SUMMARY OF THE NATIONAL ATMOSPHERIC DEPOSITION PROGRAM/NATIONAL TRENDS NETWORK INTERSITE-COMPARISON PROGRAM, NOVEMBER 1978-NOVEMBER 1989

By John D. Gordon, Timothy C. Willoughby, and LeRoy J. Schroder

---

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 91-4061

Denver, Colorado
1991

U.S. DEPARTMENT OF THE INTERIOR

MANUEL LUJAN, JR., Secretary

U.S. GEOLOGICAL SURVEY

Dallas L. Peck, Director

For additional information
write to:

Chief, Branch of Quality Assurance
U.S. Geological Survey
Box 25046, Mail Stop 401
Federal Center
Denver, CO 80225-0046

Copies of this report can
be purchased from:

U.S. Geological Survey
Books and Open-File Reports Section
Box 25425, Mail Stop 517
Federal Center
Denver, CO 80225-0425

CONTENTS

FIGURES

TABLES

CONVERSION FACTORS AND ABBREVIATED WATER QUALITY TERMS

Multiply	*By*	*To obtain*
liter (L)	1.0567	quart, liquid
milliliter (mL)	0.03381	ounce, fluid

The following terms and abbreviations also are used in this report:

microsiemens per centimeter at 25 degrees Celsius (μS/cm)
megohm (MΩ)

SUMMARY OF THE NATIONAL ATMOSPHERIC DEPOSITION PROGRAM/NATIONAL TRENDS NETWORK INTERSITE-COMPARISON PROGRAM, NOVEMBER 1978-NOVEMBER 1989

By John D. Gordon, Timothy C. Willoughby, and LeRoy J. Schroder

ABSTRACT

Intersite-comparison studies have been used since November 1978 to assess the accuracy and precision of pH and specific-conductance measurements made by National Atmospheric Deposition Program and National Trends Network site operators. Between November 1978 and June 1980, the first four intersite-comparison studies were administered by the Illinois State Water Survey Central Analytical Laboratory, Champaign, Illinois. The U.S. Geological Survey assumed responsibility for the program in October 1981. Twenty-four intersite-comparison studies were completed as of November 1989. The pH and specific-conductance measurements made during intersite-comparison studies were primarily evaluated using two methods of analysis. The first method of analysis evaluated site-operator reported values in terms of the measurement-accuracy criteria established by the Network Operations Subcommittee. The measurement-accuracy criteria for pH are ±0.10 unit if the most probable pH value is 5.0 or less or ±0.30 unit if the most probable pH value exceeds 5.0. In every intersite-comparison study except study 22, the median pH was less than 5.0. The measurement-accuracy criteria for specific conductance are ±4.0 microsiemens per centimeter of the most probable value. The second method of analysis evaluated the differences between site-operator reported values and network median values.

The percentage of site operators reporting pH values within the measurement-accuracy criteria established by the Network Operations Subcommittee increased from an average of 61 percent for intersite-comparison studies 1 through 10 to an average of 78 percent for studies 11 through 24. The operators of 12 percent of the sites that participated in studies 1 through 24 accounted for 29 percent of the pH determinations that failed to meet the measurement-accuracy criteria.

Excluding study 22 because of the different pH measurement-accuracy criteria that were applied, a nonparametric correlation analysis indicated a statistically significant association between the median pH of the reference solution and the percentage of site operators reporting pH values that met the Network Operations Subcommittee measurement-accuracy criteria in studies 11 through 24 (α=0.05, r=-0.71). No statistically significant correlation between the median pH of the reference solution and the percentage of site operators reporting pH values that met the measurement-accuracy criteria was determined for studies 1 through 10.

The percentage of site operators reporting specific-conductance values within the measurement-accuracy criteria varied from 88 to 97 percent when the median specific conductance of the reference solution was less than 25 microsiemens per centimeter at 25 degrees Celsius. When the median specific-conductance value of the reference solution was between 30 and 60 microsiemens per centimeter at 25 degrees Celsius, 38 to 82 percent of the site operators reported values that met the measurement-accuracy criteria.

A nonparametric correlation analysis indicated a statistically significant association between the median specific conductance of the reference solution and the percentage of site operators reporting specific-conductance values that met the measurement-accuracy criteria in studies 1 through 24 (α=0.05, r=-0.82).

INTRODUCTION

The Association of State Agricultural Experiment Stations established the National Atmospheric Deposition Program (NADP) in 1978 to study atmospheric deposition and its effects (Galloway and others, 1978). In 1980, Congress passed the Acid Precipitation Act (Public Law 96-294) authorizing the creation of a National Acid Precipitation Assessment Program (NAPAP). The National Trends Network (NTN) was established in 1980 by a NAPAP interagency task force (Schroder and Malo, 1984). The NADP and NTN share common operational procedures and site-location criteria. The NADP and NTN also use the same analytical laboratory to assess wet-deposition sample chemistry. Since 1982, a Network Operations Subcommittee (NOS) has coordinated the operation of both the NADP and the NTN. Data from the NADP and NTN are compiled in common computer files using the joint designation NADP/NTN. Authors of reports published previously describe the background of the NADP/NTN in further detail (Lindberg and others, 1982; See and others, 1988).

Monitoring the acidity of wet deposition is a major objective of the NADP/NTN. Because physical and chemical reactions may cause the pH of a wet-deposition sample to change over time (Hem, 1985), the weekly onsite measurements of pH are potentially the best measurements of the acidity of the wet-deposition samples collected by the NADP/NTN. Analysis of data collected in the Electric Power Research Institute Sulfate Regional Experiment determined that differences between onsite and laboratory measurements of sample pH are frequently statistically significant (Topol and Schwall, 1987).

The intersite-comparison program is an ongoing, external quality-assurance program that has been operated for the NADP/NTN by the U.S. Geological Survey since 1981. The Central Analytical Laboratory (CAL) in Champaign, Ill. initiated the intersite-comparison program to assess the ability of NADP/NTN site operators to make accurate and precise pH and specific-conductance measurements. The accuracy of a measurement refers to the extent to which it agrees with the accepted value for that quantity, whereas precision is the degree of similarity among independent measurements of the same quantity, without reference to the known or true value (V.R. Schneider, U.S. Geological Survey, written commun., 1990). Site operators are required to make onsite measurements of pH and specific conductance each week in conjunction with the collection of wet-deposition samples that have a volume of 50 mL or greater. Site-operator performance in the intersite-comparison program is the primary indicator of the quality of these field analyses. Weekly measurements of a 4.30 pH check solution and performance during site audits are additional indicators of measurement quality.

During each intersite-comparison study, reference samples were distributed to all site operators with instructions that they measure the pH and specific conductance using standard NADP/NTN procedures. The reference samples were distributed as single blind samples (that is, they were known to be reference samples, but the pH and specific conductance were not known to the site operator). Reference solutions for the intersite-comparison studies typically consisted of aliquots of homogeneous dilute nitric-acid solutions with target pH and specific-conductance values within the normal range for wet deposition. The reference solutions were usually prepared by adding a small amount (less than 0.50 mL) of concentrated nitric acid to ultrapure-deionized water (greater than 16.7 MΩ) in a 50-liter polyethylene carboy. The exact amount of concentrated nitric acid depended upon the selected target pH and the volume of solution prepared. The preparation of reference solutions is discussed in detail in a previous report (See and others, 1990b).

Purpose and Scope

Intersite-comparison studies have been used since November 1978 to assess the accuracy and precision of pH and specific-conductance measurements made by NADP/NTN site operators. A total of 24 intersite-comparison studies was completed as of November 1989. This report summarizes the results of the first 24 studies and evaluates the measurement-accuracy criteria. The pH and specific-conductance measurements made during intersite-comparison studies were primarily evaluated using two methods of analysis. The first method of analysis evaluated site-operator reported values in terms of the measurement-accuracy criteria established by the Network Operations Subcommittee. The second method of analysis evaluated the differences between site-operator reported values and network median values.

History

Because of cost constraints, laboratory-certified values for pH and specific conductance were not obtained for the reference samples used in the intersite-comparison program. In place of certified values, the mean pH and specific conductance for all responding site operators originally were used as the most probable values.

In response to growing concern among members of the NOS that outlying values were affecting the mean values for pH and specific conductance, Brooks and others (1988) confirmed that outliers in the data set significantly affected the mean values in some studies. Beginning with study 16, the median pH and median specific-conductance values for all responding site operators have been used as the most probable values, and the F-pseudosigma has been used to measure the spread of the data. Experience has shown that the median value of about 200 analyses by the site operators is a better estimate of the most probable value than either the limited number of in-house laboratory analyses that are performed or the theoretical target value (See and others, 1990a). For consistency, the data analysis for this report uses the median pH and specific-conductance values as the most probable values for all 24 intersite-comparison studies.

During the 11 years between the completion of intersite-comparison study 1 and study 24, the operating protocol of the intersite-comparison program was revised repeatedly (fig. 1). The first four intersite-comparison studies were administered by the CAL between November 1978 and June 1980. The U.S. Geological Survey assumed responsibility for the program in October 1981. Additional changes in the operating protocol included the initiation of semiannual site-operator training classes, the decision by the NOS to provide site operators with pH electrodes, and the creation of a site-visitation program.

