at the top of the chamber were erupted first. Consequently, those elements that were concentrated in the upper magmas are in higher abundance at TAs to the east of the Laboratory because the Bandelier Tuff subunit exposed in that region represents the earlier-erupted magmas. Thus, soils derived from tuff located to the east of the Laboratory are likely to have a higher abundance of certain elements than Laboratory-wide background soils. For example, at TA-33 uranium in background soils is at higher levels than Laboratory-wide background because it is more abundant in the stratigraphically lowest units of the Tschirege Member of the Bandelier Tuff. Thus, it is important to develop site-specific background for some sites.

Guidelines for Comparisons

All users of background data must follow guidelines to ensure that site data are being compared with Laboratory-wide background data in a scientifically valid manner. A primary requirement is that PRS data and background data be collected from soil or tuff having the same physical properties: if PRS data are collected from sediments, background data should also come from sediments. If site data were collected exclusively from a soil horizon naturally enriched in a metal, these data statistically would appear to represent a release from a PRS, but in fact are consistent with the range of background concentration for that metal. A second requirement is that PRS samples be measured using the same analytical methods as used for background samples.

Each data analysis report should demonstrate that the above guidelines were considered before using Laboratory-wide background distributions. Failure to meet these guidelines may be an

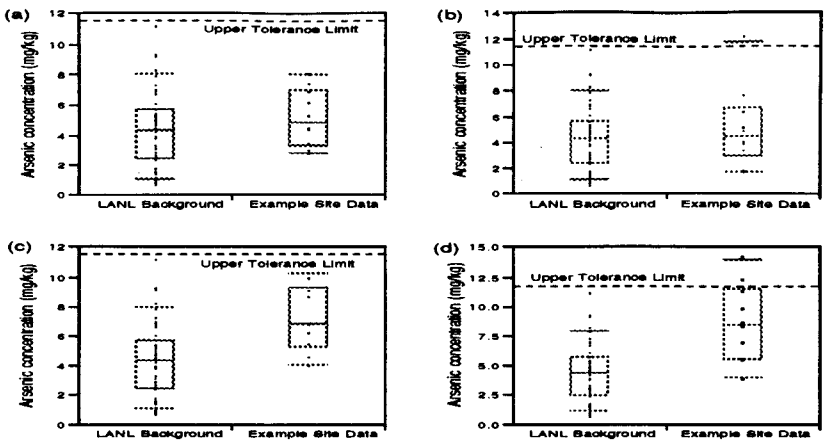
indication that site-specific background data should be collected or that a subset of Laboratory-wide background data should be compared with PRS data. Lack of data for a particular analyte could be justification for the collection of site-specific background data. Before recommending the collection of site-specific background data for an analyte, data analysts should review historical information to confirm whether the analyte in question could be present Conversely, lack of defensible historical information might justify the collection of site-specific background data.

If the collection of site-specific background data seems warranted, data analysts should seek technical advice from the background specialists designated by the Laboratory ER Project Assessments and Earth Sciences Council Chairpersons. These specialists will provide detailed guidance for the location and number of samples required for site-specific background comparisons.

RECOMMENDED STATISTICAL METHODS FOR BACKGROUND COMPARISONS

Because background comparisons are used to make decisions throughout the RCRA process, from site screening to corrective measures implementation, data analysts must have statistical methods that can be used for a broad range of decisions. This guidance defines two statistical methods for background comparisons. Both methods meet the requirements for RCRA decision making. In the first method, the hot measurement test, site concentration data are compared with a statistic representing the upper percentile of background concentrations. In the second method, the distributional shift test, the mean of site data is compared with the mean of background data to determine whether the former is statistically greater than the latter. Used together, or separately, these tests help demonstrate whether a release occurred at a PRS and help define what risk consequence the release may have. Figure 1 illustrates the differences between site data and background data detected by the two methods.

The decision to be supported by the background comparison determines which test is appropriate. When, in an initial RFI screening, a single high value requires further analysis, the hot measurement test is typically recommended. When extensive data are collected to support a risk assessment and a change in the average concentration should lead to further action at the site, the distributional shift test may be more appropriate. The rationale for selecting a statistical method that differs from those presented in this guidance should be clearly indicated in a data analysis report.



