Journal of Research of the National Bureau of Standards

Volume 91 **Number 1** **January–February 1986**

special issue

pH measurements in rainwater

Foreword
Harry S. Hertz and Stanley D. Rasberry 1

Ruggedness Testing—Part I: Ignoring Interactions
Robert C. Paule, George Marinenko, Melissa Knoerdel, and William F. Koch 3

Ruggedness Testing—Part II: Recognizing Interactions
Robert C. Paule, George Marinenko, Melissa Knoerdel, and William F. Koch 9

Effect of Variables on pH Measurement in Acid-Rain-Like Solutions as Determined by Ruggedness Tests
George Marinenko, Robert C. Paule, William F. Koch, and Melissa Knoerdel 17

An Interlaboratory Test of pH Measurements in Rainwater
William F. Koch, George Marinenko, and Robert C. Paule 23

Development of a Standard Reference Material for Rainwater Analysis
William F. Koch, George Marinenko, and Robert C. Paule 33

ISSN 0160-1741 **Library of Congress Catalog Card No.: 63-37059**

The Journal of Research of the National Bureau of Standards features advances in measurement methodology and analyses consistent with the NBS responsibility as the nation's measurement science laboratory. It includes reports on instrumentation for making accurate and precise measurements in fields of physical science and engineering, as well as the mathematical models of phenomena which enable the predictive determination of information in regions where measurements may be absent. Papers on critical data, calibration techniques, quality assurance programs, and well characterized reference materials reflect NBS programs in these areas. Special issues of the Journal are devoted to invited papers in a particular field of measurement science. Occasional survey articles and conference reports appear on topics related to the Bureau's technical and scientific programs.

Issued six times a year. Annual subscriptions: domestic $17.00; foreign $21.25. Single copy, $3.00 domestic; $3.75 foreign.

United States Government Printing Office, Washington: 1986
Order all publications from the Superintendent of Documents
U.S. Government Printing Office, Washington, DC 20402

The Secretary of Commerce has determined that the publication of the periodical is necessary in the transaction of the public business required by law of this Department. Use of funds for printing this periodical has been approved by the Director of the Office of Management and Budget through April 1, 1986.

FOREWORD

By now virtually every technologist is aware that serious environmental consequences result from acidic rainfall. Indeed every American, aware or unaware, is impacted by its deleterious effects on our forests, lakes, lands, property, and personal health. Neither the technical nor political consequences stop at national borders. This is particularly the case in North America and in Europe; both, places where the problems with acid rain are being given serious attention.

This issue of the NBS *Journal of Research* is devoted to the topic of measurement methods, Standard Reference Materials, and associated statistical considerations for effective analysis of rainwater. Such methods and standards are essential to accurate assessment of current environmental conditions and will be needed in the future to assess trends that result from corrective measures. It goes almost without saying that equitable regulation will only be possible with fair (accurate) measurement. We think it is important to note that the authors of this issue, Drs. Koch, Marinenko and Paule, and Ms. Knoerdel, are all from NBS. Because NBS has no regulatory role, the work they report here is free from any particular pressure to favor the regulator or regulated.

The authors, together with more than a dozen of their colleagues in the Center for Analytical Chemistry, have provided an excellent measurement benchmark in SRM 2694, Simulated Rainwater. However, their contributions on the problem of acid rain extend beyond those described in this issue. The Center also makes available to the measurement community several Standard Reference Materials of coal and oil, certified for sulfur content, an important precursor to acid rain. We commend this issue to every reader concerned with the assessment of acidic rainfall.

Harry S. Hertz
Director, Center for Analytical Chemistry

Stanley D. Rasberry
Chief, Office of Standard Reference Materials

Journal of Research of the National Bureau of Standards
Volume 91, Number 1, January-February 1986

Ruggedness Testing—Part I: Ignoring Interactions

Robert C. Paule, George Marinenko, Melissa Knoerdel, and William F. Koch
National Bureau of Standards, Gaithersburg, MD 20899

Accepted: August 29, 1985

A straightforward explanation of the statistical technique of ruggedness testing is presented. Efficient Plackett-Burman designs are used in ruggedness tests. These designs involve the simultaneous change of levels of a number of variables. The designs allow the ruggedness test user to determine the effect of the separated variables on the measurement process. This paper (Part I) deals with the common situation where two-factor and higher order interactions can be safely ignored. A method is presented for evaluating the experimental uncertainties. A detailed example of glass electrode measurements of pH of dilute HCl solutions is used to illustrate ruggedness testing procedures.

Key words: interactions; main effects; orthogonal designs; pH measurements; Plackett-Burman designs; ruggedness tests.

Introduction

The purpose of a ruggedness test is to find the factors that strongly influence measurement results, and to determine how closely one needs to control these factors. Ruggedness tests do not determine optimum conditions for a test method.

In the testing of a protocol, it is frequent occurence that the coordinating scientist is dismayed by the large variabilities observed between different laboratory results. The scientist may have developed the protocol being tested and has taken great care and pride in that development. His laboratory has documented "proof" of high precision and accuracy for the method. What has gone wrong? How can the other laboratories get such wild results?

A large part of the answer may be that the coordinating scientist has been unrealistically consistent in his own laboratory work. He may have always used fixed equipment such as a furnace that was set at 60.0 °C and that did not vary by more than ±0.5 °C. Even though the furnace dial read 60.0 °C, the furnace temperature may in reality have been 64.2±0.5 °C. The constant bias of 4.2 °C did not affect his precision, but it may have affected his accuracy. Other constant errors will, likewise, not affect his precision. In regard to accuracy, these additional errors may partially cancel each other. It is the nature of protocol development that work will continue until the errors do cancel, and the "right" answer is obtained. Thus, the laboratory that has developed the protocol will eventually show both good precision and accuracy. In an interlaboratory experiment, however, conditions are different. The other (individual) laboratories do not have the same biases, and the rather complete cancelling of systematic errors does not occur. Differences in laboratory conditions can result in

About the Authors: Robert C. Paule is a physical scientist assigned to the NBS National Measurement Laboratory (NML). George Marinenko and William F. Koch are chemists in NML's Inorganic Analytical Research Division in which Melissa Knoerdel, a student, serves the Division during summer vacations.

large variabilities between different laboratory results. In frustration, the coordinating scientist may tighten the protocol specifications. One can see that if temperature is important, then even a tightened protocol specification of 60.0±0.1 °C will not be effective unless the biases *between* laboratories are eliminated. A true temperature of 60.0±0.5 °C may be quite satisfactory, but large biases cannot be tolerated.

To work towards perfecting a test method one must first determine if a factor such as temperature is important, and then decide if a true ±0.5 °C tolerance is acceptable. Such matters are best investigated in a single laboratory rather than in multiple laboratories since, here, we are interested in the effect of *changes* in temperature. A constant bias within a single laboratory will not interfere in the investigation of changes of temperature. Other factors associated with the protocol must also be evaluated. How do we proceed?

The coordinating scientist may believe that the protocol contains seven factors (variables) that could influence the measurement results. Suppose it is decided to investigate the effect of each factor at only two levels: at a high level and at a low level. A full factorial investigation of the seven factors at each of the two levels would require $2^7=128$ measurements, and this does not include replicate measurements. Fortunately, one does not have to make this many measurements. One can use a class of experimental designs called Plackett-Burman designs [1].[1] It is possible, by using these designs, to study up to $N-1$ factors using only N measurements.

A Mathematical Model

A brief review of a mathematical model used to describe a measurement result may be helpful in understanding details that are associated with the use of Plackett-Burman designs. For simplicity, consider an experiment with only three factors at each of two levels (eight measurements).

$$Y_{ijk}=Y_{\ldots}+A_i+B_j+C_k+AB_{ij}+AC_{ik}+BC_{jk}+ABC_{ijk}$$

where

Y_{ijk} = a single measured value ($i,j,k=1,2$ — the low and the high levels)

$Y_{\ldots}$ = the overall average for all eight measurements

A_i, B_j, C_k = the estimated main effects (the main factors affecting the measurement results)

AB_{ij}, AC_{ik}, BC_{jk} = the estimated two-factor interactions (systematic effects not explained by the main effects)

ABC_{ijk} = the estimated three-factor interactions (systematic effects not explained by the main effects and the two-factor interactions).

There are some restrictions on the main effects and interaction terms in the model. The restrictions will not be given here since they only have to do with the "centering of the data" for the evaluation of the terms. In ruggedness testing we do not center the data about some midpoint, but rather redefine the effects as differences between the results at the high and at the low levels. We will also do away with the subscripts of the above model. We simply recognize that measurement results are affected by various main effects and interactions.

