

RESPONSES TO EPA COMMENTS
BACKGROUND GEOCHEMICAL CHARACTERIZATION REPORT
DATED FEBRUARY 1990

GENERAL COMMENTS

Comment 1

The material in the report is presented in extremely awkward fashion. A great deal of review time was spent trying to organize the data. The well numbering system adds more digits, but doesn't improve clarity at all, in trying to relate the type of well, location, etc.

For some inscrutable reason, the data printouts were reduced in size, but no effort was made to keep similar data on one page. In many cases, data for all wells but one or two are found together; the remaining data are found later on separate pages. Some compression of information would have been extremely helpful. One must question why, for example, the statistical analyses for Rocky Flats alluvium dissolved radiochemistry are presented in Section 1, Section 4 and Appendix B, with the data in Section A. This makes review of the data extremely difficult and will not be acceptable in the future.

Response:

The rationale for the well numbering system will be provided within the text. Data and computational results will be presented in the text as it is logically developed and explained, and will not be repeated in subsequent sections. Appendices will be consolidated if proposed statistical analyses indicate fewer background populations exist. Data presentations will otherwise be compressed to the extent possible. Because of the incorporation of seven additional rounds of surface-water data, and of two additional quarters of ground-water data, the actual size of appended data will increase.

Comment 2

No information was provided on the analytical methods used. This is necessary to substantiate the detection limits reported. In Appendix A-5, it appears that a considerable amount of information was repeated.

Response:

Information on analytical methods used in this study is found in the Environmental Restoration Program Quality Assurance/Quality Control Plan (QA/QC) which has been previously submitted to the EPA. The QA/QC Plan was referenced explicitly in the text.

Comment 3

It was difficult to identify which well data were grouped as spatial clusters, although this will probably be discussed further in later reports. Since there appear to have been multiple purposes involved in locating the wells, spatial comparisons may not always be possible. For example, EPA needs to know whether the northern well data for the Rocky Flats alluvium is statistically similar to the Woman Creek data. However, the Woman Creek wells are clustered fairly close together, while the northern wells were situated in a linear direction along the line of expected increases in TDS (downgradient). The document notes a gradual increase in common cations and anions from these northern wells. Is a statistical comparison of the average values for these two clusters still reasonable?

In any scientific study, the location of samples dictates the kinds of comparisons desired. Since we don't have access at present as to how the well locations were chosen, it is hard to make any judgments about spatial comparisons. The pattern of well locations does appear to have a high degree of pre-selection, so some a priori comparisons must have been in the minds of the persons who selected the locations.

Response:

The text will be revised to further address the rationale for the locations of background sampling sites. For example, well groupings and general locations were based on surficial geology and groundwater flow directions relative to SWMUs. Potentiometric surface and surficial geology maps will be presented to facilitate this discussion.

Multivariate analysis of variance (MANOVA) will be applied to background data to evaluate the appropriateness of including background data collected both north and south of the plant within a single background population.

Comment 4

Cation/anion and TDS (measured vs calculated) balances were performed on selected ground and surface water data from the report; a more complete presentation follows. Since the background waters are generally low in total dissolved solids, a higher degree of uncertainty is found with these balances. Where HCO_3 is the chief anion, there appears to be a much higher level of uncertainty in the measured value of this analyte, as well as its effect on the TDS balance.

Overall, the balance analyses suggest that there may be some problems with the gross inorganic chemistry data so far collected. EPA attempts to develop reasonable balances generally failed. Without supportive information on the analytical techniques used and how the data were reported, it is not possible to assess why there was such great variation. In order for DOE to improve their analysis DOE must improve the limits of detection for major cation parameters, specify the filter size used for dissolved analyses, identify the procedure for gravimetric TDS analysis, and add dissolved silica and temperature to the major wet chemistry analyses.

Response:

Cation/anion and TDS (measured versus calculated) balances will be incorporated as part of a preliminary data evaluation for all ground-water and surface-water samples. Discrepancies will be evaluated for cause, e.g., potential presence of ions not currently being measured. The data validation process of the QA/QC program is used to determine the acceptability of the data. Analytical methods are provided in the QA/QC Plan. It is not practical to reduce the detection limits for major cations. Silica has been added to the suite of parameters for all wells effective April 1990.

Comment 5

The most important finding in the report may be the presence of significant metals and radioactive materials, particularly manmade varieties, in the seeps at surface water quality locations SW-104 and SW-80. The suspended solid material associated with these springs appears to carry the predominant burden of radioactivity and metals, more so than either the dissolved portions, the immediate sediments, or other sampling strata. A separate section is devoted to the analysis of the relevant data for these stations.

Response:

The background data collected at SW-104 and SW-80 will be compared with the data collected at other surface water stations. If part or all of the data collected at SW-104 or SW-80 are deemed anomalous, this anomalous data will not be incorporated within the surface-water background population.

Comment 6

Although the statistical approach may be generally satisfactory, except as noted above, DOE needs to provide more information about the data for selected analytes, the method of reporting radioactivity data, and the use of the Poisson distribution for infrequently occurring compounds.

Response:

The need for additional information for selected analytes is addressed later in the response to Specific Comment 1. Response to Specific Comment 3b addresses the method of reporting radioactivity data, and response to Specific Comment 3d addresses the use of a statistical procedure based on the Poisson distribution.

SPECIFIC COMMENTS

Comment 1 Cation/Anion and TDS Balances

Approach. An effort was made to check the consistency of the major cation and anion and pH and TDS parameters for selected data. Typically in a cation/anion analysis, data reported in mg/l are converted to milliequivalents per liter, and the difference between the sum of positive and negative ions determined. Since gravimetric TDS data were collected, a calculated estimate of the TDS from the available cation/anion data can be compared with it. Specific conductance data could also have been used as a check on the cation/anion/TDS data, but not enough samples had specific conductance data simultaneously collected to perform this latter analysis, and available specific conductance data were highly variable between sampling episodes.

The USGS criterion for a reasonable cation/anion balance is $\pm 1-2\%$ for waters in the 250-1000 mg/l range. Perhaps as high as $\pm 5-10\%$ would be tolerable for waters in the range of 100-250 mg/l, in which many of the Rocky Flats surface and alluvial waters fall. This USGS criterion recognizes that missing important analytes or the presence of significant organic colloidal material may reduce the accuracy of the results. The USGS also notes that the alkalinity titration (used herein, we believe, to determine the bicarbonate concentration) can lead to errors.

For comparisons between calculated and gravimetric TDS, the USGS considers a difference of a few to perhaps 20-30 mg/l to be adequate.

The intent of such analyses is to identify any major transcription or analytical errors, and to check the overall consistency of the analyses performed.

One immediate problem occurred with the interpretation of bicarbonate (HCO_3^-) data. While most cations and anions were presented as mg/l X of ionic species, bicarbonate was given as mg/l of CaCO_3 . In order to convert this value to milliequivalents of HCO_3^- , the reported values were multiplied by a factor of meqts per 1- HCO_3^- /meqts per 1- $\text{CaCO}_3 = 61/50 = 1.22$. As results show, this tends to result in larger cation/anion balance errors than if the reported value was treated as HCO_3^- .

The USGS notes in its Study and Interpretation of the Chemical Characteristics of Natural Water, WSP 2254 that where there are relatively constant Mg^{++} values, with fluctuating Ca^{++} and HCO_3^- values, the dominant mechanism may be deposition/solution of CaCO_3 . The act of removing the ground waters for sampling could result in changed conditions that might favor deposition or re-solution of precipitated CaCO_3 .

