



Project Summary

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

G. E. Byers, R. D. Van Remortel, J. E. Teberg, M. J. Miah, C. J. Palmer, M. L. Papp, W. H. Cole, A. D. Tansey, D. L. Cassell, and P. W. Shaffer

The Northeastern soil survey was conducted during 1985 as a synoptic physical and chemical survey to characterize a statistical sampling of watersheds in a region of the United States believed to be susceptible to the effects of acidic deposition. This document addresses the implementation of a quality assurance program and the verification of the analytical data base for the Northeastern Soil Survey. It is focused 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 the soils of the Northeastern region of the United States. The quality assurance program is evaluated in terms of its success in identifying potential problems that could have an effect on the quality of the data. Verification procedures used to analyze laboratory data are described. Quality is assessed by describing the detectability, precision, accuracy (interlaboratory differences), representativeness, completeness, and comparability of the data for the quality assurance samples used throughout the soil survey. Detection limits and two-tiered precision data

quality objectives were established for most of the parameters. A step-function statistical approach was used to assess precision.

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 Northeastern Soil Survey was conducted during the Spring and Summer of 1985

and the sampling took place during the Summer and Fall of 1985. These activities initiated the first full-scale regional survey of the DDRP. The physical and chemical properties that were measured in the soil samples are listed below.

Data Quality Objectives and Assessment

The quality assurance (QA) program, soil sampling techniques, sample preparation, and analytical methods, were designed to satisfy the data quality objectives (DQOs) for field and analytical data and to assess the variability of sampling, preparation, and analytical performance. The DQOs for this survey were directed toward detectability, precision, and completeness. Accuracy, representativeness, and comparability of the data were also assessed, although specific DQOs were not imposed on these attributes.

Detectability

The instrument detection limit is the lowest value that an analytical instrument can reliably detect above instrument background concentrations and was calculated for each analyte as three times the standard deviation of a low level quality control check sample, under the assumption that its variability would approximate the variability of an analytical blank sample. The system detection limit indicates the variability of a blank sample resulting from sample collection, processing, extraction, and analysis. The system detection limit was estimated for each parameter as the variability in the ten percent of field duplicates that had the lowest concentration of the analyte of interest. Contracts with the analytical laboratories established maximum allowable instrument detection limits, defined as contract-required detection limits; however, no DQOs were established for system detectability.

Precision

Precision is the degree of scatter of repeated measurements. Measurement imprecision is distinct from the overall variability in the population itself. Determination of measurement imprecision and its sources in the Northeastern Soil Survey relied strongly on the analyses of the QA samples and was a function of the intralaboratory within-batch precision DQOs defined in the QA Plan. Overall variability (measurement and population) was

estimated from the routine data. No DQOs were established for the sampling or preparation phases of the survey.

The precision DQOs were characterized by use of a two-tiered system. Below a specific concentration, called the knot, precision is defined as a standard deviation in absolute reporting units; above the knot, precision is defined as a percent relative standard deviation. In order to address the issue of concentration-dependent variance, the range of soil analyte concentrations was divided into appropriate intervals (windows) within which the error variance was relatively constant. A step function was fitted across the windows to represent the error variance for the entire concentration range. Different step functions were used to assess the variability for each QA sample type. The data uncertainty in the routine samples due to collection error was also measured.

Accuracy (Interlaboratory Differences)

Accuracy is the ability of a measurement system to approximate a true value. Accuracy could not be determined because the audit samples used in the survey were natural soil samples with chemical composition and physical properties that were not known with any confidence. Due to this lack of acceptable values for accuracy estimates, the data were assessed only for interlaboratory differences.

Three types of comparisons were made: (1) the use of a pair-wise statistical test for significance between laboratories; (2) the pooling of audit sample data for each laboratory for direct interlaboratory comparisons; and (3) the pooling of laboratory data for each audit sample for comparison of laboratory performance among audit samples.

