

3.4 Validation and Data Quality Assessment

Data validation and verification (V&V) during CY 2010 was performed by Legacy Management Support contractor personnel at the Grand Junction, Colorado, office. Data quality assessment (DQA) is performed by personnel at the Site. The following section distinguishes DQA from data validation and discusses the technical basis, equations, and criteria used in the DQA of the water sampling analytical data.

3.4.1 General Discussion

Data validation is the principal means of assessing the usability of water analytical data. Validation also improves overall data quality by allowing the laboratory coordinator to closely monitor laboratory performance and to provide feedback to each laboratory regarding its ability to produce quality data that meets subcontract requirements. The laboratory coordinator may also use the results of data validation to direct analytical work to laboratories that demonstrate superior performance by generating timely, high-quality analytical data for the Site.

Data validation is a rigorous data review performed by the laboratory coordinator or designee on all of the water analytical data generated by the Site. Additionally, the Site lead may request a secondary detailed validation on a case-by-case basis. Data validation is currently performed as specified in Stoller *Environmental Procedures Catalog* (LMS/PRO/S04325), “Standard Practice for Validation of Laboratory Data.” This procedure is based on the following EPA documents:

- EPA 2010b, *USEPA Contract Laboratory Program National Functional Guidelines for Inorganic Data Review*, USEPA-540-R-10-011, October;
- EPA 2008, *USEPA Contract Laboratory Program National Functional Guidelines for Organic Data Review*, USEPA-540-R-08-01, June;
- EPA 2001, *USEPA Contract Laboratory Program National Functional Guidelines for Low Concentration Organic Data Review*, EPA-540-R-00/006, June; and
- EPA 1997, *Evaluation of Radiochemical Data Usability*, Office of Environmental Management, ES/ER/MS-5, April.

All water analytical data collected by the Site are considered valid unless data validation identifies analytical problems that require the data to be qualified. When it is necessary to qualify individual data records, standard qualifier codes (alphanumeric validation codes) are applied.

Common data qualifiers used by LM are defined below. Refer to *Environmental Procedures Catalog*, “Standard Practice for Validation of Laboratory Data” for formal definitions.

- U For organic and inorganic analytes, the analyte was not detected at a concentration greater than the method detection limit. For radiochemistry, the analyte was not detected at a concentration greater than the decision-level concentration.
- J The associated numerical value is an estimated quantity.
- R The data are unusable (analyte may or may not be present). Resampling and reanalysis are necessary for verification.

Data validation includes the evaluation of laboratory Quality Control (QC) data such as method blanks, laboratory control samples (LCSs), and spike recoveries. Adherence to sample and extract holding times, standard analytical methods, contractual requirements, and proper documentation are also verified.

Although DQA and data validation examine some of the same QC data, they do so from different perspectives. DQA (in this report) looks at the overall quality of an entire year of water data, in contrast to validation, which looks at the analytical details of individual data packages. Data validation focuses on laboratory performance, while DQA focuses on interpretation of data describing QC samples that originated in the field, such as field duplicate samples and equipment rinsate samples.

In contrast to data validation, the DQA performed by personnel at the Site does not assign data qualifiers to individual analytical results or data packages. DQA is a second level of QA intended to be a general assessment of how well the water data collection program is operating. The DQA is performed by evaluating water quality data in terms of the precision, accuracy, representativeness, completeness, and comparability (PARCC) parameters.

3.4.2 PARCC Parameters

Use of the PARCC parameters for DQA has been promoted by EPA guidance documents. Accuracy and precision are quantitative measures. Representativeness and comparability are qualitative measures. Completeness is a combination of both quantitative and qualitative measures.

Site personnel evaluate the PARCC parameters by following guidelines published in these former QC documents:

- RMRS 1998, *Procedure for Evaluation of Data for Usability*;
- RMRS 2000b, *Quality Assurance Program Plan for the Automated Surface-Water Monitoring Program*, RF/RMRS-2000-013, Revision 0 ; and
- RMRS 2001, *Quality Assurance Program Plan for the Groundwater Monitoring Program Rocky Flats Environmental Technology Site*.

The following sections discuss the PARCC parameters and the types of data available to assess them.