Analyses. A cation/anion and TDS balance was run on the 11 Round One ground water samples for the Rocky Flats alluvial waters. Results are presented as the percentage of cations to anions excess in the table below; the percent TDS is the percentage of the calculated TDS in excess of the measured value. The "mg/l diff" is the actual difference between the calculated and measured TDS value.

The first columns show the calculations assuming the HCO_3^- as the measured value; the second columns show the balances corrected for the CaCO_3 reported value.

Additional balances were run for a selected sample of surface water quality and deep ground water samples. The deeper ground water samples were chosen to evaluate the balance on higher TDS concentrations (420-1300 mg/l). The surface water quality samples were chosen to look at two sites suspected of contamination (SW-104 and SW-80) versus two relatively clean sites. In addition, the two suspected contamination sites had HCO_3^- values in excess of TDS, as did well sample B400389. Results are found in the tables below:

CATION/ANION AND TDS BALANCE FOR ROCKY FLATS ALLUVIAL WATERS

Well No.	%Cat/An	HCO ₃ as given		HCO ₃ converted from CaCO ₃		
		%TDS	mg/l diff	%Cat/An	%TDS	mg/l diff
B400189	-10.0	-48	-88	-24.0	-44.9	-76
B400289	-5.5	-37	-56	-19.5	-33	-49
B400389*	-34.5	+15.8	+47	-43.1	+31.1	+93
B400489	-4.7	-24.9	-55	-15.9	-18.7	-41
B200589	+8.9	-60.4	-127	-5.4	-57.9	-122
B200689	+3.8	-29.0	-46	-14.2	-18.6	-30
B200789	+1.5	-32.4	-68	-6.5	-25.7	-54
B200889	+1.2	-42.5	-119	-12.	-32.6	-91
B405689	-3.8	-17.3	-43	-13.3	-9.7	-24
B405789	-5.1	-32.5	-56	-17.0	-27.9	-50
B405586	-8.7	-58.3	-152	-21.5	-55.7	-145

CATION/ANION AND TDS BALANCE FOR SURFACE AND DEEP GROUND WATERS

Well No.	%Cat/An	HCO ₃ as given		HCO ₃ converted from CaCO ₃		
		%TDS	mg/l diff	%Cat/An	%TDS	mg/l diff
Surface Waters						
SW-04	-1.5	+1.1	+2	-8.9	+7.9	+14
SW-06	+7.2	-26	-52	-.2	-19.5	-39
SW-104*	-6.4	-7.7	-22	-15.6	+4.2	+12
SW-80*	-22.8	+16.3	+20	-30.3	+30	+36
Deep Unweathered Sandstone Ground Water						
B204089	-6.2	-7.1	-54	-7.8	-6.8	-52
B204189	+1.6	+8.3	+108	+3	+9.0	+117
B204289	-7.8	-1.7	-7	-15.1	+3.5	+15

* Indicates sample where the HCO₃ > gravimetric TDS

Analyses for the 11 alluvial wells showed a significant disparity between cations and anions. Using the unchanged bicarbonate value, results were generally within ± 10%. Since measured TDS values ranged from 150-300 mg/l, this might be a reasonable finding. However, the TDS values were consistently negative (ignoring the B400389 sample).

Using the CaCO₃ conversion for bicarbonate, the cation/anion balance shows a consistent underestimation of cations to anions, with a somewhat lessened TDS balance, although still significant.

Surface water quality wells also showed a consistently negative cation/anion balance with the CaCO₃ transformation, with the samples containing the HCO₃ > TDS, at greater variance. The range of TDS values for the surface water quality stations was 120-290 mg/l. These data behaved very similarly to the alluvial waters, except that the calculated TDS was generally higher than the gravimetric TDS.

As expected, the situation improves in the higher TDS unweathered bedrock samples. The percent differences are much smaller, while the actual TDS differences remain about the same. Well B204189 with the highest TDS showed the best cation/anion balance.

A number of possibilities suggest themselves here. Dissolved silica up to as high as 20-80 mg/l might explain some of the difference in TDS (generally underestimated in the calculated TDS). The fairly high detection limits for potassium and magnesium (<5 mg/l), might contribute to the underestimation of cations. At 2.5 mg/l each for K⁺ and Mg⁺⁺, for example, the cation/anion balance in alluvial sample B400189 decreases from -10 (-24 for CaCO₃ adjusted) to -7.8% (-14.6%). The corresponding differences are -2.9% (09.9%) if the assumed values are 4 mg/l.

A further possibility, of course, is that most of the difference in the low TDS water quality samples and inaccurate cation/anion balances are due to CaCO_3 precipitation/dissolution in the samples themselves, or an inaccurate measurement of HCO_3^- .

Those three samples which contained a measured HCO_3^- in excess of TDS, had significantly greater variances of cation/anions. This phenomenon could also be due to the CaCO_3 precipitation/dissolution in the sample, though for some reason exacerbated in these locations.

To help resolve the latter problem, we need to know the exact methods used to determine HCO_3^- ion (we assumed alkalimetric titration). The report should contain a detailed description or reference to the exact conditions assumed for this test. EPA also needs to know how the values were calculated for final reporting.

This analysis does underscore the difficulty in using a parameter like HCO_3^- for statistical purposes. At present, there is too much uncertainty in the data to use it for testing purposes.

Recommendations. The methods for analysis of potassium and magnesium should be identified, and if possible, have the detection limits lowered to about 1 mg/l. Occasional measurement of dissolved silica would also identify if it contributes significantly to the overall TDS. The cation/anion balance could also be expanded to evaluate trace metals (primarily Mn), although the contribution would be relatively small.

Response:

In the Draft Background Geochemical Characterization Report, there was a discrepancy in the reporting of HCO_3^- concentrations. Within the text, tabulated data identified reported concentrations as HCO_3^- (as CaCO_3) while tabulated data in the appendices identified reported concentrations as just HCO_3^- . Discussions with the laboratory originating these data have determined that the reported data are in fact concentrations of HCO_3^- . The text will be revised to reflect this. Methods for analysis of HCO_3^- and other major cations and anions are identified in the QA/QC Plan.

As discussed above in General Comment 4, Cation/Anion and TDS balances will be calculated as part of a preliminary data evaluation and discrepancies will be analyzed for cause.

Beginning first quarter 1990, filter silica, filtered alkalinity and unfiltered alkalinity have been evaluated in the field for eleven background wells: B400289, B400489, B200589, B200689, B200789, B200889, B401989, B203789, B405689, B405789, and B405889. Also, beginning second quarter 1990, unfiltered alkalinity is being measured in the field. These data will be included in this revised report, if available. It is not practical to lower the detection limits for potassium and magnesium for the sole purpose of possibly improving the cation/anion balance.

Comment 2 Potential Contamination in Surface Water Stations SW-104 & SW-80

Approach. Although the major inorganic chemistry data for these two sampling locations are somewhat unusual (excess HCO_3^- , elevated TDS), there are other more significant measurements indicative of potential contamination. These include heavy metals, pH, suspended solids, and especially the radiochemical data.

Surface water locations SW-104 and SW-80 differ from the remaining sites in being springs or seeps, rather than stream runoff. The two sites are also in fairly close proximity to each other and drain a common area in the southwestern portion of the Rocky Flats environs.

In order to better view the information, two tables are presented below. The first presents both total and dissolved materials for the metals and selected inorganic information; the second compares the radiochemical data.

In the background study, nine surface water quality stations were chose; at each station, sediment samples were also analyzed. Two rounds of sampling were conducted on the surface water quality, although the second sampling was far less complete than the first (dry wells and missing data).

In order to assemble this information into presentable form, the following procedures were used. The data for seven of the nine stations are aggregated (when available), and presented in comparison to the two surface water quality stations in question (SW-104 and SW-80). The manner of presentation differs somewhat, depending on whether some or all of the reported values were below detection.

