Radionuclide Data Evaluation Guidance – Implementation and Lessons Learned – 10258

Brian J. Tucker*, David Evans**
*Shaw Environmental & Infrastructure, Inc., Salem, NH 03079
**U.S. Army Engineer District, Kansas City, MO 64106

ABSTRACT

In 2009, the author submitted a paper which discussed the new Radionuclide Data Evaluation Guidance [1] (the Guidance) based upon MARLAP. Several quality control (QC) criteria were evaluated and many of the control limits depend upon the required and relative method uncertainties at the action level for a given radionuclide of concern. In this paper, the implementation of the Guidance is discussed, including the challenges, required modifications and lessons learned. The paper will discuss possible reasons why some limits may be too narrow or broad, and the validity, or lack thereof, of making changes to these limits. The effect of bias and its effect on the selection of the theoretical value or mean value for certain parameters shall also be discussed. The Guidance is implemented using an interactive project database that allows data generators and data users to accept and qualify data in real time and to review the data.

INTRODUCTION

This document is a follow-up to last year's paper, Radionuclide Data Quality Evaluation Guidance [1], which provided guidance for evaluation of radionuclide laboratory data quality based upon the Multi-Agency Radiological Laboratory Analytical Protocols Manual (MARLAP)[2]. The scope of that document included discussions of the quality criteria, corrective actions, and data qualifications that should be used or applied during the data acquisition, verification and validation processes. Many of the data quality criteria are based upon the required measurement uncertainty, μ, which is the uncertainty or variability in a measurement at the action level of a given radionuclide, or the relative method uncertainty, φ, which is used when results are greater than or equal to the action level. By linking the quality control acceptance criteria to the variability at the action level, the quality control (QC) indicators used in a measurement process become more meaningful and defensible.

The USACE FUSRAP Maywood Laboratory (UFML) has used the Radionuclide Data Evaluation Guidance [1] since January 2009. UFML generates data with a variety of methods utilizing gamma spectrometry, alpha spectrometry, and gas proportional detection. They qualify the data using qualification forms on the project's database, which contain the acceptance criteria established by the Guidance. This year's paper will discuss the lessons learned from employment of the MARLAP-based Guidance, specifically the recommended changes to the criteria for some of the following QC parameters: Calibration, Blank Analysis, Tracer Recoveries, Laboratory Control Sample Analysis, Matrix Spike Sample Analysis, Standards and Reagents, Laboratory Replicate Analysis, Field Duplicate Analysis, Spectrometry Resolution, Radionuclide Quantitation and Detection Limits, and Matrix Density.

For the radiological data evaluation guidance prepared last year, we followed the MARLAP guidance for the batch QC parameters such as method blank, laboratory control standard, matrix spike and laboratory replicate, all of which were described in detail with examples in MARLAP. We chose MARLAP because it addresses the need for a nationally consistent approach to producing radioanalytical laboratory data that meet a project's or program's data requirements. For instrument QC such as daily blanks, weekly backgrounds, and some calibration parameters, as well as tracer recoveries, we used more traditional statistical control limits such as the mean ± 3 sigma. Finally, for several calibration parameters such as full width at half-maximum (FWHM), energy, and activity, we used default values suggested by the manufacturer or used historically. This paper will address the merits and drawbacks associated with each category.

BATCH QC – MARLAP APPROACH

Results obtained from batch QC parameters are most closely tied to the results obtained for field samples in the same batch since the batch QC samples are typically prepared and analyzed in the same manner as the field samples. We used MARLAP-derived QC control limit guidelines for the batch QC parameters. The following discussion provides some perspective on the value and limitations of the MARLAP guidelines, and in some cases provides suggestions for modifications to existing QC acceptance criteria based upon results obtained over the past year. The control limits for each of these parameters are directly or indirectly proportional to the required method uncertainty, μ, or the relative method uncertainty, φ. In some cases, it would be helpful to also consider how close sample results, with which these QC parameters are associated, are to their action levels when considering qualification of the data.

