A Simple Approach for Assessing Data Quality Under a Performance-Based Measurement System

Kevin Coats and Chung-Rei Mao
U.S. Army Corps of Engineers
HTRW Center of Expertise
12565 West Center Road
Omaha, NE 68144-3869

SUMMARY

A simple approach is proposed for assessing laboratory performance and data quality under a Performance-Based Measurement System (PBMS). The proposed approach compares performance data with project-specific Measurement Quality Objectives before selecting a laboratory for sample analysis. The approach emphasizes documented performance under specified protocols. A laboratory must establish and implement detailed standard operating procedures (SOPs) for all major operations and document the process and results. A laboratory demonstrates its performance through Method Detection Limit (MDL) studies, Laboratory Control Samples (LCS) analysis, and frequent analysis of blind real-world performance evaluation (PE) samples. Data generated and reported under the proposed approach have an estimated uncertainty that meets the reporting requirement for uncertainty of the new International Organization of Standardization (ISO) Guide 17025. This presentation discusses why and how to use the proposed approach to assess performance and data quality under a PBMS.

INTRODUCTION

The production of data of known and acceptable quality that meet project-specified Data Quality Objectives (DQOs) is a primary goal of every environmental sampling and analysis activity. EPA's Environmental Monitoring Management Council (EMMC) recommends using PBMS for environmental sample analysis. EMMC defines PBMS as "A set of processes wherein the data quality needs, mandates or limitations of a program or project are specified, and serve as criteria for selecting appropriate methods to meet those needs in a cost-effective manner". To determine data quality needs, EPA developed a seven-step DQO process that provides project-specific limits on decision errors. Based on the data quality needs, before the fact one determines if a laboratory is qualified to perform the analysis and after the fact determines if the data produced is of acceptable quality.

It is noted that the data quality in many data packages is ambiguous or unknown so that the data usability may be judged limited or questionable. Typical data packages report analyte concentrations for all hits, "ND" or "<" signs along with quantitation/reporting limits for non-detects, and associated quality control (QC) data and the control limits. Typical QC data include analysis of calibration

verification samples, blank samples, laboratory control samples (LCS), matrix duplicates (MD), matrix spikes (MS), and matrix spike duplicates (MSD), depending on the contract's specifications. These QC data and associated control limits should inform data users of the quality of sample data. However, the origin or determination procedures of the quantitation/reporting limits and the QC control limits are often not clear or appropriate such that the data quality is unknown or misleading.

This presentation proposes a simple approach for assessing laboratory performance and data quality. The approach is based on the existing QA/QC platform that is adopted by most environmental laboratories and is applicable to both definitive and screening methods. Although both sampling and analysis errors affect the quality of environmental data, the following discussions focus on laboratory analytical errors on precision and bias.

THE PROPOSED APPROACH

The proposed approach emphasizes four key elements of conventional laboratory QA/QC operations.

(1) SOP Preparation
(2) MDL Study
(3) LCS Analysis
(4) Proficiency Testing

First, a laboratory must establish and implement detailed SOPs for all key laboratory operations that affect data quality and document the results of key operations. The SOPs and documentation provide an important aspect of scientific evidence and legal defensibility for reported data.

Second, a laboratory shall follow 40 CFR 136 Appendix B to establish MDLs for all target analytes. If all laboratories use the same procedure to determine MDLs, the MDLs would be a good universal indicator for evaluation of laboratory performance under known conditions. Based on MDLs, a laboratory determines the method quantitation limits (MQLs) and the concentration of the lowest, allowable calibration standards. The uncertainty of analytical data increases as analyte concentrations decrease and approach MDLs. The estimated relative uncertainty of analytes measured at a concentration of *N* times MDL would be:

$$\pm \frac{2\sqrt{2}}{N \times t_{(n\text{-}1,\ 0.99)}} \times 100\%$$

where $t_{(n\text{-}1,\ 0.99)}$ is the Student's *t* factor for a 99% confidence level and a standard deviation estimate from an MDL study with *n*-1 degrees of freedom [1]. At MDLs, the relative

uncertainty would be about ±100%. The estimated relative uncertainty will be exceeded by the uncertainty of LCS recovery at higher concentrations. Due to the large uncertainty near MDLs, data biases at concentrations below MQLs would be assumed to be equal to the mean of LCS recoveries.

