8

Quality Assurance

Lucinda M. Clark
Donald H. MacQueen

Introduction

Quality assurance (QA) is a system of activities and processes put in place to ensure that products or services meet or exceed customer specifications. Quality control (QC) consists of activities used to verify that deliverables are of acceptable quality and meet criteria established in the quality planning process. Lawrence Livermore National Laboratory conducted environmental monitoring activities during 2004 in accordance with the Environmental Protection Department Quality Assurance Management Plan (Revision 4), which is based on DOE Order 414.1A. This order sets forth policy, requirements, and responsibilities for the establishment and maintenance of plans and actions that assure quality in DOE programs using a risk-based, graded approach to QA. This process promotes the selective application of QA and management controls based on the risk associated with each activity in order to maximize effectiveness and efficiency in resource use.

LLNL and commercial laboratories analyze environmental monitoring samples using U.S. Environmental Protection Agency (EPA) standard methods when available. When EPA standard methods are not available, custom analytical procedures, usually developed at LLNL, are used. LLNL uses only State of California-certified laboratories to analyze its environmental monitoring samples. In addition, LLNL requires all analytical laboratories to maintain adequate QA programs and documentation of methods. The radiochemical methods used by LLNL laboratories are described in procedures created and maintained by the laboratory performing the analyses.

Quality Assurance Activities

Nonconformance reporting and tracking is a process used for ensuring that problems are identified, resolved, and prevented from recurring. EPD reports and tracks problems using Nonconformance Reports (NCRs).

The LLNL Environmental Protection Department (EPD) generated 17 NCRs related to environmental monitoring in 2004. Four of the NCRs were related to problems with analytical laboratories, five documented minor equipment malfunctions that did not result in lost samples, and the remaining eight documented errors made by sampling technologists.

LLNL addresses internal documentation, training, and procedural errors by conducting formal and informal training. These errors generally do not result in lost samples, but may require extra work on the part of sampling and data management personnel to resolve or compensate for the errors.

LLNL addresses analytical laboratory problems with the appropriate laboratory as they arise. Many of the documented problems related to analytical laboratories concerned minor documentation or paperwork errors, which were corrected soon after they were identified. Other problems—such as missed holding times, late analytical results, and typographical errors on data reports—accounted for the remaining analytical laboratory issues. These problems were corrected by reanalysis, resampling, reissued reports, or corrected paperwork, and associated sample results were not affected.

QA staff also track and report planned environmental monitoring samples that are not collected. A summary of these lost samples appears in **Table 8-1**.

Analytical Laboratories

LLNL awarded new Blanket Service Agreements (BSAs) to six analytical laboratories in 2004. LLNL works closely with these analytical laboratories to minimize the occurrence of problems.

Analytical Laboratory Intercomparison Studies

LLNL uses the results of intercomparison program data to identify and monitor trends in performance and to draw attention to the need to improve laboratory performances. If a laboratory performs unacceptably for a particular test in two consecutive performance evaluation studies, LLNL may choose to select another laboratory to perform the affected analyses until the original laboratory can demonstrate that the problem has been corrected. If an off-site laboratory continues to perform unacceptably or fails to prepare and implement acceptable corrective action responses, the LLNL Procurement Department will formally notify the laboratory of its unsatisfactory performance. If the problem persists, the off-site laboratory's BSA could be terminated. If an on-site laboratory continues to perform unacceptably, use of that laboratory could be suspended until the problem is corrected.

Two laboratories at Lawrence Livermore National Laboratory participated in the Mixed Analyte Performance Evaluation Program (MAPEP) sponsored by the U.S. Department of Energy (DOE) during 2004. (The Environmental Monitoring Laboratory intercomparison studies program for which data were reported in previous versions of this report was cancelled after the 2003 studies.) The two LLNL laboratories that participated in MAPEP are the Environmental Monitoring Radiological Laboratory (EMRL) and the Hazards Control Department's Analytical Laboratory (HCAL).

Table 8-1. Sampling completeness in 2004 for the Livermore site and Site 300

Environmental medium	Number of analyses planned	Number of analyses completed	Completeness (%)	Reason(s) for lost samples
Air particulate				
Radiological parameters (Livermore site)	1188	1152	97	No power at location (29), GFI tripped (4), motor problems (1), no access (1), not explained (1)
Beryllium (Livermore site)	96	96	100	
Radiological parameters (Site 300)	717	704	98	No access (11), no power (1), filter saturated with water (1)
Beryllium (Site 300)	46	46	100	
Air tritium				
Livermore site	489	480	98	Pump failure (5), insufficient flow (4)
Site 300	31	31	100	
Soil and Sediment				
Livermore site	42	42	100	
Site 300	30	30	100	
Arroyo sediment (Livermore site only)	21	21	100	
Vegetation and Foodstuffs				
Livermore site and vicinity	64	64	100	
Site 300	20	20	100	
Wine	12	12	100	
Thermoluminescent dosimeters (TLDs)				
Livermore site perimeter	56	53	95	Fence removed (3)
Livermore Valley	88	87	99	TLD found burned (1)
Site 300	52	49	94	Missing (2), lost in controlled burn (1)
Rain				
Livermore site	53	52	99	Bucket missing (1)
Site 300	10	8	80	No access (2)
Storm water runoff				
Livermore site	320	320	100	
Site 300	231	164	71	No flow at location (66), sample not analyzed by lab (1)

