9. Quality Assurance

Donald H. MacQueen • Gene Kumamoto

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.

9.1 Quality Assurance Activities

Nonconformance reporting and tracking is a formal process used to ensure that problems are identified, resolved, and prevented from recurring. The LLNL EPD tracks problems using nonconformance reports (NCRs). NCRs are initiated when items or activities are identified that do not comply with procedures or other documents that specify requirements for EPD operations or that cast doubt on the quality of EPD reports, integrity of samples, or data *and* that are not covered by other reporting or tracking mechanisms. There were no laboratory nonconformances documented. Many minor sampling or data problems are resolved without an NCR being generated.

LLNL averts sampling problems by requiring formal and informal training on sampling procedures. Errors that occur during sampling generally do not result in lost samples but may require extra work on the part of laboratory or sampling and data management personnel to correct the errors.

LLNL addresses commercial analytical laboratory problems as they arise. Many of the documented problems concern minor documentation errors and are corrected soon after they are identified. Other problems, such as missed holding times, late analytical results, incorrect analysis and typographical errors on data reports, account for the remaining issues and are not tracked as nonconformances. These problems are corrected by the commercial laboratory reissuing reports or correcting paperwork and do not affect associated sample results.

LLNL participates in the Department of Energy Consolidated Auditing Program (DOECAP). Annual, on-site visits to commercial laboratories under contract to LLNL are part of the auditing program to ensure that accurate and defensible data are generated. The audit program is based on National Environmental Laboratory Accreditation Program (NELAP) requirements. All commercial laboratories used by LLNL EPD are DOE-qualified vendors. LLNL has qualified auditors under the DOECAP program in the areas of quality assurance, organic chemistry, inorganic chemistry, radiochemistry, laboratory information management, and hazardous material management. Audit reports, checklists, and Corrective Action Plans are maintained under the DOECAP program for qualified commercial labs. In FY2008, the laboratories certified by the State of California operating at LLNL as government owned and contractor operated were not internally assessed or qualified by EPD use due to budgetary and staff limitations.

9.2 Analytical Laboratories and Laboratory Intercomparison Studies

In 2008, LLNL had Blanket Service Agreements (BSAs) with nine commercial analytical laboratories and used two on-site analytical laboratories. All analytical laboratory services used by LLNL are provided by facilities certified by the State of California. LLNL works closely with these analytical laboratories to minimize problems and ensure that QA objectives are maintained.

LLNL uses the results of intercomparison performance evaluation program data to identify and monitor trends in performance and to draw attention to the need to improve laboratory performance. If a laboratory performs unacceptably for a particular test in two consecutive performance evaluation studies, LLNL may select another laboratory to perform the affected analyses until the original laboratory has demonstrated 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 formally notifies 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.

Although laboratories are also required to participate in laboratory intercomparison programs, permission to publish their results for comparison purposes was not granted for 2008. To obtain Mixed Analyte Performance Evaluation Program (MAPEP) reports that include the results from all participating laboratories, see http://www.inl.gov/resl/mapep/reports.html. MAPEP is a DOE program and the results are publicly available.

9.3 Duplicate Analyses

Duplicate (collocated) samples are distinct samples of the same matrix collected as close to the same point in space and time as possible. Collocated samples that are 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 that are 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 identify errors such as mislabeled samples or data entry errors.

Tables 9-1, **9-2**, and **9-3** present summary statistics for collocated sample pairs, grouped by sample matrix and analyte. Samples from both the Livermore site and Site 300 are included. **Tables 9-1** and **9-2** are based on data pairs in which both values are detections (see **Section 9.4**). **Table 9-3** is based on data pairs in which either or both values are nondetections.

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 9-1**. 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 9-2**. The mean ratio should be between 0.7 and 1.3. When either of the results in a pair is a nondetection, the other result should be a nondetection or less than two times the detection limit. **Table 9-3** 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.