From the general mathematical model one can infer that experiments with a larger number of factors will have a very large number of higher-order interactions. It is generally believed that main effects tend to be most important in describing (or controlling) the measurement results, that two-factor interactions are less important, and that higher order interactions are even less important. Plackett-Burman designs are well suited for measurement processes that have negligible interactions.

Use of Plackett-Burman Designs

The most common use of Plackett-Burman (PB) designs with N measurements allows one to get the most important (main effects) information. With N measurements, however, the $N-1$ main effects are confounded with the two-factor and with higher order interactions. If the interactions are relatively small, then we may be satisfied in making only N measurements and obtaining slightly contaminated estimates for the $N-1$ main effects. Experience has tended to show that one gains more useful information by examining additional factors than by evaluating the interactions.

Numerous PB-designs are available [1]. A PB-design for seven factors and eight measurements is given in table 1. A (+) for a given factor indicates that the measurement is made with that factor set at the high level, and a (−) indicates the factor is to be at the low level. All seven factors are set for each measurement and a single result is obtained from each of the eight measurements. The measurements should be made in a random order. Typical measurement results are shown at the far right of the design. Scanning down each column of the design one sees that there are equal numbers of (+) and (−) factor settings.

---

[1] Figures in brackets indicate literature references.

Table 1. A Plackett-Burman design for N=8.

Run	Factor A	B	C	D	E	F	G	Results
1	+	+	+	−	+	−	−	1.1
2	−	+	+	+	−	+	−	6.3
3	−	−	+	+	+	−	+	1.2
4	+	−	−	+	+	+	−	0.8
5	−	+	−	−	+	+	+	6.0
6	+	−	+	−	−	+	+	0.9
7	+	+	−	+	−	−	+	1.1
8	−	−	−	−	−	−	−	1.4

The effect of any factor such as A, for example, is simply calculated as the average of the measurements made at the high level minus the average of the measurements made at the low level.

Effect of A

$$=\Sigma\frac{A(+)}{N/2}-\Sigma\frac{A(-)}{N/2}=2/N\times[\Sigma A(+)-\Sigma A(-)]\,. \quad (1)$$

Effect of A

$$=2/8\times[(1.1+0.8+0.9+1.1)-(6.3+1.2+6.0+1.4)].$$
$$=-2.75$$

The PB-design (see table 1) is constructed such that the $\Sigma A(+)$ and the $\Sigma A(-)$ terms will *each* contain an equal number of $B(+)$ and $B(-)$ terms. Thus, the A effect is orthogonal, i.e., is not affected by the B effect. In the PB-designs all main effects (columns) are orthogonal to all other main effects (columns). This orthogonality, however, does not extend to the interactions. The orthogonality of the main effects and the acceptance of a slight contamination of estimates for the main effects (by the interactions) are the major characteristics of ruggedness testing. For many practical problems this is all that is needed.

For the PB-design, the standard deviation for an effect, such as A, is obtained by using eq (1) and the standard deviation of a single measurement σ.

$$\sigma_{\text{effect }A}=\sqrt{(4/N^2)\times\text{Var}\,[\Sigma A(+)-\Sigma A(-)]}$$
$$=\sqrt{(4/N^2)\times N\sigma^2}$$
$$\sigma_{\text{effect }A}=2\sigma/\sqrt{N} \quad (2a)$$

The same equations for the PB-design apply when the standard deviation σ is replaced by a sample estimate, s.

$$s_{\text{effect }A}=2s/\sqrt{N} \quad (2b)$$

Two methods for determining a sample estimate of the standard deviation of a single measurement, s, will be presented.

PB-Design Considerations

Equation 2b shows that the standard deviation of an effect is inversely proportional to $\sqrt{N}$, the number of measurements made. One is therefore tempted to use large PB-designs. Practical experience, however, favors moderate size designs. Overly large designs require the correct setting of too many factors, and this increases the chance for blunders. In addition, large designs require more time to complete and one becomes concerned that other factors not being considered in the design can change and distort the results. The effects of incorrect factor settings and of shifting experimental conditions are propagated into *all* of the calculated results (see eq 1). The above listed ($N=8$) PB-design is a suitable size for most experiments. If more factors need to be studied, they can be handled by using a second ($N=8$) PB-design. This latter procedure may even involve the repeated testing of some of the more important factors from the first design. The ($N=8$) PB-design can also be conveniently used to study two-factor interactions (see Ruggedness Testing—Part II: Recognizing Interactions).

In general, the size of all effects in a PB-design will increase with increased separation of the high and low factor settings. We have implicitly assumed that the main effects are linear. It seems prudent to only use moderate separations of the high and low settings so that the measured effects will be relatively linear and, at the same time, large relative to the measurement error. For the high and low settings of the factors it is suggested that one use the extreme limits that one may expect to observe between different qualified laboratories.

Judging the Effects

How can one judge if any of the estimated main effects are too large? Since the main effects are expressed in the units of the measurement, one can simply make a direct judgment whether the change associated with a factor shift from a high level to a low level is too large, or not. Other, more quantitative methods of judgment which analyze the variance of measurements are given below. We should recognize that these quantitative methods still only give tentative answers and that follow-up or confirmatory experiments are frequently needed.

If n auxiliary replicate measurements are available, one can estimate the within-laboratory measurement variability, s. A t-test (with $n-1$ degrees of freedom) can be used to judge if a main effect is statistically significant relative to the measurement variability. Note that the n from the auxiliary replicate measurements will not generally be the same as the N of the ruggedness test.

$$t_{n-1}=\frac{\text{effect } A}{S_{\text{effect } A}}$$

Using eq 2b, this t-test can be written in the following form:

$$t_{n-1}=\frac{\text{calculated } A}{2s/\sqrt{N}}. \tag{3}$$

Action should be taken if the effect of a factor is statistically significant, and if the size of the effect is of practical importance; we should then tighten the protocol specification for that factor. This will help reduce the interlaboratory variability.

One may wish to repeat the complete PB-experiment so as to obtain better estimates of the factors and to get a current estimate of the within-laboratory measurement variability, s. In estimating the measurement variability one needs to guard against the occurrence of a possible measurement shift between the running of the two designs. This can be handled mathematically. Let us now work through a real example.

This ruggedness testing example deals with factors that may influence the determination of the pH in dilute acid solutions when measurements are made by use of a glass electrode. Table 2 gives the seven factor (N=8) PB-design which was used. This convenient design was first suggested by F. Yates [2]. It was frequently used by W. J. Youden [3] who did much of the pioneering work in ruggedness testing.

The above Yates-Youden design can be obtained from the seven-factor PB-design of table 1 by relabelling the PB-columns A-G to read C, F, G, D, E, B, A, and the PB-rows 1-8 to read 2, 3, 5, 4, 7, 8, 6, and 1. One then rearranges the columns and rows to be in the usual alphabetic and numeric order. The above operations are perfectly acceptable since the assignment of column and row labels is arbitrary and the rearrangement of the columns and rows has no effect on the overall arithmetic operations. Such rearrangements are, in fact, one means of randomizing the assignment of variables.

Table 2. The seven-factor PB design.

	Factor						
Run	A	B	C	D	E	F	G
1	−	−	−	−	−	−	−
2	−	−	+	−	+	+	+
3	−	+	−	+	−	+	+
4	−	+	+	+	+	−	−
5	+	−	−	+	+	−	+
6	+	−	+	+	−	+	−
7	+	+	−	−	+	+	−
8	+	+	+	−	−	−	+

A number of pH measurement experiments were run using six different dilute acid solutions. For simplicity of presentation, Part I discusses only the results from one of the solutions, an HCl solution with a known pH of 2.985. Subjects of more involved PB-testing and comparisons between the different acid solutions are described in Part II. The seven factors that were studied are listed below. The first listed level for each factor has been arbitrarily assigned the positive sign in the above table.

A. Temperature: 25 °C or 30 °C.

B. Stirring during the pH measurement: Yes or No

C. Dilution (0.5 mL distilled H_2O/20 mL of solution:
Yes or No

D. Depth of electrode immersion: 1 cm or 3 cm below liquid surface

E. Addition of $NaNO_3$ (0.033 mol/L of solution):
Yes or No

F. Addition of KCl (0.067 mol/L of solution):
Yes or No

G. Electrode equilibration time before reading the pH: 10 or 5 minutes

The above is only a partial list of factors that will change the observed value of the pH. Obviously, all other factors that are not listed above need to be kept constant. The particular, constant levels of these other factors will result in some specific offset in the pH measurements. In the ruggedness test, however, this fixed offset need not concern us since we are only interested in the measurement changes (the effects) that occur when the above seven factors ($A-G$) are changed.

Results from the ruggedness test are given in table 3. The complete experiment was also repeated on a second day. A different random order of measurement was used for each day. The two sets of measurement results are given at the far right of the design.

