♻EPA

Project Summary

Direct/Delayed Response Project: Quality Assurance Report for Physical and Chemical Analyses of Soils from the Mid-Appalachian Region of the United States

G. E. Byers, R. D. Van Remortel, M. J. Miah, J. E. Teberg, M. L. Papp,
B. A. Schumacher, B. L. Conkling, D. L. Cassell, and P. W. Shaffer

The Direct/Delayed Response Project was designed to address the concern over potential acidification of surface waters by atmospheric sulfur deposition in the United States. The purpose of these synoptic soil physical and chemical surveys was to characterize watersheds in regions of the United States believed to be susceptible to the effects of acidic deposition. This document describes the implementation of a quality assurance program and the verification of the analytical data base for the Mid-Appalachian Soil Survey. It is directed primarily towards the users of the data base who will be analyzing the data and making various assessments and conclusions relating to the effects of acidic deposition on watersheds of the region.

The results show that the measurement quality objectives for detectability, precision, accuracy, representativeness, and completeness were generally satisfied. Measurement uncertainty was generally low in relation to overall data uncertainty. A series of conclusions and recommendations are provided at the end of the report. The recommendations will be useful in the planning of future projects of this nature.

This Project Summary was developed by EPA's Environmental Monitoring Systems Laboratory, Las Vegas, NV, to announce key findings of the research project that is fully documented in a separate report of the same title (see Project Report ordering information at back).

Introduction

The U.S. Environmental Protection Agency (EPA), as a participant in the National Acid Precipitation Assessment Program, has designed and implemented a research program to predict the long-term response of watersheds and surface waters in the United States to acidic deposition. Based on this research, a sample of watershed systems will be classified according to the time scale in which each system will reach an acidic steady state, assuming current levels of acidic deposition. The Direct/Delayed Response Project (DDRP) was designed as the terrestrial component of the EPA Aquatic Effects Research Program.

The mapping for the DDRP Mid-Appalachian Soil Survey was conducted in portions of Pennsylvania, Virginia, and West Virginia during the spring and summer of 1988 and the sampling took place during the fall of 1988. These

activities initiated the third full-scale regional survey of the DDRP. The physical and chemical properties that were measured in the soil samples are listed on the next page.

Quality Assurance Program

Quality Assurance Optimization

As a result of conclusions and recommendations from the DDRP Northeastern region and Southern Blue Ridge Province soil surveys, a number of improvements were made in the quality assurance program for the Mid-Appalachian soil survey. A single preparation laboratory was established in Las Vegas, Nevada, in close proximity to the EPA Environmental Monitoring Systems Laboratory - Las Vegas (EMSL-LV) quality assurance (QA) staff. This allowed the development of, and adherence to, strictly defined sample preparation protocols and the ability to track and control progress at the laboratory on a real-time basis.

Mineral and organic samples were placed in separate batches because of differences in analyte concentrations and in the soil:solution ratios required for analysis. As a result, the analytical laboratories were able to perform instrument calibration and sample analyses within narrower linear dynamic ranges, allowing the QA staff to make more reliable assessments of the resulting data quality.

Several changes in analytical methods and procedures were initiated in the Mid-Appalachian survey. For example, the exchangeable aluminum in 1M potassium chloride parameter was replaced by exchangeable aluminum in 1M ammonium chloride as part of the measurement of exchangeable cations. This change in extraction solution reduced the analysis time for exchangeable aluminum while retaining similar experimental conditions. The quantity of mineral soil used in the exchangeable cation analyses was increased relative to the volume of extracting solution in order to increase the cation concentrations in solution which facilitated instrumental analysis. The amount of cellulose filter pulp used in extractions was decreased and a pulp prewash step was added to reduce contamination by exchangeable cations in the pulp material. The determination of extractable cations in 0.002M calcium chloride using a mechanical extractor was changed to a 3 overnight extraction using a mechanical shaker. Also, pH was measured on this extract rather than on a separate extract as used in previous surveys. Other changes included various solution modifications in the measurement of total exchangeable acidity and the use of vanadium pentoxide as a combustion accelerator in the determination of total carbon and nitrogen. Total carbon, nitrogen, and sulfur were analyzed by a single analytical laboratory in response to the Mid-Appalachian soil survey quality assurance requirements for tighter measurement quality and the specialized nature of the analytical instrumentation required for these analyses.