The first row shows the detection limit for each analyte; directly below it [identified as (N)], the number of samples which were reported below the detection limit is given. Duplicate sample values were averaged. If at least one value was reported above the detection limit, a maximum value is shown. Where all the data are above detection (shown by an asterisk), the arithmetic mean and sampling standard deviation is presented. Individual values for Stations SW-104 and 80 are placed below the other sampling station data.

The intent here is to present a simple method of comparing divergence of the two station values from the other seven (or six) stations.

A similar approach was used for the radiochemical data. The raw data was presented in the form $X \pm L$; where X is the mean value of the measurement and L represents a confidence interval for the mean value of the individual radiochemical measurements after presumably subtracting out a background correction.

Since radiochemical concentration data cannot be negative, it was assumed that whenever a value of L was greater than or equal to the absolute value of X, the reported value was less than some detectable level L'. The detectable level L' was chosen as the highest L in a data group, where L was greater than the measured value X. Where X was greater than L, the value is reported as X.

This approach proved to work quite well for almost all of the radiochemical data evaluated. Generally, the effective detection limits within a sample set of "non-detects" varied by no more than one to three units. Reportable data, by contrast, was often orders of magnitude higher than the detection limit. This approach is discussed in more detail in the section on statistical approaches.

The seven station radiochemistry data were then assembled in identical fashion to the metals/inorganics data. No means and standard deviations were calculated, since every analyte had at least one below detection value.

The intent here was to use an exploratory approach to the data, without making any assumptions about aggregating the data. Except for the first round of iron samples, the data where all values were above detection appeared to be quite normal (generally a coefficient of variation less than 1.0).

Where there appeared to be significantly different values for the two stations, the values were highlighted. The reasons for such choices will be discussed below.

Evaluation of the Tabular Data. In Table A, 17 less common metals were compared, along with pH and total suspended solids (TSS). The first entry of "Total Metals" represents analysis of the sample surface water without filtration. Two rounds of sampling were performed, but in the second round, one of the seven wells and SW-80 were dry.

The "Dissolved Metals", presented only for the first sampling round, are shown towards the bottom of the page. TDS values replace the TSS values; the pH values were reported only once, and are shown for both total and dissolved metals.

The first general indication that stations SW-104 and SW-80 may be unusual or unrepresentative sites, can be seen in the pH and TSS data for the Round 1 Total Metals Sampling event. The TSS at both stations are more than an order magnitude higher than the remaining stations. pH is significantly lower in SW-80, and somewhat lower for SW-104 than the range of the other stations. These results seem to fit well with increased metals solubility or perhaps the presence of significant amounts of organic colloidal or suspended material.

Since elevated metals are found in the suspended material rather than in the soluble (dissolved metals) form and the pH is somewhat lower than the other areas, the suspended solids might be organic or claylike materials. Another possibility might be the presence of suspended ashlike materials constituting a major fraction of the TSS. It does seem somewhat surprising to see such high TSS

TABLE A. COMPARISON OF ANALYTE DATA FOR NINE SURFACE WATER QUALITY STATIONS
 (Values in milligrams/liter or kilogram, except pH)

LOCATION	PARAMETER	ANALYTE																	pH	TSS	
		Al	As	Ba	Be	Cd	Cs	Cr	Cu	Cu	Fe	Pb	Mn	Mn	Ni	Ag	Sn	Y			
Round 1- Total Metals																					
Seven Stations	Limit (N)	<.2	<.01	<.2	<.005	<.005	<1	<.01	<.05	<.025	*	<.005	<.015	<.1	<.04	<.01	<.1	<.05	*	<5	
	Maximum	1.0										2.3	.375		.011					7.25	18
	X										.49									7.6	
	s										.80									.3	
SW-104	Value	8.4	.011	<.2	<.005	<.005	<1	<.01	<.05	<.025	22.6	.02	.72	<.1	<.04	.017	<.1	<.05		7.1	340
SW-80	Value	64.1	.116	4.5	.01	.07	2.5	.06	.07	.18	651	.23	.48	.20	.25	.15	.97	.36		6.5	4200
Round 2- Total Metals																					
Six Stations (04 DRY)	Limit (N)	<.2	<.01	<.2	<.005	<.005	<1	<.01	<.05	<.025	*	<.01	*	*	<.04	<.01	<.1	<.05			
	Maximum	4.9									5.5										
	.019	.216	.018																		NO DATA
	X										2.3		.13								
	s										2.0		.08								
SW-104	Value	4.0	<.01	.29	<.005	<.005	<1	.012	<.05	<.025	7.9	<.01	.84	.02	<.04	<.01	<.1	<.05			
SW-80	Value	DRY																			
Round 1- Dissolved Metals																					
Seven Stations	Limit (N)	<.2	<.01	<.2	<.005	<.005	<1	<.01	<.05	<.025	<.1	<.005	<.015	<.1	<.04	<.01	<.1	<.05	*	*	
	Maximum	.37										1.3	.006	.35						7.25	200
	X																			7.6	101
	s																			.3	62
SW-104	Value	<.2	<.01	<.2	<.005	<.005	<1	<.01	<.05	<.025	4.7	<.005	.48	<.1	<.04	.013	<.1	<.05		7.1	320
SW-80	Value	<.2	<.01	<.2	<.005	<.005	<1	<.01	<.05	<.025	2.0	<.005	.15	<.1	<.04	<.01	<.1	<.05		6.5	120

NO ROUND 2 DISSOLVED METALS

* Indicates that all values were above detection; X and s calculated.
 (N) below Limit indicates number of samples less than this detection value

values in a spring or seep, unless the drainage to the site was recently disturbed or the water passes through loosely packed unconsolidated materials.

The two stations show a considerable number of detectable trace or unusual metals. Ignoring SW-80 for the moment, SW-104 had the highest value in comparison to the other seven stations for 6 of the 17 total metals; in no case, were the other stations higher than SW-104. The levels of reported metals were in some cases 8-10 times the background or highest reported level in the seven stations.

SW-80 is an even more extreme case. For all seventeen metal analytes, SW-80 had the highest value, compared to the seven stations, and 16 higher than SW-104. In some cases, the values were as much as 1-300 times the concentration of the highest seven-station value. Clearly, both stations were anomalous with regard to metals suspended in the surface water.

The second round to total metals sampling suffered because it did not report inorganics and SW-80 was dry. However, the data for SW-104 tended to reconfirm the results of the first round of sampling. SW-104 still had significantly higher values for 4 of the seventeen metals, although different from the first sampling round but found in SW-80. Values for barium and chromium at SW-104 were the only reportable results for any station. Because pH and TSS were not reported, no comparison was possible.

A look at the dissolved metals picture presents a much different picture. If just dissolved metals data were considered, there would be little or no justification for considering these two stations any different from the other seven. The TDS values are within normal limits for the other stations. Only iron and manganese are somewhat higher, but probably within tolerance limits for these other stations. Apparently then, the metals found in the total metals analysis are associated with the suspended material, and very little in soluble form. It is actually somewhat surprising not to have found somewhat higher dissolved metals in the sample for SW-80, since the pH is significantly lower and favoring dissolution. Either the metals are in a very insoluble form, or else may be preferentially sorbed on the suspended materials.

Table B provides even greater evidence for the anomalous character of stations SW-104 and SW-80. The radiochemical results for the total water analysis showed very significant amounts of gross alpha and beta radiation, plutonium-239, americium-241, cesium-137, radium-226 and radium-228 in SW-80. Somewhat lower but significant results were found in SW-104 for gross alpha and beta, plutonium-239, and radium-226. It does not appear that there is a similar source of contamination for both springs, with SW-80 containing the preponderantly higher concentrations. The results appear to be consistent in that primarily man-made radionuclides are detected, and the overall radioactivity levels are much higher than the other stations.

