Chapter 1

Laboratory Quality Assurance Programs

Bruce Hoskins

A quality assurance program is necessary for all laboratories to document analytical uncertainty and to promote confidence in analytical results. Quality assurance (QA) can be divided into two parts: quality control and quality assessment. Quality control (QC) is comprised of those laboratory practices that are undertaken specifically to achieve accurate and reliable analytical results. Quality assessment is comprised of those processes undertaken to monitor and document the effectiveness of the quality control program. A regular assessment of quality control will document both accuracy and precision. Accuracy is defined as closeness of a measurement to the known or expected value. Precision is defined as the agreement or repeatability of multiple measurements on the same sample (Garfield, 1991; Western States, 1998). Accuracy and precision together characterize analytical uncertainty.

A formal QA plan can be a useful foundation document from which to derive quality control and assessment guidelines for all methods run within a lab operation. In addition to QC guidelines, a QA plan should contain a laboratory mission statement, overall QA objectives, an organizational chart, a code of ethics, training and safety practices and procedures. A complete listing of QA plan components can be found in the SSSA/NAPT QA/QC Model Plan (2005) and in EPA SW-846 (EPA, 1986).

The initial additional overhead of implementing a QA program should be more than offset by an improved ability to pinpoint problems early, resulting in a streamlining of lab operations. An effective QA program will also improve customer confidence in analytical results. The relative cost/benefit ratio of individual QC components or techniques should be considered when implementing or modifying a QA program (Garfield, 1991).

The scale of a QA program should be determined primarily by the end-use of the analytical results. It is not the purpose of this chapter to mandate QA standards for all laboratories, but to delineate common components and practices. Specific QA program components and guidelines should be determined within each laboratory operation, with input from laboratory personnel, clients, and other stakeholders.

The purpose of any soil testing laboratory is to provide a consistent index of soil fertility and to identify soil properties which may affect plant growth or potentially harm the environment. The end-use for this information may not be the same in all cases. The accuracy and precision needed to generate consistent lime and fertilizer recommendations may be different than that needed for the purpose of regulating trace element application from regulated waste, for example. In all cases, the goal should be to provide consistent quality analytical results from the laboratory resources available.

Components of a Quality Control Program

A good quality control program includes documentation, training, and implementation of good laboratory practices and procedures. Many of the QC procedures suggested here may already be in use or require only slight alterations of existing processes used in most laboratories.

A complete listing of standard operating procedures (SOP's) is one of the most important QC practices. Since slight alterations in soil testing procedures can cause surprisingly large differences in the final results, detailed SOP's will help insure that procedures are run consistently, minimizing variability in results. SOP's can also be very useful for troubleshooting problems. Documentation of SOP's is also required by many contractors, as well as by most laboratory certification agencies.

Individual SOP's for sample receipt and logging, sample preparation, extraction, calibration solution preparation, and instrument setup/operation/maintenance, should be specified in detail (Thiex, et al, 1999; SSSA, 2004). Quality assessment guidelines should be included within SOP's, spelling out what types of reference samples are to be run, at what frequency, and with general guidelines for allowable ranges of results. Numbers and frequency of reagent blanks (see below) should be specified within applicable SOP's.

Sample preparation and (where applicable) solution analysis procedures within SOP's should be referenced wherever possible to published standard methods to demonstrate method conformity and to inform customers of the exact methodology in use. One of the purposes of this bulletin is to provide a methodology reference for all soil testing laboratories in the Northeast Region.

Another useful QC technique, which can especially benefit new employees, is a written summary of known sources of error in the lab operation. These include, but are certainly not limited to the examples listed in Table 1-1.

Keeping a log of known errors encountered over time for each instrument or process can be an invaluable tool when troubleshooting laboratory problems. An error log also promotes continuity within a succession of technicians or operators, as well as more consistent operation over time for any individual technician.

Table 1-1. Known Sources of Error in Soil Testing Laboratories

Example of Sources of Error	Corrective Action
Segregation or stratification of soils in storage. Heterogeneous samples.	Rehomogenize samples prior to sub-sampling for analysis. Run replicate analysis
Contamination of samples or equipment by lab environment.	Store samples, reagents, and equipment separately.
Sample carryover on extraction vessels or other apparatus.	Rinse with cleaning solution between samples.
Samples weighed, processed or analyzed out of order.	Verify sample ID's during subsampling. Run a known reference sample at regular intervals.
Inaccurate concentrations in solutions used to calibrate instruments.	Check new standards against old standards before use.
Mismatch between sample and calibration solution matrices.	Make calibration standards in the extracting solution used for the soil samples.
Drift in instrument response.	Use frequent calibration and QC checks. Use instrument internal standards, if applicable.
Poor instrument sensitivity or high detection limits.	Optimize all operating parameters. Check for blockages in sample delivery system.
Faulty data handling or transcription errors.	Proofread input. Automate data transfer.