For the first set of the above reported measurements, the effect of factor A is calculated from eq 1 as the difference of the average value when 25 °C is used and the average value when 30 °C is used, i.e., $(2999+3055+3049+2949)/4-(2904+3015+3006+2964)/4=3013-2972=+41$. The averages and differ-

Table 3. Design and test results.

Factor							Observed pH (milli-pH units)	
A	*B*	*C*	*D*	*E*	*F*	*G*		
30	N	N	3	N	N	5	2904	2895
30	N	Y	3	Y	Y	10	3015	3017
30	Y	N	1	N	Y	10	3006	2990
30	Y	Y	1	Y	N	5	2964	2935
25	N	N	1	Y	N	10	2999	2983
25	N	Y	1	N	Y	5	3055	3053
25	Y	N	3	Y	Y	5	3049	3044
25	Y	Y	3	N	N	10	2949	2949
					Average		2993	2983

ences of the averages (the effects) are given for factors $A-G$ in the third and fourth columns of table 4. Similar calculations for the second set of measurements are given in the fifth and sixth columns of the table.

Testing the Effects From Repeated (pH) Experiments

Generally good agreements are observed between the calculated effects from the two sets of measurements. Effects A, D, E, and F are relatively large and are of interest. The average C effect is $(6+11)/2=+8.5$. To help decide if the C effect value is real, or if it might simply be due to imprecisions in the measurements, let us make a t-test.

$$t=\frac{\text{effect of avg. } C}{s_{\text{effect of avg. } C}}$$

Table 4. The effects for factors $A-G$. (milli-pH units)

		First Data Set		Second Data Set		
Factor	Level	Average	Effect	Average	Effect	Differences (d) betw. effects
A	25	3013		3007		
A	30	2972	+41	2959	+48	−7
B	Y	2992		2980		
B	N	2993	−1	2987	−7	+6
C	Y	2996		2989		
C	N	2990	+6	2978	+11	−5
D	1	3006		2990		
D	3	2979	+27	2976	+14	+13
E	Y	3007		2995		
E	N	2979	+28	2972	+23	+5
F	Y	3031		3026		
F	N	2954	+77	2941	+85	−8
G	10	2992		2985		
G	5	2993	−1	2982	+3	−4

Since the estimate for each effect is now the average of two experiments the t-test, derived in the form of eq 3, must be modified as follows:

$$t=\frac{\text{calculated avg. } C}{2s/\sqrt{2\mathrm{N}}}. \tag{4}$$

The estimate of the standard deviation, s, and the associated degrees of freedom for the t-test are obtainable from our measurements. Since the two sets of measurements were run on different days, we should be concerned that one set of measurements could be offset relative to the other set. Let us therefore calculate the s value by a method that is not vulnerable to an offset between the two sets of measurements.

We first note that an offset between the sets of measurements will not influence the values of the calculated effects. Let us therefore consider the differences between the effects as calculated for the above example (see table 4, column 7). Since we are considering the same effects from the two sets of experiments, the statistically expected values of the differences between the effects are zero. The variance of the difference is therefore the expected value of the squared differences.

$$\text{Variance of } (d)=\text{Expected value of } (d^2)$$
$$\approx\Sigma d^2/(\mathrm{N}-1) \tag{5}$$

An estimate of the expected value of (d^2) is obtained by simply averaging the squares of the differences listed in table 4, column 7. Our calculated estimate is $384/7=54.9$.

We next note that the variance of the difference (between the duplicated effects) is the sum of the variances of the two effects. The variances of the two effects should be the same since the two sets of experiments were done in the same laboratory. Equation 2b described the sample estimate for the square root of the variance of an effect. Therefore:

$$\text{Estimated variance of } (d)=4s^2/\mathrm{N}+4s^2/\mathrm{N}=8s^2/\mathrm{N}. \tag{6}$$

By combining eqs (5) and (6) and rearranging we obtain an estimate of the standard deviation of a single measurement that has $\mathrm{N}-1$ degrees of freedom associated with it.

$$s=\sqrt{[\Sigma d^2/(\mathrm{N}-1)]\times\mathrm{N}/8} \tag{7}$$

The desired t-test is obtained by combining eqs 4 and 7.

$$t_{N-1}=\frac{\text{calculated avg. } C}{2\sqrt{[\Sigma d^2/(N-1)]\times N/8}/\sqrt{2N}} \qquad (8)$$

In the current example, N equals eight so we get:

$$t_7=\frac{\text{calculated avg. } C}{2\sqrt{\Sigma d^2/7}\,/\sqrt{2\times 8}}$$

$$=\frac{+8.5}{\sqrt{384/7}\,/\sqrt{16}}=+2.30\,.$$

This quantity, in absolute value, it is slightly less than the 5% critical t-value of 2.36. It is not quite statistically significant. The C factor describes the effect of a small dilution, as one might get from not properly wiping dry the glass electrode.

As mentioned above, if the effect of any factor is too large one may wish to tighten the specification for that factor. The goal, of course, is to reduce the interlaboratory variability. More detailed discussions of the pH measurement experiments are presented in Part II.

Other PB-Designs

Numerous Plackett-Burman designs [1] are available. The following is a method for constructing the designs for various numbers of measurements, N=4, 8, 12, 16, and 20. The first row of each design is given opposite the N-value. Each row specifies the N−1 high [+] and low (−) factor settings.

$N=4 \quad + + -$
$N=8 \quad + + + - + - -$
$N=12 \quad + + - + + + - - - + -$
$N=16 \quad + + + + - + - + + - - + - - -$
$N=20 \quad + + - - + + + + - + - + - - - - + + -$

For any selected N-value, the corresponding set of (+) and (−) signs is written down as the first row of the design. The second row of the design is obtained by copying the first row after shifting it one place to the right and putting the last sign of row 1 in the first position of row 2. This type of cyclic shifting should be done a total of N-2 times, after which a final row of all minus signs is added. The result of this procedure for the N=8 Plackett-Burman design is given in table 1.

Some ruggedness test studies may not involve exactly N−1 factors. If we believe, for example, that only five instead of seven factors might influence the measured results, we might use two dummy factors. For one of the dummy factors we might pour a solution with our left hand for the (+) level and with our right hand for the (−) level. The calculated "effect" for the dummy factor should be small and should simply reflect our random errors of measurement.

Conclusions

A straightforward explanation of the statistical technique of ruggedness testing has been presented. Orthogonal Plackett-Burman designs allow the ruggedness test user to efficiently evaluate the effects of the separated variables on a measurement process. The present article (Part I) deals with the common situation where two-factor and higher order interactions can be safely ignored.

References

[1] Plackett, R. L., and J. P. Burman, The Design of Optimum Multifactorial Experiments, Biometrika, Vol. 33, 305-325 (1946).
[2] Yates, F., Complex Experiments, J. Roy. Statistical Soc. (Supplement), Vol. 2, 181-247 (1935).
[3] Youden, W. J., Designs for Multifactor Experimentation, Industrial and Engineering Chemistry, Vol. 51, 79A-80A (1959).
[4] Diamond, W. J., Practical Experimental Designs for Engineers and Scientists, pp. 103 and 110, Lifetime Learning Publications, Belmont, CA (1981).
[5] Marinenko, George; Robert C. Paule, William F. Koch, and Melissa Knoerdel, Effect of Variables on pH Measurement in Acid-Rain-Like Solutions as Determined by Ruggedness Tests, J. Res. Natl. Bur. Stand. **91-1** (1986).

Journal of Research of the National Bureau of Standards
Volume 91, Number 1, January-February 1986

Ruggedness Testing– Part II: Recognizing Interactions

Robert C. Paule, George Marinenko, Melissa Knoerdel, and William F. Koch
National Bureau of Standards, Gaithersburg, MD 20899

Accepted: August 28, 1985

This paper is a continuation of the preceding article which introduced the reader to the general concepts of ruggedness testing. The current paper describes the effects of interactions on the measurement process, and presents procedures for the separation of the main effects from the two-factor interactions. The general characteristics of interactions are described in some detail. A short-cut procedure is presented for the calculations. A number of examples of glass electrode measurements of pH of dilute acid solutions are used to illustrate ruggedness testing procedures.

Key words: glass electrodes; main effects; orthogonal designs; pH measurements; ruggedness testing; short-cut calculations; two-factor interactions.

Introduction

This paper is a continuation of the preceding (Part I) article which introduced the general principles of ruggedness testing. To be read in conjunction with Part I, it describes the effects of interactions on a measurement process and presents procedures for separating main effects and two-factor interactions.

