DETERMINING AND DOCUMENTING THE SUITABILITY OF ANALYTICAL PROCEDURES USED FOR ANALYSIS OF ENVIRONMENTAL SAMPLES

Jerry L. Parr
Catalyst Information Resources, L.L.C.
1153 Bergen Parkway, # 238
Evergreen CO 80439
303/670-7823
catalyst@eazy.net

David Friedman
U.S. Environmental Protection Agency
1200 Penn. Ave. NW, Ariel Rios Bldg (8101R)
Washington DC 20460
202/564-6662
friedman.david@epamail.epa.gov

Jerry L. Parr has a BS in Chemistry from the University of Texas (1972) and over 25 years of professional experience in environmental chemistry including sampling, sample preparation, analysis, data evaluation, and regulatory affairs. Mr. Parr is currently a member of the Environmental Laboratory Advisory Board and the Chair of the Technical Committee for the Environmental Sciences Section of ACIL. Mr. Parr is a Principal with Catalyst Information Resources, L.L.C. where his efforts involve providing information services to laboratory professionals on CD-ROM. Mr. Parr's current professional interests are in the establishment of a performance-based measurement system for environmental testing and in a national program for laboratory accreditation.

David Friedman has been involved in environmental testing for over 25 years, initially with the Food and Drug Administration and for the past 13 years with EPA. He was responsible for the development of both the RCRA Toxicity Characteristic and SW-846. Mr. Friedman has been heavily involved in the establishment of a national accreditation program for environmental laboratories, in the Agency's efforts to adopt PBMS and to promote the development and use of innovative approaches to environmental monitoring. Mr. Friedman earned his BS and MS degrees in chemistry from Brooklyn College and Lehigh University, respectively.

ABSTRACT

One of the major barriers identified in preventing a full implementation of a Performance Based Measurement System (PBMS) relates to the activities a laboratory would need to perform to demonstrate that it can competently perform a method correctly. These activities are a verification that key performance measures – bias, precision, sensitivity, and selectivity – can be achieved. The performance measures are those that have been established based on the project goals, not the typical performance published in the method from the validation study, unless the project goals use the validation data as the basis for the project measurement objectives. Various approaches to this method verification activity, including the analysis of Certified Reference Materials, comparison to the results of an independent method, and the analysis of spiked samples are appropriate, depending on the level of confidence needed in the determination.

1.0 INTRODUCTION

U.S. Environmental Protection Agency (EPA) programs have historically specified required promulgated analytical methods to be used by the regulated community in the analysis of environmental samples for regulatory compliance purposes. For example, Title 40 of the Code of Federal Regulations (CFR) Parts 136 and 141 contain promulgated methods to be used for compliance monitoring in the wastewater and drinking water regulations. Although not required, because of the regulatory acceptability of these "EPA-approved" methods, they have also been specified in many other monitoring programs, including ambient water quality monitoring.

EPA plans to change the current approach to compliance monitoring to emphasize the performance that must be achieved rather than the methods that must be used to collect the required data (Newman, 1996). EPA believes that this more flexible approach will reduce the regulated community's compliance monitoring costs and will encourage innovation in analytical technology while improving the quality of compliance monitoring. On October 6, 1997 the agency announced its intent to implement a Performance Based Measurement System (PBMS) approach for environmental monitoring in all of its media programs to the extent feasible to accomplish this goal (EPA, 1996).

In PBMS, a regulated entity may use any appropriate analytical technique to demonstrate compliance with regulatory requirements. The regulated entity is responsible for demonstrating and documenting that its chosen technique meets whatever performance criteria are established for the particular application (e.g., regulation or permit). These criteria focus on the quality of data needed for the particular program or project (the performance) rather than on the particular analytical method (the technology). PBMS allows a regulated entity to select the least costly, simplest, or most practical method that can meet the specified performance requirements.

Under PBMS, EPA would establish performance criteria without prescribing specific procedures, techniques, or instrumentation. Performance criteria would be published in regulations, permits, or technical guidance documents, depending on the individual program. The intent of these criteria would be to ensure that data would be appropriate for the intended application. Performance criteria may be based on either data quality objectives (DQO), which define the statistical confidence required in conclusions drawn from data, or on measurement quality objectives (MQO), which establish measurement system performance requirements such as sensitivity, precision, or bias. DQO and MQO performance criteria depend on the question or decisions to be addressed by the subject measurement, the level of uncertainty that is acceptable, the ease with which method performance can be verified, and other factors (GIES, 2000).

A performance-based approach thus permits the use of any appropriate laboratory method that can demonstrate compliance, whether or not the method has received prior review or approval by EPA. The fundamental characteristics of PBMS are:

- A regulated facility may use any method that is scientifically appropriate without EPA approval;
- The performance of the measurement technique is application-specific;
- The regulated entity is ultimately responsible for demonstrating compliance; and
- The performance standards remain the purview of the regulatory authority.