Quality Assessment

The second part of a QA program is quality assessment. Quality assessment checks the effectiveness of the QC practices used in the laboratory and is used to determine if an analytical process is in compliance with QA guidelines. Quality assessment is achieved through systematic measurement and documentation of bias, accuracy, and precision.

Documenting and Eliminating Bias

The most common technique used to detect and quantify analytical bias in soil testing is the inclusion of process or reagent blanks. One or more empty sample containers are carried through the entire preparation process, with the same reagents added and final dilution applied. Blank solutions are analyzed with actual samples, using the same calibration. Blanks should be run at regular intervals with each batch of samples to determine if any analyte concentration is consistently above method detection limits (MDL) and also to determine the variability of blank content. Blanks are more likely to be significant for those analytes present at relatively low concentrations, as in trace element or micronutrient analysis.

The inclusion of blanks will quantify any contribution of containers, reagents, and the laboratory environment to the content of prepared samples or solutions. A consistent blank value, if the source cannot be eliminated, should be subtracted from the concentration values for that analyte in the samples run in association with the blanks. Blank subtraction is used to correct for systematic sources of contamination, not random ones. In this way, systematic bias in the process can be corrected in order to improve accuracy.

Groups of process or reagent blanks can also be used to calculate detection limit and quantitation limit for each analyte, typically defined as 3 times and 10 times the standard deviation of the blank values, respectively, for each analyte (Taylor, 1987) (Thiex, et. al., 1999). Blanks should be run at relatively high frequency until valid mean and standard deviation statistics can be generated and a determination made as to whether blank values are consistent within an analytical process. Blank values should also be checked at increased frequency after any change in procedure or reagents.

Documenting Accuracy

Accuracy of analytical results can be documented by analyzing reference samples of known content. A reference sample is a homogenized sample, as similar as possible to the routine samples being tested. Several standard reference materials (SRM's) can be purchased from commercial or government sources, such as the National Institute of Standards and Technology (Standard Reference Materials Catalog, NIST Special Publication 260, Gaithersburg

MD 20899-0001). Analysis of an SRM is considered the most unbiased way to document accuracy in a laboratory QA program (Delavalle, 1992). There are several drawbacks, however. SRM's are quite expensive and typically of limited quantity. Many analytes of interest will not be reported or are reported, but not certified. SRM's usually do not list extractable content based on soil fertility testing methodology. Reference soils which are available are typically guaranteed for total content only.

The most useful reference samples are those that have become available through Proficiency Testing programs, such as the North American Proficiency Testing Program (NAPT). In these programs, samples of homogenized soils are sent to all cooperating laboratories, which analyze them by specified methods and protocols. Analytical results for soil fertility testing methods which are not typically reported for purchased SRM's can be obtained in this way. Median and median absolute deviation (MAD) values are typically determined and reported for each analyte and for each method, based on the data returned by participating labs. While this is not technically a certified or guaranteed analysis, the median value obtained from several laboratory sources can be considered closer to the "true" values than results derived solely from one laboratory. Proficiency testing program reference samples are available through the North American Proficiency Testing Program (Utah State University Analytical Lab, Logan UT 84322) and through the International Soil Exchange Program (P.O. Box 8005, 6700 EC Wageningen, The Netherlands) at a reasonable cost.

Documenting Precision

Precision of analytical results can be documented through replicate testing of routine samples or by routine analysis of internal reference samples. Replicate analysis typically involves two or more analyses of routine sample unknowns at some specified frequency, such as every fifth or every tenth sample. A relatively high frequency of replication should be used initially. Replication frequency can be reduced after the minimum number of replicates has been generated to produce valid statistics (see R-Chart section) and once QA precision standards for the method are being met. Replicate analysis is especially useful where appropriate standard reference materials are unavailable (Garfield, 1991). Since actual sample unknowns are being used, the final solution matrix and the concentration ranges of each analyte will automatically match those of the samples being run. Matrix and concentration range mismatch can be a concern when running internal or external reference samples (Delavalle, 1992). Since all analytical results are generated internally, no determination of accuracy is provided by sample replication.