The second round of total water sampling demonstrated the reliability of the first sampling for SW-104. Unfortunately, no information was available for SW-80. In both samplings, the same analytes were detected at very significant levels for SW-104 – gross alpha, gross beta, plutonium-239, and radium-226. The levels measured are quite uniform between samples.

The dissolved radiochemical results parallel those for the metals. Little if any of the radioactivity is associated with the filtered water. The species shown to be possibly significant in SW-104 – uranium-233/244 and uranium-238 – were not significantly different in the total sample. SW-80 is virtually free of any detectable dissolved radiochemical species.

It would then be expected that since the elevated metals and radioactivity concentrations are associated with the suspended materials, the sediment samples should show some evidence. At this time, only the radiochemistry data for the sediments is presented.

The sediment data resembles the dissolved water quality data in not indicating very significant radioactivity. Only cesium-137 in sediment sample SD-19 (which corresponds to the SW-104 surface water quality station), might be considered significant. However, atmospheric fallout of cesium-137 might account for this finding. At this point, however, the sediment results must be qualified, at least until this question is resolved: where was the sediment sample taken relative to the surface flow?

TABLE B. COMPARISON OF RADIOCHEMISTRY ANALYTE DATA FOR NINE SURFACE WATER QUALITY STATIONS

(Values in picocuries/gram or liter)

LOCATION	PARAMETER	Gross Al	Gross Be	U233	U235	U238	Sr89,90	Plut239	Amer241	Cs137	Tr11	Ra226	Ra228
		ANALYTE											
<u>Total Water Sample-- Round 1</u>													
Seven Stations	Limit	<3	<3	<3	<2	<2	<5	<.02	<.03	<.6	<150	NR	NR
	(N)	6	5	5	7	4	6	7	7	7	6	6	7
	Maximum	6	7.5	.9	.4	1.25					200	<.2	
SM-104	Value	25	.5	<.2	.7	1.1	.08	<.03	<.6	<150	.7	NR	
SM-80	Value	250	200	<.3	<.2	.3	1.1	1.0	.10	12	190	5.5	11
<u>Total Water Sample-- Round 2</u>													
Six Stations (04 DRY)	Limit	<4	<3	<.2	<.1	<.2	<.4	<.01	<.02	<.8	<230	NR	NR
	(N)	6	3	1	6	2	5	5	6	6	5	6	6
	Maximum	7.0	.7	.4	.8	.02					550		
SM-104	Value	78	58	.5	<.1	.3	<.4	.08	<.02	<.8	<230	1.4	NR
SM-80	Value	DRY											NR- Not reported
<u>Dissolved Water Sample-- Round 1</u>													
Seven Stations	Limit	<4	<3	<.3	<.2	<.2	<.5	<.02	<.03	<.6	NR	NR	NR
	(N)	7	4	3	7	3	6	7	7	7	7	7	7
	Maximum	5.5	.8	.6	.95								
SM-104	Value	<4	<3	2.6	.3	1.7	.8	<.02	<.03	<.6	NR	NR	NR
SM-80	Value	<4	<3	<.3	<.2	<.2	<.5	<.02	<.03	<.6	NR	NR	NR
<u>Sediment Total Samples</u>													
Seven Stations	Limit	<13	*	*	<.1	*	<.7	<.02	<.02	<.1	<.15	*	*
	(N)	2	7	7	7	7	6	6	7	6	1	7	7
	Maximum	40	40	1.2	1.3	.8	.08	.4	.32	1.0	2.1		
SD-19	Value	39	29	.9	<.1	1.1	<.7	.04	<.02	1.4	.18	.8	1.0
SD-18	Value	35	23	.7	<.1	.5	<.7	.03	<.02	.4	<.15	.8	1.2

* Indicates that all values were above detection; X and s calculated.

Since the springs may be located on fairly steep slopes, sediment erosion might be more characteristic than sediment deposition.

***Conclusions.** The evidence provided above is a very strong indication that surface water stations SW-104 and SW-80 are not representative of a relatively undisturbed background for the Rocky Flats plant. By relatively undisturbed, we are presuming only human activity unrelated to the major operations of the plant. For example, incidental cattle grazing, farming runoff, etc., properly fit the idea of "background" for this site. However, any past activities associated with the plant (incinerator operations, ash burials, landfills, etc.) should be not included.*

The data is suggestive of some kind of subsurface contamination (or perched ground water contaminated by some surface activity) upstream of the two springs. The springs appear to drain a common higher area between them. It is recommended that the area be investigated further for the possibility of some prior contamination. Aerial photographs may indicate past activity in the vicinity of these stations. It would be useful to have at least one GC/MS organics analysis for each station (along with TOC and/or oil and grease) for common hazardous constituents, to determine if synthetic organics may be playing a role in the absorption of the metals and radionuclides.

From the standpoint of utility of the data from these stations for background characterization, we would recommend their removal from the database. Even in the absence of locating a specific contaminant source, it is difficult to explain the presence of such significant, predominantly man-made radioactive concentrations, in background. The host of metals associated with the suspended material from SW-80 are also very unusual in their concentrations and number. Although detection limits in soils are higher, it is uncommon to see so many detectable trace metals in area soils, as was found in the suspended materials. Clearly, the site is anomalous. It may be wiser to develop other stations, rather than try to fit these erratic data into otherwise rather well-behaved background distributions.

Response:

A map of all SWMUs for the Rocky Flats Plant will be incorporated in the revised text. This map has been based upon a review of aerial photography available from 1953 through 1986. No activity indicative of a SWMU has been detected in the vicinity of SW-80 and SW-104. In the absence of independent evidence of contamination, this area should be deemed background. We recognize many of the reported total metals and total radionuclide analytes for stations SW-104 and SW-80 may be outliers relative to the other surface water data. However, if total metals and radionuclides data is normalized using the suspended solids concentrations, the results for SW-80 and SW-104 are typical of the other surface water stations. Also, the sediment radionuclide concentrations at these stations are typical of the other sediment stations. Lastly, subsequently collected data also show the concentrations of these analytes vary in proportion to the suspended solids concentration. When suspended solids are at levels typical of the other surface water stations, the water quality of these seeps is also typical. Varying suspended solids concentrations may be an artifact of seep sampling, i.e., at very low flows, the station may be physically disturbed during collection of the sample. A Standard Operating Procedure is being developed to address this issue.

In general, the distribution of the data collected at SW-80 and SW-104 will be compared to the distribution of data collected at other surface-water sites. If part or all of the data collected at SW-80 and SW-104 are found anomalous, the anomalous data will not be included as part of the surface-water background population.

One round of VOAs (GC/MS organics analysis) will be collected for all surface water stations in 1990.

Comment 3 Statistical Analyses and Selection of Analytes

Comment 3a

Application of statistical tests to present background data and recommendations for modifications to sampling. The selection of tolerance intervals, test of proportions, mean plus 3 standard deviations, criteria for log-normal transformations, and a test based on a Poisson distribution make sense as an overall conceptual plan approach. However, use of these tests and criteria at present for the limited

data in this report are somewhat premature. In particular, we are concerned about applying such criteria where further station data are to be collected, such as for ground water or surface water stations. With only one or two samples per location, it is difficult to make comparisons between different locations, even within the same sampling strata, until more background data are collected.