Method Blanks

The method blank control limit criterion, ±3μ, makes sense as long as the required method uncertainty, μ, is not too large relative to the action level for a given analyte. For example, the sum of the activities of Ra-226 and Ra-228 has a fairly low action level of 5 pCi/L. The ±3μ limits for Ra-226 and Ra-228 for data generated at UFML are ±0.825 pCi/L and ±0.86 pCi/L, respectively. Thus, method blank levels of 0.7 pCi/L each for Ra-226 and Ra-228 are within control limits. If a sample associated with this method blank was found to have activities of 2.7 pCi/L each for Ra-226 and Ra-228, the sum of the two, 5.4 pCi/L, would indicate an exceedance. The conclusion that it is an exceedance could be a false one since a significant percentage ((1.4/5.4) X 100, or 26%) of the sample activity may be due to the blank.
Recommendation: The ±3μ control limit should have a ceiling value equal to some percentage of the action level; e.g., 20% of the action level if it applies to a single analyte; 10% of each action level if it applies to the sum of two analytes.

Laboratory Control Sample (LCS)

The range of the established LCS limits illustrates the degree of bias associated with a measurement. For example, if the control limits for analyte X are -25% to +5%, the mean %D is -10% and therefore indicates a negative bias of 10%. Bias is a direct measure of the inaccuracy of a measurement. Thus, an upper limit should be placed on the amount of bias that will be tolerated before the lab has to improve method performance by modifying the method, or employing a new method altogether.
As far as the formula for the LCS control limits, no change is recommended. That formula is:

$$\%D = \frac{SSR - SA}{SA} x100 \qquad \text{(Eq. 1)}$$

Where:

%D	is the percent deviation
SSR	is the measured result (spiked result)
SA	is the spike activity (or concentration) added.

The author reinforces the requirement, however, that the LCS spiking level be as close to the action level as possible so that the uncertainty in the measurement, which is used to establish control limits, is representative of the uncertainty at the action level.
Recommendation: Establish an upper limit for the bias associated with LCS control limits

Matrix Spike Sample

The matrix spike sample is only employed for gross alpha (GA) and gross beta (GB) measurements at UFML. The MARLAP guidance provides matrix spike control limits of ±3 for the Z score, which is expressed as:

$$\text{Z-Score} = \frac{SSR - SR - SA}{\phi_{MR}\sqrt{SSR^2 + \max(SR, UBGR)^2}} \qquad \text{(Eq. 2)}$$

Where:
max (x,y) denotes the maximum of x and y,
φ_{MR} is the maximum allowable (relative) standard deviation at the UBGR
SSR is the spiked sample result, in pCi
SR is the sample result, in pCi
SA is the activity spiked into the sample, in pCi
UBGR = 15 pCi/L X (sample volume in L) for GA and 50 pCi/L X (sample volume in L) for GB

The Z score is a good acceptance criterion unless data is very precise and therefore has a relatively small φ_{MR} value. A small φ_{MR} value leads to larger Z scores, and it is not unusual to obtain failing Z scores for reasonably good spike recoveries such as 80% to 85%.

Recommendation: If the Z score is <-3 or >3, check first for any obvious problems with the sample preparation and/or measurement. If there are no obvious problems, check the %R. If the %R is between 80% and 120% and the activities of all associated samples in the batch are less than 10 pCi/L each for GA and/or less than 30 pCi/L each for GB, the MS should be deemed acceptable.

Laboratory Replicates and Field Replicates

MARLAP provides "Duplicate Analysis" criteria in Section C.4.2.2 of Appendix C and does not distinguish between laboratory replicate and field replicate criteria. The control limits provided for replicate results are 4.24μ for the absolute difference between the two results when the average of the two results is less than the action level; and 100 X 4.24 φ for the relative absolute difference between the two results when the average of the two results is greater than the action level. The replicate criteria for laboratory replicates work best if the sample that is selected as the laboratory replicate has activity levels for the radionuclides of interest that fall near their respective action levels, especially if the change in uncertainty with activity is significant. At UFML, many of the gamma spectroscopy soil results are less than the action levels of the three radionuclides of concern. However, the soils are dried and ground so that the samples are well blended and therefore relatively homogeneous. The homogeneity improves the precision of the measurement such that the acceptance criteria are usually met. For water samples, the analytes are soluble and are therefore distributed homogeneously also leading to very good precision between laboratory replicate samples. Field replicate soil samples are also homogenized in the field, however the homogenization process is not as thorough as that for laboratory replicates. Therefore, the precision of field replicate results would not be expected to be quite as good as that of laboratory replicates. Nevertheless, since they are both homogenized, the same criteria are applied to both laboratory replicates and field replicates.