During studies 5-21, all site operators were asked to return the remaining part of the study reference solution for possible re-analysis by the U.S. Geological Survey. In studies 5-15, re-analysis was done when (1) reported pH values were not within "re-analysis limits" established by the U.S. Geological Survey of the mean of all site operator values ±0.4 pH unit; or (2) reported specific-conductance values differed by more than 20 percent from the mean of all specific conductance values reported by site operators. The "re-analysis limits" were established by the U.S. Geological Survey to select samples for re-analysis only if the site operator failed to meet the measurement-accuracy criteria and reported an outlying value. Re-analysis determined if the sample had been contaminated or undergone a change in chemistry. Re-analysis also helped to determine if the erroneous measurement was because of site-operator error or an equipment malfunction (See and others, 1990a).

In studies 16-21, nonparametric statistics were used as the basis for sample re-analysis. Samples were re-analyzed when the reported values exceeded a new "re-analysis limit" established by the U.S. Geological Survey of the median ±1.5 times the F-pseudosigma value calculated from all responding site operators. The F-pseudosigma is a nonparametric statistic analogous to the standard deviation. Because the F-pseudosigma is calculated using the 25th and 75th percentiles in a data set, it is resistant to the effect of extreme outliers (Hoaglin and others, 1983). More than 99 percent of the re-analyses by the U.S. Geological Survey of pH and specific-conductance samples were within the measurement-accuracy criteria established by the NOS. These results confirmed that sample contamination prior to shipment to the site operator or sample deterioration during shipment and storage were extremely rare explanations for a site operator to fail to achieve the measurement-accuracy criteria. For this reason, the practice of requesting that the remaining parts of intersite-comparison solutions be returned to the U.S. Geological Survey for re-analyses was discontinued. Beginning with study 22, only those site operators reporting a value or values outside the measurement-accuracy criteria or the 1.5 times F-pseudosigma range that they believed to be valid were encouraged to return the remaining part of the reference solution to the U.S. Geological Survey for re-analysis. Authors of the 1988 external quality-assurance summary report discuss re-analyses of reference solutions in further detail (See and others, 1990a).

November 1978	First intersite-comparison study completed by the Illinois State Water Survey Central Analytical Laboratory (CAL).
June 1980	Semiannual site operator training initiated. Classes held at CAL.
October 1981	U.S. Geological Survey assumes responsibility for the intersite-comparison program and begins conducting studies on a semiannual basis.
October 1981 to May 1988	U.S. Geological Survey reanalyzes aliquots of test samples when site operators fail to meet measurement "re-analysis limits".
February 1983	U.S. Geological Survey completes pH and specific-conductance stability experiments indicating dilute nitric-acid solutions used in intersite-comparison studies are stable for at least 10 weeks.
December 1983	CAL begins supplying pH electrodes to site operators.
October 1984	Majority of site operators now using pH electrodes supplied by CAL.
December 1984	Site-visitation program initiated. Sites are visited by designated quality-assurance personnel to examine data-collection techniques.
February 1985	Frequency of intersite-comparison studies increased from semiannual to quarterly.
July 1985	Intersite-comparison study 10 completed.
July 1987	Semiannual intersite-comparison studies resumed.
November 1989	Intersite-comparison study 24 completed.

Figure 1. History of the intersite-comparison study program.

RESULTS OF INTERSITE-COMPARISON STUDIES

During each study, about 3 percent of the site operators reported that their pH or specific-conductance equipment was inoperable. About 1 to 2 percent of the site operators typically reported their results after the scheduled closing date of each study. Non-responding site operators (that is, operators that did not return results or an explanation that their equipment was inoperable) constituted another 4 to 7 percent of the site operators in most studies. Only data that were submitted by the closing date for each study were evaluated in this report. The maximum and minimum reported values, median (the 50th percentile), quartiles (the 25th and 75th percentiles) and other percentiles for pH and specific-conductance results in intersite-comparison studies 1-24 are listed in tables 1 and 2. Percentile rankings provide one indication of distribution shape and can be used to assess the degree of similarity among independent values (Hoaglin and others, 1983).

Evaluation of pH Determinations Using Network Operations Subcommittee Measurement-Accuracy Criteria

The NOS measurement-accuracy criteria for pH determinations are ±0.10 pH unit of the most probable value if the pH of the reference solution is 5.0 or less. When the most probable pH value of the reference solution exceeds 5.0, the accuracy criteria increase to ±0.30 pH unit. In every intersite-comparison study except study 22, the median pH was less than 5.0 (table 1).

During intersite-comparison studies 1-24, 3,124 pH determinations were done. The NOS measurement-accuracy criteria were met for 2,319 of these pH determinations (74 percent). Table 3 depicts the percentage of site operators that met the measurement-accuracy criteria for pH, the number of site operators responding, and the F-pseudosigma for each intersite-comparison study. The F-pseudosigma provides a measure of the degree of similarity among the independent determinations made in each study. Relatively smaller F-pseudosigma values are indicative of less spread in the reported values.

Because of significant changes in the operating protocol of the network, it is meaningful to compare the results of studies 1-10 with the results of studies 11-24. Beginning with study 11, most of the site operators were using electrodes supplied by CAL. Additional changes designed to improve the quality of the field measurements also went into effect immediately prior to study 11, including semi-annual site operator training classes and a site-visitation program. The percentage of site operators reporting pH values within the measurement-accuracy criteria increased from an average of 61 percent for intersite-comparison studies 1-10 to an average of 78 percent for studies 11-24. Although a definitive relation cannot be established, it can be inferred that supplying site operators with uniform-quality pH electrodes, providing instruction at training classes, and initiating a site-visitation program positively affected the results of pH determinations in the intersite-comparison program.

Table 1.--*Maximum and minimum reported values and percentiles for pH results obtained from intersite-comparison studies*

[pH, in units]

Intersite-comparison study number	Date of study		Minimum	Percentile					Maximum
				10th	25th	50th	75th	90th	
1	November	1978	2.63	2.63	2.70	3.02	3.05	3.43	3.43
2	August	1979	3.30	3.62	3.95	4.08	4.28	4.67	8.81
3	November	1979	3.60	4.11	4.26	4.47	4.60	5.19	5.95
4	June	1980	3.50	3.98	4.22	4.29	4.32	4.58	7.20
5	October	1981	3.30	3.99	4.19	4.29	4.36	4.50	5.42
6	April	1982	3.73	4.28	4.44	4.57	4.62	4.73	6.35
7	November	1982	3.32	3.87	3.95	4.03	4.09	4.19	4.88
8	May	1983	3.00	3.55	3.76	3.84	3.89	4.00	4.40
9	November	1983	3.45	4.38	4.52	4.61	4.67	4.79	5.97
10	July	1984	.38	3.95	4.08	4.12	4.18	4.27	7.50
11	February	1985	3.97	4.37	4.46	4.55	4.60	4.71	6.60
12	May	1985	3.39	3.93	4.08	4.13	4.18	4.25	6.30
13	August	1985	3.61	4.59	4.71	4.79	4.86	4.99	7.50
14	November	1985	3.65	4.47	4.55	4.61	4.67	4.74	8.10
15	February	1986	3.33	4.48	4.57	4.60	4.66	4.73	7.25
16	May	1986	3.87	4.82	4.89	4.96	5.02	5.09	5.90
17	August	1986	3.29	3.81	3.89	3.92	3.95	4.00	5.91
18	November	1986	3.64	4.24	4.29	4.32	4.35	4.40	5.60
19	July	1987	3.71	4.35	4.45	4.50	4.54	4.58	5.04
20	October	1987	3.65	4.56	4.67	4.73	4.79	4.85	6.52
21	May	1988	3.70	4.34	4.38	4.40	4.44	4.48	13.0
22	November	1988	3.47	4.90	4.98	5.06	5.13	5.19	5.50
23	May	1989	3.42	4.55	4.61	4.68	4.72	4.76	5.40
24	November	1989	3.70	3.94	3.98	4.01	4.04	4.08	6.10

Table 2.--*Maximum and minimum reported values and percentiles for specific-conductance results obtained from intersite-comparison studies*

[specific conductance, in microsiemens per centimeter at 25 degrees Celsius]

Intersite-comparison study number	Date of study	Mini-mum	Percentile					Maxi-mum
			10th	25th	50th	75th	90th	
1	November 1978	0.30	0.30	93.40	192	420	422	422
2	August 1979	26.0	29.0	31.6	35.1	37.6	40.9	42.6
3	November 1979	.31	30.9	36.8	39.6	40.8	44.5	68.8
4	June 1980	10.0	19.7	20.9	22.3	24.1	25.3	42.3
5	October 1981	.30	18.2	19.3	21.6	23.1	24.8	63.5
6	April 1982	.12	10.0	11.2	12.4	13.2	13.9	94.8
7	November 1982	3.20	28.8	34.8	38.0	41.0	43.3	71.6
8	May 1983	5.67	46.4	53.0	60.0	63.8	66.0	90.4
9	November 1983	.91	7.48	9.20	10.8	11.4	12.2	66.0
10	July 1984	7.10	25.5	28.3	31.8	34.0	35.5	44.0
11	February 1985	.90	8.55	10.3	11.4	12.4	13.8	92.5
12	May 1985	2.70	24.3	27.0	30.0	31.6	33.2	45.0
13	August 1985	.62	5.00	5.78	6.60	7.01	8.14	64.0
14	November 1985	.94	7.74	8.89	9.80	10.5	11.2	55
15	February 1986	1.00	8.45	9.69	10.4	11.0	12.0	655
16	May 1986	2.79	3.79	4.10	4.40	4.80	6.00	98.3
17	August 1986	3.98	45.7	48.4	51.4	52.9	55.0	78.4
18	November 1986	1.07	19.0	19.8	20.9	21.7	22.3	26.1
19	July 1987	9.80	12.4	12.9	13.5	14.0	14.9	47.2
20	October 1987	.73	6.56	7.10	7.42	7.9	8.90	8,450
21	May 1988	1.68	15.5	16.4	16.9	17.4	18.3	85.5
22	November 1988	5.90	18.4	19.0	19.5	20.1	20.9	38.2
23	May 1989	6.30	7.74	8.30	8.64	8.9	9.40	90.4
24	November 1989	4.00	38.4	41.5	43.2	44.4	46.0	157.9