As we have discussed above regarding the surface water quality stations SW-104 and SW-80, a preliminary screening of the data within such stratified groups might be a first major step. If there are truly anomalous sampling sites which do not fit the definition of background, these early rounds of data can be used to filter the sampling points. Incidentally, we did not find any other such obvious outliers or anomalies in the other sampled media. What we have done, in essence, is an order-of-magnitude multivariate screening. There is a fundamental assumption that variation in the data is random, or capable of depth or areal stratification if not random. However, obvious outliers do need to be recognized at this time, which include locations where many of the analyte suite are uncharacteristically high.

We see a progression of statistical considerations in time as follows:

- a. Preliminary screening of background data as it becomes available for obvious outliers from analytical error, inappropriate sampling sites, etc.;*
- b. Evaluation of distributions of data after enough data has been collected within a well or surface point (for water quality analyses); evaluation of soil distribution data after enough data has become available to characterize geographically distinct strata (e.g. Rocky Flats alluvium in general), or desired substrata (North Creek Rocky Flats alluvium vs Woman Creek);*
- c. Interwell or interstation water quality comparisons within sampling strata (or geographical distinctions) evaluated with ANOVA or other tests after collection of enough data;*
- d. Decisions about aggregating sampling data for common or distinct background populations;*
- e. Decisions on the kinds of statistical tests to be used, from the proposed suite of tests or others.*

At any stage of this process, feedback to plans for continued collection of background data is important. Certain analytes may be added or dropped. Representative sites might need to be changed. Obviously, the changes should be done as early as possible, so that relatively consistent and useful data sets can be generated. The need for sample size will continue to be a function of the kind of information desired -- e.g., the extent of definition of spatial and temporal variability the sample size(s) necessary to perform certain kinds of statistical tests.

The question of temporal variability is still viable, although it may not be so important in the cases of the inorganic analytes which are generally above detection. The cations and anions, TDS, pH, etc., might well show a definite time variation. The temporal differences could also be overshadowed by spatial variation.

Metals and radiochemistry data present a somewhat different picture, in that values for many of the most important analytes are often below detection (e.g., As, Ba, Be, Cr, Pb, Hg, Cd, Ag, Th, Li, etc.). Even if there is seasonal variation, it will be correspondingly more difficult to define these differences with the presence of any non-detectable values. The number of samples needed to evaluate the tests of proportions or the Poisson test may be considerably greater, for the less-frequently occurring metals.

Finally, the analytes TOC and TOX, although often limited in their ability to finely discriminate instances of contamination, are useful surrogates, especially if there is considerable contamination above background. However, these analytes normally have detectable background concentrations (in the case of TOX, from interferences from non-carbonaceous chloride and other materials); it might be useful to include these analytes in a water quality background sampling scheme. In the case of station SW-80, for example, the TOC and TOX measurements might have been able to indicate if there were either significant carbonaceous and/or chlorinated solvents present in the spring.

It is recommended, that TOC, TOX and GC/MS organics analyses be done as part of the background study. EPA recommends sampling TOC and TOX for all water quality analyses. GC/MS could be done either in a single round once a year, or at a sub-sample of locations. Concurrent sampling generally has the advantage of greater data comparability between sampling sites.

Response:

DOE is committed to the goal of characterizing background and delineating background variability for a range of analytes. We recognize that this is an ongoing process and that the shape of the investigation may need to be modified as more information about background becomes available. VOAs (GC/MS organics analysis) were added to the suite of analytes for ground water samples beginning in January 1990. No VOA data, however, are available for the revised report. TOC and TOX will not be added to the list of analytes for water samples because they are less sensitive and subject to interferences relative to GC/MS analysis, and will never be used in lieu of GC/MS analysis.

Within the context of this report and to the extent possible with the data at hand, the text of this report will be revised, (1) to review formally data quality, including a preliminary screening of outliers, (2) to evaluate statistically the interwell and interstation variation of water quality, and (3) to establish common background populations which are statistically defensible. In addition, the text will be revised to delete the statistical test based on the Poisson distribution.

Temporal variability will be examined with time series plots in the revised text.

Comment 3b

Radiochemistry measurement reporting results and use in this review. As indicated above, the radiochemical results analyzed in this review, transformed that in the form $X \pm L$ to a form where the datum was either \leq an L or a reportable X value. In addition to making the radiochemical data more comparable with trace metals and synthetics data, we also see such a transformation as helping to avoid certain statistical conflicts. Apparently in the report, the mean measurement values were simply analyzed as given, even where negative numbers were involved. EPA has a number of concerns with this approach.

As EPA understands the development of radiochemical data, the statistical confidence limits and estimation of error are based on the Poisson distribution for discrete counting data. On page 2-8, the report indicates that the confidence limit reported is two standard deviations, which is close to a 95% confidence level. However, the principle of the confidence interval around an individual measurement seems to be identical to that for other quantifiable measurement data – can a potential value be distinguished from background with a certain degree of statistical confidence. If a background correction has been subtracted from a nominal measurement, the issue then becomes whether the resulting value is significantly different from zero.

For metals and organic chemistry measurements, the analyte signal must be distinguishable from background noise. Generally, a series of replicate blanks are measured, and their variance calculated. Three standard deviations above zero is often used as a level of statistical confidence that above this level, the measured analyte value is greater than zero.

The $X \pm L$ value seems to serve much the same purpose. This statement indicates that the true measured value M is $X-L \leq M \leq X+L$ at roughly the 95% confidence level. However, if zero is contained within this range of the likely true value for X, one cannot conclude that the value X is significantly different from zero. Presuming that the reported value $X \pm L$ includes both a background and "dead time" correction in the measurements, a report that $X \pm L$ where $X \leq L$, is very similar to saying that a value $X \leq$ some detectable limit near or equal to L.

Although our approach in the analysis above may not be statistically rigorous, it follows generally accepted censoring principles in analytical chemistry. Taking the largest L for a given analyte and characteristic sampling medium data set for which $X \leq L$, sets an effective detection limit comparable to an MDL or PQL. Since Poisson data have a characteristically greater variance as the measured value increases (but which becomes a proportionately smaller relative standard deviation of the mean),

those values of $X \geq L$ can be reported as the best estimate X of M . This is also comparable to GC/MS and other organics data where the variance is proportional to concentration over orders of magnitude.

By treating the radiochemical data in the fashion described above, another statistical problem is avoided. In the statistical analyses for many of the radiochemicals, very large coefficients of variation were observed (3.0-5.5) as, for example, in Table 4-5, even though most of the values were at or below detection. The data were manipulated as if every radiochemical measurement was above detection (including negative values probably a result of the background correction). The inordinately high coefficients of variation, however, are an artifact of the data manipulation, and certainly not an indication of wide variance (by contrast to data which vary over orders of magnitude). Where a mean value can approach zero, there is the possibility of an infinitely high coefficient of variation (witness the Z-distribution itself!). This is certainly no criterion for a log-transformation, however. One runs into the extremely tacky problem that there is no such thing as a logarithm of zero or for a negative number.

For concentration measurement data, as the mean approaches zero, so should the variance. Realistically, no negative measurements are possible. A possible solution is to set a level of detection. Then the radiochemical data can be handled in similar fashion to infrequency occurring metals or organics data - i.e., tests of proportion, Poisson, etc.

Response:

DOE investigated the use of a standard approach to determining the average and standard deviation of data, where each datum is a mean (X_i) and has a variance ($\text{var } X_i$). The approach to calculation of the average is to take a weighted average using the reciprocal of the estimated variance as a weighting factor so that the weighted average and variance is:

$$X = \frac{\sum \frac{X_i}{\text{var}(X_i)}}{\sum \frac{1}{\text{var}(X_i)}}$$

$$\text{var}(X) = \frac{1}{\sum \frac{1}{\text{var}(X_i)}}$$

Because the variance of a datum, with a relatively large magnitude, is generally greater than the variance of a datum with a relatively small magnitude, larger (more accurate) data were weighted less than smaller (less accurate) data. If the variance were a measure of relative error, the method would work. However, DOE did not wish to alter a standard procedure for use in this document, even if the alteration was logical. It was therefore decided that the error term will be ignored for a "first cut" determination of tolerance intervals. This approach is suitable where much of the data is significantly greater than the minimum detectable activity (MDA). This is not the case as it turns out for several radionuclides. Therefore, as you correctly point out, the error term should be factored into determination of the tolerance intervals for radionuclides. Your "censoring" approach to handling the error term for a radionuclide analysis for statistical analysis is a good suggestion.