3.4.2.1 Criteria for Precision

The precision of a measurement is an expression of the agreement between duplicate measurements of the same property taken under similar conditions. Precision can be expressed quantitatively by the relative percent difference (RPD) between real and field duplicate samples for metals, VOCs, polychlorinated biphenyls, and water quality parameters (WQPs) as defined by the following equation:

$$RPD = \frac{|(S - D)|}{(S + D)/2} \times 100$$

where: S = Concentration of analyte in the real sample,
D = Concentration of analyte in the duplicate sample, and
RPD = relative percent difference.
Nondetects are not included.

The Site uses the duplicate error ratio (DER) to quantify the precision of radionuclide activity data:

$$DER = \frac{|(S - D)|}{\sqrt{[(TPU_S)^2 + (TPU_D)^2]}}$$

where: DER = Duplicate error ratio,
S = Sample result,
D = Duplicate (or lab replicate) result,
TPU_S = Total propagated uncertainty of the sample,
TPU_D = Total propagated uncertainty of the duplicate.

Because total propagated uncertainty is seldom reported with radionuclide activity data, the two-sigma error or random counting error has been substituted for total propagated uncertainty in the U, Am, and Pu calculations made for this report.

The Site QC criterion for water RPDs is that individual RPDs should be ≤30 percent. The analogous criterion for DERs is ≤1.96. The overall goal for the water data set is to have 85 percent of the RPD and DER values comply with the QC criteria.

3.4.2.2 Criteria for Accuracy

Accuracy is the degree of agreement for a measurement with an accepted reference or true value and is a measure of the bias in a system. The closer the measurement is to the true value, the more accurate the measurement. The Site validation process is the principal means for evaluating the accuracy of analytical results.

Because the Site V&V process compares the actual analytical methods used by each laboratory to the contract-required analytical methods, the Site does not repeat this evaluation.

Matrix spike (MS) and matrix spike duplicate (MSD) recoveries are reported by the analytical laboratories for most nonradionuclide analytical suites. Criteria for acceptable MS recoveries vary between laboratories, depending on the analyte and the analytical method. The Site criterion for acceptable MS results ranges from 75 to 125 percent recovery.

LCS recoveries for radionuclides are often available for water quality data. Laboratories in practice will commonly accept LCS values in the range of 70 to 130 percent. LCS percent recoveries between the 70 to 130 percent laboratory range and the 75 to 125 percent QC range required by the Site laboratory contracts are examined by data validators for acceptability on an analyte-by-analyte basis. The Site criterion for acceptable LCS recoveries ranges from 75 to 125 percent recovery.

Because some laboratories reported LCS results in picocuries per liter, while others calculated percent recovery, the Site uses the “relative bias” reporting criterion. The relative bias criterion is

defined in the Basic Ordering Agreement by the following formula (see Page J-6 of the National Basic Ordering Agreement, Section 2.3.2.5):

$$\text{Relative Bias} = (\text{Observed} - \text{Known}) \div \text{Known}$$

where: Observed = measured activity of LCS standard (pCi/L),
Known = known activity of LCS standard (pCi/L).

Acceptable values for relative bias results range from -0.25 to +0.25.

3.4.2.3 Criteria for Representativeness

Representativeness in DQA is limited to an evaluation of whether analytical results for field samples are truly representative of environmental concentrations, or whether they may have been influenced by the introduction of contamination during collection and handling. The potential introduction of contamination is commonly evaluated by examination of the analytical results for equipment rinsates.

Equipment rinsates are used to assess the efficacy of the process used to clean and decontaminate water sampling equipment. Analytes detected in rinsate samples indicate possible cross-contamination between environmental samples. Rinsates are samples of analyte-free distilled or deionized water that has been poured over or through decontaminated sampling equipment and subsequently handled in the same manner as environmental samples. For flow-paced composite samples that are collected over time in carboys, a location-specific “rinse carboy” is prepared using distilled water. This carboy is treated the same as other surface-water samples from that location and analyzed for the same parameters. Analytical data for these rinse carboys are used to assess how well the carboys were cleaned between field deployments and to determine whether contamination was introduced during sample preparation.

Although rinsates are used specifically as indicators of cross-contamination from improper decontamination of equipment, they are carried through the entire sampling, shipping, and laboratory process. Therefore, they are good indicators of potential contamination introduced during any of these steps.