Table 8-1. Sampling completeness in 2004 for the Livermore site and Site 300 (continued)

Environmental medium	Number of analyses planned	Number of analyses completed	Completeness (%)	Reason(s) for lost samples
Drainage Retention Basin				
Field measurements	208	208	100	
Samples	72	72	100	
Releases	89	88	99	Samples not collected, no explanation (1)
Livermore site wastewater				
B196	926	926	100	
C196	314	312	99	Sampler error (2)
LWRP[a] effluent	48	47	98	Sample lost at laboratory (1)
Digester sludge	80	74	93	Digester #2 was closed October–December (6)
WDR 96-248				
Surface impoundment wastewater	58	58	100	
Surface impoundment groundwater	145	144	99	Not sampled (1)
Sewage ponds wastewater	34	34	100	
Sewage ponds groundwater	84	84	100	
Miscellaneous aqueous samples				
Other surface water (Livermore Valley only)	51	51	100	
Cooling towers (Site 300 only)	24	24	100	

a LWRP = Livermore Water Reclamation Plant

The results of EMRL's participation in the studies are presented in Table 8-2. According to the results, 33 of 38 reported results were determined to be acceptable, 2 results were acceptable with warning, and 3 results were unacceptable, based on established control limits.

Unacceptable results for plutonium-238 and plutonium-239/240 in the 04-RDF12 air filter study were the result of switching the active filter with a blank filter that was shipped with it. In the future, the active filter will be labeled upon receipt and the two blank filters that accompany the active filter will be dissolved along with the active filter to eliminate loss of activity. The three dissolved filters will be analyzed as one sample.

Table 8-2. EMRL performance in the MAPEP Intercomparison Program Studies for 2004

Study	Analyte	Result	Ref Value	Flag[a]	Acceptance Range[b]	Uncertainty Value
Air filter (Bq/sample)						
MAPEP-04-GrF12	Gross alpha	0.0520	0.37	A	>0.0-0.8	0.00289
MAPEP-04-GrF12	Gross beta	1.28	1.21	A	0.6-1.8	0.00983
MAPEP-04-RdF12	Cesium-134	2.95	2.9	A	2.03-3.77	0.225
MAPEP-04-RdF12	Cesium-137	2.42	1.96	W	1.40-2.60	0.328
MAPEP-04-RdF12	Cobalt-57	2.49	2.44	A	1.68-3.12	0.231
MAPEP-04-RdF12	Cobalt-60	2.35	2.35	A	1.61-2.99	0.191
MAPEP-04-RdF12	Manganese-54	2.42	3.03	W	2.10-3.90	0.441
MAPEP-04-RdF12	Plutonium-238	0.0172	0.13	N	0.09-0.17	0.00129
MAPEP-04-RdF12	Plutonium-239/240	0.012	0.09	N	0.06-0.12	0.00101
MAPEP-04-RdF12	Zinc-65	4.64	4.11	A	2.80-5.20	0.608
Aqueous (Bq/L)						
MAPEP-03-W11	Americium-241	0.0188	0.0144	A	_[c]	0.00157
MAPEP-03-W11	Cesium-137	127	124	A	86.80-161.20	13.4
MAPEP-03-W11	Cobalt-57	187	173	A	121.10-224.90	18.4
MAPEP-03-W11	Cobalt-60	130	121.8	A	85.26-158.34	7.83
MAPEP-03-W11	Manganese-54	162	155	A	108.50-201.50	19.8
MAPEP-03-W11	Plutonium-238	1.30	1.49	A	1.04-1.94	0.0722
MAPEP-03-W11	Plutonium-239/240	2.04	2.39	A	1.67-3.11	0.11
MAPEP-03-W11	Zinc-65	384	320	A	224.00-416.00	38.7
MAPEP-04-GrW12	Gross alpha	0.542	1.24	A	0.0-2.5	0.0172
MAPEP-04-GrW12	Gross beta	3.51	4.07	A	2.0-6.2	0.0744
MAPEP-04-MaW12	Americium-241	0.558	0.59	A	0.42-0.78	0.0241
MAPEP-04-MaW12	Cesium-134	189	208	A	145.60-270.40	15.6
MAPEP-04-MaW12	Cesium-137	250	250	A	175.00-325.00	37.9
MAPEP-04-MaW12	Cobalt-57	189	185	A	129.50-240.50	15.8
MAPEP-04-MaW12	Cobalt-60	165	163	A	114.10-211.90	11
MAPEP-04-MaW12	Manganese-54	250	267	A	186.90-347.10	24.8
MAPEP-04-MaW12	Plutonium-238	1.10	1.24	A	0.84-1.56	0.0608
MAPEP-04-MaW12	Zinc-65	217	208	A	145.60-270.40	29.6
Soil (Bq/kg)						
MAPEP-04-MaS12	Americium-241	73.1	67	A	46.88-87.06	3.62
MAPEP-04-MaS12	Cesium-134	350	414	A	290.08-538.72	18.4
MAPEP-04-MaS12	Cesium-137	830	836	A	585.34-1087.06	78.5
MAPEP-04-MaS12	Cobalt-57	403	400	A	279.72-519.48	29.4
MAPEP-04-MaS12	Cobalt-60	524	518	A	362.60-673.40	27.6
MAPEP-04-MaS12	Manganese-54	830	485	N	339.29-630.11	53.3

