New Quality Control Assessment Kits for Data Quality Objectives

L. H. Keith, M. A. Re, G. L. Patton, and P. G. Edwards, Radian Corporation, P. O. Box 201088, Austin, Texas 78720-1088

The purpose of environmental sampling and analysis is to assess a small, but informative, portion of a population and then draw an inference about that population from the data gathered. There are an almost infinite number of samples that could be taken at any given site, so environmental samples must be collected in such a way as to be representative of the environmental area of interest. Typically, environmental samples may be taken from matrices that include water (surface waters, drinking water, ground water, industrial wastewater, etc.), soils, aqueous sediments, vegetation, air, or manufactured products (e.g., paper, waste oils, etc.). Quality control (QC) samples are used to provide an assessment of the kinds and amounts of bias and/or imprecision in the data that is obtained from the environmental samples. Thus, QC samples are used to assess the collection and measurement system in a similar way that environmental samples are used to assess the portion of the environment from which they come. Therefore, representative environmental samples are collected and analyzed to form conclusions about a particular site, and representative QC samples are analyzed to form conclusions about system that measures the environmental samples. This similarity in environmental sample usage and QC sample usage is often not appreciated or even recognized.

There are many different types of QC samples and each is designed for a specific purpose. Some provide an assessment of bias while others provide an assessment of imprecision. In addition, some are designed to assess laboratory-based variability and others are designed to assess overall variability (both sampling and analysis). An expert system named ***"Practical Environmental QC Samples"*** (1) provides answers for the question of what kinds of QC samples to use for specific purposes but it doesn't calculate how many QC samples are needed to assure specific confidence levels. A new computer program named ***DQO-PRO*** compliments ***Practical Environmental QC Samples*** and calculates the numbers of samples (both QC samples and environmental samples) needed to resolve individual project needs. For example, ***DQO-PRO*** calculates numbers of samples needed to assure, at a selected confidence level, that a localized area of contamination ("hot spot") is not missed. It also calculates numbers of samples needed, at a selected confidence level, to estimate the average concentration of a pollutant in samples and the standard deviation or the relative standard deviation (coefficient of variation) of the method used for its analysis.

One of the most basic QC data assessments is to determine the presence of false positives and false negatives in environmental analytical data. An analyte that is incorrectly concluded to be present in a sample is a false positive; these can cause regulatory and financial consequences for a laboratory's clients. One cause of false positives is misinterpretation of the identity of interfering analytes for the target analytes. When interferents are present in a sample, the method must be modified to eliminate them, but when they are present in the materials used to prepare or analyze samples (e.g., bottles, solvents, reagents, filters, columns, detectors, etc.), their sources must be determined and the interferent removed if possible. Various kinds of QC samples (e.g., as determined from the ***Practical QC*** program) can be used to determine where, in the chain of events, the interferents are contributed but the first step is to recognize their presence. Method blanks, which consist of a blank matrix similar to the samples, but without the target analytes, are used to determine overall if false positives are present in the materials and/or the process used to prepare and analyze samples (but they don't identify the source of error).

A false negative occurs when an analyte is concluded to be absent in a sample while, in reality, it is present at detectable levels. False negatives commonly occur from poor recovery of target analytes from a matrix, or from interferences that mask the target analytes. They are especially troublesome to government and regulatory personnel and also to scientists who work with risk assessments because they result in analytes being concluded to be absent when, in fact, they are present.

Most environmental analyses for PCBs, PCDDs and PCDFs are conducted in "batch" modes to facilitate more cost effective analyses. In doing so, one method blank (also called a lab blank) and one or two method spikes (or matrix spikes) are typically analyzed along with about 20 environmental samples. The resulting data for all of the environmental samples in that batch are accepted or rejected on the basis of those QC samples.

When used this way, the QC data of a batch does not provide a statistically sufficient amount of information for the environmental samples. One or two QC samples, which is how these QC samples are grouped, does not provide enough information to predict the reliability of the other environmental samples that are grouped with them. An implicit assumption that the environmental samples analyzed in conjunction with a method blank and one or two spiked method blanks (or matrix spikes) do not contain false positives or false negatives because the accompanying one or two QC samples did not contain them is not necessarily correct. *Thus, the present way of assessing QC data contains a basic flaw that is not usually recognized.*

How can method blanks and method spikes (i.e., spiked method blanks) be used as representatives for the environmental sample population? The answer is to use a statistically valid number of QC samples. That number depends on the Data Quality Objectives (DQOs) of a particular sampling and analysis project. As an example, the number of QC samples needed can vary from 6 (for an 80% probability that the data will not contain more than 25% false positives or false negatives) to 458 (for a 99% probability that the data will not contain more than 1% false positives or false negatives).