Table 9-1. 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 beta	98	15.8	0.958	0.94	1.94×10^{-5} (Bq/m^3)
	Beryllium[(e)]	14	69.4	0.48	0.48	2.1 (pg/m^3)
	Uranium-235	12	7.94	0.968	0.97	1.65×10^{-8} (µg/m^3
	Uranium-238	12	8.73	0.962	0.98	2.61×10^{-6} (µg/m^3)
	Tritium	22	23.6	0.959	0.96	0.001 (Bq/m^3)
Dose (TLD)	90-day radiological dose	31	2.94	0.955	0.92	0.786 (mrem)
Groundwater	Gross alpha[(d)]	11	40.6	0.29	0.19	0.16 (Bq/L)
	Gross beta[(e)]	47	16.6	0.641	0.76	0.0884 (Bq/L)
	Arsenic	25	13.5	1.01	0.99	–0.000696 (mg/L)
	Barium[(e)]	15	3.11	0.207	0.13	0.0468 (mg/L)
	Nitrate (as NO_3)	19	2.36	1	1	0.0734 (mg/L)
	Tritium	15	7.41	0.992	1	12.4 (Bq/L)
	Uranium-234+ uranium-233	18	9.44	1.04	0.98	–0.00307 (Bq/L)
	Uranium-235	15	29.7	0.897	0.92	0.000121 (Bq/L)
	Uranium-238	18	7.52	0.947	0.98	0.00108 (Bq/L)
Sewer	Gross beta[(d)]	52	16.1	0.479	0.25	0.000348 (Bq/mL)
	Acetone[(d)]	13	30.1	0.199	0.1	177 (µg/L)
	Chloroform[(d)]	13	20.2	0.789	0.79	0.442 (µg/L)

(a) Number of collocated pairs included in regression analysis.

(b) 75th percentile of percent relative standard deviations (%RSD) where %RSD = $\left(\frac{200}{\sqrt{2}}\right)\frac{|x_1 - x_2|}{x_1 + x_2}$ where x_1 and x_2 are the reported concentrations of each routine–collocated 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 9-2. 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
Air	Gross alpha	5	1.3	0.41	2.7
Groundwater	Radium-226	4	1.2	0.83	1.5
	Thorium-228	1	1.3	1.3	1.3
OW	Gross alpha	1	0.88	0.88	0.88
	Gross beta	1	1	1	1
Runoff (from rain)	Gross beta	1	1.3	1.3	1.3
Soil	Cesium-137	3	1.1	0.82	1.6
	Potassium-40	3	1	1	1
	Plutonium-238	1	1.2	1.2	1.2
	Plutonium-239+240	2	0.94	0.82	1.1
	Radium-226	3	1	0.97	1
	Radium-228	3	1	0.98	1.1
	Thorium-228	3	0.98	0.92	1
	Uranium-235	3	0.92	0.74	1
	Uranium-238	1	1	1	1
Sewer	Gross alpha	3	1.2	0.76	1.8
	Tritium	2	0.98	0.68	1.3
Vegetation	Tritium	5	1.3	0.46	2.1

(a) Number of collocated 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 9-1** are the 75th percentile of the individual precision values. Routine and collocated sample results show good %RSD—90% of the pairs have %RSD of 29% or better; 75% have %RSD of 16% or better.

Table 9-3. 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	No. inconsistent pairs[(a)]	No. pairs	Percent inconsistent pairs
Air	Gross alpha	1	99	1
	Gross beta	1	6	17
Groundwater	Gross alpha	7	38	18
	Arsenic	2	11	18
	Cadmium	1	41	2.4
	Chromium	1	20	5
	Lead	1	41	2.4
	Nitrate (as NO_3)	1	7	14
	Zinc	3	39	7.7
Sewer	Gross alpha	7	50	14
	Toluene	1	12	8.3

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

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 9-1**. Allowing for normal analytical and environmental 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.

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 if more than one agency is involved and each sets its own analytical criteria.