Representativeness

The representativeness objectives of the survey were qualitative in nature. The general objectives were that: (1) the soil samples collected by the field crews be representative of the soil sampling class characteristics, (2) the samples be homogenized and subsampled properly by the preparation laboratory, and (3) the QA duplicate samples adequately represent the range of analyte concentrations found in the routine samples.

Completeness

The 100 percent completeness objectives of the survey were to determine whether (1) all pedons designated for sampling were actually sampled by the field crews, (2) samples received by the preparation laboratories were prepared and analyzed, and (3) all samples received by the analytical laboratories were analyzed. That 90 percent of the required measurements were made on samples. Enough data were provided to allow statistically significant conclusions to be drawn. Data qualifiers, or flags, for completeness were inserted in the database to indicate any missing values.

Comparability

Data comparability objectives were qualitative in nature. The goal was comparability of data from the surveys within the DDRP and for DDRP surveys to be comparable to other similar programs. The stated objective was the uniform use of known, accepted and documented procedures for location, soil sample collection, sample preparation, extraction, analysis, and standard reporting units for the Northeastern Soil Survey. Uniform QA/QC protocols and on-site inspections assured that these procedures were implemented properly. The result was that analytical data should be comparable to those from other surveys.

Quality Assurance and Quality Control Samples

Quality assurance samples were used to independently assess data quality and to monitor the internal quality control procedures. The composition and identification of the QA samples were unknown to the analyst. Three types of QA samples were used in the Northeastern Soil Survey: field duplicates (soil samples were collected by each sampling crew from one horizon of one pedon each day and were placed randomly in the same batch with the other samples from the same pedon); (2) preparation duplicates (one soil sample per sample batch was selected and split by the preparation laboratory and placed randomly in the sample batch); and (3) natural audit samples (one duplicate sample pair from a homogenized bulk soil sample representing one of five soils typical of the eastern United States were placed randomly in the sample batch).

The composition of QC samples was known and the analytical results from each laboratory were required to

Physical and Chemical Properties Measured in the Direct/Delayed Response Project Northeastern Soil Survey

Air-dry Moisture Content
Specific Surface Area
Total Sand
Very Coarse Sand
Coarse Sand
Medium Sand
Fine Sand
Very Fine Sand
Total Silt
Coarse Silt
Fine Silt
Total Clay
pH in Deionized Water
pH in 0.002M CaCl₂
pH in 0.01M CaCl₂

Ca in 1.0M Ammonium Chloride
Mg in 1.0M Ammonium Chloride
K in 1.0M Ammonium Chloride
Na in 1.0M Ammonium Chloride
Ca in 1.0M Ammonium Acetate
Mg in 1.0M Ammonium Acetate
K in 1.0M Ammonium Acetate
Na in 1.0M Ammonium Acetate
CEC in 1.0M Ammonium Chloride
CEC in 1.0M Ammonium Acetate
Ex. Acidity by 1.0M KCl
Ex. Acidity by BaCl₂-TEA
Ext. Aluminum in 1.0M KCl
Ca in 0.002M Calcium Chloride
Mg in 0.002M Calcium Chloride
K in 0.002M Calcium Chloride
Na in 0.002M Calcium Chloride
Fe in 0.002M Calcium Chloride

Al in 0.002M Calcium Chloride
Ext. Fe in Pyrophosphate
Ext. Al in Pyrophosphate
Ext. Fe in Ammonium Oxalate
Ext. Al in Ammonium Oxalate
Ext. Fe in Citrate Dithionite
Ext. Al in Citrate Dithionite
Ext. Sulfate in Deionized Water
Ext. Sulfate in Sodium Phosphate
Sulfate Isotherm 0 mg sulfur/L
Sulfate Isotherm 2 mg sulfur/L
Sulfate Isotherm 4 mg sulfur/L
Sulfate Isotherm 8 mg sulfur/L
Sulfate Isotherm 16 mg sulfur/L
Sulfate Isotherm 32 mg sulfur/L
Total Carbon
Total Nitrogen
Total Sulfur