(a) Site data are within the range of background: no distributional shift or hot measurements (i.e., values greater than the UTL).
(b) Site data fail hot measurement test: one of eleven arsenic concentrations exceeds the UTL of 11.6 mg/kg.
(c) Site data show a distributional shift: the Wilcoxon rank sum test shows that site data tend to be greater than the background data.
(d) Site data show a distributional shift and fail the hot measurement test: two of ten arsenic concentrations exceed the UTL of 11.6 mg/kg and the site data tend to be greater than the background data.

Figure 1. BOX PLOT COMPARISONS OF EXAMPLE SITE DATA WITH LABORATORY BACKGROUND DATA.

Because the selection of a particular statistical method depends on the statistical distributions of site and background data, analysts are encouraged to prepare graphical data displays to facilitate the communication of the results of data comparisons. Box plots (see Figure 1), in which background data and site data can be compared side-by-side, are most useful. Analysts should also consider using histograms and probability plots. These graphs provide tangible evidence of the similarities or differences between site data and background data.

The level of effort spent to evaluate potential differences between PRS data and background data should be related to the site-specific information available. For example, if historical information indicates that beryllium was released at a firing site, the potential differences between the beryllium concentration data from the firing site and Laboratory-wide or site-specific background data should be carefully evaluated and presented in the data analysis report.

Hot Measurement Test

The hot measurement test defines a threshold value that represents high background concentrations. No matter what parameters are chosen to define the threshold, there exists a probability that a background measurement will exceed the hot measurement threshold. The frequency of false positive results is minimized by using a threshold statistically related to higher background concentrations. The confidence limit on a percentile of the distribution, termed the tolerance limit, is such a value and is one of the background comparison methods recommended by EPA (1989, 1141). The Laboratory has selected the 99th percentile for calculating the UTL, based on the general guidance in the RCRA groundwater document. If the underlying distributional model is correct, the UTL based on the 99th percentile is rarely exceeded. EPA recommends calculating an upper 95% confidence limit for the target percentile (EPA 1989, 1141). The UTL for the 99th percentile at 95% confidence can be calculated using Equation (1).

$$UTL_{0.99,0.95} = \text{mean} + \text{standard deviation} * k_{0.99,0.95} \qquad (1)$$

The k-factor depends on the number of background samples; complete tables of k-factors are published in the RCRA groundwater statistical analysis document (EPA 1989, 1141) and Gilbert (1987, 0312). Table 1 presents k-factors selected to represent the range of values used to compute UTLs for Laboratory background soil samples. To apply Equation (1), the background data must be normally distributed or transformed to normality (e.g., by using log-transformation). If data deviate sufficiently from normality, nonparametric methods for calculating tolerance limits should be considered (e.g., as described in Gilbert [1987, 0312]). Alternatively, when appropriate, the data analyst may trim outliers from the distribution and calculate the UTL based on the trimmed mean and standard deviation.

TABLE 1
SELECTED K-FACTORS USED TO CALCULATE UTLS*

Number of Background Samples	$k_{0.99,0.95}$
45	2.897 *
46	2.8902
47	2.8834
48	2.8766
49	2.8698
50	2.863 *

*reprinted from Gilbert (1987, 0312)

The Laboratory-wide soil background data were used to calculate the UTL values presented in Table 2. As discussed earlier in this document (see Use of Laboratory-Wide Background Data, page 3), use of Laboratory-wide background data as the default is recommended, and, the use of these data should be justified for the specific background comparison being performed. Table 2 UTL values should not be used without considering the guidelines for comparisons discussed earlier in this document (see Guidelines for Comparisons, pages 4-5). A minimalist approach to data preparation was used for the initial UTL calculations. Because some soil concentration data were significantly skewed, a log-transformation was applied to improve the fit of these data to a normal distribution (the details of these data distributions are included in the Appendix of this policy paper). The UTLs calculated for the lognormal distribution were backtransformed into original units to simplify comparisons with site data. No values were trimmed from the distributions used to calculate the mean and standard deviation. If four or fewer nondetects were reported for an analyte, values below the laboratory detection limit were replaced by one-half of the detection limit (EPA 1989, 1141). The UTL was not calculated for any analyte having more than four nondetect values (>10% nondetects).