The equations for making these calculations have long been known but they are complicated and, therefore, not often used (2 - 5). We have written a computer program, named ***DQO-PRO***, that incorporates them with the user interface of a simple calculator. ***DQO-PRO*** provides answers for three objectives: (1) determining the rate at which an event occurs, (2) determining an estimate of an average within a tolerable error, and (3) determining the sampling grid necessary to detect "hot spots". Thus, informed decisions can quickly and easily be made on the most basic of QC data assessments which, before now, were complicated and difficult for non-statisticians. ***DQO-PRO*** can be equally useful when planning for desired levels of data quality in sampling and analysis projects, or when assessing the levels of data quality present in QC data sets on hand.

When *DQO-PRO* is used to optimize a study design so that statistical confidence levels planned with sampling and analysis projects can be achieved, *all significant analytical parameters must be maintained without change during the period of time that the QC samples are being accumulated to assess the* measurement process for the environmental sample population. Significant parameters that can affect analytical method performance include the instrumentation, the analyst, and the matrix.

- Changing or modifying instruments can affect instrument detection levels and many other measurement parameters.
- Analysts with varying degrees of experience and different analytical techniques can also affect results of the measurement system.
- Different matrices may have different artifacts, interferences, and also affect the recovery of target analytes differently.

Laboratories can readily document the consistent use of instrumentation and the analyst for performance of a given method with environmental samples. Environmental matrices, however, are more inconsistent; this is especially true with soils. Thus, a consistent source of representative matrices is also important for an assessment of false positive and false negative conclusions from the sample measurement system.

In order to facilitate the use of statistical assessments with QC data, we are providing *DQO-PRO* at no cost to people who wish to use it. In addition, we have packaged representative soils in convenient QC Assessment Kits. Using these kits provides ongoing control of the third major parameter (the matrix) needed to maintain consistency among a statistically relevant population of QC samples over time. The QC Assessment Kits contain 10 units of conveniently packaged soil for method blanks using any desired method for PCBs, PCDDs, PCDFs or any other target analytes. Some QC Assessment Kits also contain soils from the identical lot of homogenized soil that are pre-spiked with PCBs, PCDDs and PCDFs and thoroughly homogenized. Alternatively, two QC Assessment Kits with blank soils can be purchased and one of them spiked with custom prepared target analytes at any desired concentrations. The soils used in these kits were selected from pristine areas in North Carolina and California so they represent both East Coast and West Coast regions. Both soils are sandy loam; this type of soil was selected because it commonly occurs throughout the world and also because most organic pollutants spiked onto this type of soil typically give average recoveries (not high as with sand and not low as with clays).

The more kits that are used over any given time period, where all significant parameters remain constant, the higher the statistical probability becomes that low rates of false positives or false negatives can be identified in the associated environmental samples. **Since similar QC samples would be analyzed anyway, analyzing a group or batch of samples from a QC Assessment Kit will not significantly increase costs, but it will significantly improve the assumption of measurement process consistency because it removes the variability associated with unknown matrices and poorly homogenized samples.** Time limitations of 3 to 6 months are recommended as reasonable lengths of time over which to accumulate statistical populations of QC data from these kits.

Documented method parameters should be consistent in laboratories that frequently use a given method for several weeks to several months. Table 1 provides an example of potential benefits, in terms of increasing statistical confidence to detect a low error rate, that can be gained by using QC Assessment Kits over a controlled period of time.

Table 1 Numbers of QC Samples Versus Confidence Levels (Probability) of Not Exceeding Selected Average Error Rates

Number of Kits	Number of QC Samples	Confidence Level With 20 % Error Rate	Confidence Level With 10% Error Rate	Confidence Level With 5% Error Rate	Confidence Level With 1% Error Rate
1	10	89%	65%	40%	10%
2	20	99%	88%	64%	18%
5	50	100%	99%	92%	39%
10	100	100%	100%	99%	63%
15	150	100%	100%	100%	78%
20	200	100%	100%	100%	87%
30	300	100%	100%	100%	95%
50	500	100%	100%	100%	99%
100	1000	100%	100%	100%	100%

References

1. Keith, L. H. "Practical QC Samples" an expert system program in ***Practical QC***, Instant Reference Sources, Inc., 7605 Rockpoint Dr., Austin, TX 78731 (1994).

2. Cochran, W. G., "*Sampling Techniques*", John Wiley & Sons Inc., New York, 3rd Edition, 1977.

3. Grant, Eugene L. and Richard S. Leavenworth. "*Statistical Quality Control*", Sixth Edition. McGraw-Hill, Inc., New York, pp. 201-208 (1988).

4. Hahn, Gerald J. and William Q. Meeker. "*Statistical Intervals: A Guide for Practitioners.*" John Wiley & Sons, Inc., New York, pp. 104-105 (1991).

5. Gilbert, R.O., in "*Statistical Methods for Environmental Pollution Monitoring*", Van Nostrand Reinhold Company, Inc., pp. 119-131 (1987).