Table 3.--*Summary statistics obtained from analysis of site-operator results for pH measurements obtained from intersite-comparison studies*

Intersite-comparison study number	Number of site operators responding	Median for all responding site operators	F-pseudo-sigma[1]	Percentage of site operators that met the measurement-accuracy criteria[2]
1	7	3.02	0.26	57
2	32	4.08	.25	41
3	32	4.47	.25	34
4	57	4.29	.07	65
5	73	4.29	.12	59
6	85	4.57	.13	53
7	93	4.03	.10	66
8	101	3.84	.10	64
9	118	4.61	.11	60
10	120	4.12	.07	71
11	163	4.55	.10	65
12	162	4.13	.07	72
13	172	4.79	.11	66
14	165	4.61	.09	73
15	173	4.60	.07	78
16	170	4.96	.10	65
17	170	3.92	.04	83
18	168	4.32	.05	88
19	175	4.50	.07	83
20	167	4.73	.09	71
21	177	4.40	.04	88
22	176	5.06	.11	94
23	182	4.68	.08	75
24	186	4.01	.04	91

[1]F-pseudosigma = $\frac{\text{75th - 25th percentile of all responding site operators}}{1.349}$

[2]Accuracy criteria for pH determintations are ±0.10 unit if median pH is 5.0 or less; if median pH exceeds 5.0, the accuracy criteria are ±0.30 pH unit.

A nonparametric correlation analysis was used to assess the relation between the ability of site operators to meet measurement-accuracy criteria and intersite-comparison study number. Nonparametric regression analysis was used to assess the trend in site-operator ability to meet measurement-accuracy criteria. Nonparametric tests are appropriate for data that do not consistently meet all the assumptions necessary for using parametric statistical tests, including equal variances and normal distributions (Conover, 1980). The Spearman rank correlation test was used as a nonparametric measure of the association between two variables; nonparametric rank regression analysis provided a measure of their linear relation (Conover, 1980). Using the intersite-comparison study number as a substitute for time, a statistically significant positive Spearman rank correlation coefficient between the percentage of site operators successfully meeting measurement-accuracy criteria for pH determinations and time was determined (α=0.05, r=0.88). The nonparametric rank regression equation between the percentage of site operators successfully meeting measurement-accuracy criteria for pH determinations and the intersite-comparison study number is depicted in figure 2.

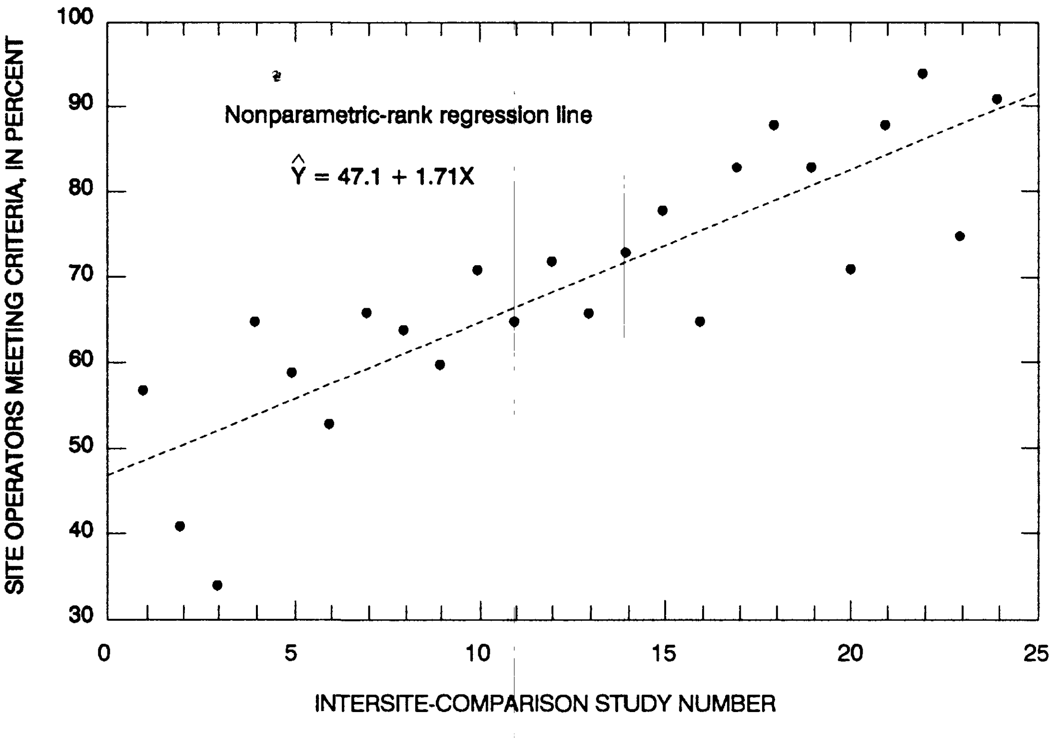

Figure 2.--Relation between site operators meeting measurement-accuracy criteria for pH determinations, in percent, and intersite-comparison study number.

pH Determinations Outside Measurement-Accuracy Criteria

A total of 805 of the 3,124 pH determinations (26 percent) did not meet NOS measurement-accuracy criteria in studies 1-24. Analysis of these determinations yields the following observations:

(1) Of the pH determinations that did not meet the measurement-accuracy criteria, 61 percent missed the goals by 0.10 pH unit or less; 74 percent missed the goals by 0.25 pH unit or less (fig. 3).

(2) A total of 13 percent of the pH determinations that did not meet the measurement-accuracy criteria missed the goals by 0.50 pH unit or more.

(3) The operators of a small percentage of the NADP/NTN sites accounted for a disproportionately large percentage of determinations that failed to meet the measurement-accuracy criteria: the operators of 32 sites, about 12 percent of the sites that participated in intersite-comparison studies 1-24, accounted for 29 percent of the pH determinations outside the acceptable criteria for measurement accuracy.

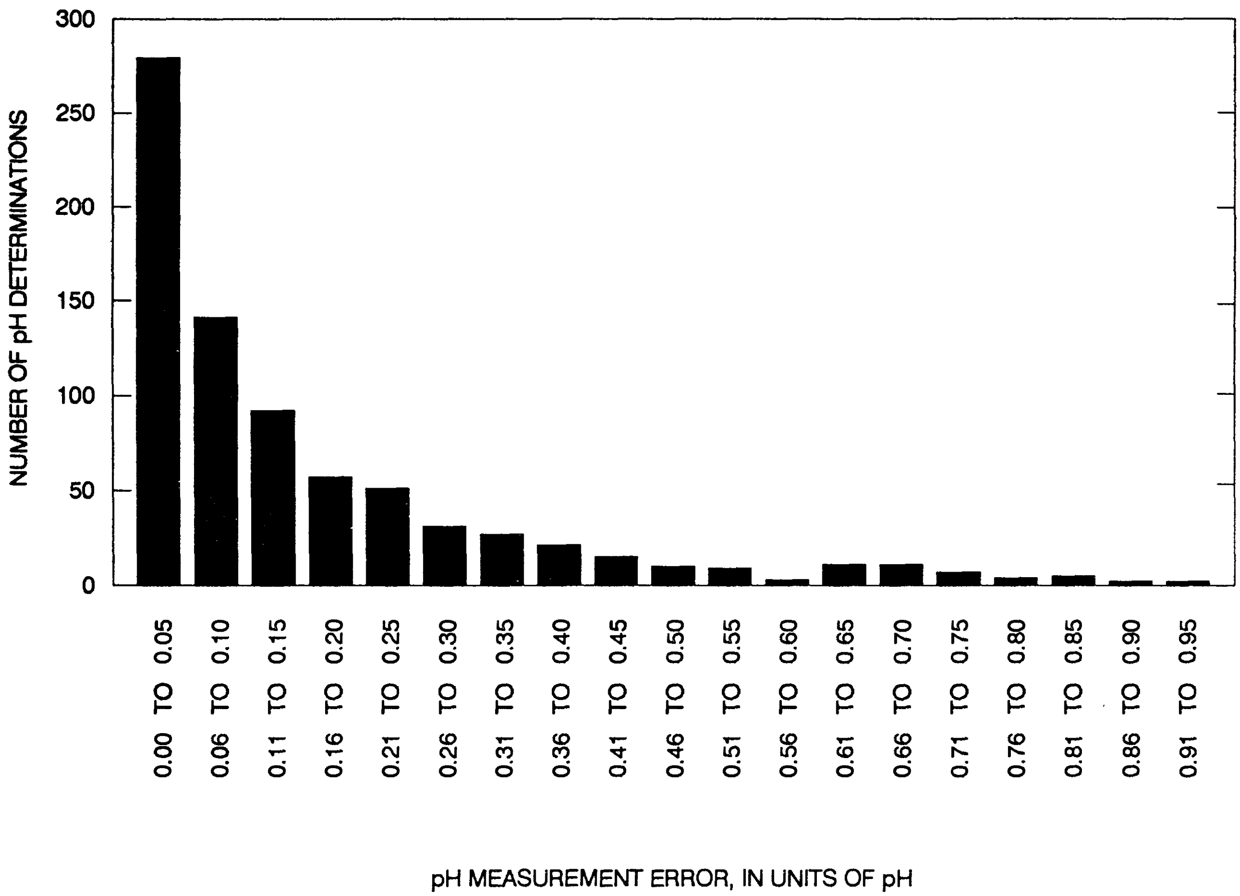

Figure 3.--Frequency distribution of measurement error for pH determinations that failed to meet the measurement-accuracy criteria in studies 1-24. pH measurement error was computed as the absolute difference between a reported pH value and the pH value corresponding to the nearest measurement-accuracy limit as defined by the Network Operation Subcommittee.