Interactions and That Confounded Confounding

From Part I we know that an N measurement experiment can be used to determine N-1 main factors, provided the interactions are small. It is usually the case, in experiments involving well-behaved functions of the measurement variables, that when the main effects are small the associated interactions are very small. The interactions are, in effect, the non-ideal departures from a simple additive model consisting of only constant main effects. Nevertheless, situations occasionally arise in which interactions are important.

In an eight-run, seven-factor experiment each main effect is confounded with 15 different possible interactions. Of the 15 interactions, the number and types are as follows: 3 two-factor, 4 three-factor, 4 four-factor, 3 five-factor, and 1 six-factor. Table 5,[1] which corresponds to the Yates-Youden design (see table 2 of part I), shows each of the main effects and the associated two- and three-factor interactions.

About the Authors: Robert C. Paule is a physical scientist assigned to the NBS National Measurement Laboratory (NML). George Marinenko and William F. Koch are chemists in NML's Inorganic Analytical Research Division in which Melissa Knoerdel, a student, serves the Division during summer vacations.

Table 5.[1] Interactions associated with the main effects.

A	*B*	*C*	*D*	*E*	*F*	*G*
–*BD*	–*AD*	–*AE*	–*AB*	–*AC*	–*AG*	–*AF*
–*CE*	–*CF*	–*BF*	–*CG*	–*BG*	–*BC*	–*BE*
–*FG*	–*EG*	–*DG*	–*EF*	–*DF*	–*DE*	–*CD*
BCG	*ACG*	*ABG*	*ACF*	*ABF*	*ABE*	*ABC*
BEF	*AEF*	*ADF*	*AEG*	*ADG*	*ACD*	*ADE*
CDF	*CDE*	*BDE*	*BCE*	*BCD*	*BDG*	*BDF*
DEG	*DFG*	*EFG*	*BFG*	*CFG*	*CEG*	*CEF*

---

[1] This paper (Part II), being a continuation of Part I, extends Part I's numbering system for tables and equations; its citations are listed in Part I's References section.

The Yates-Youden design (and the Plackett-Burman designs of a size such that $N = 2^k$, where k is a positive integer) allow a relatively easy separation and determination of the more important confounding interactions. These designs allow one to use the Multiplication Rule for signs. The Multiplication Rule [4][1] states that the pairwise multiplication of like signs produces a (+) and that of unlike signs produces a (−). Thus, looking at table 2 of Part I, the row pairwise multiplication of the signs for columns B and D produces the following column for the BD interaction:

BD
+
+
+
+
−
−
−
−

Note that this column is the exact opposite of the signs of column A, given in table 2. Thus, $-BD$ is the same as A. It is inseparable from A in the eight-run, seven-factor experiment since the values of the eight measurements are combined in an identical manner. Similar multiplications of signs shows that $A = -CE = -FG$. Multiplication of signs of the rows of columns "BC" and G produces the three-factor interaction BCG which is observed to be the same as factor A. Column "BC" can be simply obtained by using minus column F (see table 5). The confounding of all higher order interactions can be obtained by an extension of this general procedure.

If we wish to protect ourselves from misinterpretations due to large interactions, we must make more than N measurements for determining the N-1 main factors. To evaluate the main effects and all interactions, we must do the full factorial experiment. For seven factors this requires 128 measurements. Usually, however, one does not have to go this far. A reasonable compromise experiment consists of making two sets of N measurements which allow the separation of each of the main effects from the two-factor interactions. This compromise, however, does not separate *among* each of the two-factor interactions, and in addition it assumes that three-factor and other higher order interactions are unimportant. If we demand more information, then we have no choice but to make more measurements!

Let us again consider a seven factor pH experiment involving 2N (= 16) measurements. This time we will use the previously reported first set of pH measurements (see table 3 of Part I), *and* a third set of pH measurements which was made with all levels of the design reversed. Let us now consider the combined results from the first and third sets of measurements.

An examination of the signs of table 6, and the use of the Multiplication Rule, will show that the two-factor interactions $-BD$, $-CE$, and $-FG$ (which were grouped together in column 1 of table 5) still have an identical sign pattern but that this pattern is now different from the A main effect. The $-BD$ interaction has the following sign pattern (− − − − + + + + − − − − + + + +). One can see that the last half of the interaction sign pattern is a repetition of the first half whenever an even-number of factors is multiplied together, but that the last half has a sign reversal whenever an odd-number of factors is multiplied together.

Table 6. Designs and test results.

Factor							Observed pH
A	*B*	*C*	*D*	*E*	*F*	*G*	(*milli-pH units*)
−	−	−	−	−	−	−	2904
−	−	+	−	+	+	+	3015
−	+	−	+	−	+	+	3006
−	+	+	+	+	−	−	2964
+	−	−	+	+	−	+	2999
+	−	+	+	−	+	−	3055
+	+	−	−	+	+	−	3049
+	+	+	−	−	−	+	2949
+	+	+	+	+	+	+	3040
+	+	−	+	−	−	−	2931
+	−	+	−	+	−	−	2978
+	−	−	−	−	+	+	3030
−	+	+	−	−	+	−	2967
−	+	−	−	+	−	+	2911
−	−	+	+	−	−	+	2874
−	−	−	+	+	+	−	2979

From table 6, and the Multiplication Rule, one can see that the three two-factor interactions within each column of table 5 are not separable from one another, but that they are separable from the main effects. The three-factor interactions are not separable from the main effects. A further consideration of table 6 will show that the main effects and their odd-numbered interactions are not separable from one another, but that they are separable from all of the even-numbered interactions. The nearest higher order interaction contamination for either the odd- or the even-numbered interactions is now two-factor multiples distant. If the magnitude of the interactions decreases as one goes toward higher order interactions, then one has achieved a practical separation (isolation) of the main effects and of the groups of two-factor interactions.

Main effects *A–G* can be calculated from the data of table 6 by use of eq (1) (from Part I). The calculated respective effects are +51, −2, +4, +6, +27, +79, and −0.4 milli-pH units. The two-factor interactions are calculated in the same manner as the main effects. Note that the value of the "new N" in eq (1) is the combined N from both sets of measurements (new N = 16). As shown above, the sign pattern for the $-BD$ interaction is (− − − − + + + + − − − − + + + +). The value for the combined (*BD*, *CE*, *FG*) interactions is +11 milli-pH units. The other two-factor interactions can be calculated in a similar manner. Finally, we note from table 6 that if an offset had occurred between the first and third set of measurements, it would not affect the calculations of the main effects or the interactions. This immunity to offsets between the different sets of measurements is a consequence of using the Plackett-Burman based design. The PB-design will always have an equal number of positive and negative signs within each set so that the absolute level of the sets of measurements will not affect the calculations.

Short-Cut Calculations

All of the ruggedness testing calculations are conceptually quite simple, but are tedious to perform. Hand calculators that have at least nine memory registers allow short-cuts that minimize the arithmetic operations and the keying of the data. We will assume here that our calculations are made on sets of eight measurements. Let the average of these measurements be $\bar{X}$ Starting from eq (1) of Part I, the derivation of the short-cut method is as follows:

$$\text{Effect } A = \frac{\Sigma A(+)}{\text{N}/2} - \frac{\Sigma A(-)}{\text{N}/2}$$

$$= \frac{\Sigma A(+)}{\text{N}/2} + \left[\frac{\Sigma A(+)}{\text{N}/2} - \frac{\Sigma A(+)}{\text{N}/2}\right] - \frac{\Sigma A(-)}{\text{N}/2}$$

$$\text{Effect } A = \frac{4\Sigma A(+)}{\text{N}} - 2\bar{X} \qquad (9)$$

For N = 8,

$$\text{Effect } A = \Sigma A(+)/2 - 2\bar{X} \qquad (9\text{a})$$

Let us now rewrite the table 2 design of Part I for the first set of eight pH measurements, substituting the ordered measurement numbers for the positive signs. For the set of eight measurements, one keys the measurements into memory registers 1–8, respectively, and then calculates the last term of eq (9a) which is two times the average of the eight measurements. This quantity is stored in memory register 9. In order to minimize the chance of error, it is advisable to use the measurement results that are stored in memory registers 1–8 to calculate this latter quantity. One then simply uses eq (9a) and the columns of table 7 to calculate the various effects:

$$\text{Effect } A = (\text{Registers } 5+6+7+8)/2 - \text{Register } 9$$
$$= +40.75 \text{ milli-pH}$$
$$\text{Effect } B = (\text{Registers } 3+4+7+8)/2 - \text{Register } 9$$
$$= -1.25 \text{ milli-pH.}$$

Table 7. Relabeled design and test results.