An alternative or supplement to replicate analysis is to run internal reference sample(s). An internal reference is typically an in-house homogenized sample, subsamples of which are run at regular or irregular intervals in the routine sample stream. Bulk samples can be prepared relatively easily and with minimal expense. It is important that the bulk sample be finely sieved and thoroughly homogenized before use and remixed at regular intervals (weekly, for example) to prevent sample stratification. Internal reference sample content can be validated by running it in tandem with purchased SRM or PT samples. Once validated and checked for homogeneity, an

internal reference sample can be used as an inexpensive surrogate SRM and is often the primary daily QC assessment used in soil testing labs.

Known and Blind Checks

Quality assessment samples can be run with the full knowledge of the technical staff or as single or double blind samples. Check samples of known composition run at known intervals can be used by technicians to monitor the quality of analytical results as they are being produced. A blind sample is known to the technical staff as a check sample, but the composition is unknown. A double blind sample is completely unknown to the technical staff and is used to eliminate any possible bias in the results, from knowing the location or composition of the check sample. Blind and double blind samples are best reserved for formal quality control appraisals (Taylor, 1987).

Descriptive Statistics and Control Charts

Descriptive statistics used to quantify a laboratory QA program can be presented in a variety of ways. Accuracy is measured in terms of the deviation or relative deviation of a measured value from the known or assumed value. Precision is presented in terms of standard deviation (SD) or relative standard deviation (RSD) from the mean of repeated measurements made on the same sample. Together, accuracy and precision document the systematic and random errors which constitute the analytical uncertainty in laboratory results. Besides documenting uncertainty, descriptive statistics from an established QA program can be used for other purposes. Accuracy and precision statistics are the performance criteria used to determine if a methodology is in "statistical control", that is whether quality assurance standards are being met over the long term. Check sample statistics can also be used by technicians and managers as daily decision-making tools during sample analysis to determine if expected results are being generated and if the analytical system is functioning properly at any given time. Determining that a problem exists as soon as it happens can save a great deal of lost time in running samples over again at a later date (Delavalle, 1992).

X-Charts

Quality assessment statistics can be presented graphically, through control charts, for ease of interpretation. X-charts can be used to present both accuracy and precision data. Repeated measurements of external or internal reference samples are graphed on a time line. A minimum of 7 measurements is needed, though 15 are recommended for valid statistical calculations (Taylor, 1987). Superimposed on the individual results is the cumulative mean (in the case of an internal reference sample) or the known value (in the case of an external SRM or

PT sample). Control levels which typically represent +/- 2 SD (upper and lower warning limits: UWL & LWL) and +/- 3 SD (upper and lower control limits: UCL & LCL) from the mean are also superimposed (Figure 1-1). In a normally distributed sample population, +/- 2 SD represents a 95 % confidence interval (CI) and +/- 3 SD corresponds to a 99 % CI.

An individual value between UWL and UCL or LWL and LCL is considered acceptable, though two or more in a row are unacceptable. A single value outside UCL or LCL is considered unacceptable. If statistical control is considered unacceptable based on either standard, all routine sample unknowns between the unacceptable check sample(s) and the last check sample which was in control should be rerun. Check sample results which fall within the warning limits, but which are exhibiting a trend toward the UWL or LWL can signal a potential problem in the process which needs to be addressed (Delavalle, 1992; SSSA, 2004). X-charts are especially useful as a day to day tool to monitor for ongoing or emerging problems.