The statistical approach in the Mid-Appalachian survey made use of a balanced hierarchical design that allowed various components of measurement uncertainty to be estimated with respect to the larger population uncertainty. A step function technique was developed by the QA staff to evaluate data quality in terms of predefined objectives and in relation to the routine sample data.

An effort was made to use measurement quality samples in the field, at the preparation laboratory, and at the analytical laboratories in such a way as to provide optimal real-time control and evaluation of data quality. Of particular importance was the addition of field audit samples at the sampling phase to allow the estimation of sampling and system-wide measurement uncertainties.

A Laboratory Entry and Verification Information System (LEVIS) was developed for use by the QA staff and by the analytical laboratories. The LEVIS program facilitated the entry, edit, and review of raw data and the calculation of final data values. The program also performed verification checks for the measurement quality samples and produced QC summary reports.

Many significant changes in the batch acceptance criteria were initiated. Among them were the tightening of contract-required detection limits and precision requirements. Several new measurement quality samples were introduced to check both precision and accuracy. Acceptable accuracy windows for laboratory audit samples were established and used as contractual requirements for the laboratories. A template was also developed by the QA staff to assist in setting major and minor flags, and helped to eliminate the subjectivity present in the previous DDRP surveys in regard to reanalysis requests.

Data Quality Assessment

The quality assurance (QA) program for soil sampling, sample preparation, and sample analysis were designed to satisfy measurement quality objectives (MQOs) for the resulting data and to assess the variability of sampling, preparation, and analytical performance. The MQOs for this survey were primarily directed toward the attributes of detectability, precision, accuracy, and completeness. Representativeness and comparability of the data were also assessed, although quantitative MQOs were not imposed.

Detectability

The instrument detection limit is the lowest value that an analytical instrument can reliably detect above instrument background concentrations. During sample analysis, the overall instrument detection limits for a given parameter were defined as three times the pooled standard deviation of at least 15 nonconsecutive calibration blanks run on three separate days. Acceptable initial instrument detection limits were established prior to any sample analysis and subsequent values were determined and reported on a batch-by-batch basis. Contracts with the analytical laboratories specified maximum allowable instrument detection limits and, if a reported batch detection limit was invalid, the batch was reanalyzed for that parameter.

Secondary checks on the reliability of the instrument detection limits were made using independently determined values. These independent instrument detection limits were calculated for each parameter as three times the standard deviation of a series of low level quality control check samples.

System detection limits were estimated using low concentration field audit samples which reflected the uncertainty introduced during soil sampling, preparation, extraction, and analysis. These limits allowed data users to identify soil samples which had a measured concentration that was statistically different from the reagent or calibration blanks. The system detection limit for each parameter was calculated as three times the standard deviation of the low-range field audit samples.

Precision

Precision is the level of agreement among multiple measurements of the same soil characteristic. Measurement imprecision is distinct from the overall variability in the population itself. Determination of measurement imprecision and its sources in the Mid-Appalachian soil survey relied strongly on analysis of the measurement quality samples and was a function of the

Physical and Chemical Properties Measured in the Direct/Delayed Response Project
Mid-Appalachian Soil Survey