Replicate criteria may not work well if the relative uncertainty is high and the action level is low. For example, for Ra-226 and Ra-228 in water, the action level requirement is that the sum of the two activities must be less than 5 pCi/L. The Maywood lab replicate control limits for Ra-226 and Ra-228, for the condition of the activities being less than the action level, are 1.17 pCi/L and 1.93 pCi/L, respectively, for the absolute difference of the activities of a given replicate pair. If the regular sample activities are 2 pCi/L and 1.5 pCi/L, respectively for Ra-226 and Ra-228, and the replicate activities are 3 pCi/L and 3 pCi/L, the replicate precision criteria are met even though the precision is not good enough to determine whether or not the action level is exceeded (the sum of the regular sample activities are less than the action level, at 3.5 pCi/L, while the sum of the replicate sample activities are greater than the action level, at 6.0 pCi/L. In a case such as this, the task manager would likely require additional analyses to improve the overall precision and allow for a more informed decision.

Recommendation: No changes recommended.

INSTRUMENT QC AND TRACER RECOVERIES

Daily blanks are analyzed for gamma spectrometry and gas proportional detector (GPD) analyses only. Weekly backgrounds are analyzed for gamma spectrometry and GPD parameters; monthly backgrounds are analyzed for alpha spectrometry test methods. Presently, the weekly backgrounds for gamma spectrometry and gas proportional detectors are long counts (typically 1000 minutes), while daily blanks are short counts of 30 minutes. Both counts use an empty container (Marinelli or tunacan for gamma, and planchet for GPD). The daily blanks represent a check on the weekly backgrounds, so both counts are performed with an empty container.

A tracer compound is typically an isotope of the same element as the isotope of interest, a known amount of which is added to every field and QC sample. Since the chemical behavior of the tracer is identical to that of the isotope of interest, any method anomalies such as analyte losses, interferences, etc. that affect the isotope of interest also affect the tracer. Thus the tracer recovery provides an indication of how the activity of the analyte of interest may have changed due to method anomalies. Tracers are only used for alpha spectrometry parameters.

For all of these QC parameters, the Guidance recommends an acceptance criterion of the mean ± 3σ. This criterion is not based upon MARLAP since MARLAP does not provide specific guidance for daily blanks, weekly backgrounds, and tracer recoveries. The mean ± 3σ criterion assumes that the data of interest is Gaussian, which for some of the analytes is not the case, as explained in greater detail below.

Spectrometry resolution for gamma spectrometers and alpha spectrometers is discussed briefly in this section.

Gamma Spectrometry

Plots of daily blank activity versus the frequency of occurrence of that activity, and weekly background activity versus the frequency of occurrence of that activity, were prepared for gamma spectrometry detector-analyte combinations for three detectors and three analytes. A discussion of the findings is summarized in Table I below.

Table I. Description of Gamma Spectrometry Daily Blank and Weekly Background Data Distributions for Ac-228, Pb-214, and Th-234

Analyte	Daily Blank	Weekly Background
Ac-228	approximately Gaussian with some degree of binomial distribution (secondary peak in the negative region)	approximately Gaussian
Pb-214	approximately Gaussian	approximately Gaussian for detectors 1 and 2; slight binomial distribution for detector 3
Th-234	Peak maximum at low activity (0.03 to 0.12 pCi/g) with frequency tailing to near zero (between 0.3 to 0.5 pCi/g)	For detectors 1 and 2, broad somewhat Gaussian peaks with small maxima superimposed on each Gaussian shape; for detector 3, the plot shows a series of successively smaller maxima from 0.02 to 0.085 cpm

For gamma spectrometry resolution, the Guidance states that the target radionuclide energy should be within 2 keV of the observed peak. This criterion does not apply for isotopes present at a value less than the MDA. No changes are recommended to this criterion.

Gas Proportional Detectors

Plots of daily blank activity versus the frequency of occurrence of that activity, and weekly background activity versus the frequency of occurrence of that activity, were prepared for gas proportional detector-analyte combinations for gross alpha and gross beta on each of eight detectors. A discussion of the findings is summarized in Table II below.