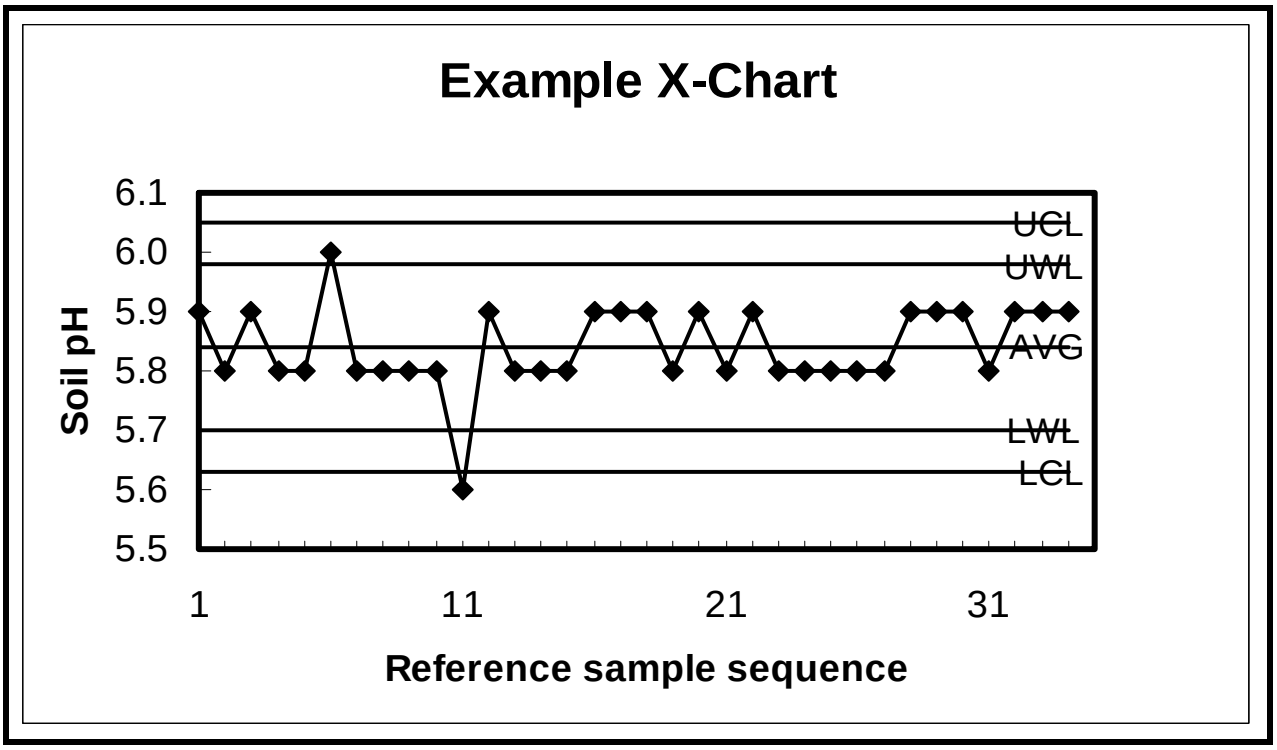

Figure 1-1. Typical X-Chart used in a QA/QC program.

R-Charts

Another graphical display is the R-chart or range chart. When two or more replicate analyses are run on a routine sample or a reference sample, the difference between the lowest and highest values in a set of replicates (or just the difference between replicates when there are only two) is called the replicate range. The R-chart maps individual replicate ranges for a given analyte over time. The replicated samples should ideally be within an acceptable total range of concentration, for the same analytical process or methodology (Delavalle, 1992). A cumulative mean range is calculated and superimposed on the individual range values. Warning and control limits are calculated as 2.512 times (95 % CI) and 3.267 times (99 % CI) the mean range (Taylor, 1987). Since replicate ranges are absolute, only one warning and control limit are displayed (Figure 1-2). Since R-chart data consist solely of replicate ranges, they can only be used to document precision. A minimum of 15 replicated samples is recommended for producing an R-chart (Taylor, 1987).