Air-dry moisture content	Na in 1.0M ammonium chloride	Ext. Fe in acid oxalate
Total sand	Al in 1.0M ammonium chloride	Ext. Al in acid oxalate
Very coarse sand	Ca in l.OM ammonium acetate	Ext. Si in acid oxalate
Coarse sand	Mg in 1.0M ammonium acetate	Ext. Fe in citrate dithionite
Medium sand	K in 1.0M ammonium acetate	Ext. Al in citrate dithionite
Fine sand	Na in 1.0M ammonium acetate	Ext. sulfate in deionized water
Very fine sand	CEC in 1.0M ammonium chloride	Ext. sulfate in sodium phosphate
Total silt	CEC in 1.0M ammonium acetate	Sulfate isotherm 0 mg sulfur
Coarse silt	Ex. Acidity by barium chloride-TEA	Sulfate isotherm 2 mg sulfur
Fine silt	Ca in 0.002M calcium chloride	Sulfate isotherm 4 mg sulfur
Total clay	Mg in O.002M calcium chloride	Sulfate isotherm 8 mg sulfur
pH in deionized water	K in 0.002M calcium chloride	Sulfate isotherm 16 mg sulfur
pH in 0.002M calcium chloride	Na in 0.002M calcium chloride	Sulfate isotherm 32 mg sulfur
pH in 0.0M calcium chloride	Fe in O.002M calcium chloride	Total carbon
Ca in 1.01M ammonium chloride	Al in 0.002M calcium chloride	Total nitrogen
Mg in 1.0M ammonium chloride	Ext. Fe in pyrophosphate	Total sulfur
K in 1.0M ammonium chloride	Ext. Al in pyrophosphate	

intralaboratory within-batch precision MQOs defined in the QA Plan. Overall variability stemming from measurement and population sources was estimated from the routine data.

The precision MQOs were a two-tiered system. Below a specific concentration, called the "knot", precision was defined as an absolute standard deviation in parameter reporting units; above the knot, precision was defined as a percent relative standard deviation. To address the issue of concentration-dependent variance, the range of soil analyte concentrations was partitioned into appropriate intervals within which the error variance was relatively constant. A step function was fitted across the intervals to represent the error variance for the entire concentration range. Different step functions were used to assess variability in selected measurement quality samples, and variability in the routine population of soil samples collected was also estimated.

Accuracy

Accuracy is the level of agreement between an observed value and the "true" value of a soil characteristic. Data from the laboratory audit samples were used to estimate analytical accuracy and

data from the field audit samples were used to assess accuracy from a system-wide measurement perspective. Each audit sample type was assigned a range of acceptable values for each parameter in the form of an accuracy window, which was derived from an MQO-based confidence interval placed around a weighted estimate of the mean calculated using analytical data from the previous DDRP surveys.

The three aspects of accuracy investigated were bias, laboratory differences, and laboratory trends. Analytical bias was considered to be the quantitative measure of accuracy used in the estimation of measurement uncertainty. Laboratory differences were assessed in relation to known reference values and, in conjunction with laboratory trends, served as quantitative and qualitative evaluations of analytical laboratory performance.

Representativeness

The representativeness objectives of the survey were qualitative and quantitative in nature. The general objectives were that: (1) the pedons sampled by the field sampling crews be representative of the soil sampling class characteristics, (2) the samples that were

collected would be homogenized and subsampled properly by the preparation laboratory personnel, and (3) the field duplicate samples adequately represent the range and frequency distribution of analyte concentrations found in the routine samples.

Completeness

The completeness objectives of the survey were to ensure that (1) all soil pedons designated for sampling were actually sampled, (2) all samples received by the preparation laboratory were processed, and (3) all samples received by the analytical laboratories were analyzed and that 90 percent or more of the required measurements were made on all of the samples. Enough data were provided to allow statistically significant conclusions to be drawn. Data qualifiers, or flags, for completeness were inserted in the data base to indicate any missing values.

Comparability

Comparability of data from the three DDRP surveys was approached as a complex issue having several levels of detail which should be considered. Level 1 comparability was established on the basis of statistical evaluation methods,

measurement quality samples, and the sample collection, preparation, and analysis methods used. Level 2 comparability was established by the acceptability and useability of the verified data bases as defined by the data users. Level 3 comparability allowed the direct quantitative comparison of data for each parameter of interest.