Using the observed maximum concentration in the background data is an alternative to using the UTL as the hot measurement threshold. However, when few background samples are available, using the maximum concentration will result in an underestimation of the upper background percentile. In general, the sample maximum concentration (for "n" samples) is an estimate of the

th percentile. Thus, if 10 samples are collected, the sample maximum concentration is an estimate of the 95% percentile. As more Laboratory-wide background data become available, the maximum value will increase, but the UTL typically will not change. Because the maximum is extremely sensitive to background sample size, it is not recommended for use as a hot measurement threshold. Rarely detected analytes, which include antimony, cadmium, mercury, selenium and thallium, are an exception to this general recommendation. For this limited subset of rarely detected analytes, the maximum detected background concentration can be used as the hot measurement threshold.

Exceeding the UTL does not prove that a release occurred at a PRS. Assuming the PRS is at background and the statistical model is correct, there is a 1% probability that the 99th percentile will be exceeded by each sample collected from the PRS. Furthermore, a typical metals suite requires comparison of 23 analytes with background. If the concentrations of the 23 metal analytes vary independently, the 1% probability that each PRS sample exceeds the 99th percentile increases to a 21% probability that at least one of the 23 nintyninth percentiles will be

exceeded in a single sample. Additionally, given that the probability values for these multiple comparisons have not been adjusted, the overall level of confidence for 23 analytes will be substantially less than 95%. Consequently, the results of a hot measurement test must be carefully evaluated. The possibility of exceeding a UTL due to an unusual, but naturally occurring, soil matrix is a further consideration.

TABLE 2
UTLS FOR LABORATORY SOIL BACKGROUND DATA

Analyte	SAL (mg/kg)	Mean (mg/kg)	Standard Deviation (mg/kg)	UTL 99%,0.95 (mg/kg)	N	N > DL*
Aluminum (LT)		19000	13800	123000	47	47
Antimony	32	2.45	0.36	2.5 (MAX)	46	2
Arsenic		4.4	2.5	11.6	46	46
Barium (LT)	5,600	161	129	1140	47	47
Beryllium		1.15	0.75	3.31	47	47
Cadmium	80	0.39	0.54	2.7 (MAX)	47	5
Calcium (LT)		5790	12500	54400	47	47
Chromium (Total)		11.7	7.8	34.2	47	47
Cobalt (LT)		15.2	7.6	51.1	47	47
Copper	3,000	5.3	3.6	15.7	47	45
Iron		14500	7320	35600	47	47
Lead	400	15.0	8.3	39.0	47	44
Magnesium (LT)		2920	2150	16100	47	47
Manganese	11,000	343	238	1030	47	47
Mercury	24	0.05	0.01	0.1 (MAX)	48	4
Molybdenum	400	NA	NA	NA	NA	NA
Nickel	1,600	9.7	5.9	26.7	47	45
Potassium		2420	1304	6180	47	47
Selenium	400	0.43	0.41	1.7 (MAX)	46	23
Silver	400	NA	NA	NA	NA	NA
Sodium (LT)		577	453	3320	47	47
Strontium	48,000	NA	NA	NA	NA	NA
Thallium	6.4	0.27	0.24	0.9 (MAX)	45	21
Vanadium	560	25	14	66	47	47
Zinc	24,000	41	21	101	47	47
^{40}Potassium (1)		21.6	5.07	36.1	50	50
^{232}Thorium (1)	5	1.71	0.34	2.68	50	50
^{234}Uranium (1)	86	1.21	0.29	2.03	50	50
^{235}Uranium (1)	18	0.052	0.012	0.088	50	50
^{238}Uranium (1)	59	1.14	0.27	1.90	50	50

* Concentration values < detection limit (DL) were replace by 1/2 of the DL.
† SAL for Chromium III is 80,000 mg/kg; for Chromium VI, 400 mg/kg.
LT UTL is based on log-transformed data.
NA Data is not available for Laboratory background.
MAX Maximum value is reported, rather than the UTL.
(1) Data are converted from elemental concentrations reported in the Laboratory background report. Units are in pCi/g.

The results of the UTL comparison should also be evaluated based on potential human health risk or ecological risk screening levels. Some analytes, arsenic and beryllium in particular, represent a special case. Because background levels at the Laboratory for beryllium or arsenic exceed risk-based screening levels, no screening action levels (SALs) have been calculated for these two analytes for the Laboratory ER Project. Seven metals (antimony, barium, chromium, lead, manganese, thallium, and vanadium) and one radionuclide (thorium-232) have background UTLs that appear close to their SAL values. The UTLs of this group represent a significant fraction of the SAL (8 to 50%). If, in a comparison similar to a multiple constituent test, each of these metals had a concentration equal to the SAL, the total of the metal UTLs divided by the corresponding SAL (in effect normalizing the UTL-to-SAL ratio) would equal 89%. Thus, the concentrations of most naturally occurring metals are significantly lower than their respective SALs.