compared with the accepted values as the samples are analyzed. This immediate feedback on the functioning of the analytical system allowed analytical and sample processing problems to be resolved quickly, with the result that error from that source was minimized. Six types of QC samples were used in the Northeastern Soil Survey: (1) calibration blanks were used as a check for sample contamination and for baseline drift in the analytical instrument immediately after calibration; (2) reagent blanks underwent the same treatment as the routine samples and served as a check for reagent contamination; (3) 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 calibration of the instrument throughout the analysis of the sample batch; (4) detection limit QC check samples were low concentration samples that eliminated the necessity of determining the detection limit every day and allowed accuracy to be determined at the low end of the calibration range; (5) matrix spikes were sample aliquots to which known quantities of analyte are added for determining the sample matrix effect on analytical measurements; and (6) analytical duplicates were splits of a single sample and were used to estimate analytical within-batch precision.

In addition to the use of QC samples for quality control, two system audits, or on-site evaluations were also conducted, one immediately after award of the contract to the laboratories and the second after sample analysis had begun.

Internal Consistency Checks

An internal consistency computer program provided a meaningful check of routine data by identifying values that differed from the majority of observed values and that might have gone unnoticed had the check not been made. The checks uncovered errors in data entry and transcription as well as errors that occurred on an analytical batch basis. In this study, a 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 DFFITS statistics were calculated to identify extreme data values that could be considered outliers.

Outliers determined by the computer program and representing only about 1 percent of the data, were checked for transcription errors. For a few analyses, a significant number of outliers were present. There were some parameters that did not correlate well with any of the other parameters.

Data Management

The field sampling and analytical data were entered into the Northeastern survey data bases at Oak Ridge National Laboratory in Tennessee. Both data bases progressed through three phases: raw, verified, and validated. The QA staff at the EPA Environmental Monitoring Systems Laboratory at Las Vegas, Nevada, verified the two data bases. The field sampling data were entered into data sets from specialized forms, visually

checked, and frozen as the official raw data base. The analytical data were entered into data sets and visually checked, thereby allowing errors in transcription to be identified corrected. The verification stage was accomplished by a systematic evaluation of completeness and coding accuracy, and flags were used in the data base to note discrepancies. The verified data base was used to assess data for this QA report. The validation stage identified, confirmed, and flagged data values that warrant special attention or caution when used for data analysis.

Results and Discussion

Detectability

The calculated instrument detection limits were less than the contract-required detection limits in the majority of the cases of parameters for which detection limits were established. The DQOs established for detectability were not met for the cation exchange capacity parameters in ammonium chloride and ammonium acetate, aluminum in potassium chloride, total carbon, and total nitrogen. The calculated instrument detection limits were an order of magnitude above the DQOs for the cation exchange parameters.

Precision

The analytical within-batch precision objectives were satisfied for most of the parameters. These included clay, the pH parameters, exchangeable cations, cation exchange capacity and acidity,

extractable iron and aluminum, the extractable sulfates and sulfate isotherm parameters, and total carbon, nitrogen, and sulfur. Occasionally, an objective was not met for an upper or lower tier of a parameter. There were some instances where the DQOs were slightly exceeded for these parameters either above or below the knot. The DQOs were not met for total sand and silt, extractable cations in calcium chloride, and sulfate zero isotherm parameters. When the two tiers were pooled over the total concentration range for each of the parameters, a precision index showed that the particle size and extractable cations in calcium chloride did not meet the overall DQOs.

Although no DQOs were set for the preparation laboratory phase, the preparation laboratory within-batch precision also met the DQOs set for analytical within-batch precision for all parameters except sand and silt, calcium and magnesium in ammonium acetate, the two cation exchange capacities, acidity in barium chloride, cations in calcium chloride, extractable sulfate in water, sulfate zero isotherm, and total nitrogen. This indicates that the preparation laboratories performed relatively well in subsampling the bulk samples.