Factor *A*	 *B*	 *C*	 *D*	 *E*	 *F*	 *G*	Observed pH (*milli-pH units*)
−	−	−	−	−	−	−	2904
−	−	2	−	2	2	2	3015
−	3	−	3	−	3	3	3006
−	4	4	4	4	−	−	2964
5	−	−	5	5	−	5	2999
6	−	6	6	−	6	−	3055
7	7	−	−	7	7	−	3049
8	8	8	−	−	−	8	2949
						Average	2992.625

The reverse sign PB-design listed in the bottom of table 6 can be similarly rewritten and used with eq (9a) to again calculate effects *A–G*.

Table 8 lists the calculated effects from the three sets of eight pH measurements that have been previously reported in tables 3 and 6. The actual, chronological order used for making our measurement sets consisted of the table 2 design of Part I, the reverse-sign design, the repeat table 2 design, and occasionally a repeat of the reverse-sign design. The labeling from our previously referenced data sets 2 and 3, will henceforth be reversed to conform to the chronological order. Thus, set 2 will now refer to the reverse-sign design, and set 3 to the repeat table 2 design.

Table 8. Calculated effects.

			(milli-pH units)				
Factor	*A*	*B*	*C*	*D*	*E*	*F*	*G*
Set 1	41	−1	6	27	28	77	−1
Set 2	62	−3	2	−16	26	80	0
Set 3	48	−7	11	14	23	85	3

We see that the set 2, reverse-sign PB-design, gives slightly different results. Note that the use of the Multiplication Rule on the reverse-sign PB-design (listed in the bottom half of table 6), results in positive (rather than negative) two-factor interactions which are confounded with the main effects. For example, $A=+BD=+CE=+FG$. Let us take averages for the table 8 results for sets 1 and 2. For each of the averages of sets 1 and 2, the two-factor (and other even-number factor interactions) drop out. By similar reasoning the differences between sets 1 and 2, when divided by two, yields the separated even-number factor interactions.The set 1 and 2 averages and average differences are listed in table 9.

Table 9. Averages and average differences for effects.

	For Data Sets 1 and 2 (milli-pH units)						
Factor	*A*	*B*	*C*	*D*	*E*	*F*	*G*
Average	52	−2	4	6	27	79	−1
Avg. Difference	11	−1	−2	−22	−1	2	1

The table 9 results for sets 1 and 2 are the same as the results obtained by the more tedious, direct calculations. The calculated effects from sets 2 and 3 could also be used to produce results comparable to table 9. It should be obvious that better (more stable) estimates can be obtained by first pre-averaging the effects from sets 1 and 3 before making the combined calculation with the set 2 data to produce the averages and average differences. The results from such calculations are given in table 10. For simplicity of presentation, these averages and average differences will hereafter be called the Main Effects and the Two-Factor Interactions, respectively.

Let us now summarize the short-cut calculations: for each data set use eq (9) or (9a), and its associated design table (such as table 7) to calculate the "contaminated" effects (as shown in table 8). Where possible, for like-sign designs, calculate the pre-averages. Also, for the like-sign designs calculate the squared differences of the "contaminated" effects (their use will be described in the next paragraph). The pre-averaged effects from the like- and reverse-sign designs are used to calculate averages and average differences (as shown in table 10). These latter averages and average differences are the separated, and relatively uncontaminated, main effects and interactions.

For PB-designs with N = 8, the standard deviation of a single measurement is obtained by taking the square-root of the average of the above calculated squared differences of the "contaminated" effects (see eq (7) of Part I). For the current experiment s equals $\sqrt{384/7}$ or 7.4 milli-pH units. It has 7 degrees of freedom associated with it.

Judging the Main Effects and Interactions

To help decide if the main effects and two-factor interactions are real, or if they may simply be due to imprecisions in the measurements, let us once again use the t-statistic.

$$t=\frac{\text{effect under test}}{s_{\text{effect under test}}}$$

To determine the t-value we must evaluate the denominator of the equation. Since the main effects (ME) and the two-factor interactions (2FI) are calculated by taking either the averages or average differences from the same sets of data, the standard deviation of the ME and the 2FI will be the same. For the current example, the ME and the 2FI are calculated as follows:

$$\frac{1}{2}\left(\frac{\text{Sets } 1+3}{2}\pm\text{ Set } 2\right)$$

The recognition of the form of the above calculations, and the use of the square of eq (2b), allow the evaluation of the standard deviation of the ME or the 2FI.

$$s(\text{ME or 2FI})=\frac{1}{2}\left[\frac{4s^2/8+4s^2/8}{2^2}+\frac{4s^2/8}{1}\right]^{\frac{1}{2}}=\sqrt{3}\,s/4 \qquad (10)$$

$$t=\frac{4\times(\text{effect under test})}{\sqrt{3}\,s}.$$

Table 10. Main effects and two-factor interaction.

	For Data Sets [Avg. (1 and 3)] and 2 (milli-pH units)							
Factor	*A*	*B*	*C*	*D*	*E*	*F*	*G*	Legend for + Factors
Main Effects	53	−4	5	2	26	81	1	*A* = 25 °C
Two-Factor Interactions	9	1	−3	−18	0	−1	−1	*B* = solution stirred
								C = 0.5 mL dilution
								D = 1 cm electrode immersion
								E = $NaNO_3$ added
								F = KCl added
								G = pH measured at 10 min

For the current example s equals $\sqrt{384/7}$ or 7.4 milli-pH units. The observed t-value is as follows:

$$t_7 = 0.31 \times (\text{effect under test})\,.$$

For the main effect $A (=53)$ we have:

$$t_7 = 0.31 \times 53 = 16\,.$$

We see from table 10 that the main effects A, E, and F, and the two-factor interactions under columns A and D are statistically significant. This is very similar to our prior conclusions in Part I, except that we now distinguish between the main effects and the two-factor interactions.

Results From the Other Dilute Acid Solutions

A total of six different dilute acid solutions covering a pH range from 3.0 to 5.0 was tested. The purpose of these tests with the six solutions was to further evaluate the measurement procedures and to determine a practical upper limit for the pH measurements. Toward the end of the tests, the original glass electrode was broken and a second glass electrode of the same model and manufacturer was used as a replacement. This accidental breakage gave additional practical insight into ruggedness testing.

Some information regarding the six acid solutions is given in table 11.

The main design (table 2 of Part I) and the reverse-sign design were run with all six solutions. A standard pH=4.00 buffer was used to calibrate the pH meter before making each set of eight measurements. In many cases the designs were repeated and the standard deviations for a single measurement, s, were calculated. In accordance with the chronological order of the measurements, the main designs (table 2 of Part I) are labeled as sets 1 and 3, and the reverse-sign designs are labeled as sets 2 and 4. The calculated pre-averages for the like-sign designs are listed in table 12.

While making the measurements on sets 1, 2, and 3 of HCl solution 6, it was clear that there was great instability in the measurements. This instability is reflected in the table-displayed standard deviation of 107 milli-pH units for solution 6. The listed pre-averages for solution 6 are also wild. The measuring equipment was operating properly. The problem appears to be associated with solution 6. Let us temporarily set the solution 6 results aside and examine the rest of table 12.

For glass electrode #1 there are rather large effects for the addition of either $NaNO_3$ or KCl ⟨Factors E and F⟩. The pre-averages appear to be larger for the H_2SO_4 solutions (1,2) than for the HCl solutions (3–5). For glass electrode #2, the pre-averages for $NaNO_3$ and/or KCl appear to be much smaller and are the same magnitude for both the H_2SO_4 and the HCl solutions.

The main effects and two-factor interactions were next calculated from the pre-averages by the short-cut procedures (see table 13). Average values for the standard deviation of a single measurement are also shown.

The general calculations described by eq (10), and the table-displayed s values, provide a method by which we can judge the main effects and the two-factor interactions. The standard deviations for either the ME or 2FI are 3.0, 3.2, 3.0, 6.5, and 7.4 milli-pH units for solutions 5, 3, 2, 1, and 4, respectively. From this we can conclude that an effect or interaction of more than 20 milli-pH units is not a chance happening.

Looking at table 13, we see that main effects A (lowering temperature), E (adding $NaNO_3$), and F (adding KCl) are large. The $NaNO_3$ main effect is about 30 for HCl solutions 5, 3, and 4, and is about 55 for H_2SO_4 solutions 2 and 1. The KCl ME is about 85 for HCl solutions 5 and 3, and 115 for the H_2SO_4 solutions. We note that the glass electrode #2 gives a KCl ME of only 18 for HCl solution 4. This result is not a fluke, but is the combined result from four separate experiments.

Two-factor interactions of appreciable size are observed under the column headings C, D, and G. Table 5 lists the possible 2FI interactions. It may be reasonable to assume that the 2FI associated with column C is due to the AE and/or BF interactions (temperature with $NaNO_3$ and/or stirring with KCl). It is known from experience that stirring can influence pH measurement

Table 11. Properties of acid solutions tested.