Table 8-2. EMRL performance in the MAPEP Intercomparison Program Studies for 2004 (continued)

Study	Analyte	Result	Ref Value	Flag[a]	Acceptance Range[b]	Uncertainty Value
MAPEP-04-MaS12	Plutonium-238	35.1	35.4	A	24.78-46.02	2.53
MAPEP-04-MaS12	Plutonium-239/240	47.4	41.8	A	29.27-54.35	3.19
MAPEP-04-MaS12	Potassium-40	627	604	A	422.80-785.20	82.4
MAPEP-04-MaS12	Zinc-65	769	699	A	489.51-909.09	76.7

a Acceptable (A flag) results have bias ≤20%. Results acceptable with warning (W flag) have bias >20% and bias ≤30%. Results with bias >30% (N flag) are not acceptable.

b Significant figures shown are those of the MAPEP program.

c Acceptance range not provided for this analysis.

The manganese-54 result for the 04-MaS12 soil study was unacceptable due to a transcription error. The actual result, 507 Bq/kg is in the acceptable range. A new report format was developed to prevent similar errors from occurring in the future.

The results of HCAL's participation in the 2004 MAPEP studies (see **Table 8-3**) indicate that five of five sample results fell within the 3σ acceptance control limits.

Table 8-3. HCAL performance in the MAPEP Intercomparison Program Studies for 2004

Study	Analyte	Result	Ref Value	Flag[a]	Acceptance Range	Uncertainty Value
Air filter (Bq/sample)						
MAPEP-04-GrF12	Gross alpha	0.19	0.37	A	>0.0-0.8	0.04
MAPEP-04-GrF12	Gross beta	1.40	1.21	A	0.6-1.8	0.09
Aqueous (Bq/L)						
MAPEP-04-GrW12	Gross alpha	1.09	1.24	A	0.0-2.5	0.21
MAPEP-04-GrW12	Gross beta	3.59	4.07	A	2.0-6.2	0.35
MAPEP-04-MaW12	Hydrogen-3	82.5	82.9	A	58.1-108	5.5

a Acceptable (A flag) results have bias ≤20%. Results acceptable with warning (W flag) have bias >20% and bias ≤30%. Results with bias >30% (N flag) are not acceptable.

HCAL also participated in three Environmental Resource Associates (ERA) performance evaluation studies in 2004. The results of these studies are presented in **Table 8-4**. Fourteen of 15 analytes reported by HCAL in these studies fell within acceptable limits. HCAL was unable to determine the cause of the unacceptable result for iron. No problems were identified in a review of the raw data, and the results for a duplicate sample as well as a follow-up sample were both in the acceptable range.

Table 8-4. HCAL performance in the ERA Intercomparison Program Studies for 2004

Study	Analyte	Reported Value	ERA Assigned Value	Control Limits	Warning Limits	Performance Evaluation
Radiological (pCi/L)						
RAD-59	Gross alpha	34.4	31.7	18.0-45.4	22.5-40.9	Acceptable
RAD-59	Gross beta	38.1	36.3	26.7-45.0	30.5-42.1	Acceptable
RAD-59	Tritium	20300	20700	17100-24300	18300-23100	Acceptable
Nonradiological (µg/L)						
WP-116	Aluminum	3090	3100	2670-3500	2810–3360	Acceptable
WP-116	Arsenic	299	299	248-352	266–335	Acceptable
WP-116	Beryllium	350	353	300-399	316–382	Acceptable
WP-116	Cadmium	709	729	622-827	657–793	Acceptable
WP-116	Chromium	336	322	279–365	294–351	Acceptable
WP-116	Copper	135	139	123-155	129–150	Acceptable
WP-116	Iron	253	216	187–250	197–239	Not Acceptable
WP-116	Lead	130	124	102-146	110–138	Acceptable
WP-116	Nickel	946	922	834–1030	866–997	Acceptable
WP-116	Silver	81.2	83.4	71.0–91.6	75.1–91.6	Acceptable
WP-116	Zinc	583	559	494–630	516–607	Acceptable
WP-118	Iron	104	102	85.7-122	91.7-116	Acceptable