3.4.2.4 Criteria for Completeness

A qualitative measure of completeness is the rate of successful sampling. The DQA verifies that all planned samples were collected, unless insufficient water was available for sampling. The completeness goal for successful sampling is the collection of at least 90 percent of the planned samples. However, the availability of water is outside the control of the Site. If all required stations were visited, sampling completeness is considered acceptable.

Completeness as a quantitative measure of data quality may be expressed as the percentage of valid or acceptable data obtained from a measurement system. The Site tracks analytical laboratory performance through both the shipment of samples to the laboratory and the receipt of data from the laboratory. The Site also evaluates data completeness using the following formula:

$$\text{Completeness} = DP_u = \frac{DP_t - DP_n}{DP_t} \times 100$$

where: DP_u = Percentage of usable data points
 DP_t = Total number of data points
 DP_n = Nonusable (rejected) data points

The completeness criterion is having ≥ 90 percent valid samples.

3.4.2.5 Criteria for Comparability

Comparability is a qualitative parameter. Consistency in the acquisition, handling, and analysis of samples is necessary for comparing results. Samples are collected in accordance with Site standard operating procedures, transported according to Site standard operating procedures and U.S. Department of Transportation shipping regulations, and analyzed using standard EPA or nationally recognized analytical methods. These criteria help to ensure comparability of results with other analyses performed in a similar manner.

The laboratory coordinator or designee verifies that laboratory analyses are performed according to the standard protocols specified by the Site subcontract to each laboratory. Therefore, the analytical results should be comparable to data produced by similar methods.

3.4.3 Water DQA Results for CY 2010

Data used to evaluate the PARCC parameters are included in the available CY 2010 analytical data generated by the laboratories. These include analyses of field duplicate and rinsate QC samples submitted to the laboratory, and laboratory-generated QA/QC samples such as LCSs. ***This PARCC evaluation is limited to analyses at routine RFLMA locations, for samples collected by routine protocol, and for analytes that are listed in Table 1 of RFLMA Attachment 2.***²³ By limiting the evaluation to RFLMA locations, sample protocols, and analytes, more targeted and accurate assessment is made for analytes that have water quality standards applicable to the Site. The DQA of these analyses is discussed below by each PARCC parameter.

During CY 2010, 114 locations were sampled one or more times. This resulted in a total of 497 water samples collected.²⁴ During CY 2010, 1,237 bottles of water were submitted to analytical laboratories for analysis. Table 95 breaks this data down by sample type.

Table 95. CY 2010 Sample Type Breakdown

	Unique Water Samples	Unique Bottle Codes
Primary samples (REALs)	469	1,098
Field duplicates (DUPs)	28	80
Rinsates (RNSs)	20	59
Totals	517	1,237

²³ Hardness and total suspended solids are also included, though these analytes are not listed in Table 1 of RFLMA Attachment 2.

²⁴ This is the sum of real and duplicate samples for unique sampling events.

3.4.3.1 Precision During CY 2010

DERs are indicators of precision for radionuclide analyses. The QC criterion for precision requires that individual DER values should be ≤ 1.96 , and overall the data set should have ≥ 85 percent compliance with the criterion. Appendix Table B-1 is a tabulation of the DER values for CY 2010 radionuclide analyses. The table has been sorted by the DER parameter so that the range of values is apparent. The DER range is from 0.06 to 1.47.

Table 96 summarizes the DER findings of Table B-1 and indicates if the 85 percent goal has been met. Overall, 100 percent of the DER data are in compliance with the criterion, indicating excellent precision for radionuclide analyses.

Table 96. Summary of DER Values

Analyte Group	Total Number of DER Results	Number of Unacceptable Results DER >1.96	Number of Acceptable Results	Percentage Acceptable	Goal Met
Radionuclides	20	0	20	100	Yes

The RPD between real and field duplicate sample results is an indicator of precision for nonradionuclide analyses. Individual RPD values should be ≤ 30 percent, and at least 85 percent of the RPDs should comply with the criterion. Appendix Table B-2 tabulates RPD values and is sorted first by analyte suite, then by RPD, in order to highlight the RPD range of each suite. RPD values ranged from 0.0 percent to 95.8 percent for metals, 0.5 percent to 115 percent for WQPs, and 0.0 percent to 72.7 percent for VOCs/SVOCs.