Third, a laboratory shall establish control charts for the recovery of LCSs (i.e., blank spikes). If all laboratories use equivalent LCSs, empirically established in-house control limits would be another good universal indicator for laboratory performance and data quality. Because of clean matrices of LCSs, the control limits should be treated as the minimal uncertainties for field samples. LCS control limits may also be used to estimate the uncertainty of analyte recovery from field samples. The uncertainty of the mean recovery of field samples could be estimated as:

$$\pm \frac{t_{(n-1,\ 1-\alpha/2)} \times \sigma_{LCS} \times 100}{\%R} \%$$

where $t_{(n-1,\ 1-\alpha/2)}$ is the Student's t factor with n-1 degrees of freedom at 1-α/2 confidence level; σ_{LCS} is the standard deviation of the percent LCS recovery; and %R is the mean LCS recovery [2, 3]. Because of wide variations of matrix interferences, data biases at concentrations above MQLs will be assumed to be equal to the bias of mean LCS recoveries.

Last, because LCSs are prepared with interference-free matrices, the laboratory performance on field samples must frequently be verified with blind real-world PE samples. Double blind PE samples are preferred to single blind PE samples. If a laboratory is able to pass double blind PE samples on a routine basis, the laboratory demonstrates its performance on field sample analysis.

The four elements provide a foundation for the proposed approach. Using MDLs and LCS recoveries to assess data quality of field samples assumes consistency and comparability among different laboratories. The procedures used for determination of MDLs and the control limits of LCS recoveries will affect the values of MDLs and control limits, and hence the estimate of data uncertainties. It is noted that many laboratories do not exactly follow 40 CFR 136 to determine MDLs and use different procedures to establish QC control limits. The remaining discussions address those variations and propose standardized protocols for determination of MDLs and control limits.

MDL STUDY AND USAGE

Although most laboratories follow Appendix B of 40 CFR 136 to determine MDL, there are some variations, which may affect the MDL values. According to 40 CFR 136, the determination procedures involve spiking seven replicate aliquots

of reagent water or sample matrix with analytes of interest at a concentration within one to five times the estimated MDL. The seven aliquots are carried through the entire analytical process; the standard deviation of the seven replicate analyses is calculated; and the MDLs is determined as a product of the standard deviation and a one-tailed Student's *t* factor.

A common deviation of MDL studies is spiking too high to yield MDLs that are biased low. According to Appendix B of 40 CFR 136, the spike concentrations of the seven MDL spikes should be one to five times the estimated MDL for reagent water matrix and one to ten for clean solids or sample matrices. Otherwise, adjust the spike concentrations and repeat the study until the ratios are within these ranges. Because MDLs are based on the variances at the measured concentrations, the validity of the ratios between spike concentrations and estimated MDLs should be verified by comparing the mean of the seven measured concentrations, instead of the nominal spike concentration, with the determined MDL.

Because MDLs in real-world matrices could be elevated, the validity of MDLs in sample matrices should be verified. Based on the definition of MDL presented in 40 CFR 136, there is a 1% probability that a sample with no analyte will produce a concentration greater than or equal to the MDL. However, there is a 50% probability that a sample with a true concentration at the MDL will be measured as less than the MDL. For this reason, the validity of MDLs in other sample matrices shall be checked with MDL check samples in sample matrices at Reliable Detection Limits (RDLs) [(4)]. RDLs are based on the 1% probability of false negatives and are equal to twice the MDLs.

A laboratory shall establish its method quantitation limits (MQLs) based on the determined MDLs. At MQLs, the analytical errors should be no less than calibration errors, which are equal to the acceptance criteria for initial calibration verification (ICV) or continuing calibration verification (CCV). The acceptance criteria are usually ±10% for inorganic or classical analyses and ±20% for organic analyses. MQLs should therefore be set at ten times MDLs for inorganic and classical analyses and five times MDLs for organic analyses. MQLs also determine the concentration levels of the lowest, allowable calibration standards.

Because of the large uncertainty and bias associated with measured concentrations near the MDL, EPA did not specify acceptable limits for analyte recovery in MDL studies. However, if there is an excessively low or high recovery, the determined MDLs may not be meaningful and an MDL check sample should be used to estimate the MDL. For example, an MDL of 5µg/L based on 100µg/L MDL spikes and 10% recovery is not acceptable, because one could not reliably detect a 10µg/L spike if the recovery is 10%.