Although contract laboratories are also required to participate in laboratory intercomparison programs, permission to publish their results for comparison purposes was not granted for 2004. See the following website to obtain MAPEP reports that include the results from all participating laboratories:
http://www.inel.gov/resl/mapep/reports.html

Duplicate Analyses

Duplicate or collocated samples are distinct samples of the same matrix collected as closely to the same point in space and time as possible. Collocated samples processed and analyzed by the same laboratory provide intralaboratory information about the precision of the entire measurement system, including sample acquisition, homogeneity, handling, shipping, storage, preparation, and analysis. Collocated samples processed and analyzed by different laboratories provide interlaboratory information about the precision of the entire measurement system (U.S. EPA 1987). Collocated samples may also be used to identify errors such as mislabeled samples or data entry errors.

Tables 8-5, 8-6, and 8-7 present statistical data for collocated sample pairs, grouped by sample matrix and analyte. Samples from both the Livermore site and Site 300 are included. Tables 8-5 and 8-6 are based on data pairs in which both values are detections (see "Data Presentation"). Table 8-7 is based on data pairs in which either or both values are nondetections.

Table 8-5. Quality assurance collocated sampling: Summary statistics for analytes with more than eight pairs in which both results were above the detection limit

Media	Analyte	N[a]	%RSD[b]	Slope	r^{2}[c]	Intercept
Air	Gross alpha (variability)[d]	76	56.1	0.436	0.22	2.06×10^{-5} (Bq/m^3)
	Gross beta	102	20.1	0.964	0.87	6.24×10^{-6} (Bq/m^3)
	Beryllium (outliers)[e]	10	12.9	8.28	0.68	–41.2 (pg/m^3)
	Uranium-235 + 236	12	11.8	0.767	0.96	3.08×10^{-8} (μg/m^3)
	Uranium-238	12	13.1	0.748	0.94	5.16×10^{-6} (μg/m^3)
	Uranium-235/238 (outliers)[e]	12	1.08	1.01	0.54	–0.000201 (ratio)
	Tritium	20	14	1.07	0.99	–0.0235 (Bq/m^3)
Dose (TLD)	90-day radiological dose (outliers)[e]	30	3.53	0.597	0.39	5.88 (mrem)
Groundwater	Gross beta	16	16	0.98	0.88	–0.0233(Bq/L)
	Arsenic	16	4.64	0.995	1	0.000268 (mg/L)
	Barium	9	1.89	1	1	–0.00132 (mg/L)
	Bromide (outliers)[e]	10	13.6	0.596	0.69	0.22 (mg/L)
	Chloride	11	0	1.01	1	–0.864 (mg/L)
	Nitrate (as NO_3)	19	1.18	1.01	1	–0.246 (mg/L)
	Ortho-Phosphate	10	2.86	1	1	–0.000588 (mg/L)
	Potassium	18	4.52	1.02	1	–0.116 (mg/L)
	Sulfate	11	0.358	1	1	–0.636 (mg/L)
	Uranium-234+233	11	2.4	1.01	1	0.000177 (Bq/L)
	Uranium-238	11	5.04	1.01	1	–0.000586 (Bq/L)
Sewer	Gross beta	51	14.1	0.999	0.85	4.23×10^{-5} (Bq/mL)
	TDS	9	17.2	1.06	0.97	51.4 (mg/L)
	TSS	9	14.1	0.772	0.9	53.2 (mg/L)

a Number of collocated pairs included in regression analysis

b 75th percentile of percent relative standard deviations (%RSD) where $\%RSD = \left(\frac{200}{\sqrt{2}}\right)\frac{|x_1 - x_2|}{x_1 + x_2}$ and x_1 and x_2 are the reported concentrations of each routine-duplicate pair

c Coefficient of determination

d Outside acceptable range of slope or r^2 because of variability

e Outside acceptable range of slope or r^2 because of outliers

Table 8-6. Quality assurance collocated sampling: Summary statistics for selected analytes with eight or fewer pairs in which both results were above the detection limit