Table 97 summarizes the RPD findings of Table B-2 and indicates if the 85 percent goal has been met. During CY 2010, the RPD goal was met for all analyte groups. Overall, the nonradionuclide data had 94.2 percent acceptable RPDs and therefore exceeded the 85 percent goal.

Table 97. Summary of RPD Values

Analyte Group	Total Number of RPD Results	Number of Unacceptable Results RPD >30%	Number of Acceptable Results	Percentage Acceptable	Goal Met
Metals	39	2	37	94.9	Yes
WQPs	7	1	6	85.7	Yes
VOCs/SVOCs	40	2	38	95.0	Yes
Totals	86	5	81	94.2	Yes (overall)

3.4.3.2 Accuracy During CY 2010

MS recoveries provide another measure of accuracy. Appendix Table B-3 displays recoveries for 1,536 MS and MSD analytical records for metals, VOCs/SVOCs, and WQPs. These data are summarized in Table 98. All individual suites met the goal with greater than 90 percent of their spike recoveries falling in the acceptable range. Overall, across all analytical suites, the percentage of acceptable MS/MSD results was 96.7 percent.

Table 98. Summary of MS and MSD Recovery Data

Analyte Group	Total Number of MS & MSD Results	Number of Low Results Below 75%	Number of High Results Above 125%	Number Acceptable	Percentage Acceptable	Goal Met
Metals	578	4	0	574	99.3	Yes
WQPs	118	5	0	113	95.8	Yes
VOCs/SVOCs	840	33	8	799	95.1	Yes
Totals	1,536	42	8	1,486	96.7	Yes (overall)

Appendix Table B-4 contains 156 relative bias values for LCSs. These are used by the Site to evaluate the accuracy of radionuclide analyses. The QC criterion for the acceptable range of relative bias values is from -0.25 to +0.25. During CY 2010, the bias ranged from -0.25 to +0.140. All of the data met the QC criterion.

LCS results for nonradionuclide suites were available for metals, VOCs/SVOCs, and WQPs (including anions). These LCS recoveries are tabulated in Appendix Table B-5, which is sorted by analyte group, then by percent recovery. There are 430 LCS data records for metals. The LCS recoveries for metals fell in the range 80.4 percent to 116 percent and were all within the 75 percent to 125 percent acceptable QC range. There are 1,077 LCS data records for VOCs/SVOCs. LCS recoveries for VOCs/SVOCs fell between 34.0 percent and 130 percent. Fifty-three records are outside the 75 percent to 125 percent acceptable QC range (95.1 percent acceptable). There are 164 LCS data records for WQPs. LCS recoveries for WQPs fell between 88 percent and 114 percent and were all acceptable. Overall for nonradionuclides, 96.8 percent of the LCS recoveries indicate that CY 2010 water analytical data for metals, VOCs/SVOCs, and WQPs are of high accuracy.

Another aspect of accuracy is “rejected data.” Out of 10,308 analytical records representing reals, duplicates, and rinsates during CY 2010, one record was rejected (R qualified) during data V&V. Another way to state this is that 99.99 percent of the analytical data collected during the year were considered to be valid and usable.

3.4.3.3 Representativeness During CY 2010

As defined earlier, representativeness is an evaluation of the sampling procedure for its ability to reflect the true concentrations of contaminants in water. The Site uses equipment rinsate samples (and “rinse carboys”) to determine whether contamination is introduced from improper or incomplete decontamination of the sampling equipment.

During CY 2010 a total of 435 rinsate analytical records were generated for metals, radionuclides, VOCs/SVOCs, and WQPs. The majority of these records lack evidence of contamination. The remaining 10 records are tabulated in Appendix Table B-6. Three of these are “B”-qualified metals data, which constitute only weak evidence of contamination. The B qualifier for inorganics indicates that the concentrations are above the instrument detection limit but below the contract required detection limit. All the remaining records are “J”-qualified, indicating an estimated value.

Overall, there is very little evidence of introduced contamination during CY 2010 water sampling and shipping activities. Most of the 435 rinsate records appear to be clean. Therefore, water quality data for the year are judged to be representative of the actual water concentrations.

Because all required sampling locations were visited, and the samples that could be collected were analyzed, analyses for the year are judged to be representative with respect to spatial coverage.