Success Rates For Achieving pH Measurement-Accuracy Criteria

Because many NADP/NTN sites have had several site operators, a success rate for achieving pH measurement-accuracy criteria was determined for each site whose operators actively participated in the intersite-comparison program, rather than for individual site operators. The success rate was calculated as:

$$\text{Success Rate} = \frac{\text{Number of times the operators of a site met accuracy criteria}}{\text{Number of times the operators of a site submitted data}} \times 100 \qquad (1)$$

The success rates for sites with actively participating operators in studies 1-10 (sites whose operators participated at least five times) were compared with the success rates for sites with actively participating operators in studies 11-24. The percentage of sites whose operators achieved success rates between 80 and 100 percent increased from 29 percent in studies 1-10 to 58 percent in studies 11-24. Whereas the operators of many sites have good success rates of achieving the pH measurement-accuracy criteria, the operators of 32 sites have success rates of 50 percent or less. The median success rate for the operators of this group of sites was 39 percent. These sites were distributed among all geographic regions of the United States, including the northeast where concern about acidic wet deposition and its effects is greatest.

Of the 32 sites whose operators met the measurement-accuracy criteria 50 percent of the time or less, 27 registered brief yet distinct periods when their performance was acceptable. For example, it was not uncommon for a site whose operators met the measurement-accuracy criteria in only 7 of 20 studies in which they participated to have recorded 6 of these successes consecutively. A statistical test known as a runs test verified this non-random pattern of success and failure (Dixon and Massey, 1969). Changes in site-operator personnel are the most likely explanation for this pattern of performance.

Differences Between Measured and Median pH

Determining the success or failure of individual site operators to meet the NOS measurement-accuracy criteria and determining success rates for each site are both useful means of evaluating the accuracy of pH determinations made in the intersite-comparison program. Additional insight regarding measurement precision can be gained by analyzing the differences between the values reported by the individual site operators and the network median values.

The frequency distribution of differences between measured and median pH for all pH determinations that were made during intersite-comparison studies 1-24 is shown in figure 4. For all 24 studies, 49 percent of the determinations was within 0.05 pH unit of the median value, 82 percent was within 0.15 pH unit of the median, and 90 percent was within 0.25 pH unit of the median. The interquartile range of measured minus median pH for each

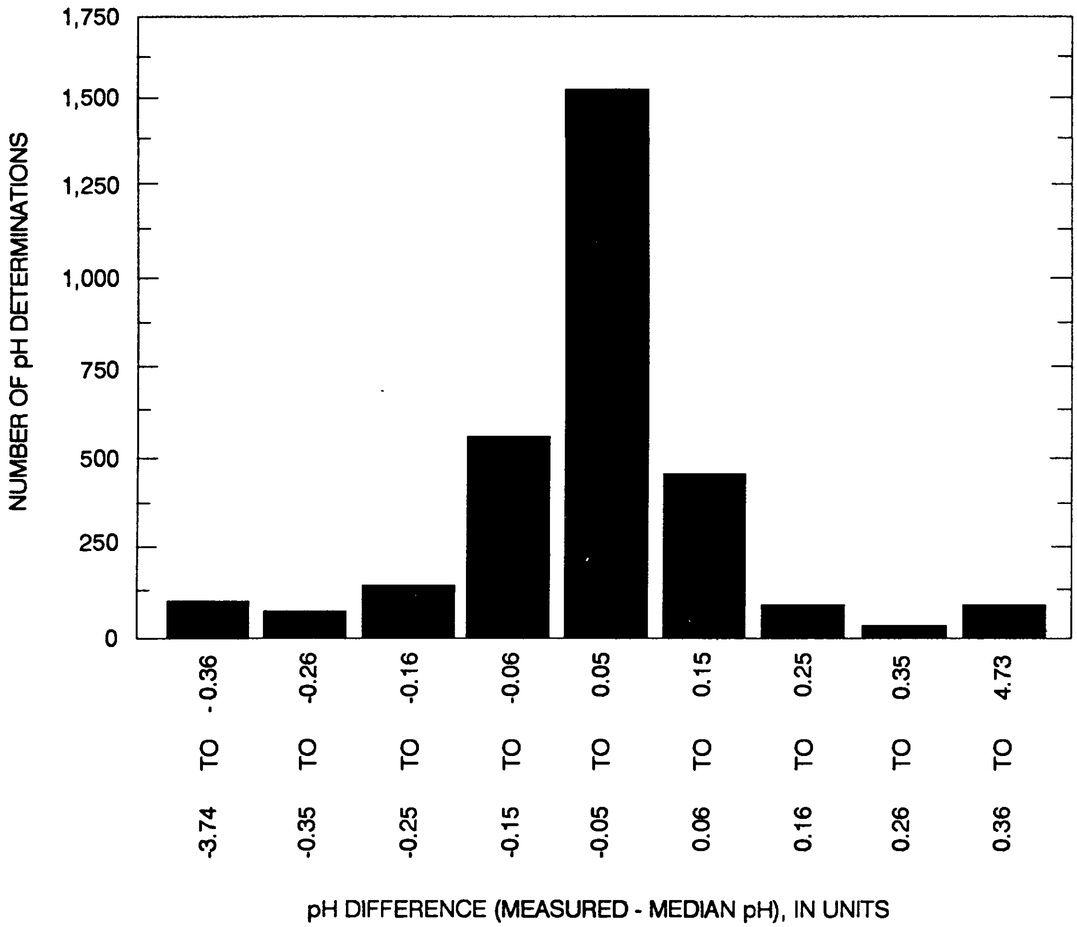

Figure 4.--Frequency distribution of differences between measured and median pH for studies 1-24.

intersite-comparison study is shown in figure 5A. The interquartile range is defined as the distance from the first to the third quartile. The first quartile is equal to the 25th percentile of the data, and the third quartile is equal to the 75th percentile (Chambers and others, 1983). The size of the interquartile range provides a measure of the degree of similarity among independent measurements of the same quantity, and is indicative of the overall precision of the site operators participating in each study. The median pH for each intersite-comparison study is shown in figure 5B.