Table II. Description of Gas Proportional Detector Daily Blank and Weekly Background Data Distributions for Gross Alpha and Gross Beta

Analyte	Daily Blank	Weekly Background
Gross Alpha	Distinctive binomial distribution for all detectors; first maximum between 0.0 and 0.02 cpm and the second maximum between 0.45 and 0.5 cpm	Approximately Gaussian
Gross Beta	Slight binomial distribution with the primary peak between 0.3 and 0.7 cpm and a smaller secondary peak between 0.75 and 0.95 cpm for bank 1 detectors; the position and relative intensities of the primary and secondary peaks varies between detector banks.	Somewhat Gaussian peaks with small maxima superimposed on each Gaussian shape

For the gamma spectrometry and gas proportional detector daily blanks and weekly backgrounds, the plotted data may be used to:

- Establish more accurate control limits;
- Identify data trends for a given detector that may indicate a detector performance problem, and
- Identify unique data distributions that may be typical for a particular instrument detector-analyte combination.

Alpha Spectrometry Detectors

Due to the very low background count rates for alpha spectrometer detectors, an absolute total count value of six counts is used as a control limit for each detector for the monthly background measurements. Six counts in 12 hours, or 720 minutes, is a count rate of 0.0083 counts per minute (cpm), which is one to two orders of magnitude less than the not-to-exceed MDA values for all alpha spectroscopy analytes.

For tracer recoveries, plots of tracer activity versus the number of points within activity data ranges were prepared. One plot was prepared for each radiotracer and so included results for all detectors. The total number of points ranged from 105 to 138. A discussion of the findings is summarized in Table III below.

Table III. Description of Alpha Spectrometry Tracer Activity Data Distributions for Ba-133, U-232 and Th-229

Tracer	**Number of Points**	**Description of Tracer Activity Plot**
Th-229	105	Non-parametric; primary peak at 21.8 pCi/L with smaller peaks at 19, 20.2, and 23.2 pCi/L
U-232	121	Non-parametric; primary peak at 20.5 pCi/L with smaller peaks at 21.8 and 22.8 pCi/L
Ba-133	138	Non-parametric; broad primary peak between 685 pCi and 740 pCi; shoulder peaks at 610 pCi and 775 pCi

For alpha spectrometry resolution, the Guidance states that the target radionuclide energy should be within 40 keV of the observed peak. This criterion does not apply for isotopes present at a value less than the MDA. No changes are recommended to this criterion.

Recommendation: For the non-parametric data plots, investigate non-parametric distributions that will provide both a best fit for a given set of data, as well as revised control limits.

QUALITY CONTROL FOR CALIBRATIONS

Gamma Spectrometry Initial Energies & Efficiencies Calibration (MARLAP, Chapter Sections 15.2, 15.6 and 18.5.6)

A full energy and shape calibration, Peak-to-Compton ratio calibration, and efficiency calibration, the three of which are called an initial calibration, is performed annually and after hardware replacement or significant instrument repairs. The initial calibration is performed with a NIST-traceable mixed gamma standard, typically a radioactive source with 9 to 12 radionuclides having gamma ray emissions at energies from approximately 60 keV to 2000 keV. At this time, only one initial calibration data set, performed in January 2009, has been generated under the new guidance. While all QC acceptance criteria were met, not enough data has been generated to determine whether the acceptance criteria should be modified.

Gamma Spectrometry Continuing Calibration

Continuing calibrations are performed daily, or every day that a gamma spectroscopy detector analyzes samples and/or standards. The same NIST-traceable mixed gamma standard employed for the initial calibration, is used for continuing calibrations. The following limits, described in the Guidance, were either used historically or adopted as per the recommendation of the gamma spectrometer instrument manufacturer. The activity of each radioisotope in the calibration standard must be within 10% relative of the true, decay-corrected activity. The energy of each isotope must be within 1% of the true energy. The FWHM value is monitored and must be less than 3% of the observed energy for one low energy, one mid-energy, and one high energy calibration radionuclide. These QC parameter limits are discussed below relative to the calibration data collected in 2009.