In the revised text radiochemical results will be reported with an associated minimum detectable activity (MDA). Where a radiochemical result of the form $X \pm L$ is such that $X < \text{MDA}$, that radiochemical result will be evaluated as data censored (undetected) at the value at the associated MDA. For example a radiochemical value of 0.06 ± 0.12 pCi/l with an MDA of 0.09 pC./l would be evaluated as 0.09U pCi/l.

Comment 3c

Use of the population standard deviation. For some reason, when standard deviations were calculated for quantifiable data sets the population standard deviation formula was used. This estimate, which works for very large samples, is biased for smaller samples. The correct, unbiased estimate of the population standard deviation is the sampling standard deviation (using the denominator $n-1$ rather than n).

Response:

Tables and text will be revised to provide the sampling standard deviation.

Comment 3d

Use of the Poisson statistical test. At present it is unclear with which statistical test using a Poisson distribution is proposed. EPA's understanding of the Poisson applications by Gibbons is that very large data sets were used. Where the probability of occurrence of values is less than .1, 50-100 samples might be necessary to characterize a given analyte, if done separately. One of Gibbons' applications of a Poisson distribution was to assume that the frequency of occurrence of 32 compounds done routinely in volatile GC/MS organics analyses at many background wells, behaved as expressions of a single Poisson variable. There is still some question as to how tenable such an assumption is, where certain compounds have a much greater likelihood of occurrence than others. His approach allowed an aggregation of individual compound data; is such a scheme being proposed for the metals/inorganics data collected here?

Response:

Tolerance intervals based on the Poisson distribution were proposed in the Draft Geochemical Characterization Report. As you pointed out, a large number of samples are necessary to adequately characterize the frequency of occurrence. Because of the large number of samples required by this procedure, this statistical test will not be proposed in the revised text.

Comment 3e

Criteria for Normality/Log-Normality or Other Transformations. No single approach for defining the appropriateness of one or another distribution model to a data set is really definitive. It has been shown above that the coefficient of variation approach has definite limits where the data can be zero or less. The frequency with which the coefficient of variation for very small data sets even from a normal population distribution exceeds 1.0 is expected to be much higher than for larger sample sizes. As sample size approaches 15-20 samples, a better assessment of the distributional assumptions is more tenable. Other tests for normality including chi-square, graphical plots, tests of the residuals, etc. can be used.

The importance of the distribution assumption of normality is also a function of how the data will be used, and therefore could change the criteria for deciding upon normality/log-normality. EPA generally stresses that tests like the Student-t or ANOVA are robust with respect to this assumption. This is true, so long as the test is a comparison of means, and not individual values. Where tests of single values versus a background distribution are involved (tolerance limits, control charts, etc.), the test is quite sensitive to the form of the distribution. If a data set having a positively skewed distribution is approximated by normal statistics, extreme upper values will more easily be characterized as exceeding tolerance criteria than if log-normal statistics were used. For these types of comparisons, some check on the symmetry of the data set (kurtosis, skew) would probably be appropriate. At present, however, there is still not enough background data for many of the sets of interest, to make this determination.

Response:

The coefficient of variation will not be proposed as a test of normality in the revised text. Normality will be determined using the Shapiro-Wilk test at the 95% confidence level.

Comment 3f

Soil borehole sampling. It is somewhat puzzling why only a three-foot interval was used at the soil surface for borehole sampling. Most of the important soil characteristic changes occur within the top three feet. Further, deposition of radionuclide or other atmospheric fallout is undoubtedly concentrated in the topmost surficial soils. It should also be explained how the present data can be related to the off-site plutonium study mentioned on page 3-20. Will the off-site study characterizing the surface soils be presented within this report?

Response:

The purpose of the background sampling is to provide data for comparison to site samples. Discrete site samples are collected across a two foot interval because two feet of full recovery of borehole soil is required to provide sufficient volume of materials for a full suite of analyses. Previous drilling experience has determined that full recovery could not always be expected on the two foot interval, and that this recovery often had organic material or large cobbles unsuitable for geochemical analysis. This will be explained in the revised report.

Plutonium and americium deposited at the surface was expected to be below the MDA for the three foot composite. These data show this to be the case. A report on the off-site study characterizing the surface soils will be prepared this summer, but will not be presented in this report.

SECTION 1.0

Comment 1

A detectable concentration of cyanide was reported as .0043 mg/l. This value is not found in the Appendix raw data; all were below detection. Which is correct?

Response:

A detectable concentration of cyanide was reported in Table 1-2 for Round 2 surface-water samples. Raw data for surface-water, Round 2 inorganics was inadvertently not reported in Appendix A-3. This data will be reported showing the datum in question.

Comment 2

The assessment was made that the concentration of Na and Cl decreases as one moves across the site (based on surface water quality). Na and Cl remain unchanged from SW-107 to SW-41; Na increased in Round 2 sampling for these wells. It is probably too early to make such judgments with the limited amount of sampling data.

Response:

Additional sampling data for rounds 3 through 9 will be included in the revised report. The above assessment will be reevaluated.

Comment 3

In making the judgment that surface water is being recharged by alluvial water in the more easterly dissected areas, is this likely to be true for spring areas SW-104 and SW-80, or could these be the result of perched ground water?

Response:

The springs at SW-104 and SW-80 occur at the elevation of the contact of Rocky Flats Alluvium and top of bedrock within the adjacent pediment. At the sampling location, a hillside, the top of bedrock

has been lowered by erosion and overlain by a thin veneer of colluvium. The ground water that surfaces through colluvium at this location originated within Rocky Flats Alluvium.

Comment 4

With enough within-well data (four or more samples), the ANOVA could be used on detectable analytes to determine whether a common population between wells is feasible.

Response:

Currently, the maximum within-well data available is three samples, and the maximum intrastation data available for surface water is 9 samples. Before deciding if either within-well or intrastation data are suitable for ANOVA to determine the feasibility of a common population between sampling sites, seasonal variability within these samples must be evaluated and eliminated.

Comment 5

Multivariate and principal components analysis would be a useful exercise in judging commonality of sampling groups. However, where below-detection data are involved, concentration magnitude needs to be tested in addition to a simple test of proportions (frequency of occurrence data). The preliminary screening does both, but in a more heuristic fashion.

Response:

Multivariate analysis of variance (MANOVA) is a statistical procedure which may provide a basis for evaluating the commonality of proposed background populations. It is not a substitute for preliminary screening which will be incorporated within the study prior to MANOVA. A test of proportions is a statistical procedure for examining the commonality between an already established background population and some other population. This test is proposed only for those instances where the occurrence of censored data in both populations is in excess of 50 percent. This condition precludes the application of other common statistical procedures.

Comment 6

EPA commented on the submittal of the Background Hydrogeochemical Characterization and Monitoring Plan. In those comments EPA raised concern over the delay in evaluating temporal variations in water quality caused by collecting samples quarterly. As it stands, after two years, quarterly temporal averages would be based on only two samples. This is too long to wait for a temporal average based on two samples.