Within-batch imprecision estimates increased, as expected, from analytical to sample preparation to field sampling. The between-batch precision estimates were generally low.

Accuracy (Interlaboratory Differences)

The Scheffe's pair-wise multiple comparison test showed that about 5 percent of the interlaboratory differences were significantly different and 0.8 percent were highly significantly different. The latter could be considered to be of concern to the data user. Highly significant differences were shown mostly for the cations in calcium chloride. About half the cases were in the Bw audit sample.

The lowest interlaboratory differences were shown for pH and the highest differences were shown for the cation exchange capacity parameters, calcium in calcium chloride, and cations in ammonium chloride. The mean interlaboratory differences for laboratories 1,2,3, and 4 were 9.5, 9.3, 11.6, and 6.9 percent, respectively.

Among the audit samples, the laboratories showed the highest differences overall for the C horizon audit sample and the lowest differences for the

A audit sample. The mean interlaboratory differences for the Oa, A, Bs, Bw, and C audit samples were 13.8, 9.3, 14.5, 10.7, and 15.8 percent, respectively.

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. This resulted in a patchwork of differences on a parameter group basis.

Representativeness

The field duplicates were representative of the range of concentration in the range of the routine samples for most parameters. The natural audit samples and the preparation duplicates were rarely representative of the routine samples. The preparation duplicates consistently represented only the extreme lower range of routine sample concentrations for a given parameter. The preparation laboratory personnel apparently selected bulk samples for duplication based not on random selection but on the quantity of each sample available. The sampling crews were able to collect the largest amount of sample from the thicker horizons normally found in the lower portion of the soil pedon. Horizon type selection appears to be highly skewed (73 percent) toward the transitional B and C horizons which typically have very low analyte concentrations in their extractions.

Completeness

Ninety-six percent of the designated pedons were sampled. Although this does not fully satisfy the DQO of 100 percent for sampling completeness, a sufficient number of pedons were sampled to enable estimates and conclusions to be drawn from the data. The requested soil sampling and sample processing tasks were performed on 100 percent of the samples received by the preparation laboratory which satisfies the DQO of 100 percent. The analytical completeness level exceeded 99 percent for all parameters. Sufficient data were generated to make conclusions for each parameter in the data bases, with the possible exception of iron in the calcium chloride extract.

Comparability

The verified data bases were used for the assessment of data quality for the Northeastern and the Southern Blue Ridge Province soil surveys. The data bases from each survey therefore can be compared to each other. Flags were applied consistently.

Sufficient audit sample data were available from the DDRP control laboratory analyses to provide an estimate of the audit sample composition for use in the assessment of precise interlaboratory differences and comparability. Data from each audit sample can be compared between the two surveys for any given parameter. Significant differences can thus be attributed to differing amounts of measurement error. Reanalyses have corrected all data significantly affected by methods amendments which occurred during the survey progression.

Identical soil preparation methods were used in preparing soil samples for the two surveys. The procedure for selecting the preparation duplicate was refined during the Southern Blue Ridge survey, resulting in better representativeness of the preparation duplicate.

Due to an inconsistent application of the sampling of the field duplicates in the Northeastern survey, the variances of the field duplicates tend to fluctuate among the pedons. Overall within-batch variability was greater in the Southern Blue Ridge survey than in the Northeastern survey. This suggests that the measurement error in the Northeastern survey may have been somewhat underestimated. However, this does not mean that the routine data between regions are not comparable. The same methodology was used in the routine soil sampling for each survey. There were no deviations from the sampling protocols that would compromise the integrity of the routine data. These field sampling discrepancies that could affect data comparability were documented during the Northeastern survey and were resolved. The field sampling audit team did not report any deviations from the sampling protocols that would compromise the integrity of the routine data for the Southern Blue Ridge Province Soil Survey.