Data sets that do not meet LLNL regression analysis criteria fall into one of two categories: outliers and high variability. Outliers can occur because of data transcription errors, measurement errors, or real but anomalous results. Of the 18 data sets reported in **Table 9-1**, three did not meet the criterion for acceptability because of outliers. **Figure 9-2** illustrates a set of collocated pairs with one outlier.

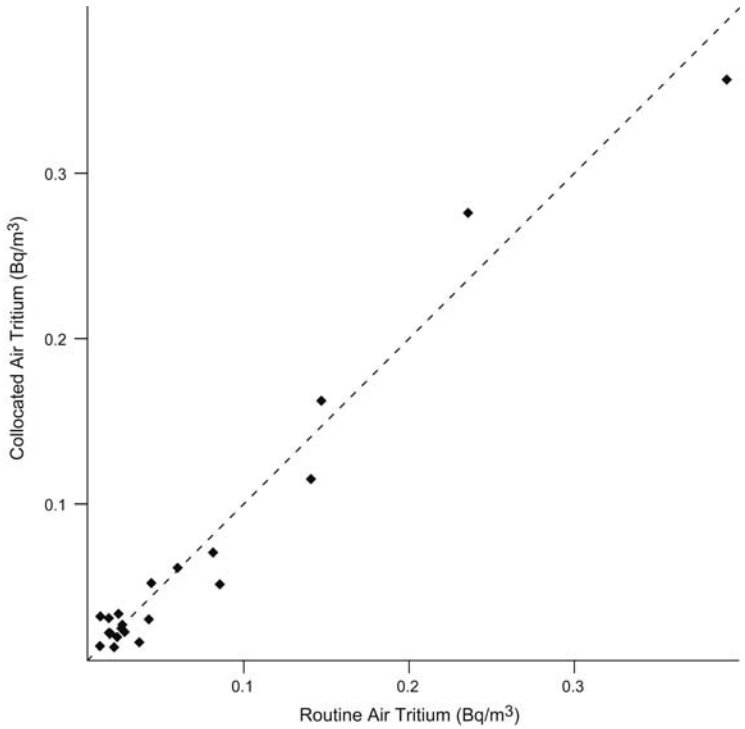

Figure 9-1. Example of data points that demonstrate good agreement between collocated sample results using tritium concentrations in air

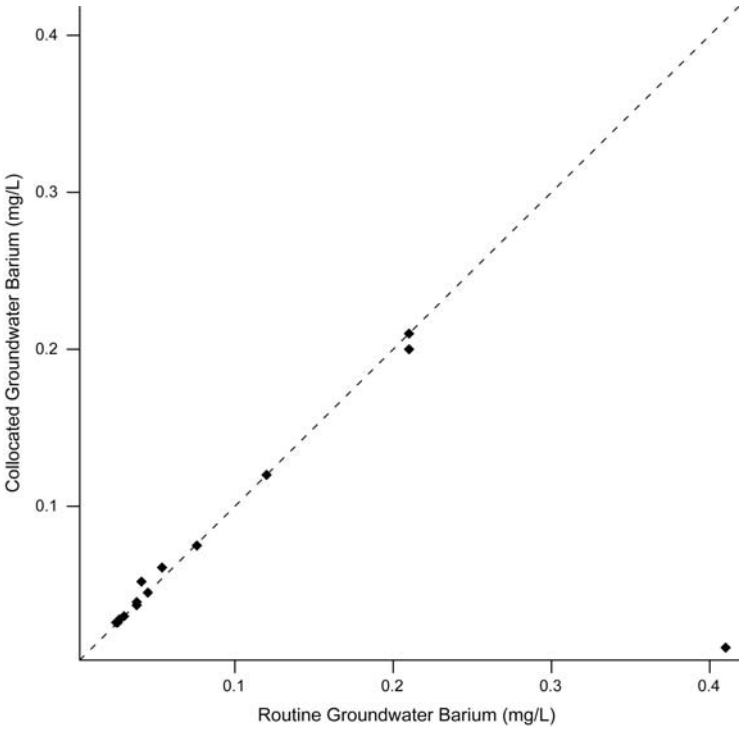

Figure 9-2. Example of data with one outlier using collocated groundwater barium concentrations