Response:

DOE does not propose more frequent sampling to assess seasonal variability because data will tend to be serially correlated and will not necessarily provide more information. It also provides no information on long-term changes, e.g., a wet or dry year. Most importantly, the fundamental premise of the proposed statistical technique for assessing contaminant releases (tolerance levels) is not affected by seasonal variations, i.e., it is a point-in-time analysis.

Comment 7

In general, the report does not address the concerns raised through the comments submitted to DOE on the background Hydrogeochemical Characterization and Monitoring Plan.

Response:

The comments submitted by the EPA on the Background Hydrogeochemical Characterization and Monitoring Plan were received by DOE on October 11, 1989. By this time sampling sites had been established, and ground water sampling had begun for fourth quarter, 1990.

Concerns identified within those comments which are applicable to the Background Geochemical Report have in general been addressed by the draft report or by the responses to similar comments directly addressed by the EPA to the Background Geochemical Report in this document.

Comment 8

The analysis of background data and calculation of tolerance intervals must not include anomalous or suspicious data points when evaluating the mean and standard deviation. (Even though tolerance intervals would be large for strata and analytes containing few samples, the tolerance interval could be presented prefaced with the knowledge that these intervals would change. Tolerance intervals should be calculated for alpha, beta, Sr 89, 90, and Cs 137 in unweathered sandstone ground water.)

Response:

Prior to calculations of means, standard deviations, and tolerance intervals, background populations will be examined for normality and for outliers.

Tolerance intervals can be calculated for data sets with as little as three values. As described in Section 2.4, however, tolerance intervals based on very few samples may not be sensitive enough to detect contamination.

Comment 9

It seems that no effort was made to follow the "outlier" determination process defined in the plan (i.e. Dixon's or Rosner's tests). As a result, maximum detected levels are presented when in fact these concentrations may be outliers (nitrate in valley fill alluvium, cyanide in Rocky Flats alluvium, many metals in the Rocky Flats alluvium, metals in the valley fill alluvium, and metals in the unweathered sandstone).

Response:

For the draft report, we felt exclusion of outliers was premature with the limited data set. We do note, however, that the assessment of contamination in the Phase II RFI/RIFS Work Plan for Operating Unit 2 excluded surface water stations, SW-80 and SW-104, to be conservative. Background populations will be examined with Dixon's, Rosner's or some other appropriate outlier test in the revised report. We further note that outlier tests are not appropriate where over half the data is censored.

Comment 10

In accordance with Section 3.3.2.6 of the plan, the Cohen procedure is to be followed when many of the analytical results are below detection limit. This has not been done for some of the information presented as tolerance limits (nitrate in valley fill alluvium, cyanide in rocky flats alluvium, many metals in the rocky flats alluvium, metals in the valley fill alluvium, and metals in the unweathered sandstone, background surface water data, background sediment data, background borehole data).

Response:

Cohen's procedure was not applied to analytes which were reported with greater than 50 percent censored data. This will be made clear in the text.

Comment 11

If maximum detected levels are provided where there is insufficient data and the number of samples for a specific strata is low by default (i.e. dry holes resulting in 2 samples for colluvium, 2 samples for weathered sandstone), then additional wells should be completed to prevent systematic non-determination of tolerance intervals. Section 2 delineates the logic and methods for determination of upper tolerance limits. The utilization of maximum detected levels is not presented as an option to the determination of tolerance limits. Does the logic and methodology for determination of tolerance limits not address all the possible data deficiency scenarios?

Response:

Maximum detected levels were presented in Tables 1-1, 1-2, 1-3, and 1-4 for reference only and were not presented as an alternative to the use of tolerance intervals. Alternatives to the determination of tolerance limits include ANOVA and test of proportions, and the criteria for the use of these statistical tests are developed in Section 2. The lack of water in colluvium and weathered sandstone is indicative of a general hydrologic condition in these units. Drilling additional wells adjacent to "wet" wells does not adequately characterize spatial variability, and a major drilling program to locate saturated conditions does not seem appropriate at this time. More recent data indicate at least three wells within a unit contain water for a given quarter, and MANOVA will be used to ascertain if colluvial or weathered sandstone ground water is part of a larger ground-water population.

Comment 12

How is the tritium upper tolerance limit calculated for unweathered sandstone?

Response:

A tritium upper tolerance limit should not have been presented for unweathered sandstone ground water because the number of samples for which results were reported was less than seven. This upper tolerance limit will be deleted from the text.

Comment 13

How can the tolerance intervals presented contain more significant figures than the analyses?

Response:

This is a computer format error. Tolerance intervals will no longer be presented with more significant figures than the analyses.

Comment 14

Asterisks should be placed after all of the radionuclide upper tolerance limit data for ground water and soils data in weathered claystone.

Response:

Asterisks indicate that the reported value is not a tolerance limit but rather a maximum reported value. All ground-water values for weathered claystone should have been asterisked because only maximum reported values were tabulated in Table 1-1. No maximum reported values for radionuclides were tabulated in Table 1-4 because the total number of samples (17) allowed the computation of tolerance limits for every radionuclide for which reported values were available. This should change, however, after the re-evaluation of radionuclide data per response to Specific Comment 3b.

Comment 15

The results of the determination of upper tolerance limits for "other" parameters in surface water should be presented as dissolved, not as total. Why were dissolved analyses for radionuclides not analyzed for Round 2 of the surface water samples?

Response:

The analytes listed under "other" in Table 2-1 (TDS, CO₃²⁻, HCO₃⁻, Cl⁻, SO₄²⁻, NO₃⁻, CN⁻) were determined from unfiltered samples; consequently, they are presented under total. Dissolved analyses for Round 2 or subsequent rounds were not performed because only total concentrations are relevant to risk assessments. Furthermore, radionuclide standards are established for total, not dissolved, concentrations.

Comment 16

A tolerance interval upper limits should be presented for radionuclide analyses of field blanks associated with analysis of ground water, surface water, sediments and soils. This is important as some of the man-made radionuclide analyses are very similar to the field blank analyses. Yet without this comparison, it might appear that the man-made radionuclides are present in background samples, when the presence of these radionuclides at depth, in soils or as dissolved constituents in background ground water, is unlikely, and probably the result of the random nature of the measurement as indicated by the field blank results. Greater numbers of field blanks should also be analyzed for radionuclide water samples in order to more accurately calculate the upper tolerance interval for field blanks.

Response:

Radionuclide data will be reported with an associated minimum detectable activity (MDA). Reported radionuclide values below the associated MDA will be evaluated as censored data with a detection limit equal to the MDA. This procedure should eliminate the need to determine the random nature of measurement through a review of field blank results. Field blank results will, however, be reviewed to evaluate sampling quality control.

SECTION 2.2

Comment 1

EPA has not been informed of continuing on-site disposal of hazardous and radioactive wastes. DOE must identify the presently utilized disposal sites to both the State of Colorado and EPA.

Response:

The subject phrase in the document is a misstatement. Continuing on-site disposal of hazardous and radioactive wastes is not occurring at Rocky Flats. The text will be revised to reflect this fact.

Comment 2

The lack of sufficient background characterization data has not been the result of aggressive investigation/clean-up schedules.

Response:

Because of the aggressive schedules, the need for a comprehensive background characterization program was not evident until significant data assessment was underway. Plans were prepared and implemented in a timely manner once this was recognized.

SECTION 2.3

Comment 1

Figure 2-1 should be revised to coincide with the narrative of this section. It is unclear what the second paragraph of this section is intended to describe. What is the comparison of background to non-background? How does the use of Cohen's test fit into this methodology when non-detects are present in the data population? If the methods presented are to be used for evaluation of release from a past disposal site, the second paragraph of this section appears to present a method for evaluation of release from a site which is different from that proposed within the Background Hydrogeochemical Characterization and Monitoring Plan. If this is the case, DOE must delineate these changes as addendum to the plan initially presented and submit these addendum to CDH and EPA for formal review and approval.