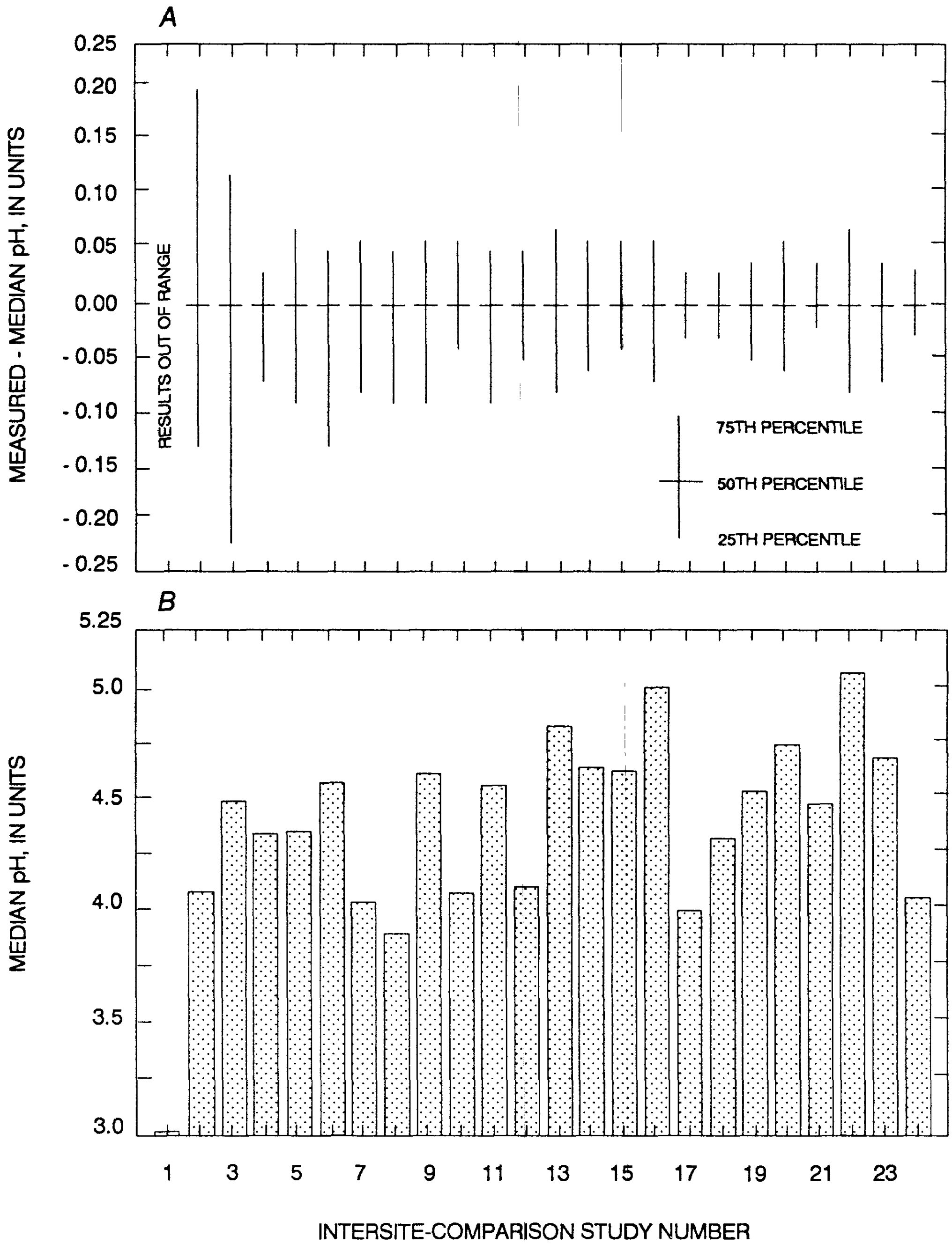

Figure 5.--*A*, Interquartile range of measured minus median pH for each intersite-comparison study; *B*, Median pH for each intersite-comparison study.

To determine if the selected target pH affected the degree of similarity among pH measurements reported by the site operators in a given study, a nonparametric correlation test was used to assess the degree of association between the measured minus median interquartile range and the median pH of the reference solution. For studies 1-10, the Spearman rank correlation coefficient between the measured minus median interquartile range and intersite-comparison study median pH was -0.05, which was not statistically significant (α=0.05). For studies 11-24, however, the Spearman rank correlation coefficient was 0.85, which was statistically significant (α=0.05). The amount of variability in reported pH values increased monotonically with median pH in studies 11-24; there was more variability in the reported pH values when the median pH of the reference solution was relatively larger.

Evaluation of Specific-Conductance Determinations Using Network Operations Subcommittee Measurement-Accuracy Criteria

The measurement-accuracy criteria established by the NOS for specific-conductance determinations are ±4.0 μS/cm. These measurement-accuracy criteria are used for all intersite-comparison studies, regardless of the median specific conductance of the reference solution. Median specific conductance of the reference solutions ranged from 4 to 60 μS/cm in intersite-comparison studies 1-24.

During intersite-comparison studies 1-24, 3,113 specific-conductance determinations were done. Of this total, 2,692 (86 percent) met the NOS measurement-accuracy criteria. Table 4 depicts the percentage of site operators that met the measurement-accuracy criteria for specific conductance, the number of site operators responding, and the F-pseudosigma for each intersite-comparison study. A scatterplot depicting the percentage of site operators that met the specific-conductance accuracy criteria on an individual study basis is shown in figure 6.

The percentage of site operators reporting specific-conductance values within the NOS measurement-accuracy criteria increased from an average of 73 percent in studies 1-10 to an average of 90 percent in studies 11-24. The increase in the percentage of site operators that met the measurement-accuracy criteria in studies 11-24 compared with studies 1-10 was to a large extent associated with differences in the median specific conductance of the reference solutions used in studies 1-10 and studies 11-24. From the start of the intersite program, site operators have had little difficulty meeting the measurement-accuracy goals for specific conductance when the median specific conductance was relatively small. The percentage of site operators reporting specific-conductance values that met the NOS measurement-accuracy criteria varied between 88 and 97 percent when the median specific conductance was less than 25 μS/cm. When the median specific conductance of the reference solution was between 30 to 60 μS/cm, only 38 to 82 percent of the participating site operators was able to meet the measurement-accuracy criteria (reference solutions with median specific-conductance values between 25 and 30 μS/cm were not used in studies 1-24). The median specific conductance was less than 25 μS/cm in only four of the first ten studies; in studies 11-24, the median specific conductance was in this range in eleven studies.

Table 4.--*Summary statistics obtained from analysis of site-operator results for specific-conductance measurements obtained from intersite-comparison studies*

[μS/cm, microsiemens per centimeter at 25 degrees Celsius]

Intersite-comparison study number	Number of site operators responding	Median for all responding site operators	F-pseudo-sigma[1]	Percentage of site operators that met the measurement-accuracy criteria[2]
1	7	192	242	14
2	29	35.1	4.45	66
3	30	39.6	3.02	67
4	57	22.3	2.41	88
5	72	21.6	2.86	88
6	86	12.4	1.49	92
7	91	38.0	4.60	62
8	99	60.0	8.01	38
9	116	10.8	1.61	90
10	120	31.8	4.19	70
11	164	11.4	1.54	88
12	162	30.0	3.45	76
13	172	6.60	.93	95
14	166	9.80	1.20	93
15	174	10.4	.98	92
16	167	4.40	.52	95
17	171	51.4	3.33	75
18	168	20.9	1.39	95
19	174	13.5	.83	95
20	163	7.42	.56	93
21	178	16.9	.77	94
22	176	19.5	.82	95
23	183	8.64	.44	97
24	188	43.2	2.13	82

[1]F-psuedosigma = $\frac{\text{75th - 25th percentile of all responding site operators}}{1.349}$

[2]Accuracy criteria for specific-conductance determinations are ±4.0 microsiemens per centimeter at 25 degrees Celsius.

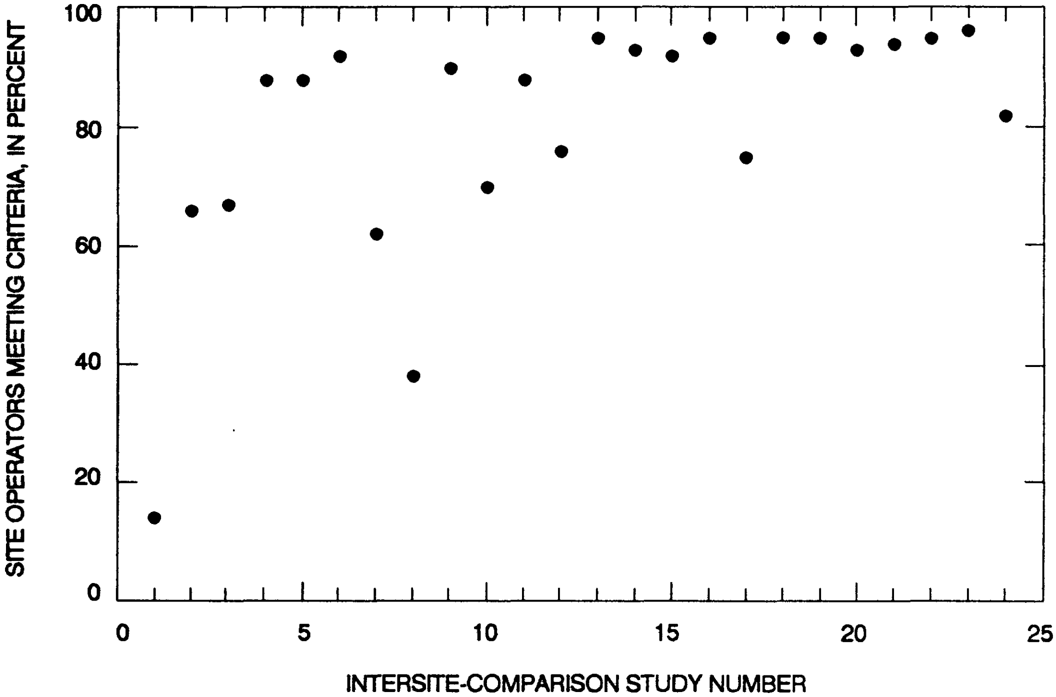

Figure 6.--Comparison of site operators meeting measurement-accuracy criteria for specific-conductance determinations and intersite-comparison study number.

Specific-Conductance Determinations Outside Measurement-Accuracy Criteria

A total of 421 of the 3,113 specific-conductance determinations (14 percent) did not meet the NOS measurement-accuracy criteria in studies 1-24. Analysis of these determinations yields the following observations:

(1) The percentage of site operators that failed to meet the measurement-accuracy criteria by a wide margin (more than 8 µS/cm) was about the same as the percentage that missed the measurement-accuracy criteria by a narrow margin (2 µS/cm or less). Of the determinations that failed to meet the measurement-accuracy criteria, 40 percent were within 2 µS/cm, whereas 32 percent of the failed determinations missed the measurement-accuracy criteria by more than 8 µS/cm.

(2) The operators of a small percentage of sites accounted for a disproportionately large percentage of determinations that failed to meet the measurement-accuracy criteria: the operators of 22 sites, less than 9 percent of the total number of sites that participated in intersite-comparison studies 1-24, accounted for 16 percent of the specific-conductance determinations that were outside the accuracy criteria.