Uncertainty Estimates

The term "uncertainty" was used to describe the sum of all quantifiable sources of error associated with a given portion of the measurement system. Uncertainty estimates, or delta values, were calculated for each parameter using the square root of the sum of the within- and between-batch variances and squared bias term. Four delta values were calculated for each parameter. The $delta_1$, values represent analytical laboratory uncertainty and were estimated using the laboratory audit samples. The $delta_2$ values represent the confounded uncertainty of sample preparation and analysis, and were estimated using the preparation duplicates. The $delta_3$ values represent the confounded overall measurement uncertainty of field sampling, sample preparation, and sample analysis, and were estimated using the field duplicates. The $delta_4$ values represent uncertainty due to the spatial heterogeneity of the routine sample population confounded with the overall measurement uncertainty of field sampling, sample preparation, and sample analysis, and were estimated using the sampling class/horizon groups of routine samples. The sampling class/horizon groups refer to configurations of the samples that were believed to have similar physical and chemical properties in relation to soil responses to acid deposition.

Measurement Quality Samples

Quality evaluation (QE) samples were used to assess overall measurement uncertainty and to provide an independent check on the quality control (QC) procedures. The QE samples were known to the QA staff but were either blind or double-blind to the sampling crews and preparation or analytical laboratory personnel. Six types of QE samples were used in the Mid-Appalachian soil survey: (1) field duplicates (soil samples were collected by each sampling crew from one horizon of every third pedon sampled and were placed randomly in the sample batch with the other samples from the same pedon); (2) field audits (duplicate, median-

concentration, mineral soil samples or triplicate, median-concentration, organic soil samples were sent by the QA staff to the sampling crews for processing as if they were routine samples); (3) low-range field audits (low concentration, mineral soil samples were sent with the field audits to the sampling crews by the QA staff); (4) preparation duplicates (a pair of preparation duplicates, one split from the field duplicate sample and one split from its associated routine sample, was created at the preparation laboratory and placed randomly in the sample batch); (5) laboratory audits (duplicate, median range, mineral soil samples or triplicate, median-range, organic soil samples, identical to the field audits were sent by the QA staff to the preparation laboratory for inclusion in each batch); and (6) low-range laboratory audits (low concentration, mineral soil samples identical to the low-range field audits were sent to the preparation laboratory by the QA staff for inclusion with each mineral soil batch).

The composition of the QC samples was known to the analyst, and the analytical results from each laboratory were required to satisfy the batch acceptance criteria as the samples were analyzed. Immediate feedback on the functioning of the analytical system allowed sample processing and analytical deficiencies to be resolved quickly, resulting in minimal error from these sources. Nine types of QC samples were used in the Mid-Appalachian soil survey: (1) QC audit samples were median-concentration mineral soil audit samples provided directly to the analytical laboratories together with respective accuracy windows, and were used to control bias and reduce between-laboratory and between-batch measurement uncertainty; (2) analytical duplicates were splits of a single sample and were used to control analytical within-batch precision; (3) calibration blanks were used as a check for sample contamination, analytical carryover effects, and baseline drift in the analytical instrument immediately after calibration; (4) reagent blanks underwent the same treatment as the routine samples and served as a check for reagent contamination; (5) QC check samples contained the analyte of interest in the mid-calibration range and served as a check on the accuracy and consistency of the instrument calibration throughout the sample batch analysis; (6) detection limit GC check samples were low concentration samples that eliminated the necessity of determining the detection limit every day and allowed

accuracy to be estimated at the low end of the calibration range; (7) matrix spikes were sample aliquots to which known quantities of analyte were added for determining the sample matrix effect on analytical measurements; (8) adsorption spike solutions were either the extraction solution used in determining calcium chloride-tractable calcium or the extraction solution standards used for generation of the sulfur adsorption isotherms, and served as a check for reagent contamination; and (9) ion chromatography resolution samples contained known concentrations of sulfate, phosphate, and nitrate, and were used to provide evidence of acceptable peak separation for the sulfate analyses.