Energy

Our experience is that the energy difference in percent ([(true energy – observed energy)/true energy] X 100) is both radionuclide-specific and detector-specific, and is always well below 1.0%, which is the not-to-exceed value provided in the Guidance. The pertinent question is: what is a reasonable upper limit for the energy difference that will maintain the system's ability to identify the peaks of interest and minimize the probability of a non-target peak being detected as a false positive? The existing limit achieves this goal; however the energy difference, true energy – observed energy, could change dramatically and still be within limits. New energy percent difference (%D) limits are therefore proposed in Table IV.

Table IV. Proposed Energy Percent Difference Limits

Calibration Isotope	Energy %D Limit	Calibration Isotope	Energy %D Limit
Am-241	0.3	**Sn-113**	0.1
Cd-109	0.15	**Cs-137**	0.1
Co-57	0.2	**Co-60**	0.1
Ce-139	0.1	**Y-88**	0.1

FWHM

To ensure proper peak resolution, the Guidance also indicates that the Full Width Half Maximum (FWHM) value must be less than 3% of the observed energy. The FWHM values in units of keV increase linearly as the energy of the radionuclide's primary gamma photon increases. The use of a default value in percent helps to reduce some of this difference, however, FWHM values, expressed as a percentage of the observed energy, decrease dramatically as one moves from low to high energies as shown in the table below. The laboratory has generated 150-200 points using the MARLAP guidance. The range of values obtained on 3 detectors for the 3 calibration radionuclides (Am-241, Cs-137, and Co-60) in both keV and percent of the energy of the radionuclide, are shown below in Table V.

Table V. Range of FWHM Values Obtained for Three Gamma Spectrometry Calibration Radionuclides

	FWHM Range	
Calibration Radionuclide	**keV**	**Percent of Radionuclide Energy**
Am-241	0.8 – 1.2	1.4% - 2.0%
Cs-137	1.26 – 1.6	0.19% - 0.24%
Co-60	1.6 – 2.27	0.12% - 0.17%

These results suggest that the 3% default value currently being used is too lenient, and that there should be specific not-to-exceed limits for each radionuclide. The author recommends 2.0%, 0.3%, and 0.2%, respectively, for Am-241, Cs-137, and Co-60. Such limits are slightly higher than the historical range of values and would therefore trigger a qualification or corrective action if a FWHM value fell outside its limit, which would indicate an abnormal instrument condition.

Activity

Based upon the percent recovery values for the activities of the calibration radionuclides, the acceptance criterion of 90-110% is retained.

Alpha Spectrometry Initial Energies & Efficiencies Calibration and Continuing Calibration (MARLAP, Chapter Section 15.4)

For both initial and continuing calibrations, a pulser check is run before the calibration on the same day of the calibration run. For the pulser check, the following QC acceptance criteria apply to each detector:

- Full Width Half Maximum (FWHM): within the mean ± 3σ; the mean and 3σ are established using 20 or more pulser check values.

- Pulser Energy Center: 5483 ± 50 keV

The original guidance used a pulser check energy center value of 5500 keV. This was an approximate value and was replaced with the actual mean value using the most current pulser energies. This value will be updated periodically to reflect the most recent pulser energy center values. The distribution of the FWHM data shall be evaluated similar to the daily blanks and backgrounds.

For the initial and continuing calibrations, the Guidance directed the laboratory to track the average efficiency, the percent recoveries of the activities and the maximum peak energies of the three calibration radioisotopes, Am-241, Pu-239, and U-238. For activities, we recommend retaining the 90-110% recovery criterion. By retaining this criterion, there is a low failure rate (approx. 5%) for Am-241; a near-zero failure rate for Pu-239, and approximately 1% failure rate for U-238.

For the maximum peak energies, the limits should be revised approximately every 6-9 months. The limits should be changed to the following based upon 2009 data: Am-241- 5500 ± 20 keV; Pu-239 – 5167 ± 20 keV; and U-238 – 4210 ± 20 keV. These limits have relative standard deviations between 0.36% and 0.48% and correspond to failure rates of between 2.6% and 5.6%.