CONTROL LIMITS OF LCS RECOVERY

Laboratories usually use control charts to demonstrate that it is under statistical control at a specified confidence. The 99% confidence intervals of the mean are routinely used as the control limits if certain statistical assumptions (e.g., independent data, normal distribution, etc.) are met. The control limits reported in a data package could be based on contract or regulatory requirements, published method performance data, or laboratory in-house empirically established control limits. Using any of those control limits is acceptable as long as the laboratory has demonstrated its ability to achieve the limits on a routine basis.

Many laboratories often report project-specified control limits in data packages. To ensure that data meet project-specified control limits, many laboratories screen LCS recovery data with the specified limits. When LCS recovery is within the specified limits, laboratories consider the LCS recovery is acceptable; otherwise, reanalyze the LCS sample. If the second analysis passes, laboratories take no further actions and report sample data and the specified control limits. If the second analysis fails, laboratories take corrective actions and reanalyze all associated samples. As long as each individual LCS recovery is within project-specified control limits, most laboratories consider their performances meet project-specified control limits. However, if those LCS recovery data are charted, the calculated control limits are frequently wider than the specified limits even though each individual data is within the specified limits. This infers that the laboratory performance does not meet project-specified limits and using project-specified control limits gives misleading information on laboratory performance and data quality. A laboratory should use statistical control limits to demonstrate its performance. The wider in-house control limits could be due to a small number of LCS recovery data, which often fail to meet statistical assumptions (i.e., normal distribution, independency, etc.) Slightly wider in-house control limits are anticipated and acceptable if the sample size is small; however, when more data points (i.e., $\geq$20) are available, the data should show a central tendency and empirically established in-house control limits should meet project-specified control limits as a proof of acceptable laboratory performance.

Technically, prediction intervals, instead of confidence intervals, should be used to establish control limits, because it is the uncertainty of the next data, instead of the existing data, that is to be determined [(5)]. Control limits based on prediction intervals are wider than those based on confidence intervals. However, most laboratories use 99% confidence intervals to set control limits that lead an impression of tighter control and cause data comparability concerns. It is often noted that some laboratories establish control limits based on very few (<10) data points and some based on several thousand data points collected over a period of several years. Very few data points will not provide reliable control limits as discussed above; however, using data points over extended time may not reflect the current laboratory performance either. In addition, many laboratories retain only acceptable LCS recovery data for control chart analysis so that the control limits are tightened over time. Eventually, the laboratories have to rerun LCSs

frequently and the control limits are misleading. Obviously, a protocol for establishing and using control limits is needed to ensure the consistency and comparability of control limits for LCS recovery. The protocol should address the requirements for LCS concentration and matrix, sample size and distribution, outlier testing and treatment, statistical hypothesis and analysis, control chart updating and usage, etc. It is recommended that the protocols be established based on ASTM or ISO guides.

CONCLUSIONS

MDL and LCS recovery are two unique analytical parameters that most environmental laboratories routinely perform using the same procedures and equivalent samples. Because of their consistency and availability, MDL and LCS recovery could be used as universal indicators for evaluation of laboratory performance and data quality. The precision and bias of LCS that are determined based on control charts of LCS recovery data may be used to estimate the precision and bias of sample data. However, to ensure the data comparability, laboratories must explicitly follow specified protocols to determine MDLs and control limits for LCS recoveries. Laboratories should frequently run MDL check samples and blind PE samples to check the validity of MDLs in the sample matrix and laboratory performance on field samples. Laboratories must prepare and implement detailed SOPs for all key operations and document the results. In-house SOPs on control charts and empirically established LCS control limits should be submitted for review of laboratory performance before sample analysis. The proposed approach is simple to implement for assessing laboratory performance and data quality.

REFERENCES

1. Taylor, J. K. “Quality Assurance of Chemical Measurements.” Chelsea, MI: Lewis Publishers, Inc., 1987.
2. Thompson, M., Ellison, S. L., Fajgelj, A., Willetts, P., and Wood, R. “Harmonized Guidelines for the Use of Recovery Information in Analytical Measurement.” Technical Report, Symposium on Harmonization of Quality Assurance System for Analytical Laboratories, Orlando, USA 1996.
3. Private communications with Thomas Georgian of USACE HTRW-CX.
4. USACE. “Shell for Analytical Chemistry Requirements.” 1998.
5. Hahn, G. J. “Finding an Interval for the Next Observation from a Normal Distribution.” *Journal of Quality Technology*, 1, 1969, pp. 168-171.