Technical system audits or on-site evaluations were also conducted. Each analytical laboratory underwent an audit after successfully analyzing a set of performance evaluation samples prior to receiving a contract. Second-round audits were performed after each laboratory had completed most of the analyses on two sample batches. During third-round auditing, two laboratories underwent two additional audits. The third laboratory underwent only one additional audit.

Data Verification

The analytical laboratories and the QA staff used LEGS for data verification. Phase one of the LEVIS program included the data entry component and two verification components. The analytical laboratories entered and evaluated data as it was produced using a QC summary report of batch data characteristics and sixteen soil chemistry relationships. The data were then transferred to the central computer at EMSL-LV. The QA staff evaluated preliminary and formally submitted batches from the analytical laboratories using precision and accuracy windows which were checked by LEVIS. The LEVIS program flagged data which did not meet the MQOs. Following the initial data evaluation for a batch, the QA staff prepared a summary document of all flagged parameters, indicating whether a flag originated from an unacceptable value for the QC, chemistry, precision, or accuracy criteria. The number and severity of the flags for each parameter were checked using the QA reanalysis template to determine if reanalysis was required.

After completion and receipt of all final batch data from the analytical laboratories, two internal consistency checks were performed to check for possible outliers in the routine data. A

4

correlation approach was used to assess internal consistency in which the coefficients of determination were obtained by performing weighted linear regressions. From the regressions, studentized residuals and difference of fit statistics were calculated to identify extreme data values that could be considered outliers. Approximately 3 percent of the data underwent confirmation using this consistency check. The second internal consistency check used the data structured into a pedon/horizon data set using the original soil profile sequence. Data for each pedon were visually scanned by soil scientists for consistency of the parameter values from one horizon to the next. Approximately 2 percent of the data underwent confirmation using this consistency check.

Data Management

The field sampling and sample preparation data were entered into SAS-AF TM raw data files on personal computers at EMSL-LV. The analytical data were entered into LEVIS on personal computers located at the laboratories. Data verification was accomplished by a systematic evaluation of completeness, precision, internal consistency, and coding accuracy. Apparent discrepancies were appended with flags unless they could be corrected. After verification was complete, the data bases were frozen and sent to the Oak Ridge National Laboratory in Tennessee to undergo validation in cooperation with personnel at the EPA Environmental Research Laboratory at Corvallis, Oregon, and at EMSL-LV. The validation procedures included a specific assessment of outlying data points for inclusion or omission in validated data sets based on assigned confidence levels.

Results and Discussion

Detectability

The calculated instrument detection limits were less than the contract-required limits for every parameter. Therefore, the MQOs for detectability in the Mid-Appalachian Soil Survey were completely satisfied. System detection limits, calculated from the low-range field audit samples, were used to assess system-wide detectability. Using a criterion of 80 percent or more of the routine sample concentrations exceeding the system detection limits as a basis for assessment, most of the parameters were suitable for all data uses throughout the concentration range. The six

exceptions were exchangeable calcium and sodium in ammonium acetate and ammonium chloride, iron extracted in calcium chloride or acid oxalate, silicon extracted in acid oxalate, and total sulfur. Data users should use caution when assessing these parameters, as significant portion of the routine sample concentrations were less than the corresponding system detection limits and may be difficult to distinguish in regard to overall system detection.

Precision

The analytical within-batch precision objectives were satisfied for most of the parameters. Occasionally, an objective was not satisfied for an upper or lower tier of a parameter, such as magnesium in calcium chloride or silicon in acid oxalate. When the two tiers were pooled over the total concentration range for each of the parameter groups, however, a precision index showed that all parameter groups satisfied the "overall" precision objectives.

The precision of the preparation duplicates was about the same, on the average, as the laboratory audit samples and satisfied the MQOs except for iron in acid oxalate. This indicates that the preparation laboratory performed very well in subsampling the bulk soil samples. For the field duplicates, only a few cases, such as calcium in ammonium acetate, exceeded the precision objectives. In addition, the relatively low precision for the field duplicates for some parameter groups suggests that the component of error from soil sampling, which includes spatial heterogeneity within the sampled horizons, is a large portion of the data collection error.