Response:

Figure 2-1 will be revised to coincide with the text.

The second paragraph will be clarified. The objective of this paragraph is to establish a method to determine where chemical results exceed expected background variability. These procedures could assist in the delineation of a plume. This paragraph does not propose an alternative method of evaluating a release from a site.

Cohen's procedure as described in Section 2.4 provides a method for the calculation of a mean and standard deviation for data sets which include censored data. This procedure requires at least 50 percent detects.

SECTION 2.4

Comment 1

It is very difficult to determine if the procedures outlined within Section 2 are followed. The procedures concerning presentation of tolerance limit data versus means plus three standard deviations do not address the presentation of maximum detected levels in Section 1.

Response:

The presentation of maximum detected levels in the Tables in Section 1 were for reference only. Also presented for reference only were the mean plus three standard deviations.

SECTION 3.0

Comment 1

Reference is made within this section of the report to 1989 documents prepared by DOE and EG&G pertaining to QA/QC plans and Standard Operating Procedures. These documents have never been submitted to EPA. As such, complete review of this report is not possible.

Response:

Copies of the QA/QC plan and SOPs are being revised and will be submitted in accordance with the IAG Schedule.

SECTION 4.0

Comment 1

It is indicated that the Appendices show whether data was collected for specific analyte groups. In Appendix A-3, inorganics data for the Round 2 surface water quality stations are shown as received. However, the data were not presented in the report.

Response:

Inorganics data for the Round 2 surface-water quality stations as well as additional data for all parameters for Rounds 3 through 9 will be included in the appendices.

SECTION 4.1.1

Comment 1

To incorporate negative numbers of radionuclide activity into calculation of the mean leads to problems when attempting to calculate the tolerance interval, the standard deviation and coefficient of variation. Incorporation of negative numbers into a calculation of the mean forces the mean to become a small number. This then may result in a coefficient of variation which is larger than one. The resulting determination of normality is then misleading and may lead one to erroneously conclude that a population is log normal. In addition, it is impossible to have a log normal population of negative numbers.

Response:

Radiochemical results will be evaluated per the response to Specific Comment 3b.

Comment 2

The number of samples utilized to calculate the background results must not systematically provide for large standard deviations. The background study must be modified to include more wells so as to be able to calculate tolerance intervals for those strata for which there is presently a lack of data points within a population. To attempt to compare results of the colluvium and rocky flats alluvium when the number of data points in the colluvium has systematically provided inordinately large standard deviations, and resulting means plus three standard deviations, is misleading.

Response:

See our response to Section 1: Comment 11.

For Rocky Flats Alluvium ground water, eleven samples have been analyzed for every analyte except Ra 226. Eleven is a sufficient number of samples to calculate tolerance intervals provided at least six of these samples have analyses above the detection limit. Where the proportion of censored (undetected) data exceeds 50% but is less than 90%, the appropriate statistical procedure is test of proportions per Statistical Analysis of Ground-Water Monitoring Data at RCRA Facilities: Interim Final Guidance (EPA, 1989). Increasing the number of sampling sites (wells) should not modify the observed ratio of non-detected to total data.

Comparisons between sample groups of different sample size will be restricted to comparisons of sample means.

SECTION 4.1.3

Comment 1

The coefficients of variation for the radionuclide data that leads one to postulate the presence of a lognormal population are most likely due to the inclusion of negative radionuclide analyses.

Response:

The coefficient of variation will no longer be used as an indicator of normality, and radiochemical results will be evaluated per the response to Specific Comment 3b. Normality will be reevaluated using the Shapiro-Wilk test at the 95% confidence level.

SECTION 4.1.4

Comment 1

Table 4-13 reports the mean lithium concentration as higher than the maximum detected value. This seems to be the result of the maximum reported detection limit being greater than the detected levels. Similarly, the table reports the coefficient of variation for zinc as potentially indicating a lognormal distribution, when this is most likely an artifact of the use of the low detection limit in the Cohen's estimate of the mean.

Response:

Cohen's procedure will be limited to data bases where the detection levels are less than or equal to the least reported value. The coefficient of variation will no longer be used as an indicator of normality. Normality will be reevaluated using a normal probability plot.

Comment 2

Table 4-14 presents a coefficient of variation for cyanide of 1.66. This is probably the result of utilizing the Cohen's method for redetermination of the mean when the detection limit is significantly lower than the detected values.

Response:

The coefficient of variation will no longer be used as an indicator of normality. Normality will be reevaluated using a normal probability plot.

Comment 3

When negative results are presented and utilized in calculating the means for radionuclide analyses, the coefficients of variation are misleading and will not provide meaningful information related to the normality of the distribution.

Response:

See response to Section 4.1.3: Comment 1.

Comment 4

It is premature to determine that weathered claystone and weathered sandstone ground water are of the same population. This determination should be reserved for when enough data is present to substantiate this statement.

Response:

The determination that weathered claystone and weathered sandstone ground water are of the same population will only be made, if indicated by appropriate statistical analysis, such as MANOVA.

SECTION 4.1.5

Comment 1

The coefficients of variation presented for chloride, nitrate and sulfate in unweathered sandstone ground water indicate that lognormal distributions may be appropriate. It is more likely that these coefficients of variation are the result of the use of Cohen's method in estimating the revised mean sites. There is also a distinct possibility of zonation of more saline portions of the unweathered sandstone that might be identified by further geologic borings/wells.

Response:

The coefficient of variation will no longer be used as an indicator of normality. Normality will be reevaluated using a normal probability plot. Unweathered sandstone ground-water samples will be examined to determine if TDS increases with depth.

Comment 2

The negative radionuclide analyses are again the reason for high coefficients of variation presented for unweathered sandstone ground water.

Response:

See response to Section 4.1.3: Comment 1.

SECTION 4.2.1

Comment 1

Of the eleven analytes for which a lognormal distribution may be appropriate, many are populations where there were non-detects requiring the use of Cohen's approximation of the revised mean. There is a definite correlation between the use of this approximation and the finding of lognormality. There is also a systematic error in this lognormality determination when evaluated data distributed around zero (i.e. the radionuclide data).

Response:

See response to Section 4.1.3: Comment 1.

Comment 2

Two of the surface water stations, SW-80 and SW-104, appear to be outliers, yet no outlier determinations have been performed on these data.

Response:

See our response to Section 1: Comment 9.

After the determination of appropriate background populations, outlier tests will be performed for those analytes with at least 50 percent detects.

SECTION 4.3

Comment 1

Again, the coefficient of variation does not appear to be meaningful indicator of normality when the distribution is centered around zero or when an estimate of the mean is developed using Cohen's approximation.

Response:

See response to Section 4.1.3: Comment 1.

SECTION 4.4

Comment 1

The coefficient of variation does not appear to be a meaningful indicator of normality when the distribution is centered around zero or when an estimate of the mean is developed using Cohen's approximation. The coefficients of variation for As, Be, K, and the radionuclide data in rocky flats alluvium appear to be affected by these variables.

Response:

See response to Section 4.1.3: Comment 1.

APPENDIX A-2

Comment 1

The field pH for the first round water quality sampling showed values much lower than the laboratory values for all stations. Could this have been due to temperature differences?

Response:

It is widely accepted that pH is an unstable measured parameter. The discrepancies between field measured and laboratory measured pH for surface-water and ground-water samples are probably a result of gas exchange and an increase in temperature. Laboratory measurements were made at 25°C, a temperature approximately 15 to 24°C warmer than that of the field measurements.

Field instruments have been upgraded to include temperature correction for pH in the field at the time of sampling.