Success Rates for Achieving Specific-Conductance Measurement-Accuracy Criteria

A success rate for achieving the NOS accuracy criteria for specific-conductance measurements was determined for each site whose operators actively participated in the intersite-comparison program (equation 1). The operators of most sites had excellent success rates of achieving NOS measurement-accuracy criteria for specific-conductance measurements. About 67 percent of the sites whose operators participated in at least five studies achieved success rates of 85 percent or better. There was, however, a small group of sites whose operators achieved success rates of 60 percent or less, with a median success rate of 47 percent. Ten of the 22 sites whose operators had success rates of 60 percent or less for specific-conductance determinations were among the 32 sites whose operators had the poorest success rates of achieving NOS measurement-accuracy criteria for pH determinations.

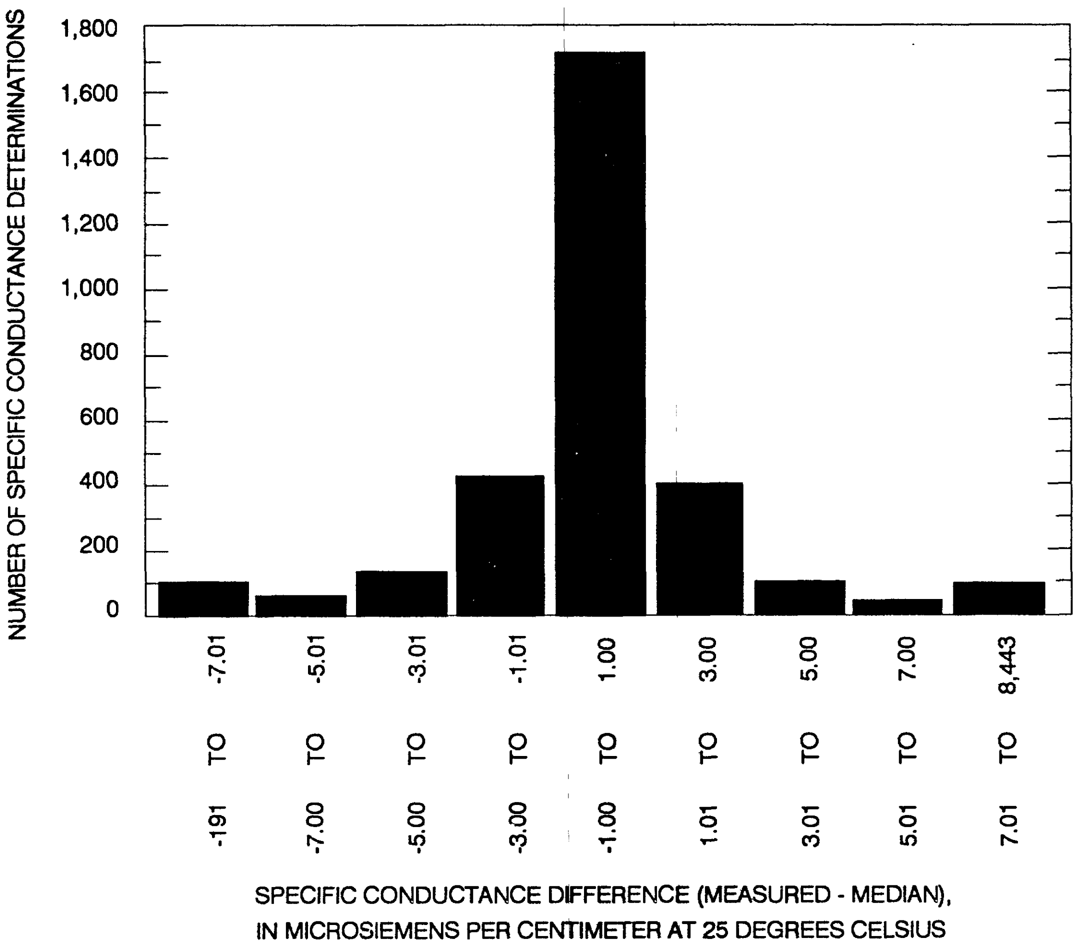

Figure 7.--Frequency distribution of differences between measured and median specific conductance for studies 1-24.

Differences between Measured and Median Specific Conductance

The frequency distribution of differences between the measured and the median specific conductance for all determinations that were made during intersite-comparison studies 1-24 is shown in figure 7. For the 24 studies, a total of 55 percent of the determinations was within ±1 μS/cm of the median values of the individual studies, and 82 percent was within ±3 μS/cm of the median. The percentage of determinations approximating the median values of the individual studies indicates a large degree of measurement precision on an overall basis. The interquartile range of measured minus median specific-conductance values was, however, strongly affected by the specific conductance of the reference solution. When the median specific conductance of the reference solution was less than 25 μS/cm, the interquartile range generally was much smaller than when the median specific conductance exceeded 25 μS/cm (figs. 8A and 8B). For example, intersite-comparison study 8, which had a reference solution with one of the largest median specific-conductance values (60 μS/cm), also had one of the largest measured minus median interquartile ranges (11 μS/cm). For studies 1-24, the Spearman rank correlation coefficient between the measured minus median interquartile range and the median specific conductance was statistically significant (α=0.05, r=0.85).

EVALUATION OF NETWORK OPERATIONS SUBCOMMITTEE MEASUREMENT-ACCURACY CRITERIA

The question of whether the established measurement-accuracy criteria are appropriate has been raised periodically by various members of the NOS. For example, Schroder and others (1987) commented that the criteria established for specific-conductance measurements should be changed. The authors noted that the percentage of site operators meeting the accuracy criteria was only 38 percent for study 8 (May 1983) when the median specific conductance of the reference solution was 60 μS/cm.

For each intersite-comparison study, a different reference solution is prepared. The target pH is selected between 3.8 and 5.3 units, with consideration given to the pH of preceding studies. The specific conductance of the reference solution is dependent upon the selected target pH and ranged between 4 and 60 μS/cm in studies 1-24. The measurement-accuracy criteria established by the NOS for pH determinations are ±0.10 pH unit of the most probable value if the pH of the reference solution is 5.0 or less. Two separate studies with median pH values of 3.84 and 4.96, therefore, applied the same measurement-accuracy criteria, even though a reference solution with a pH of 3.84 contains over ten times the hydrogen ion concentration that a reference solution with a pH of 4.96 contains. Regardless of the specific conductance of the reference solution, the measurement-accuracy criteria established by the NOS are ±4.0 μS/cm. The accuracy criteria thus allow for ± 100 percent measurement error if the target specific conductance is 4 μS/cm, but only ±6.7 percent measurement error if the target specific conductance is 60 μS/cm. Nonparametric statistical tests were used to evaluate the appropriateness of the NOS measurement-accuracy criteria for both pH and specific-conductance determinations.

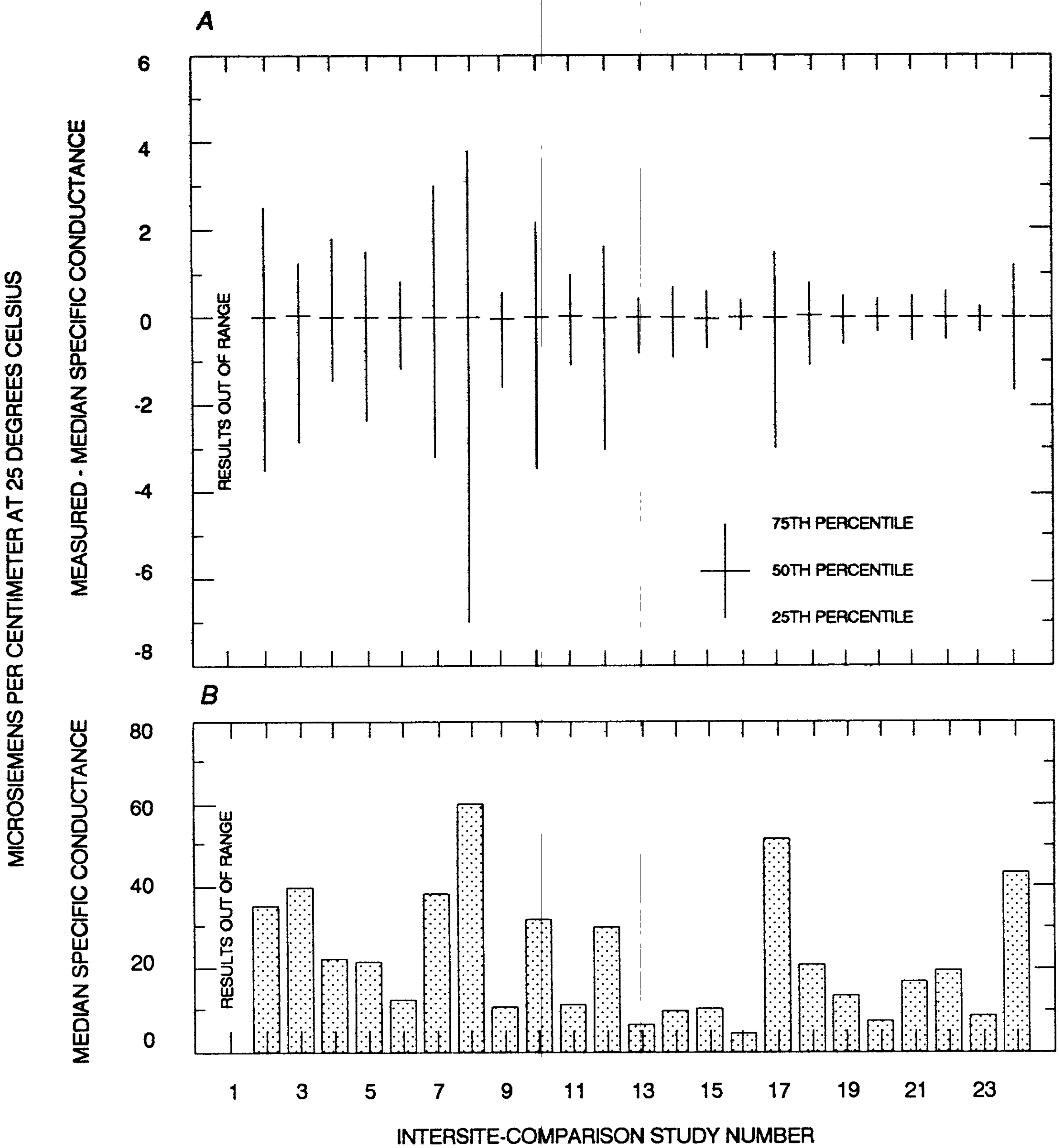

Figure 8.--*A*, Interquartile range of measured minus median specific conductance for each intersite-comparison study; *B*, Median specific conductance for each intersite-comparison study.