Accuracy

Accuracy in the Mid-Appalachian soil survey was evaluated by estimating analytical bias with respect to a reference value, defined as the mean of an accuracy window for a given parameter. Laboratory differences and trends were assessed by comparing mean values for the laboratories, combined across audit samples, to the pooled reference values.

Analytical bias was negligible when compared with the system detection limit for all parameters for which system detection limits were established. The only case of significant bias occurred for the cation exchange capacity in ammonium acetate for the two laboratories which used the titration procedure. Bias for the two newly-established parameters for the Mid-

Appalachian survey, i.e., aluminum in ammonium chloride and silicon in acid oxalate, were both much less than their respective detection limits. The percentage of observations that were outside the respective accuracy windows and the magnitude of their contribution to the overall bias estimates were also calculated. The results show a very wide range in the ratios of bias for values outside the window compared to the total analytical bias. Only 11 of the 50 parameters had values outside the window that contributed more than one-third to the bias: coarse and fine silt, magnesium and aluminum in calcium chloride, magnesium and cation exchange capacity in ammonium acetate, aluminum in sodium pyrophosphate and citrate dithionite, iron and silicon in acid oxalate, and the 32 mg S/L sulfate isotherm parameters.

Laboratory differences were expressed as a percentage from the reference value for each laboratory for each parameter. When compared across laboratories, all laboratory differences were ten percent or less except for clay, potassium and sodium in ammonium acetate, and iron and aluminum in calcium chloride. In all five cases, the concentrations were very low. When combined into parameter groups, all laboratory differences were six percent or less except for cations in ammonium acetate for Laboratory 1 and cations in ammonium chloride for Laboratory 4.

Approximately 54 percent of the parameters (25 of 46 parameters) showed significant laboratory differences by Scheffe's pairwise comparison test. Four of the 25 parameters (total silt, pH in 0.002M calcium chloride, cation exchange capacity in ammonium acetate, and silicon in acid oxalate) showed all three laboratories to be significantly different from each other. There were no general trends for any specific laboratory for most of the parameter groups. For the cations in calcium chloride, however, Laboratory 2 was significantly different from the other laboratories in all cases shown and Laboratory 4 was significantly lower for the sulfate isotherm parameters than the other laboratories in the four cases that showed differences.

Moving averages of the laboratory audit samples were plotted to identify situations when a particular laboratory showed an upward or downward trend over time for a given parameter. Generally, the trends did not show extreme divergence with respect to the accuracy window acceptance criteria. However, certain data users may find the

5

trends to be noteworthy for a specific data analysis.

No single laboratory was consistently superior to the others for all parameters or parameter groups regarding low differences. Each laboratory appears to have individual strengths for specific analytical methods. This is probably a reflection of the combination of experience, instrumentation, and laboratory management practices within each laboratory, resulting in a patchwork of differences on a parameter group basis.

Representativeness

All pedons sampled were within the range of morphological characteristics outlined in their respective sampling classes. The homogenization and subsampling procedures at the preparation laboratory produced representative analytical soil samples of known and accepted quality. Overall precision for the preparation duplicates was approximately equivalent to that of the laboratory audit samples, hence, the procedures were shown to be suitable for creating representative subsamples.

The analyte concentrations in the field duplicates generally were representative of the range and frequency distribution of routine sample concentrations. The exceptions were seen mainly in the cations and extractable sulfate parameters and were mostly representative of the concentration range, but not of the distribution within the range. Analyte concentrations in the preparation duplicates were generally representative of the corresponding field duplicates. The audit samples, as expected, were usually representative only of the overall range of data from the routine samples.

Completeness

Sampling of the specified pedons had a completeness level of 100 percent, and processing was accomplished for 100 percent of the field samples satisfying the sample receipt criteria at the preparation laboratory Analytical completeness in the verified data base was 100 percent for all parameters. Sufficient validated data were generated to make conclusions for each parameter in the Mid-Appalachian survey data bases, with a completeness level of 98 percent or higher for all parameters.