Both the multiple constituent evaluation and the UTL-to-SAL comparison will help determine what level of effort should be expended to evaluate deviations from background. For most naturally occurring metals, when only a single statistical comparison with background is performed, the UTL will be adequate because probability levels are not compromised. Under this circumstance, the UTL is the simplest comparison and is functionally most similar to comparisons of site data to target risk levels or SALs.

Distributional Shift Test

The distributional shift test is used to determine whether site data are systematically greater than background data. Several types of distributional shift tests are available.

- The Student t-test is a parametric, statistical, two-sample test that determines whether the mean concentration of site data is statistically greater than the mean concentration of background data.

- The Wilcoxon rank-sum test is the nonparametric equivalent to the t-test (Gilbert 1987, 0312; Gilbert and Simpson 1992, 0974). The Wilcoxon test pools site and background data into one aggregate set and determines whether the average rank of site data is greater than that of the background data. The Wilcoxon test is recommended when site data consist of few samples or when nondetects are frequent.

- The Quantile test (Gilbert and Simpson 1992, 0974), which compares the upper quantile of background data with that of PRS data, is more capable of detecting a difference when only a small number of PRS concentrations are elevated. The Quantile test is the most useful distributional shift test for PRSs at which samples from a release represent a small fraction of the overall data collected at the PRS because it does not artificially reduce statistical significance. For example, to detect contamination from historical spills at unknown locations, an RFI work plan may call for samples to be collected from a grid. Most sample results show no contamination, but those in or near spill locations show elevated concentrations.

Use of the distributional shift test is dependent on the number of samples available for comparison. In general, at least 10 sample concentrations for comparison with background data are needed to provide adequate confidence for detecting a shift. Frequently, during Phase I of an RFI, inadequate numbers of samples are collected to warrant a distributional shift comparison.

To infer a significant result in a distributional shift test, a 95% confidence level is recommended. Given that multiple comparisons will be performed with the distributional shift test, the same statistical interpretation issues cited above for the hot measurement test are also present. In addition, the human health and ecological consequences of PRS concentration data above background must be considered along with differences in metal concentrations between soil horizons. In particular, multiple comparison tests with SALs must be performed and ecological SAL comparisons must be made.

REFERENCES

EPA (US Environmental Protection Agency), 1989. "Ecological Assessment of Hazardous Waste Sites: A Field and Laboratory Reference," EPA/600 3-89/013, Environmental Research Laboratory, Corvallis, OR. (EPA 1989, 0303)

EPA (US Environmental Protection Agency), 1989. "Interim Final RCRA Facility Investigation (RFI) Guidance, Volume I of IV, Development of an RFI Work Plan and General Considerations for RCRA Facility Investigations," EPA/530-SW-89-031, OSWER General Directive 9502.00-6D, Office of Solid Waste, Washington, DC. (EPA 1989, 0088)

EPA (US Environmental Protection Agency), April 1989. "Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities. Interim Final Guidance," Office of Solid Waste, Waste Management Division, US Environmental Protection Agency, Washington DC. (EPA 1989, 1141)

EPA (US Environmental Protection Agency), April 1992. "Guidance for Data Useability in Risk Assessment (Part A)," Office of Emergency Remedial Response, US Environmental Protection Agency, Washington DC. (EPA 1992, 1166)

Gilbert, R.O., 1987. Statistical Methods for Environmental Pollution Monitoring. Von Nostrand Reinhold Company Inc. New York. 320 pp. (Gilbert 1987, 0312)

Gilbert, R. O., and J. C. Simpson, 1992. "Statistical Methods for Evaluating the Attainment of Cleanup Standards, Volume 3: Reference-Based Standards for Soils and Solid Media," Pacific Northwest Laboratory, Richland, Washington 99352. (Gilbert and Simpson 1992, 0974)

Longmire, P., S. Reneau, P. Watt, L. McFadden, J. Gardner, C. Duffy, and R. Ryti, 1995. "Natural Background Geochemistry, Geomorphology, and Pedogenesis of Selected Soil Profiles and Bandelier Tuff, Los Alamos, New Mexico," Los Alamos National Laboratory Report LA-12913-MS, Los Alamos, New Mexico. (Longmire et al. 1995, 1142)