Media	Analyte	N[a]	Mean ratio	Minimum ratio	Maximum ratio
Aqueous	Gross beta	1	1.2	1.2	1.2
Groundwater	Gross alpha	6	0.89	0.2	1.4
	Tritium	6	0.86	0.13	1.2
	Radium-226	8	1.3	0.34	4.3
	Uranium-235 and uranium-236	8	0.84	0.5	1.3
Runoff (from rain)	Gross alpha	5	0.99	0.69	1.6
	Gross beta	5	1	0.72	1.5
	Uranium-234 and uranium-233	2	0.96	0.92	0.99
	Uranium-235 and uranium-236	1	1.5	1.5	1.5
	Uranium-238	2	1.1	1	1.2
Soil	Gross alpha	1	0.83	0.83	0.83
	Gross beta	1	0.95	0.95	0.95
	Cesium-137	4	1.1	0.95	1.3
	Potassium-40	4	1	0.95	1.1
	Plutonium-238	3	1.6	0.83	2.2
	Plutonium-239+240	3	1.3	1.1	1.4
	Radium-226	4	0.99	0.93	1.1
	Radium-228	4	0.99	0.97	1
	Thorium-228	4	0.98	0.94	1
	Uranium-235	4	0.86	0.61	0.98
	Uranium-238	4	0.96	0.78	1.1
Sewer	Gross alpha	2	0.96	0.89	1
	Tritium	2	0.97	0.68	1.3
Vegetation	Tritium	4	4.3	0.56	14

a Number of collated pairs used in ratio calculations

Precision is measured by the percent relative standard deviation (%RSD); see the EPA's Data Quality Objectives for Remedial Response Activities: Development Process, Section 4.6 (U.S. EPA 1987). Acceptable values for %RSD vary greatly with matrix, analyte, and analytical method; however, lower values represent better precision. The results for %RSD given in **Table 8-5** are the 75th percentile of the individual precision values.

Regression analysis consists of fitting a straight line to the collocated sample pairs. Good agreement is indicated when the data lie close to a line with a slope equal to 1 and an intercept equal to 0, as illustrated in **Figure 8-1**. Allowing for normal analytical

Table 8-7. Quality assurance collocated sampling: Summary statistics for analytes with at least four pairs in which one or both results were below the detection limit

Media	Analyte	Number of inconsistent pairs	Number of pairs	Percent of inconsistent pairs [a]
Groundwater	Copper	1	17	5.9
	Total organic carbon	2	8	25
	Total organic carbon	2	8	25
	Tritium	1	16	6.3
Soil	Americium-241	1	4	25
Sewer	Gross alpha	1	50	2
	Bromoform	1	4	25
	Ethanol	1	4	25
	Freon 113	1	5	20

a Inconsistent pairs are those for which one of the results is more than twice the reporting limit of the other.

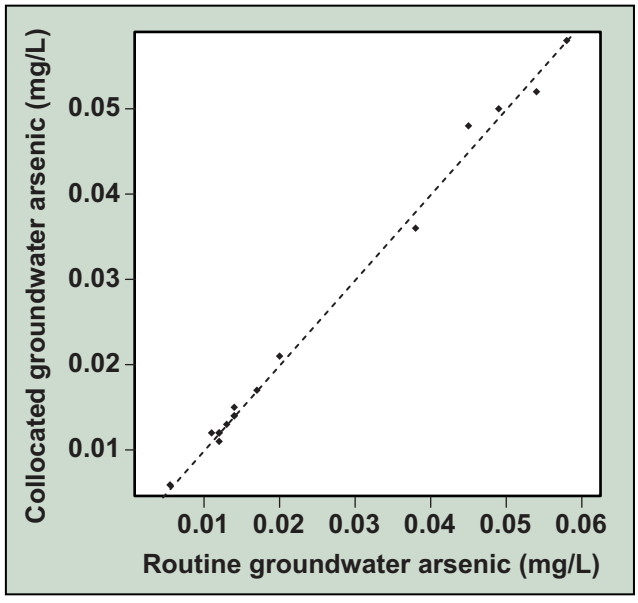

Figure 8-1. Example of data points that lie close to a line with slope equal to 1 and intercept equal to 0 using groundwater arsenic concentrations from collocated samples

variation, the slope of the fitted line should be between 0.7 and 1.3, and the absolute value of the intercept should be less than the detection limit. The coefficient of determination (r^2) should be greater than 0.8. These criteria apply to pairs in which both results are above the detection limit.

When there were more than eight data pairs with both results in each pair considered detections, precision and regression analyses were performed; those results are presented in **Table 8-5**. When there were eight or fewer data pairs with both results above the detection limit, the ratios of the individual duplicate sample pairs were averaged; the mean, minimum, and maximum ratios for selected analytes are given in **Table 8-6**. The mean ratio should be between 0.7 and 1.3. When either of the results in a pair is a nondetection, then the other result should be a nondetection or less than two times the detection limit. **Table 8-7** identifies the sample media and analytes for which at least one pair failed this criterion. Media and analytes with fewer than four pairs are omitted from the table.