The current prescriptive system has placed the burden on EPA to determine if methods are suitable for their intended use, and has led to extensive "method validation" studies by the Agency, many times involving interlaboratory comparisons. Under PBMS, laboratories will not rely on using "validated" methods; rather, each laboratory will be required to document and demonstrate that its measurement system provides data consistent with the intended purpose. This activity, termed "method verification," is one of the four critical factors identified by the U. S. Supreme Court in ensuring admissibility of scientific evidence, and is a foundational element in virtually any treatise on measurements for ensuring technical reliability (ELAB, 1999).

2.0 METHOD VERIFICATION

Method validation consist of the activities an organization such as EPA or ASTM performs to provide potential users of the method the expected performance of the method and directions for others on how to use the method. Ideally, in this context, the objectives for the validation should be consistent with the objectives of the expected measurement. However, many times the organization is simply trying to make available a standardized method for general use where these project objectives are not known.

The outcome from this validation effort is generally a statement of the precision, bias, and sensitivity of the method. However, as stated by John Taylor, "Such statements are often misinterpreted, they merely describe the results of the exercise and are, at best, estimates of typical performance expectations of the method. However such information should be obtained to the extent possible since it provides a quantitative basis for judging performance capability (Taylor, 1983)."

Validation is useful to:

- assist in method selection,
- provide an indication of potential utility,
- provide a useful guide for the best performance that can be expected to assist in establishing measurement quality objectives,
- provide a basis for comparison of alternative methods,
- help in establishing legal standing, and
- obtain EPA approval.

However, the use of a validated method cannot ensure that data of the necessary quality was achieved. Nor should the results from the validation effort be used to establish laboratory requirements for QC samples or to guarantee the quality of all future measurements. Validation is simply an exercise to establish typical performance expectations of the method.

By contrast, method verification is the activity a laboratory who uses a method does to demonstrate that it can competently perform a method correctly (MacDowell, 1999). It is a verification that key performance measures – bias, precision, sensitivity, selectivity – can be achieved. The performance measures the laboratory should use are those that have been established based on the project goals, not the typical performance published in some EPA approved method, unless the project goals use the performance data from some published, standardized method as the basis for the project measurement objectives.

The results from this verification should be compared to the needed performance of the method and thus should provide sufficient confidence that the laboratory is competent in performing the method. The results from this verification should not be used as the basis for the quality of all future measurements. The verification results simply document expected performance.

Thus, method verification can be considered the documentation clients should require laboratories to supply with the results of analyses in order for clients to be able to determine if the method: (1) is capable of achieving the requirements they established for data quality; (2) is suitable for its intended purpose; and (3) is technically reliable.

The rigor of the demonstration of selectivity, sensitivity, bias, and precision that needs to be performed to verify a method's performance for any particular set of samples is a function of the level of validation information available

for the particular measurement system, the level of information available on the properties of the particular set of samples, the concentration range of interest, and the level of quality assurance needed (Keith, 1983).

When sufficiently detailed validation data are already available on the properties of the set of samples to be analyzed, or for similar sample matrices, these data (e.g., historical data from the laboratory conducting the demonstration, data published in the literature, method developer/equipment manufacturer information) may be used to support the laboratory's efforts.

The sections below describe some of the options which should be considered for the method verification.

3.0 SELECTIVITY

Selectivity is defined as a measurement system's ability to accurately discriminate the analyte of concern from other analytes in the matrix that may interfere.

Using samples of the matrix free of the analyte of interest or containing insignificant quantities of the analyte of interest demonstrate the selectivity of the measurement system by documenting that the measurement system is free from, for example:

- overlapping chromatographic peaks in region where compound of interest elutes
- interfering ions in mass spectrum
- overlapping spectra peaks in emission spectrometry
- interfering absorbances in absorption spectrometry
- cross sensitivity in immunoassay
- overlapping wavelengths in the region of response for colorimetric techniques

Another approach would be to demonstrate freedom from unacceptable instrument or test response to interfering compounds using a split of the as received sample coupled with acceptable recovery of spikes from a split sample spiked with the analyte of interest at the level of interest.

For a candidate measurement system that is typically capable of quantifying the analyte of interest at the concentration/level of concern (e.g., 2 mg/l), spike recovery testing can be performed to determine selectivity. This will enable a determination of the concentration/level at which the measurement system will detect the analyte of concern in the matrix of concern.

4.0 SENSITIVITY

Sensitivity relates to the ability of a measurement system to accurately measure the analyte (e.g., arsenic, benzene) or property (e.g., specific gravity, vapor pressure, flash point) of interest at a specific level of interest.

A demonstration of adequate sensitivity would be one in which the measurement system shows a specified probability (e.g., 99%) of detecting the analyte at/above the specified concentration or level within a specified degree of bias (e.g., ± 10%) and precision (e.g., ±0.2 mg/l) at the specified concentration or level.

Options that might be used to document measurement system sensitivity are discussed below.

4.1 Spikes at Concentration of Interest

When determining that an analyte is present at or above a specified concentration/level, use samples of the matrix spiked with a known amount of the analyte of interest to demonstrate the sensitivity of the measurement system by showing that the system can detect the analyte of interest at the level of interest within a specified bias and precision.