Comparability

The statistical methods were comparable among the three DDRP surveys. Additional audit samples were used to optimize QE and QC activities in the Mid-Appalachian survey although they did not directly affect the comparability of the data from the surveys.

In most cases, the sampling, preparation and analytical methods and protocols for the three surveys were comparable, although there were several parameter protocol modifications in the Mid-Appalachian survey. The only modification which may affect comparability was the change of aluminum extractant from 1M potassium chloride to 1M ammonium chloride. As a result of more stringent MQOs in the Mid-Appalachian survey, overall measurement quality generally was better in the Mid-Appalachian data bases. Therefore, it is possible that some of the Northeastern region and Southern Blue Ridge Province data may not be suitable for the same data analysis procedures performed on the Mid- Appalachian data without some form of caveat. Two examples are the organic soil sulfate isotherm data and the specific surface data from the Northeastern and Southern Blue Ridge Province surveys which are generally considered to be suspect by data users.

Significant differences among the surveys were identified using the median-concentration B mineral soil audit sample common to the three surveys. Only nine parameters showed differences among surveys at the 0.01 significance level: aluminum extracted with potassium chloride versus ammonium chloride, water-extractable sulfate, the 0,2, and 4 mg S/l sulfate isotherms, calcium in calcium chloride, iron in citrate dithionite, total carbon, and coarse silt. Data users should exercise caution when using data from the nine parameters exhibiting significant differences.

As part of the DDRP, an interlaboratory methods comparison study was conducted which compared soil analysis data from two laboratories using the DDRP methods to 16 statistically randomly-chosen external laboratories. These laboratories used their own methods which were similar, but not identical to, the DDRP methods. The results of the study show the generally good comparability of the DDRP data with data from other independent laboratories.

Uncertainty Estimates

Within-batch imprecision estimates increased, as expected, from analytical ($delta_1$, values) to sample preparation ($delta_2$ values) to field sampling ($delta_3$ values) to sampling class/horizon groups ($delta_4$ values). The between-batch precision estimates were generally low.

The overall measurement uncertainty in the routine samples was based on the $delta_3$ values which were estimated from the field duplicate samples. The $delta_3$ values, in relation to the associated $delta_4$, values, were used to provide the data users with a basis for assessing the contribution of the measurement system to the data uncertainty. Using this procedure, measurement uncertainty was negligible for 90 percent (45 of the 50 parameters measured) of the data. The five exceptions were the fine silt and coarse silt fractions which had no strict precision MQOs, sodium in both calcium chloride and ammonium acetate, and silicon in acid oxalate. The sodium and silicon parameters showed high relative measurement uncertainty due to the inordinate effect of measurement error in the low concentration range near the detection limit.

Conclusions and Recommendations

As a result of previous conclusions and recommendations from the DDRP Northeastern and Southern Blue Ridge Province surveys, a number of improvements were made in the quality assurance program for the Mid-Appalachian survey. The quality assurance data are presented in a manner considered to be the most appropriate for use by the primary data users. The data quality evaluation procedures and the report format resulted from regular interactions with the data users, and the assessment by several external reviewers. Each user has a subjective concept of data quality as well as a knowledge of the specific level of data quality required for his/her own use. The users are therefore encouraged to become familiar in detail with the text, figures, and tables to facilitate the identification of data satisfying their specific requirements.

The consolidation of sample preparation facilities at a single laboratory facilitated quality assurance of the samples from the field sampling through the sample analysis phase. The separation into different batches of mineral and organic samples should be continued. In addition, measurement quality samples should continue to be distributed among batches and analytical laboratories in such a way as to provide a balanced design for assessment purposes.

The use of a computerized data entry and verification system which allowed the calculation of final data values and produced a list of flags and data entry errors for each sample batch greatly facilitated the verification and reanalysis decision-making process. A similar program should be tailored for each future survey, with the addition of a method that will identify and confirm outlying data points for real-time control purposes.