Collocated sample comparisons are more variable when the members of the pair are analyzed by different methods or with different criteria for analytical precision. For example, radiological analyses using different counting times or different laboratory aliquot sizes will have different amounts of variability. Different criteria are rarely, if ever, used with collocated sample pairs in LLNL environmental monitoring sampling. Different criteria are sometimes used in special studies when more than one regulatory agency is involved.

Routine and collocated sample results show fairly good agreement: 90% of the pairs have a precision of 38% or better. Data sets not meeting our precision criteria fall into one of two categories. The first category, outliers, can occur because of data transcription errors, measurement errors, or real but anomalous results. Of the 22 data sets reported in **Table 8-5**, four did not meet the criterion for acceptability because of outliers. **Figure 8-2** illustrates a set of collocated pairs with two outliers.

The second category is data sets that do not meet the criterion for acceptability because results are highly variable. This tends to be typical of measurements at extremely low concentrations, as illustrated in **Figure 8-3**. Low concentrations of radionuclides on particulates in air highlight this effect, because a small number or radionuclide-containing particles on an air filter can significantly affect results. Other causes of high variability are sampling and analytical methodology. Analyses of total organic carbon and total organic halides in water are particularly difficult to control. Of the 22 data sets in **Table 8-5**, one shows sufficient variability in results to make it fall outside the acceptable range.

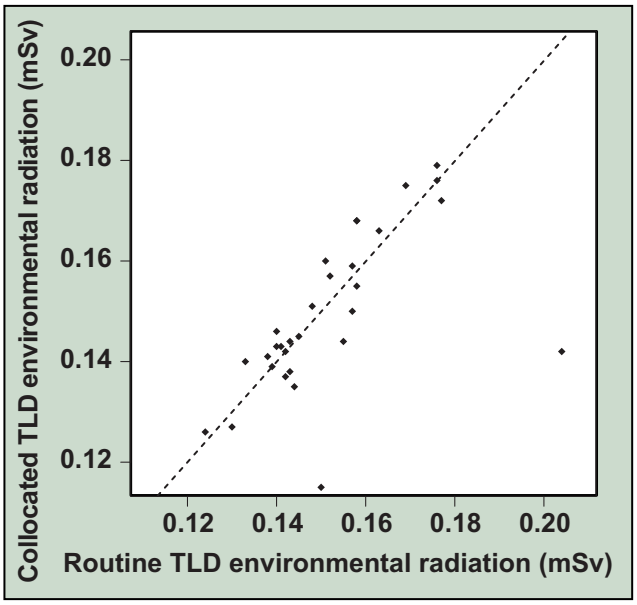

Figure 8-2. Example of data with two outliers using collocated TLD environmental radiation measurements

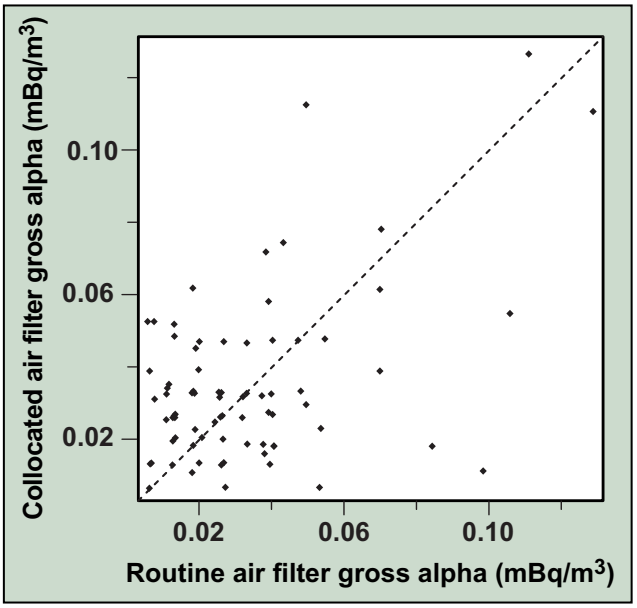

Figure 8-3. Example of variability using air filter gross alpha concentrations from collocated samples

Data Presentation

Most data tables provided in the report CD were created using computer scripts that retrieve data from the database, convert to SI units when necessary, calculate summary statistics for tables that include summary statistics, format data as appropriate, lay out the table into the desired rows and columns, and present a draft table. Final tables are included after review by the responsible analyst. Analytical laboratory data, and values calculated from analytical laboratory data, are normally displayed with two or at most three significant digits. Significant trailing zeros may be omitted.

Radiological Data

Most of the data tables display radiological data as a result plus-or-minus an associated 2σ uncertainty. This measure of uncertainty represents intrinsic variation in the measurement process, most of which is due to the random nature of radioactive decay (see also the section "Reporting Uncertainty in Data Tables" in this chapter). The uncertainties are not used in summary statistic calculations. Any radiological result exhibiting a 2σ uncertainty greater than or equal to 100% of the result is considered to be a nondetection.