Label	pH	Acid	Ionic Contaminants	Electrodes used (#1 or #2) and 1984 Measurement Dates	
Soln. 5*	3.0	HCl	Hi purity water	#1 July 16–20	#2 Aug. 10
Soln. 3	3.7	"	" " "	#1 July 24–30	
Soln. 4	4.4	"	" " "	—	#2 July 31–Aug. 2
Soln. 6	5.0	"	" " "	#1 July 19–20	
Soln. 2	3.6	H_2SO_4	20 ppm by wt in water	#1 July 12–13	#2 Aug. 21
Soln. 1	4.3	H_2SO_4	3 ppm by wt in water	#1 July 9–12	#2 Aug. 20

* This solution was used for the above examples.

Table 12. Summary of pre-averages for main designs and reverse-sign designs.

Electrode	Soln.	Sets	pH	Milli-pH Units *s*								Legend for + Factors
				s	*A*	*B*	*C*	*D*	*E*	*F*	*G*	
Main Designs												
1	5	1+3	3.0	7	45	−4	9	21	26	81	1	*A* = 25 °C
1	3	1+3	3.7	8	35	−14	0	30	44	94	−15	*B* = solution stirred
1	6	1+3	5.0	107	−34	−44	−28	30	194	158	−98	*C* = 0.5 mL dilution
1	2	1+3	3.6	7	35	−15	1	44	60	120	−10	*D* = 1 cm electrode immersion
1	1	1+3	4.3	15	10	−21	−6	39	70	116	−18	*E* = $NaN0_3$ added
												F = KCl added
2	4	1+3	4.4	14	36	−32	−4	−28	28	11	−32	*G* = pH measured at 10 minutes
Reverse-Sign Designs												
1	5	2	3.0	–	62	−3	2	−16	26	80	0	
1	3	2+4	3.7	10	40	−16	26	−22	28	86	6	
1	6	2	5.0	–	−35	−123	192	−32	70	45	134	
1	2	2	3.6	–	42	−9	23	−44	54	130	24	
1	1	2	4.3	–	17	−21	41	−31	45	90	28	
2	5	4	3.0	–	55	−6	11	−5	4	37	0	
2	4	2+4	4.4	27	29	−14	30	−2	19	26	8	
2	2	4	3.6	–	37	−7	12	−7	11	33	2	
2	1	4	4.3	–	27	−9	20	1	23	30	0	

Table 13. Summary of results—main effects and two-factor interactions.

Electrode	Soln.	pH	Acid	Milli-pH Units								Legend for + Factors
				s	*A*	*B*	*C*	*D*	*E*	*F*	*G*	
Main Effects												
1	5	3.0	HCl	7	53	−4	5	2	26	81	1	*A* = 25 °C
1	3	3.7	"	9	38	−15	13	4	36	90	−4	*B* = solution stirred
1	2	3.6	H_2S0_4	7	38	−12	12	0	57	125	7	*C* = 0.5 mL dilution
1	1	4.3	"	15	14	−21	18	4	58	103	5	*D* = 1 cm electrode immersion
												E = $NaN0_3$ added
2	4	4.4	HCl	21	32	−23	13	−15	24	18	−12	*F* = KCl added
												G = pH measured at 10 minutes
Two-Factor Interactions												
1	5	3.0	HCl	7	9	1	−3	−18	0	−1	−1	
1	3	3.7	"	9	2	−1	13	−26	−8	−4	10	
1	2	3.6	H_2S0_4	7	4	3	11	−44	−3	5	17	
1	1	4.3	"	15	4	0	24	−35	−12	−13	23	
2	4	4.4	HCl	21	−4	9	17	13	−4	8	20	

results. The temperature, $NaN0_3$, and KCl are large effects, and it seems reasonable to assume that these effects may also be involved in the larger 2FI. By similar reasoning, the 2FI associated with column *G* may be due to AF and/or BE interactions (temperature with KCl and/or stirring with $NaN0_3$). The *D* column 2FI may be due to AB and/or EF interactions (temperature with stirring and/or $NaN0_3$ with KCl). It is interesting to note that the column *D* 2FI are appreciably different between the glass electrodes (#1 and #2), and also that the KCl ME with the HCl solutions is different between the two electrodes.

We were surprised by the slightly different characteristics exhibited by the two glass electrodes. In general, each laboratory will have its own set of surprises. The point to be made is that in ruggedness testing one should widely explore for factors that can influence the measurement results. Conversely, it seems wise not to be too concerned with determining higher order interactions, or with making extremely precise determinations of the main effects and two-factor interactions.

Can a factor such as the two different glass electrodes be used as a factor in the PB-design? It is certainly not a quantitative factor, such as temperature, and we cannot set the electrodes at a high or a low level. Nevertheless, we can accept the two levels we happen to get, and can go through all of the calculations. The advantages of orthogonal effects are maintained. With 100% hindsight, we see that it would have been wise to include the two different glass electrodes as one of the factors in the PB-design. As it turned out, we stumbled onto this factor by making a sequential series of PB-experiments.

How should we react to the observation of an important nonquantitative factor such as the glass electrode? We cannot easily tighten the glass electrode specifications. Clearly, further exploratory work dealing with the basic science of the glass electrode measurement process is needed.

As a final note we present table 14 which suggests an upper limit for precise pH measurements of a fully dissociated acid such as HCl. The results shown in table 14, which were collected from previously listed tables, suggest an upper pH limit of about 4.5. Higher pH measurements, associated with solutions 6, are plagued by the solution's absorption of CO_2 from the atmosphere. Special measurement procedures are required for pH measurements above 4.5.

Table 14. Effect of pH on standard deviation.

pH	*s* (milli-pH units)
3.0	7
3.6	7
3.7	9
4.3	15
4.4	22
5.0	107

Conclusions

The principles of ruggedness testing involving the recognition of interactions have been discussed and illustrated by a study of pH measurements using glass electrodes. Separation of the main effects and the two-factor interactions was obtained. The statistical significance of individual main effects, and of groups of two-factor interactions was determined. The precision of the pH measurements was found to be a strong function of the pH level. The physical interpretation of these findings is further discussed in companion articles, [5] in particular, of the *NBS Journal of Research.*

Journal of Research of the National Bureau of Standards
Volume 91, Number 1, January–February 1986

Effect of Variables on pH Measurement in Acid-Rain-Like Solutions as Determined by Ruggedness Tests

George Marinenko, Robert C. Paule, William F. Koch, and Melissa Knoerdel

National Bureau of Standards, Gaithersburg, MD 20899

Accepted: November 4, 1985

Ruggedness Test (RT) experiments were performed to assess the significance of the various main factors which affect pH measurements in low ionic strength aqueous solutions, as well as to establish the presence of interactions between the main factors. Stirring has an adverse effect on the measurement of pH, since it not only increases the random noise but also biases the measured value.Temperature control to the nearest 0.5 °C is sufficient for maintaining measurements accurate to 0.01 pH. Addition of $NaNO_3$ or KCl can not be tolerated in accurate pH measurements. Three small two-factor interactions were also revealed.

Key words: acid rain; glass electrodes; low ionic-strength solutions; main effects; pH measurements; ruggedness tests; two-factor interactions.

Introduction

The evaluation of the performance of a practical pH measuring system is a critical component of quality assurance for such programs as wet deposition monitoring. The purpose for these ruggedness test (RT) experiments, [1,2][1] was to assess the significance of the various main factors which affect pH measurements in low ionic strength aqueous systems, as well as to detect the possible presence of interactions between the main factors.

About the Authors: George Marinenko is a chemist in the Inorganic Analytical Research Division at NBS, part of the Bureau's National Measurement Laboratory to which Robert C. Paule, a physical scientist, is assigned. William F. Koch is also a chemist in the Inorganic Analytical Research Division. Melissa Knoerdel, a student, serves the Division during summer vacations.

RT experiments were conducted using a Plackett-Burman (PB) design, with seven factors and eight different factor combinations per set of measurements. Each factor is used at one of two chosen levels, arbitrarily designated as high (+) and low (−) levels (see table 1). In this design with (N=8) measurements per set, each

Table 1. A Plackett-Burman design for N=8 runs.

	Factor						
Run	A	B	C	D	E	F	G
1	−	−	−	−	−	−	−
2	−	−	+	−	+	+	+
3	−	+	−	+	−	+	+
4	−	+	+	+	+	−	−
5	+	−	−	+	+	−	+
6	+	−	+	+	−	+	−
7	+	+	−	−	+	+	−
8	+	+	+	−	−	−	+

---

[1] Figures in brackets indicate literature references.

factor appears an equal number of times (4) in its high level and in its low level. The main effect of any factor, then, is calculated simply as the average of the measurements made at the high level minus the average of the measurements made at the low level of the factor. By conducting 2N measurements (16 in our case) one can also separate the main effects from groups of two-factor interactions.