The second category, high variability, occurs when the measurement process inherently has substantial variability (see **Figure 9-3** for an example). It also tends to occur at extremely low environmental concentrations. Low concentrations of radionuclides on particulates in air highlight this effect because a small number of radionuclide-containing particles on an air filter can significantly affect results. Analyses of total organic carbon and total organic halides in water are particularly difficult to control. Of the 18 data sets listed in **Table 9-1**, four show sufficient variability in the results to make them fall outside the acceptable range.

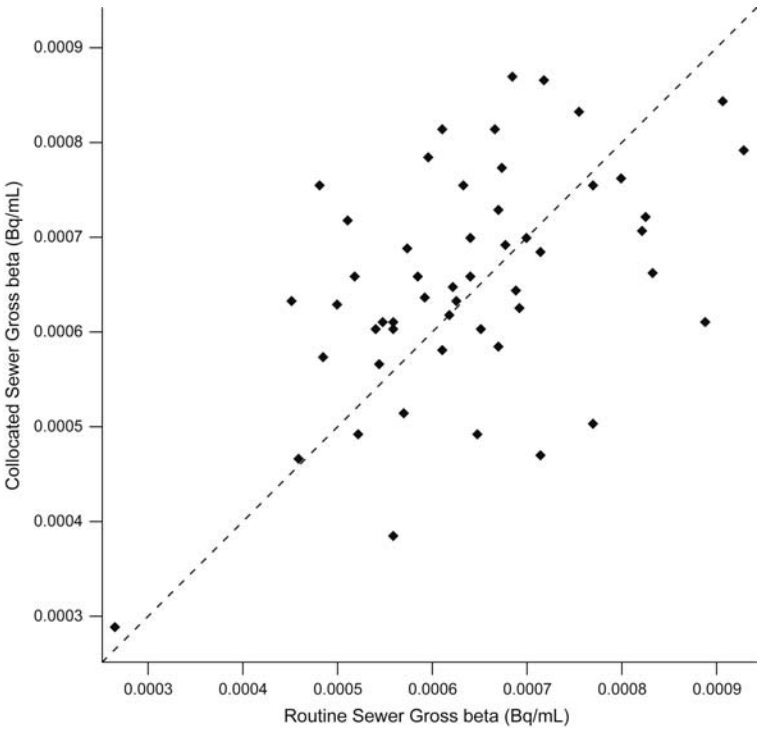

Figure 9-3. Example of variability using collocated sewer gross beta concentrations

9.4 Data Presentation

The data tables in **Appendix A** were created using computer scripts that retrieve data from a database, convert the data into Système International (SI) units when necessary, calculate summary statistics, format data as appropriate, format the table into rows and columns, and present a draft table. The tables are reviewed by the responsible analyst. Analytical laboratory data and the values calculated from the data are normally displayed with two, or at most three, significant digits. Significant trailing zeros may be omitted.

9.4.1 Radiological Data

Most of the data tables in **Appendix A** 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

Section 9.6). 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 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 concentration at or near background may have a negative value. Such results are reported in the data 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.

9.4.2 Nonradiological Data

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

9.5 Statistical Comparisons and Summary Statistics

Standard comparison techniques such as regression analysis, *t*-tests, and analysis of variance are used where appropriate to determine the statistical significance of trends or differences between means. When a comparison is made, the results are described as either "statistically significant" or "not statistically significant." Other uses of the word "significant" in this report do not imply that statistical tests have been performed but relate to the concept of practical significance and are based on professional judgment.

Summary statistics are calculated according to Gallegos (2009). 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 data tables may present other measures at the discretion of the analyst.

The median indicates the middle of the data set (i.e., 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. In this report, at least four values are required to calculate the median and at least six values are required to calculate the IQR.