Some radiological results are derived from the number of sample counts minus the number of background counts inside the measurement apparatus. Therefore, a sample with a low concentration may have a negative value; such results are reported in the tables and used in the calculation of summary statistics and statistical comparisons.

Some data tables provide a limit-of-sensitivity value instead of an uncertainty when the radiological result is below the detection criterion. Such results are displayed with the limit-of-sensitivity value in parentheses.

Nonradiological Data

Nonradiological data reported by the analytical laboratory as being below the reporting limit are displayed in tables with a less-than symbol. The reporting limit values are used in the calculation of summary statistics, as explained below.

Statistical Comparisons and Summary Statistics

Standard comparison techniques (such as regression, t-tests, and analysis of variance) have been used where appropriate to determine the statistical significance of trends or differences between means. When such a comparison is made, it is explicitly stated in the text as being "statistically significant" or "not statistically significant." Other uses of the word "significant" in the text do not imply that statistical tests have been performed. Instead, these uses relate to the concept of practical significance and are based on professional judgment.

Summary statistics are calculated according to the *Environmental Monitoring Plan* (Woods 2005). The usual summary statistics are the median, which is a measure of central tendency, and interquartile range (IQR), which is a measure of dispersion (variability). However, some tables may present other measures, at the discretion of the responsible analyst.

The median indicates the middle of the data set. That is, half of the measured results are above the median, and half are below. The IQR is the range that encompasses the middle 50% of the data set. The IQR is calculated by subtracting the 25th percentile of the data set from the 75th percentile of the data set. When necessary, the percentiles are interpolated from the data. Different software vendors may use slightly different formulas for calculating percentiles. Radiological data sets that include values less than zero may have an IQR greater than the median. To calculate the median, at least four values are required; to calculate the IQR at least six values are needed.

Summary statistics are calculated from values that, if necessary, have already been rounded (such as when units have been converted from pCi to Bq) and are then rounded to an appropriate number of significant digits. The calculation of summary statistics is also affected by the presence of nondetections. A nondetection indicates that no specific measured value is available; instead, the best information available is that the actual value is less than the reporting limit. Adjustments to the calculation of the median and IQR for data sets that include nondetections are described below.

For data sets with all measurements above the reporting limit and radiological data sets that include reported values below the reporting limit, all reported values, including any below the reporting limit, are included in the calculation of summary statistics.

For data sets that include one or more values reported as "less than the reporting limit," the reporting limit is used as an upper bound value in the calculation of summary statistics.

If the number of values is odd, the middle value (when sorted from smallest to largest) is the median. If the middle value and all larger values are detections then the middle value is reported as the median. Otherwise, the median is assigned a less-than (<) sign.

If the number of values is even, the median is halfway between the middle two values (i.e., the middle two when the values are sorted from smallest to largest). If both of the middle two values and all larger values are detections, then the median is reported. Otherwise, the median is assigned a less-than sign.

If any of the values used to calculate the 25th percentile is a nondetection, or any values larger than the 25th percentile are nondetections, then the IQR cannot be calculated and is not reported.

The median and the IQR are not calculated for data sets having no detections.

Reporting Uncertainty in Data Tables

The measurement uncertainties associated with results from analytical laboratories are represented in two ways. The first of these, significant digits, relates to the resolution of the measuring device. For example, if an ordinary household ruler with a metric scale is used to measure the length of an object in centimeters, and the ruler has tick marks every tenth centimeter, then the length can reliably and consistently be measured to the nearest tenth of a centimeter (i.e., to the nearest tick mark). However, an attempt to be more precise is not likely to yield reliable or reproducible results, because it requires a visual estimate of a distance between tick marks. The appropriate way to report such a measurement would be, for example, "2.1 cm." This would indicate that the "true" length of the object is nearer to 2.1 cm than to 2.0 cm or 2.2 cm (i.e., between 2.05 and

2.15 cm). This result is said to have two significant digits. Although not explicitly stated, the uncertainty is considered to be ± 0.05 cm. A more precise measuring device might be able to measure an object to the nearest one-hundredth of a centimeter; in that case a value such as "2.12 cm" might be reported. This value would have three significant digits and the implied uncertainty would be ± 0.005 cm. A result reported as "3.0 cm" has two significant digits. That is, the trailing zero is significant, and implies that the true length is between 2.95 and 3.05 cm; closer to 3.0 than to 2.9 or 3.1 cm.

When performing calculations with measured values that have significant digits, all digits are used. The number of significant digits in the calculated result is the same as that of the measured value with the fewest number of significant digits.