Scatterplots of the relation between the percentage of site operators that met the NOS measurement-accuracy criteria and the median pH of the reference solution are shown in figures 9A and 9B. For studies 1-10, the Spearman rank correlation coefficient between the median pH of the reference solutions used in the intersite-comparison program and the percentage of site operators that met pH measurement-accuracy criteria was not statistically significant (α=0.05, r=-0.24). Excluding study 22 because of the different pH measurement-accuracy criteria that were applied, a statistically significant Spearman rank correlation coefficient between the median pH of the reference solution and the percentage of site operators achieving the accuracy criteria was determined for studies 11-24 (α=0.05, r=-0.71). During studies 11-24, a relatively larger percentage of site operators met the pH accuracy criteria when a reference solution with a relatively smaller median pH was used. The median pH was between 4.50 and 5.00 pH units in two of the first ten studies; in studies 11-24, the median pH was in this range in eight of the studies. The increase in the percentage of site operators meeting the accuracy criteria in studies 11-24 compared with studies 1-10 was in spite of the fact that generally larger median pH reference solutions were used in studies 11-24.

For studies 1-24, the Spearman rank correlation coefficient between the percentage of site operators meeting specific-conductance criteria and the median specific conductance of the reference solution was statistically significant (α=0.05, r=-0.82). Most of the variability in the percentage of site operators that met the specific-conductance accuracy criteria in intersite-comparison studies 1-24 was associated with the median specific conductance of the reference solution. The nonparametric rank regression equation between the percentage of site operators meeting specific-conductance criteria and the median pH of the reference solution was Y = 105.3 - 0.79(X) (fig. 10).

To facilitate the evaluation of pH and specific-conductance determinations, the current NOS measurement-accuracy criteria need to be revised. In place of the current accuracy criteria of ±0.10 or ±0.30 unit for pH and ±4.0 μS/cm for specific conductance, a continuum of accuracy criteria could be established. New measurement-accuracy criteria could, for example, be based on statistical quality-control techniques (Taylor, 1987). Alternatively, adjustable ranges of acceptable percent measurement error depending on the target values of the reference solution could be established.

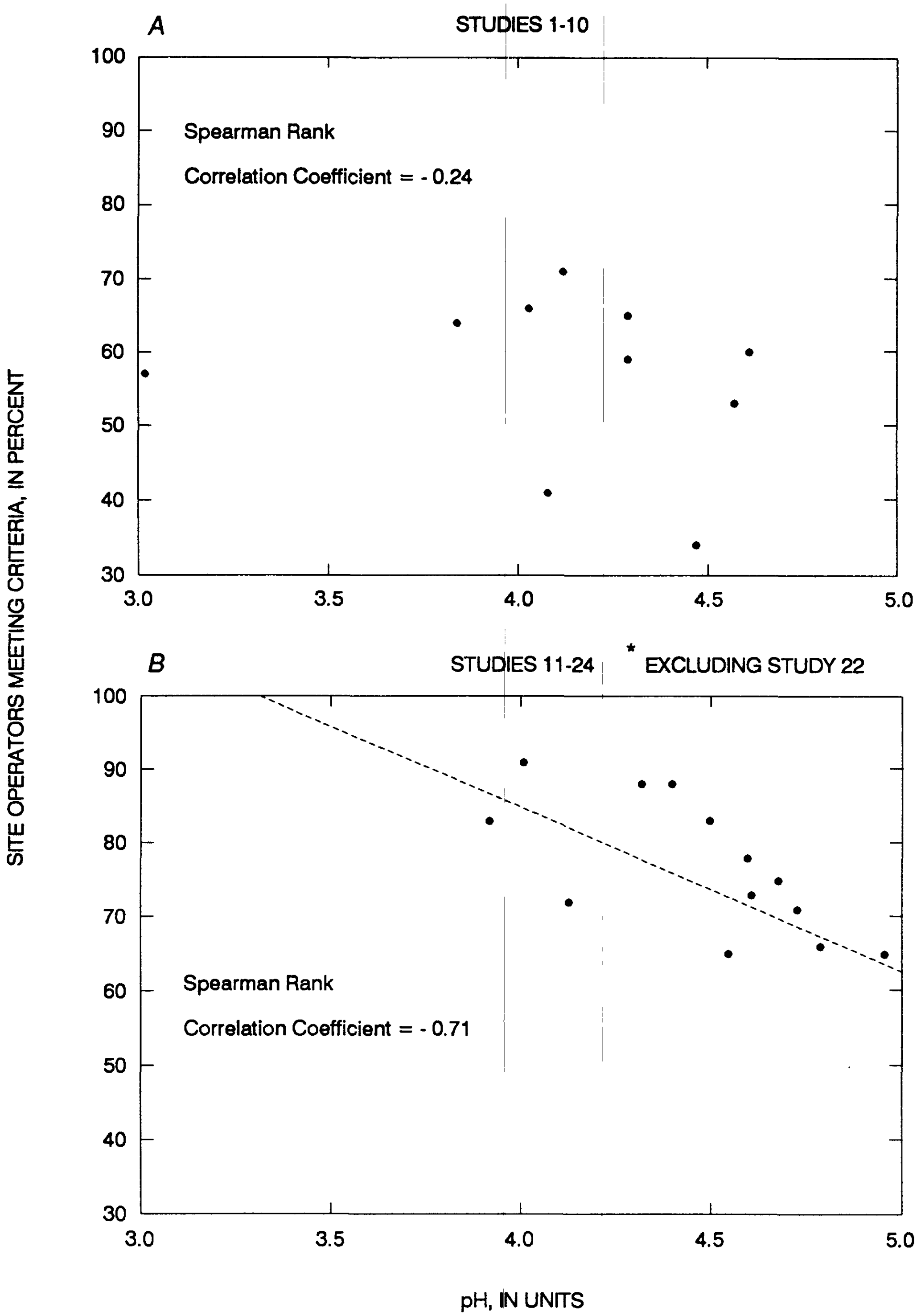

Figure 9.--*A*, Relation between site operators meeting pH measurement-accuracy criteria, in percent, and the median pH of the intersite-comparison study reference solution in studies 1-10; *B*, Relation between site operators meeting pH measurement-accuracy criteria, in percent, and the median pH of the intersite-comparison study reference solution in studies 11-24, excluding study 22.

---

*Study 22 was excluded because accuracy criteria of ±0.10 pH unit were applied for all studies except 22. Accuracy criteria of ±0.30 pH unit were used in study 22.

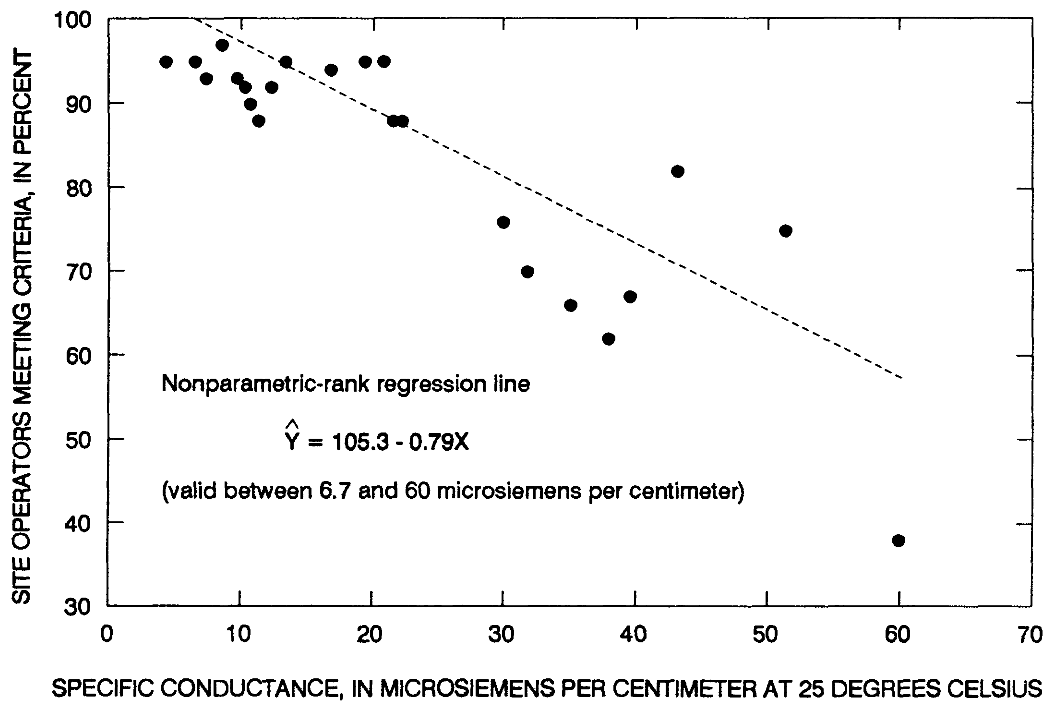

Figure 10.--Relation between site operators meeting specific-conductance measurement-accuracy criteria, in percent, and the median specific conductance of the intersite-comparison study reference solution.