The RT experiments for the pH experiment system included rather drastic variations in the experimental parameters. Such extreme variations are not usually encountered in the course of controlled pH measurements made within a single laboratory.

These variations are, however, at the upper limits of those encountered in practical pH measurements made in different laboratories. In the RT experiments a considerable efficiency is obtained by the process of changing all variables simultaneously. It is worthwhile to note, however, that RT experiments are seldom made under the conditions of final interest. Nevertheless, RT can be useful for quality assurance in wet deposition monitoring programs, by aiding in the selection of the tolerances of the test conditions. This results in more precise and accurate pH measurements.

Experimental

The measurement process first involved the calibration of a commercial combination glass/reference electrode and a digital pH meter. Calibration was accomplished by setting the slope adjustment of the pH meter to 100% and adjusting the pH meter calibration setting while measuring the pH=4.006 buffer (SRM 185e). Following the above calibration step, the pH=6.863 buffer (SRM 186d) was measured. A strong acid (HCl) reference solution denoted as SA83 1000 series (pH=3.685, prepared and determined in our laboratory in a hydrogen cell without liquid junction), was also measured. At this point the system was ready for the experimental solutions of the RT.

The experiments entailed the measurement of the pH of solutions at selected levels of seven factors. The factors initially chosen were: temperature, stirring during measurement, dilution (0.5 mL distilled H_2O per 20 mL of solution), depth of electrode immersion, pH meter slope setting[2] (90% and 100%), addition of KCl and equilibration time. After a preliminary set of experiments, two things become evident: 1) as expected, the addition of KCl had a pronounced effect on the measured pH, and 2) the magnitude of the effect of slope adjustment between 90 and 100% was so large in comparison to the effects of the other variables that it obliterated all other effects. For this reason it was decided to modify the experimental plan as follows. The slope adjustment on the pH meter was to remain fixed at 100% for all measurements. Hence, slope was no longer a variable factor. Instead of slope, a new factor was introduced: an addition of $NaNO_3$.

All subsequent experiments were conducted using the following seven factors: 1) temperature; 2) stirring; 3) dilution; 4) immersion depth; 5) addition of $NaNO_3$; 6) addition of KCl; and 7) equilibration time.

The various designated factors were adjusted to values dictated by the PB design, and the pH of the experimental solutions was measured at 1, 5 and 10 min after the immersion of the electrode into the solution. Table 2 shows an example of the measurement scheme. Only the

Table 2. One example of measurement scheme using solution No. 99-5.

Measurement No.	Temp. °C	Stirring	Dilution 0.5 mL H_2O per 20 mL soln	Immersion Depth	$NaNO_3$ 0.034 mol/L	KCl 0.068 mol/L	Equilibration Time		
							1 min	5 min	10 min
1	30	No	No	Bottom	No	No	2.905	2.904	2.912
2	30	No	Yes	Bottom	Yes	Yes	3.021	3.012	3.015
3	30	Yes	No	Top	No	Yes	2.981	3.006	3.006
4	30	Yes	Yes	Top	Yes	No	2.946	2.964	2.965
5	25	No	No	Top	Yes	No	2.993	2.996	2.999
6	25	No	Yes	Top	No	Yes	3.057	3.055	3.060
7	25	Yes	No	Bottom	Yes	Yes	3.045	3.049	3.050
8	25	Yes	Yes	Bottom	No	No	2.938	2.946	2.949

Reference Solution	Before the Experiment	After the Experiment
pH 4.01 Buffer	4.005	3.977
pH 6.86 Buffer	6.820	6.810
SA83 1000	3.610	3.634

---

[2] On a pH meter, the slope setting is equivalent to setting a gain factor, namely mV/pH. 100% slope corresponds to the theoretical Nernstian slope of 59.157 mV/pH at 25 °C, and 90% slope corresponds to 53.241 mV/pH.

underlined pH measurements in the last two columns at the right of the table were used in the RT. The other pH measurements were obtained as a matter of routine, but were only used for background information. (One of the conclusions from this background information is that an equilibration time of 1 min is inadequate for stable pH measurements.) Following the completion of a set of pH measurements on each experimental solution, the pH values of the strong acid reference solution SA83 1000, and the pH 4.01 and 6.86 buffers, were remeasured.

The measurements were conducted in 30 mL borosilicate glass beakers. For each experiment a beaker containing exactly 20 mL of the appropriate solution was placed into a constant temperature bath that was thermostated to the nearest 0.1 °C. The solution in the beaker was allowed to come to thermal equilibrium with the bath. At this point a calibrated combination glass/reference electrode was immersed into the beaker. For the purpose of the experiment, this immersion time was designated as $t=0$. The values displayed digitally by the pH meter were recorded at the appropriate time intervals. The readability of the pH meter was 0.001 pH.

The RT experiments were performed on six different solutions (see table 3) which can be conveniently subdivided into three groups. The first group of solutions (solutions 99-1 and 99-2) was prepared from coulometrically standardized H_2SO_4. In addition to sulfuric acid, solution 99-1 also contained approximately 2 ppm by weight of other anions and cations, while solution 99-2 contained approximately 10 ppm of other ions. The second group of solutions (99-3 and 99-4), consisting of two reference HCl solutions, was prepared in 1983 in our laboratory. These solutions were not intentionally spiked with any foreign ions. The third group (99-5 and 99-6) consisted of two freshly prepared pure HCl solutions, containing no added contaminants. The last column of table 3 lists our reference pH values for the test solutions. For the sulfuric acid solutions (99-1 and 99-2) coulometrically determined concentrations along with the mean activity coefficient from the Debye-Hückel equation were used to calculate the reference pH values. The mean activity coefficient value for these two solutions is 0.927. The pH values of solutions 99-3 and 99-4 were accurately measured in a hydrogen cell without liquid junction. The pH values of solutions 99-5 and 99-6 were calculated, using the coulometrically determined hydrogen ion concentrations and tabular values of activity coefficients [3].

Table 3. List of solutions used in ruggedness test.

Solution No.	Solution Description	pH [@ 25 °C]
99-1	Simulated acid rain (H_2SO_4)	4.293
99-2	Simulated acid rain (H_2SO_4)	3.586
99-3	SA83 1000 (HCl)	3.685
99-4	SA83 2000 (HCl)	4.303
99-5	HCl 1.0008×10^{-3} mol/kg	3.015
99-6	HCl 1.026×10^{-5} mol/kg	4.989

Results and Discussion

The results of the experiments were presented and statistically analyzed in our earlier publications [1,2]. For the sake of convenience, the summary data of the observed main effects and two-factor interactions, which were presented in [2], are again introduced as tables 4 and 5. Some additional reduction of data has been performed in presenting these tables. For reasons

Table 4. Main effects of various factors on pH of strong acid solutions as revealed by the ruggedness test.

Solution	Electrode No.	*A* Temp	*B* Stirring	*C* Dilution	*D* Immersion Depth	*E* $NaNO_3$	*F* KCl	*G* Equil. Time
HCl	1	0.045	−0.010	0.010	0.005	0.030	0.085	0.000
H_2SO_4	1	0.025	−0.015	0.015	0.000	0.060	0.115	0.005
HCl	2	0.032	−0.025	0.015	−0.015	0.025	0.020	−0.010

Legend to Factor Levels

Factor	+	−
A	25 °C	30 °C
B	Stirred	Not stirred
C	0.5 mL dilution/20 mL	No dilution
D	1 cm to top	3 cm to top
E	$NaNO_3$	None added
F	KCl	None added
G	10 min	5 min

Table 5. Two-factor interactions as revealed by the ruggedness test.

Solution	Electrode	AE/BF/DG	AB/EF/CG	AF/BE/CD
HCl	1	0.005	−0.020	0.005
H_2SO_4	1	0.020	−0.040	0.020
HCl	2	0.015	0.015	0.020

The letter designation of factors in table 5 is the same as in table 4. The following is the legend to the two factor interactions:

AE:	Temp. with $NaNO_3$	*AB*:	Temp. with Stirring	*AF*:	Temp. with KCl
BF:	Stirring with KCl	*EF*:	$NaNO_3$ with KCl	*BE*:	Stirring with $NaNO_3$
DG:	Immersion depth with time	*CG*:	Dilution with time	*CD*:	Dilution with immersion depth

which shall be discussed later, the results for the low ionic strength solution 99-6 are excluded from these tables. Two different combination glass/reference electrodes were used in the RT experiments. The electrode numbers are given in column 2 of tables 4 and 5. The choice of two different electrode pairs was not by design. The first electrode was broken in use.