Summary statistics are calculated from values that, if necessary, have already been rounded, such as when units have been converted from picocuries to becquerels, 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, 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, the median is reported. Otherwise, the median is assigned a less-than (<) sign.

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

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

9.6 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 one-tenth of a centimeter, the length can reliably and consistently be measured to the nearest tenth of a centimeter (i.e., to the nearest tick mark). An attempt to be more precise is not likely to yield reliable or reproducible results because it would require a visual estimate of a distance between tick marks. The appropriate way to report a measurement using this ruler would be, for example, 2.1 cm, which 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). A measurement of 2.1 cm has two significant digits. Although not 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 to micrograms 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 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 of 1.4 to 3.8 Bq/g (i.e., 2.6 – 1.2 = 1.4 and 2.6 + 1.2 = 3.8). When necessary, results are converted from pCi to Bq by multiplying by 0.037; this introduces extraneous digits that are not significant and should not be shown in data tables. For example, 5.3 pCi/g × 0.037 = 0.1961 Bq/g. The initial value, 5.3, has two significant digits, so the value 0.1961 would be rounded to two significant digits, that is, 0.20.

However, the rounding rule changes when there is a radiological uncertainty associated with a radiological result. In this case, data are presented according to the method recommended in Multi-Agency Radiological Laboratory Analytical Protocols (MARLAP) Section 19.3.7 (U.S. NRC/U.S. EPA 2004). First the uncertainty is rounded to the appropriate number of significant digits, after which the result is rounded to the same number of decimal places. For example, suppose a result and uncertainty after unit conversion are 0.1961 ± 0.05436, and the appropriate number of significant digits is two. First, 0.05436 is rounded to 0.054 (two significant digits). 0.054 has three decimal places, so 0.1961 is then rounded to three decimal places, i.e., 0.196. These would be presented in the data tables as 0.196 ± 0.054.

When rounding a value with a final digit of "5," the software that was used to prepare the data tables follows the IEEE Standard 754−1985, which is "go to the even digit." For example, 2.45 would be rounded down to 2.4, and 2.55 would be rounded up to 2.6.

The software that prepares the data tables pays careful attention to the details of rounding for significant digits. It should be noted, however, that these details are of little practical significance. For example, if a result of 5.6 is incorrectly rounded to 5.5 or 5.7, the introduced "error" is less than 2% (0.1/5.6 = 0.018). Such an error will rarely have any impact on the interpretation of the data with respect to human health or environmental impact.

9.7 Quality Assurance Process for the Environmental Report

Unlike the preceding sections, which focused on standards of accuracy and precision in data acquisition and reporting, this section describes the actions that are taken to ensure the accuracy of this data-rich environmental report, the preparation of which involves many operations and many people. The key elements that are used to ensure accuracy are described below.

Analytical laboratories send reports electronically, which 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 EPD and ERD Data Management Teams (DMTs) carefully check incoming data throughout the year to make sure that electronic and printed reports from the laboratories agree. This aspect of QC is essential to the report's accuracy. Because of this ongoing QC of incoming data, data stored in the database and used to prepare the annual environmental report tables are unlikely to contain errors.

As described in **Section 9.4**, scripts are used to pull data from the database directly into the format of the table, including unit conversion and summary statistic calculations. All of the data tables contained in **Appendix A** were prepared for this report 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 he or she has received from DMT, and that the summary calculations have been done correctly.

For this 2008 environmental report, LLNL staff checked tables and figures in the body of the report. Forms to aid in the QC of tables and figures were distributed along with the appropriate figure, table, and text, and a coordinator kept track of the process. Items that were checked included clarity and accuracy of figure captions and table titles; 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 editor, who, in collaboration with the responsible author, ensured that corrections were made.

9.8 Errata

Appendix E contains the protocol for errata in LLNL *Environmental Reports* and the errata for LLNL *Environmental Report 2007*.