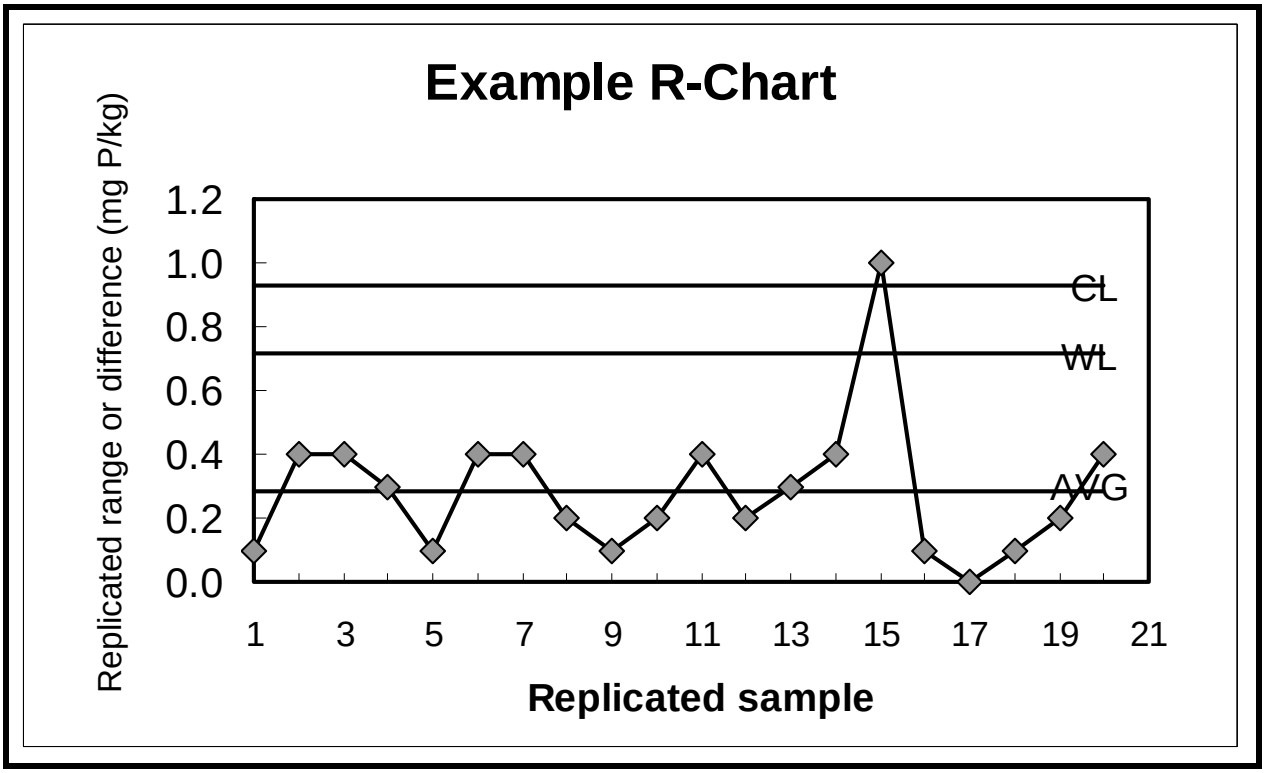

Figure 1-2. Typical R-chart used in a QA/QC program.

Establishing Control Limits

Since warning and control limits are calculated from cumulative statistical data, new quality control assessments are always viewed relative to past performance. Cumulative statistics effectively characterize the inherent capability of a laboratory to execute a given methodology. Realistic QC standards for accuracy and precision in any lab must take this capability into account. The first step should be to define attainable accuracy and precision within the normal range of sample content (Taylor, 1987). Once attainable standards are determined, they should be used to maintain consistent analytical quality over time. Allowance must be made for decreased accuracy and precision and increased analytical uncertainty as an analyte approaches MDL.

Recommended Reading

For a more thorough coverage of modern QA/QC programs for soil testing laboratories, including statistical analysis, planning, documentation, and control charting, the books by Garfield (1991) and Taylor (1992) are highly recommended, as is the SSSA/NAPT QA/QC manual (2004).

References

1. Delavalle, N. B. 1992. Handbook on Reference Methods for Soil Analysis. Quality Assurance Plans for Agricultural Testing Laboratories: p. 18 - 32. Soil and Plant Analysis Council, Inc. Athens, GA.

2. EPA, 1986. Test Methods for Evaluating Solid Waste, SW-846, Vol.1A. Chapter one: Quality Control.

3. Garfield, F. M. 1991. Quality Assurance Principles for Analytical Laboratories. Association of Official Analytical Chemists. Arlington, VA.

4. Ogden, L., P. Kane, P. Knapp, N. Thiex, L. Torma, 1998. Association of American Feed Control Officials Quality Assurance/Quality Control Guidelines for State Feed Laboratories. Association of American Feed Control Officials.

5. Peters, J, ed., 2003. Recommended Methods of Manure Analysis. Laboratory Quality Assurance Program: p. 5 - 11. University of Wisconsin Extension. Bulletin A-3769.

6. Soil Sci. Soc. Am. / NAPT, 2004. QA/QC Model Plan for Soil Testing Laboratories. SSSA and Oregon State University.

7. Schumacher, B.A., A.J. Neary, C.J. Palmer, D.G. Maynard, L. Pastorek, I.K. Morrison, and M. Marsh, 1995. Laboratory Methods for Soil and Foliar Analysis in Long-Term Environmental Monitoring Programs. Chapter 2: Good Laboratory Practices. EPA/600/R-95/077.

8. Taylor, J. K. 1987. Quality Assurance of Chemical Measurements. Lewis Publishers, Inc. Chelsea, MI.

9. Thiex, N., L. Torma, M. Wolf, and M. Collins, 1999. Quality Assurance Quality Control Guidelines for Forage Laboratories. National Forage Testing Association.

10. Western States Laboratory Proficiency Testing Program, Soil and Plant Analysis Methods, 1998, V4.10. Quality Assurance in the Agricultural Lab, p. 4-6.