4.2 Method Detection Limit Study

Another option for verifying sensitivity would be to determine the Method Detection Limit (MDL) and demonstrate that the MDL is below the concentration/level of interest. One approach to performing a Method Detection Limit study is to use the procedures in 40 CFR 136.

5.0 BIAS

Bias is the difference between the mean (i.e., the average) of the test results for the analyte of interest and the real value of the analyte concentration or value. Options that might be used to document measurement system bias are described below.

5.1 Use of a Certified Reference Material

If multiple analyses of an appropriate certified reference material containing a known quantity of the analyte of interest are performed, then the average of those analyses should be within the specified allowable range of bias from the known standard concentration/level.

5.2 Multiple Matrix Spikes

If multiple matrix samples are spiked at the same level with the analyte of concern at the concentration of concern, then the average of the spike percent recovery values should be within a specified allowable range compared to the known value of the spike. Note that if the appropriate number of multiple matrix spikes are used, the same data used to determine bias could be used in a demonstration of analytical precision. While this approach can evaluate the bias of the method, the results should be used with caution as it is well-known that such spikes may not truly represent actual sample performance due to issues associated with the spiking process. Nonetheless, this approach is the only practical approach to method verification for many methods.

5.3 Compare Candidate Measurement System with Fundamentally Different Measurement System.

This type of demonstration would compare the candidate measurement system average results for multiple analyses of the same sample to similar analytical results from a fundamentally different measurement system (based on other analytical principles or a "different science", such as a reference, or standard method) whose bias is known. An adequate demonstration would be one in which the candidate system's average data are statistically the same as the data average of the different measurement system.

5.0 PRECISION

Precision is a measure of the variability of test results obtained from applying a measurement system to samples that are ostensibly the same (e.g., taken at the same time and location, poured from the same container). The measurement system produces a scatter or variability of the data resulting from multiple measurements of the same sample that is within an acceptable range (e.g., ±10% of the average of the data) with an acceptable probability (e.g., 99% of the time).

The precision of an analytical measurement can be determined by duplicate, triplicate, or other multiple analysis of a matrix sample. If the spread of the analytical results of the multiple samples are within the specified range (e.g., one standard deviation equals ± 20% of the analytical mean), then precision is demonstrated. Specific variations of multiple matrix analysis that could be used to demonstrate and document precision are described below.

6.1 Multiple Matrix Spikes

The analysis of multiple matrix samples spiked with the same concentration of analyte might be used when the spiked samples are also being used to determine analytical bias.

6.2 Surrogates

Where precision is based on the analysis of multiple samples of a matrix similar to the matrix of concern, and the matrix has analytes similar to the analyte of concern (e.g., a surrogate for a chlorinated volatile organic compound of interest might be another chlorinated volatile organic compound), the recovery of the surrogate spikes can be used to estimate precision. This approach might be used if the analyte of interest is not commercially available or is too dangerous, toxic, or expensive to use as a spike, or is unstable (i.e., has a poor shelf life).

6.3 Certified Reference Materials or other Laboratory Control Samples (LCS)

These materials, which are designed to provide a sample that would have analytically similar properties to that of the actual samples to be analyzed, can also be used to assess precision. LCS can be actual samples of the materials in question (e.g., Laboratory Replicates, Matrix Spikes) or synthetic samples, but should be representative of the field samples. This approach might be used to demonstrate the precision of the measurement device or instrument rather than the precision of the entire measurement system.

6.4 Multiple Sample Analyses

If the bias of the measurement system is known, the analysis of multiple samples can be used to estimate precision. This approach is only appropriate if the analytes are present at measurable concentrations.

SUMMARY

For PBMS to be successful, laboratories will need to accept responsibility for demonstrating that the methods they use are adequate for their intended purpose. This demonstration will need to evaluate the selectivity, sensitivity, bias, and precision of their method, compared to the data quality needs of the measurement effort. There is no one approach which can be defined; rather, the approach is a balance of the level of quality assurance effort, the concentration in the sample, the uncertainty in the measurement and the data quality need.

WORKS CITED

Newman, A. (1996) *EPA Plans New Approach to Analytical Methods*. Anal. Chem. 68: 733A-737B.
EPA (1996) *Performance Based Measurement System Notice of Intent*. Federal Register 62 FR 52098.
ELAB (1999) *Recommendations for the Implementation of Performance-Based Measurement Systems (PBMS)*, <www.epa.gov/ttn/nelac>
GIES (2000), *Implementing a Performance Based Measurement System*, <www.CatalystInfoResources.com>
Keith, L.H., et. al, (1983) *Principles of Environmental Analysis*, Anal Chem. 55: 2210-2218.
MacDowell, R.D., (1999) *Who Developed this Method?*, LC-GC, 17 (2), 112-116.
Taylor, J.K. (1983) *Validation of Analytical Methods*, Anal Chem, 55: 600A-608A.