Considerable effort was expended throughout the three surveys to improve detectability of various parameters and significant improvement was made for the exchangeable cations and sulfur parameters. Additional methods research is needed to improve detectability further. In addition, both instrument and system-wide detectability should be defined in the MQOs. As part of the implementation of such MQOs, it is recommended that low concentration audit samples, entered into the system during the sampling phase, continue to be utilized as substitutes for soil blank samples.

The precision results indicate that the analytical within-batch precision objectives were satisfied in most cases. Occasionally an objective was not satisfied for an upper or lower tier of a parameter. It is recommended that the lower and upper tier precision objectives for exchangeable aluminum in ammonium chloride and extractable silicon in acid oxalate be modified as described in the report. The preparation and field duplicate samples generally satisfied the precision objectives. A precision index of parameter groups revealed that all groups were of acceptable precision.

Audit sample accuracy windows were developed for use in the Mid-Appalachian survey from data collected in the previous two surveys. The use of a quality control audit sample should be continued in future surveys. In addition

liquid audit samples should be incorporated into the quality assurance program and be used to differentiate soil extraction error from instrument error. Emphasis should be placed on the use of audit sample control charts by the QA staff to identify abnormal scatter outside the accuracy windows during the batch analysis.

No single laboratory was consistently superior to the others for all parameters or parameter groups. To control inter-laboratory differences in future surveys, it is recommended to continue the selection of a specific laboratory to perform analysis on a parameter basis for those parameters or parameter groups that reveal inherently high differences of where specialized instrumentation is used, e.g., total elemental analysis. Also, a stringent performance evaluation process should be continued to select the best available contract analytical laboratories.

Quality is a continuum and the need for improved data quality dictates the data quality objectives chosen. These objectives may or may not be attainable with the current technology. It is recommended that the analytical procedures for specific parameters be revised, tested, modified, and clarified where appropriate.

In evaluating representativeness of the quality assurance samples, it is evident that the field duplicates and pre-paration duplicates were representative of the routine sample concentration range for most parameters. It is recommended that the field sampling and preparation laboratory protocols continue to specify statistically valid methods for selecting the field and preparation duplicates. The method should be reiterated to the sampling and laboratory personnel during the pre-sampling training session. The QA field auditor should ensure that a sufficient amount of soil is collected for each bulk sample in the field to allow a

preparation duplicate to be subsampled.

Data comparability across the three surveys was generally good. It is recommended, however, that a methods comparison be performed for the two soil extraction methods used for exchangeable aluminum, i.e., 1M ammonium chloride and 1M potassium chloride.

The step function statistical approach has been shown to be an effective procedure for evaluating measurement quality issues in environmental data spanning a wide concentration range. It is recommended that additional research and development be undertaken to identify an optimal step function procedure that is fully compatible with the measurement quality sample design.

G. E. Byers, R. D. Van Remortel, M. J. Miah, J. E. Teberg, M. L. Papp, and B. A. Schumacher are affiliated with the Lockheed Engineering & Sciences Company, Las Vegas, Nevada. *B. L. Conkling* is affiliated with the Environmental Research Center of the University of Nevada at Las Vegas. *D.L. Cassell and P.W. Shaffer* are affiliated with the NSI Technology Services Corporation, Corvallis, Oregon.

L. J. Blume and D. T. Heggem are the EPA Project Officers (see below).

The complete report, entitled *"Direct/Delayed Response Project: Quality Assurance Report for Physical and Chemical Analyses of Soils from the Mid-Appalachian Region of the United States,"* (Order No. PB 204710/AS; Cost: $39.00, subject to change) will be available only from:

> National Technical Information Service
> 5285 Port Royal Road
> Springfield, VA 22161
> Telephone: 703-487-4650

The EPA Project Officers can be contacted at:

> Environmental Monitoring Systems Laboratory-Las Vegas
> U.S. Environmental Protection Agency
> Las Vegas, NV 89193-3478

United States
Environmental Protection
Agency

Center for Environmental Research
Information
Cincinnati OH 45268

Official Business
Penalty for Private Use $300

EPA/600/S4-90/001