As part of the DDRP, an interlaboratory methods comparison study was conducted which compared the analysis of soils for two laboratories using the DDRP methods to 16 statistically chosen external laboratories. These laboratories used their own methods which were

similar, but not identical, to the DDRP methods. The results of the study will show the comparability of the DDRP data with other similar surveys using other laboratories.

Conclusions and Recommendations

Analysis of data from the Northeastern Soil Survey indicate areas where improvement is needed in the QA program. The quality assurance data are presented in a manner considered to be the most appropriate for use by the primary data users. The development of this approach resulted from regular interactions with the data users. In addition, the statistical approaches taken and the formats used were assessed in depth by several external reviewers to ensure agreement in the final presentations. A great deal of information is included in the many figures and tables. Each user has a subjective conception for data quality as well as a need for a specific level of data quality desired for his/her own use. The user is therefore encouraged to become familiar in detail with the text, figures, and tables in order to best assess the data for his/her specific needs.

A computerized data entry and verification system should be developed that will calculate the final data values and produce a list of flags and data entry

errors. A computer link between the laboratories and the quality assurance staff should be established that will enable the transfer of preliminary and final data. The verification program should be designed to evaluate the quality control checks and other contractual requirements, thereby inducing the laboratories to assume much of the responsibility for identification and correction of errant data. Evaluation of the blind audit samples should be made part of the computer verification system and a better procedure for checking these values should be developed. The internal consistency checks should become an integral part of the verification process.

Attention should be given to improving detectability in future surveys. Both instrument detection and system detection limits should be addressed in the data quality objectives for future surveys.

The DQOs for total sand and total silt should be increased from 1.0 percent to 3.0 percent. A two-tiered precision objective should be defined for the extractable cations in calcium chloride. Specific data quality objectives should be defined for system wide measurement.

Low concentration audit samples, entered into the system during the sampling phase, should be utilized as substitutes for soil blank samples. A quality control soil audit sample should

be incorporated into the quality assurance program to better monitor the analytical results of the laboratories. The laboratories should be required to report the analytical results of the analyses on a batch-by-batch basis to the QA staff immediately after the analysis of each batch. The laboratory protocols should specify a statistically valid method for selecting the preparation duplicate.

An effort should be made to locate an uncontaminated filter material for the determination of the basic cations or modify the pretreatment procedure for the filter material used. Additional methods details should be reviewed and provided where appropriate in order to reduce within-laboratory analytical variability.

The DDRP staff should consider the possibility of choosing laboratories to perform analyses on a parameter basis for those parameters or parameter groups that revealed inherently high differences or where specialized instrumentation is used, e.g., total carbon, nitrogen, and sulfur. A more stringent laboratory selection procedure should be adopted in the pre-evaluation process for the selection of contract laboratories.

The issue of accuracy should be addressed because the current approach using interlaboratory differences has limited utility. Data quality objectives for accuracy should be established.

G. E. Byers, R. D. Van Remortel, J. E. Teberg, M. J. Miah, M. L. Papp, W. H. Cole and A. D. Tansey are with Lockheed Engineering and Sciences Company, Las Vegas, NV 89119; C. J. Palmer is with Environmental Research Center, University of Nevada, Las Vegas, NV 89114; and D. L. Cassell and P. W. Shaffer are with NSI Technology Services Corporation, Corvallis, OR 97333.

L. J. Blume is the EPA Project Officer (see below)

The complete report, entitled "Direct/Delayed Response Project: Quality Assurance Report for Physical and Chemical Analyses of Soils from the Northeastern United States," (Order No. PB90-219 395/AS; Cost: \$31.00, 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 Officer can be contacted at:
Environmental Monitoring Systems Laboratory
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/S4-89/037



U.S. OFFICIAL MAIL

PENALTY
FOR
PRIVATE
USE \$300
* METER
6090444

U.S. POSTAGE

0.35

000085833 PS
U S ENVIR PROTECTION AGENCY
REGION 5 LIBRARY
230 S DEARBORN STREET
CHICAGO IL 60604