3.4.3.4 Completeness During CY 2010

If sufficient water is available for sampling, the goal is to have 100 percent successful sampling of all required locations. However, the availability of water is beyond the control of the samplers. Surface-water monitoring during CY 2010 targeted sampling at 17 RFLMA surface-water sampling locations. In actuality, samples were collected at 16 sites and were submitted to the laboratory for analysis; one location, PLFPONDEFF was not required to be sampled in CY 2010 according to the RFLMA monitoring protocols. Groundwater monitoring during CY 2010 targeted sampling at 91 wells. In actuality, samples were collected at 90 wells and were submitted to the laboratory for analysis. One location, Sentinel well 95299, was dry. Treatment system monitoring during CY 2010 targeted sampling at eight locations; samples were collected at all eight locations and were submitted to the laboratory for analysis. Because dry locations do not count against sampling success rates (being beyond the control of samplers), success rates for surface water, groundwater, and treatment system sampling are all 100 percent.

V&V completeness is summarized in Table 99. This table compiles, by analyte group, the total number of data points for reals, duplicates, and rinsate samples. It then subtracts rejected data points (one for 2010) as well as points that lack validation qualifiers (zero for 2010). The result is the net number of usable validated or verified data points, and this is expressed as percent usable data, or percent V&V completeness. The QC goal for completeness is ≥ 90 percent.

Table 99. Summary of V&V Data Completeness

Analyte Group	Number of Data Points	Number of Unvalidated Points	Number Rejected	Net Usable Points	Percent Completeness	Goal Met
Metals	1,229	0	0	1,229	100.0	Yes
Radionuclides	268	0	0	268	100.0	Yes
WQPs	225	0	1	224	99.6	Yes
VOCs/SVOCs	8,586	0	0	8,586	100.0	Yes
	Sum of Number of Data Points	Sum of Number of Unvalidated Points	Sum of Number Rejected	Sum of Net Usable Points	Overall Completeness	Goal Met
Totals	10,308	0	1	10,308	99.99	Yes

Validation completeness for all suites was nearly 100.0 percent and exceeded the completeness goal. Therefore, from the perspective of V&V completeness, the CY 2010 water data are acceptable.

Another measure of completeness is that an adequate number of QC samples (field duplicates and equipment rinsates) must be collected to meet QC requirements. The recommended frequency for collecting duplicate samples is 1 duplicate (DUP) per 20 or fewer primary (REAL) water samples. In other words, duplicates should be collected at a 5 percent or greater frequency per REAL sample. Like duplicates, rinsate samples (RNS) are also to be collected at a 5 percent or greater rate.

The sample collection frequencies of REAL, DUP, and RNS samples are tabulated by analyte group in Table 100. The ratios of REAL/DUP samples shown meet water program QC goals with 1 DUP per 13.4 REALs. Across all analyte suites and samples collected during the year, the overall frequency of duplicates was 7.46 percent, exceeding program goals (≥ 5 percent).

The ratios of REAL/RNS samples shown in Table 100 did not quite meet water program QC goals with 1 RNS per 21.1 REALs. Across all analyte suites and samples collected during the year, the overall frequency of rinsates was 4.73 percent, just missing program goals (≥ 5 percent). This deficiency is due to the use of dedicated sampling equipment at many wells; since the equipment is dedicated, RNSs are not required to indicate effective equipment decon between sampling events.

Table 100. Summary of Field QC Samples and Data Records

Analyte Group	Number of Locations Sampled for REALs	Number of Locations Sampled for DUPs	Ratio REALs/DUPs (Goal <20)	Ratio REALs/RNSs (Goal <20)	Number REAL Records	Number DUP Records	Number RNS Records	Total Records
Metals	75	17	9.68	13.8	1,045	108	76	1,229
Radionuclides	11	8	11.2	9.33	224	20	24	268
WQPs	42	7	23.4	42.2	211	9	5	225
VOCs/SVOCs	94	11	14.1	23.4	7,708	548	330	8,586
		Totals	13.4	21.1	9,188	685	435	10,428
		Percentages				7.46%	4.73%	

3.4.3.5 Comparability During CY 2010

No significant changes were made to water sampling or analytical procedures during CY 2010. Therefore, the analytical data generated during the year should be generally comparable to corresponding analyses from previous years.