MARLAP recommends verification of manufacturer's specifications for point source efficiency. For measuring a detector's counting efficiency, one should count the source long enough so that the relative uncertainty of the alpha peak count is <3%. At UFML, the relative uncertainty of the median activity values are as follows: Am-241, 3.6%; Pu-239, 3.5%; and U-238, 3.5% Review of the 2009 alpha spectrometer detector efficiencies indicates that the average efficiency acceptance criteria should remain unchanged at 20.72 – 23.20%

Recommendation: Increase the counting time of the alpha spectrometer calibration radionuclides to reduce the relative uncertainty of their measured activities to < 3%

Gas-Flow Proportional Counter Calibration Checks

The means of eight detectors were calculated for both gross alpha (GA) and gross beta (GB) for 2009 calibration check data. The average and standard deviation of the eight mean values are as follows: GA – 29.96 ± 0.38; GB – 6428 ± 53. The precision of the calibration results is excellent at about 1%. No change to the QC acceptance limits of the mean ± 3σ is recommended.

Radionuclide Quantitation and Detection Limits

The only recommended change in this Section is based upon qualification of a sample result as non-detect when it is less than the critical level. The current version of the Guidance states that the critical level is the upper 95% confidence interval of the background. Section 20.4.1.1 of MARLAP [2] indicates that if the distribution of many (>20) net instrument signal measurements is approximately normal, then the critical level of the net signal S_C is defined as:

$$S_C = Z_{1-\alpha}\,\sigma_0 \qquad \text{(Eq. 3)}$$

Where: $Z_{1-\alpha}$ = the (1-α) quantile of the normal distribution.
σ_0 = standard deviation of the standard normal distribution

The gamma spectrometry and gas proportional detector weekly backgrounds are approximately normal as noted in the **INSTRUMENT QC AND TRACER RECOVERIES** section above. The standard deviation, σ_0, may be estimated by the standard deviation of 20 or more replicate blank measurements, σ_B, using the following equation.

$$\sigma_0 = \sigma_B \sqrt{1 + 1/n} \qquad \text{(Eq. 4)}$$

where n = the number of replicate blank measurements. Therefore, the critical level of the net signal S_C can be determined using the following equation.

$$S_C = Z_{1-\alpha}\,\sigma_B \sqrt{1 + 1/n} \qquad \text{(Eq. 5)}$$

Equation 5 is only valid if the sample measurements and blank measurements are made under the same conditions and in the same manner.

Recommendation: Change the definition of critical level in the Radionuclide Quantitation and Detection Limits section of the Guidance to that described above and within Section 20.4.1.1 of MARLAP.

CONCLUSIONS

As of January 2010, the new Radionuclide Data Evaluation Guidance has been used at the UFML for approximately one year. The quality control guidelines proposed within this Guidance have been evaluated in this paper based upon quality control data generated within the past year. The MARLAP guidelines provided for several batch QC parameters are generally good. However, modifications to some batch QC parameters are recommended when the following conditions are encountered:

- Method blank result uncertainties are a significant percentage of the action level for a given analyte;
- There is significant measurement bias in a LCS result;
- Matrix spike Z scores fail but the associated data indicates good spike recoveries;

Data for daily blanks, weekly backgrounds, and tracer recoveries were plotted. The plotted data may be used to:

- Establish more accurate control limits;
- Identify data trends for a given detector that may indicate a detector performance problem, and
- Identify unique data distributions that may be typical for a particular instrument detector-analyte combination.

For gamma spectrometry calibrations, more realistic control limits may be established for the FWHM and energies of calibration radionuclide peaks based upon 2009 data. The new limits will be better suited for recognizing quality problems with gamma spectrometry calibrations. For alpha spectrometry calibrations, it is recommended that the calibration standard be counted long enough so that the relative uncertainty of the alpha peak count be less than 3%. No changes to the gas proportional calibration is recommended.

Within the Radionuclide Quantitation and Detection Limits section, the definition of critical level, below which a result is qualified non-detect, U, has been changed to be consistent with Chapter 20.4 of MARLAP [2].

While MARLAP control limit guidelines, and traditional control limits such as the mean ± 3σ provide a very good starting point for quality control acceptance criteria for batch QC and instrument QC parameters, modifications to these limits are recommended based upon project action levels, sample result bias and precision, and blank levels relative to project action levels.

REFERENCES

1. U.S. Army Corps of Engineers, Kansas City District, Radionuclide Data Quality Evaluation Guidance, May 2009.

2. MARLAP. *Multi-Agency Radiological Laboratory Analytical Protocols Manual*, NUREG -1576, EPA 402-N-04-001A, NTIS PB2004-105421. July 2004.