SUMMARY

The intersite-comparison program is an ongoing, external quality-assurance program used to assess the ability of NADP/NTN site operators to make accurate and precise pH and specific conductance measurements. During intersite-comparison studies 1-24, 3,124 pH determinations were made. Of this total, NOS measurement-accuracy criteria were met for 2,319 pH determinations (74 percent). Because of significant changes in the operating protocol of the network, it is meaningful to compare the results of studies 1-10 with the results of studies 11-24. These changes included the decision by the NOS to supply site operators with uniform-quality pH electrodes and the initiation of site-visitation and site-operator training programs. The percentage of site operators reporting pH values within NOS measurement-accuracy criteria for the network increased from an average of 61 percent for intersite-comparison studies 1-10 to an average of 78 percent for studies 11-24.

Between the completion of intersite-comparison study 1 and study 24, there was an improving trend in the ability of site operators to successfully meet the accuracy criteria for pH determinations. The percentage of site operators successfully meeting the measurement-accuracy criteria was larger in recent studies compared with the percentage meeting the criteria in earlier studies. Nonparametric correlation and regression analysis techniques were used to determine if there was a statistically significant trend in the ability of site operators to meet measurement-accuracy criteria.

To determine whether the selected target pH affected the degree of similarity (that is, the precision) among pH measurements reported by the site operators in a given study, a nonparametric correlation test was used to assess the degree of association between the measured minus median interquartile range and the median pH of the reference solution. For studies 11-24, the size of the measured minus median interquartile range increased monotonically with median pH; there was less precision among reported pH values when the median pH of the reference solution was relatively larger.

No statistically significant correlation between the median pH of the reference solution and the percentage of site operators achieving NOS measurement-accuracy criteria was determined for intersite-comparison studies 1-10. Excluding study 22 because of the different pH measurement-accuracy criteria that were applied, a statistically significant Spearman rank correlation coefficient between the median pH of the reference solution and the percentage of site operators achieving the accuracy criteria was determined for studies 11-24. A relatively larger percentage of site operators met the accuracy criteria for pH in studies 11-24 when the median pH of the reference solution was relatively smaller. The median pH was between 4.50 and 5.00 pH units in two of the first ten studies; in studies 11-24, the median pH was in this range in eight studies. The increase in the percentage of site operators that met the accuracy criteria in studies 11-24 compared with studies 1-10 was in spite of the fact that the reference solutions used in studies 11-24 generally had larger median pH values than the reference solutions used in studies 1-10.

The percentage of site operators reporting specific-conductance values within the NOS measurement-accuracy criteria increased from an average of 73 percent in studies 1-10 to an average of 90 percent in studies 11-24. Between 88 and 97 percent of the participating site operators met the accuracy criteria when the median specific conductance was less than 25 µS/cm. Site operators were much less successful meeting the measurement-accuracy criteria when the median specific conductance was between 30 and 60 µS/cm; 38 to 82 percent of the site operators met the criteria when the median specific conductance of the reference solution was in this range. The Spearman rank correlation coefficient between the median specific conductance of the reference solution and the percentage of site operators achieving the measurement-accuracy criteria was statistically significant for studies 1-24. The large increase in the percentage of site operators that met the measurement-accuracy criteria in studies 11-24 compared with studies 1-10 was mostly because of differences in the median specific conductance of the reference solutions used in studies 1-10 and studies 11-24. The median specific conductance was between 30 and 60 µS/cm in six of the first ten studies; in studies 11-24, the median specific conductance was in this range in only three studies.

For the 24 studies, a total of 55 percent of the determinations was within ±1 μS/cm of the median values of the individual studies, and 82 percent was within ±3 μS/cm of the median. The percentage of determinations approximating the median values of the individual studies indicates a large degree of measurement precision on an overall basis.

To facilitate the evaluation of pH and specific-conductance determinations, the current NOS measurement-accuracy criteria need to be revised. In place of the current accuracy criteria of ±0.10 or ±0.30 unit for pH and ±4.0 μS/cm for specific conductance, a continuum of accuracy criteria could be established. New measurement-accuracy criteria could, for example, be based on statistical quality-control techniques. Alternatively, adjustable ranges of acceptable percent measurement error depending on the target values of the reference solution could be established.

The operators of a small number of sites accounted for a disproportionately large number of pH determinations that failed to meet the measurement-accuracy criteria. The operators of 12 percent of the sites that participated in studies 1-24 were responsible for 29 percent of the pH determinations that failed to meet measurement-accuracy criteria. Ten of the 22 sites whose operators had the poorest success rates of achieving the accuracy criteria for specific-conductance determinations were also among the 32 sites whose operators had the poorest success rates of achieving accuracy criteria for pH determinations.

SELECTED REFERENCES

Brooks, M.H., Schroder, L. J., and Willoughby, T.C., 1988, Results of external quality-assurance program for the National Atmospheric Deposition Program and National Trends Network during 1985: U.S. Geological Survey Water-Resources Investigations Report 87-4219, 14 p.

Chambers J.M., Cleveland, W.S., Kleiner, Beat, Tukey, P.A., 1983, Graphical methods for data analysis: Boston, Duxbury Press, 395 p.

Conover, W.J., 1980, Practical Nonparametric Statistics (2d ed.): New York, John Wiley, 493 p.

Dixon, W.J. and Massey, F.J., 1969, Introduction to statistical analysis (3d ed.): New York, McGraw-Hill, 638 p.

Galloway, J.N., Cowling, E.B., Gorham, Eville, and McFee, W.W., 1978, A national program for assessing the problem of atmospheric deposition (acid rain): National Atmospheric Deposition Program Report to The Council on Environmental Quality, 94 p.

Hem, J.D., 1985, Study and interpretation of the chemical characteristics of natural water (3d ed.): U.S. Geological Survey Water-Supply Paper 2254, 263 p.

Hoaglin, D.C., Moesteller, F., and Tukey, J.W., eds., 1983, Understanding robust statistics and exploratory data analysis: New York, John Wiley, 447 p.

Lindberg, S., Bigelow, D., Bowersox, V., Knapp, W., and Olsen, T., 1982, National Atmospheric Deposition Program annual data summary, precipitation chemistry in the United States: National Atmospheric Deposition Program Subcommittee Number 3, Data Management and Analysis, 134 p.

Schroder, L.J., and Brennan, J.O., 1984, Precision of the measurement of pH and specific conductance at National Atmospheric Deposition Program monitoring sites, October 1983-November 1983: U.S. Geological Survey Water-Resources Investigations Report 84-4325, 12 p.

Schroder, L.J., and Malo, B.A., 1984, Quality assurance program for wet deposition sampling and chemical analysis for the National Trends Network: APCA/ASQC Specialty Conference on Quality Assurance in Air Pollution Measurements, Boulder, Colo., October, 1984, Proceedings, p. 1-7.

Schroder, L.J., Brooks, M.H., and Willoughby, T. C., 1987, Results of inter-comparison studies for the measurement of pH and specific conductance at National Atmospheric Deposition Program/National Trends Network monitoring sites, October 1981-October 1985: U.S. Geological Survey Water-Resources Investigations Report 86-4363, 22 p.

See, R.B., Schroder, L.J., and Willoughby, T.C., 1988, External quality-assurance results for the National Atmospheric Deposition Program and The National Trends Network During 1986: U.S. Geological Survey Water-Resources Investigations Report 88-4007, 15 p.

See, R.B., Willoughby, T.C., and Gordon, J.D., 1990, External quality-assurance results for the National Atmospheric Deposition Program/ National Trends Network during 1988: U.S. Geological Survey Water-Resources Investigations Report 90-4030, 26 p.

See, R.B., Willoughby, T.C., Gordon, J.D., and Brooks, M.H., 1990, Programs and Analytical Methods for the U.S. Geological Survey Acid Rain Quality Assurance Project: U.S. Geological Survey Water-Resources Investigations Report 90-4029, 57 p.

Taylor, J.K., 1987, Quality Assurance of chemical measurements: Chelsea, Mich., Lewis Publishers, 328 p.

Topol, L.E. and Schwall, R.J., 1987, Network description and measurements for 1981 and 1982, Volume 1: Camarillo, Calif., Environmental Monitoring and Services, Inc., Utility Acid Precipitation Study Program 105, 242 p.

☆U.S. GOVERNMENT PRINTING OFFICE: 1991 - 575-817/45113