Most unit conversion factors do not have significant digits. For example, the conversion from milligrams (mg) to micrograms (µg) requires multiplying by the fixed (constant) value of 1000. The value 1000 is exact; it has no uncertainty and therefore the concept of significant digits does not apply.

The other method of representing uncertainty is based on random variation. For radiological measurements, there is variation due to the random nature of radioactive decay. As a sample is measured, the number of radioactive decay events is counted, and the reported result is calculated from the number of decay events that were observed. If the sample is recounted, the number of decay events will almost always be different—because radioactive decay events occur randomly. Uncertainties of this type are reported in this volume as 2σ uncertainties. A 2σ uncertainty represents the range of results expected to occur approximately 95% of the time, if a sample were to be recounted many times. A radiological result reported as, for example, "2.6 ± 1.2 Bq/g" would indicate that with approximately 95% confidence, the "true" value is in the range 1.4 to 3.8 Bq/g (i.e., 2.6 – 1.2 = 1.4 and 2.6 + 1.2 = 3.8).

The concept of significant digits applies to both the radiological result and its uncertainty. So, for example, in a result reported as "2.6 ± 1.2", both the measurement and its uncertainty have the same number of significant digits, that is, two. When expanding an interval reported in the "±" form, for example "2.4 ± 0.44", to a range of values, the rule described above for calculations involving significant digits must be followed. For example, 2.4 – 0.44 = 1.96. However, the measurements 2.4 and 0.44 each have two significant digits, so 1.96 must be rounded to two significant digits, i.e., to 2.0. Similarly, 2.4 + 0.44 = 2.84, and this must be rounded to 2.8. Therefore, a measurement reported as "2.4 ± 0.44 Bq/g" would represent an interval of 2.0 to 2.8 Bq/g.

When rounding a value having a final digit of "5", the software that prepared the tables follows the Institute of Electrical and Electronics Engineers (IEEE) Standard 754-1985, which is "go to the even digit". For example, 2.45 would round down to 2.4, and 2.55 would round up to 2.6.

Quality Assurance Process for the Environmental Report

Unlike the preceding sections, which focused on standards of accuracy and precision in data acquisition and reporting, the following discussion deals with actions to ensure that the content of this report is accurate and has not been corrupted during the publication process. Because publication of a large, data-rich document like this site annual environmental report involves many operations and many people, the chances of introducing errors are great.

The formal QA procedure used for this report has always concentrated on ensuring that the data presented in tables and figures is the same as that reported by the analytical laboratory. Authors, contributors, and technicians have been enlisted to check the accuracy of sections other than those with which they were involved. Members of the Data Management Team (DMT) were excluded from this process because they prepared the tables. When checking values in tables and figures, checkers randomly selected 10% of the numbers and compared them to values in the reports provided by the analytical laboratories. If these values agreed with the reports, further checking was considered unnecessary. If there was disagreement, the checker compared another 10% of the data with the analytical values. If more errors were found, the entire table or figure was checked against the data in the database. This process included checking unit conversions (e.g., from English to SI units) and summary calculations (e.g., mean, interquartile range, fractions of various limits).

The above process was extremely time-consuming. In recent years, advances have been made that are tending to eliminate most of the potential for errors in simple supplementary data tables, such as are found primarily on the report CD. One of the advances is that, rather than sending printed reports that have to be hand-entered into the electronic database, the analytical laboratories now send reports electronically, and these are loaded directly into the database. This practice should result in perfect agreement between the database and data in printed reports from the laboratories. In practice, however, laboratory reporting is not perfect, so the DMT carefully checks all incoming data throughout the year, to make sure that electronic and printed reports from the laboratories agree. This aspect of QC, while not formally part of the QA process for the preparation of this environmental report, is essential to this report's accuracy. Because of this ongoing QC of incoming data, data stored in the database and available for the annual environmental report tables at the time the report is prepared are unlikely to contain errors.

Another advance is that scripts have been written to pull the data from the DMT database directly into the format of the table, including unit conversion and summary statistic calculations. All data tables found on the CD are prepared in this manner. For these tables, it is the responsibility of the appropriate analyst to check each year that the

table is up-to-date (e.g., new locations/analytes added, old ones removed), that the data agree with the data they have received from DMT, and that the summary calculations have been done correctly.

In 2005, LLNL staff checked tables and figures in the report as described above. Quality assurance of the data tables found on the report CD emphasized checking for problems with the scripts rather than for data accuracy. Forms to aid in the QC of tables and figures were distributed along with the appropriate figure, table, and text; a coordinator kept track of the process. Items to be checked included figure captions and table titles for clarity and accuracy, data accuracy and completeness, figure labels and table headings, units, significant digits, and consistency with text. Completed QC forms and the corrected figures or tables were returned to the report editors, who, in collaboration with the contributor, ensured that corrections were made.