Temperature is the first main effect listed in table 4. For electrodes 1 and 2, and for the HCl and H_2SO_4 solutions, the temperature effect in going from 30 °C to 25 °C is moderately large, about +0.035 pH. The calculated effect of the change in temperature from 30 °C to 25 °C, based strictly on the change of the Nernstian response factor, should result in a pH increase of +0.017. The observed change is of the correct sign, but the magnitude per five degrees is somewhat larger than predicted. Two possible sources of explanation for the apparently excessive temperature effect may be: 1) temperature gradients across the combination glass/reference electrode; and 2) temperature response of the residual liquid junction potential. Both of these effects are difficult to calculate theoretically. Thus for practical pH measurements, one must empirically determine the overall temperature effect. The RT experiments, in fact, performed this task.

The values for the second main effect, stirring, are about −0.015 pH for solutions 99-1 to 99-5, for the electrodes used in this RT (table 4). Solution 99-6 had such a low strength that stirring caused the results to be very erratic, and hence they were not included in table 4. In previous work [4–6], we observed for a number of different electrodes that stirring had an even greater effect than −0.015 pH. The current RT did not purposely include different electrodes as one of the variables to be studied. We believe that the small stirring effects observed in this RT are not representative of most pH electrode systems. The oscillations observed in the measured pH are caused by periodic disruption of the diffusion gradient at the liquid junction between the filling solution of the Ag, AgCl reference electrode (3.5 mol/L KCl solution) and the test solution. In the less vigorous period of the stirring cycle, the 3.5 mol/L KCl filling solution, diffuses through the ceramic plug and establishes a diffuse concentration gradient in the vicinity of the junction. In the vigorous, convective period of the stirring cycle, when the magnetic stirring bar just passes the ceramic junction region of the reference electrode, the shearing action of the stirrer creates a sharp concentration boundary at the ceramic plug-solution interface. We know that for certain geometries the oscillations of the pH readings are in phase with the rotation of the magnetic stirrer. This phenomenon is much less pronounced in the case of more concentrated acid solutions. The total ionic strength difference at the junction is not as great for the more concentrated solutions, and consequently the fluctuations are attenuated. This phenomenon is also highly dependent on the type and condition of the reference electrode junction.[3]

From both the RT and previous work, we conclude that accurate pH measurements should be performed on quiescent solutions. Of course, stirring prior to the measurement is required for homogenization of the test solution.

The third observed main effect of +0.015 pH is caused by dilution. On the basis of the 2.5% concentration change due to dilution, the calculated pH showed increase by 0.011 pH. Dilution is one of the few pH measurement effects which can be clearly calculated from theory. Excellent agreement with the RT measurements is obtained.

As can be seen, in table 4, the fourth factor, the electrode immersion depth produces little effect. No significant difference in pH values as a function of the electrode immersion depth of 1 and 3 cm could be detected.

---

[3] In the course of publishing this study, there appeared a new article [7] reporting the behavior of nine different glass electrodes with different reference electrodes and different junctions. This new study also showed large and variable liquid junction effects. The general conclusions reached are complementary to our study.

The fifth and sixth main effects are due to addition of neutral salts (0.033 mol/L $NaNO_3$ and 0.067 mol/L KCl). These effects are quite pronounced. For electrode 1, we observed that the addition of KCl has a greater effect than the addition of $NaNO_3$. It should be noted that the molar concentration of the added KCl is exactly twice that of $NaNO_3$. As one would expect from theory, and as observed for electrode 1, the additions of salts have a greater effect on the H_2SO_4 than on the HCl solutions. The presence of doubly charged sulfate ion is to a large extent responsible for the different behavior of the two acids, since activity coefficients for chloride and sulfate at a given ionic strength are quite different. For electrode 2, the $NaNO_3$ salt effect is qualitatively similar to that observed for electrode 1. The KCl salt effect, however, is much smaller. All of the observed salt effects are the result of many experiments and were reproducible over the extended (two month) period of the RT experiments [2]. Nevertheless, the observed salt effects are not in close agreement with the effects calculated from the Debye-Hückel equation. It is obvious that the glass/reference electrodes are not behaving ideally.

In attempting to rationalize the differences in the observed salt effects with electrodes 1 and 2, a limited number of auxiliary experiments were conducted a year later with electrode 2. These experiments indicate that the apparent salt effects had changed appreciably. We conclude that the use of added salts to low ionic strength solutions for pH measurements is an unacceptable practice.

The seventh factor, G, namely equilibration time was measured at three levels, 1, 5 and 10 min after the immersion of the electrode into the test solution. The analyses of two sets of data (1 and 5 min; 5 and 10 min) show that 1 min is not sufficient time for equilibrium condition to be established. Therefore, the data which are considered and discussed here for the ruggedness tests refer to 5 minute and 10 minute equilibration times as the two levels of factor G. Little difference is observed between the 5- and 10-min equilibration times.

In summary, we conclude that the following three main effects are of practical importance: 1) temperature, 2) stirring, and 3) the addition of $NaNO_3$ and KCl.

Besides the observed large main effects, there are three sets of moderate size two-factor interactions (2FI), summarized in table 5. Our intent is to examine such 2FI and ascribe, if possible, some physical interpretation to these interactions. These 2FI are of the same magnitude, on the order of 0.020 pH.

It can be seen in table 5, that the identified interactions are not unique, but rather they consist of three groups of three possible 2FI: *AE/BF/DG; AB/EF/CG; AF/BE/CD*. The first group consists of the following possible interactions: temperature with $NaNO_3$, stirring with KCl, and immersion depth with time. The first and the third 2FI do not appear to be reasonable candidates. There appears to be no logical reason for an interaction between temperature and $NaNO_3$. For the third 2FI, both the immersion depth and equilibration time are negligible main effects, and therefore the associated 2FI is not considered likely. We attribute the first group's 2FI to stirring with KCl.

The second group consists of temperature with stirring, $NaNO_3$ with KCl, and dilution with time. Temperature with stirring does not appear to be a physically viable interaction candidate, while the interaction due to the two added salts is quite plausible, and shall be discussed shortly. Dilution and equilibration time are negligible main effects, so that we choose to rule out this third 2FI candidate.

The third group 2FI candidates are: temperature with KCl, stirring with $NaNO_3$, and dilution with immersion depth. Here, as with the first group, temperature with salt does not appear to be physically reasonable, and the diluiton and immersion depth are negligible main effects. Thus, we choose stirring with $NaNO_3$ as the likely 2FI. Since these two factors have appreciable, opposing main effects, which by no means bear any simple relationship, it is not surprisingly that a two-factor interaction is observed.

The two-factor interactions observed between the two salts may be partially explained by the ionic strength effect on the mean activity coefficient. The mean activity coefficient is a nonlinear function of the molality, with decreasing slope at higher molalities [3]. Thus, the overall effect of the addition of the two salts is less than additive. Taking out a single factor (main effect) for each of the salts, overcompensates when two salts are added. Thus the opposite sign effect is anticipated. Indeed for electrode 1 (table 5) the $NaNO_3$ and KCl two-factor interaction (2FI) is negative. We do not at present know why the corresponding 2FI for electrode 2 is positive.

For electrode 2 we do note that the main effect for KCl is small and out of line with the other KCl and $NaNO_3$ main effects (see table 4). Since the reference electrode is filled with 3.5 mol/L KCl solution, this suggests that perhaps the difference in the structure of the two liquid junctions of the two reference electrodes is responsible for the observed differences. Since the two electrodes exhibit different main effects, it is not surprising that some of the 2FI for the two electrodes are also different.

From the RT we infer that one must be very cautious in choosing the calibration standards for dilute strong acid measurements. These standards should closely simulate the system of interest because there are a number of unexplained, observed effects which can only be excluded from the measured value through calibration with a similar standard. This aspect of pH measurement is treated in detail in [8].

Having established the magnitude of the influence of the various factors on measured pH, we combined the RT data with calibration data obtained by using the strong acid reference solution SA83 1000. These data are summarized in the table 6. The normalized pH results are in reasonable agreement. The deviations given in the last column of table 6 are generally small. The largest deviation of +0.063 pH is approximately equal to two standard deviations for the between the laboratory component of variability observed in our interlaboratory study [8].

Table 6. Measured normalized pH of five acid solutions.

Solution	pH Normalized to SA83 1000	pH (calc.)	ΔpH.
99-1	4.294	4.293	+0.001
99-2	3.557	3.586	−0.029
99-3	3.698	3.685	+0.013
99-4	4.366	4.303	+0.063
99-5	3.028	3.015	+0.013

Conclusions

Stirring has an adverse effect on the measurement of pH of dilute acid solutions. Stirring not only amplifiesthe random noise but also biases the measured values. Moderate temperature control (±0.5 °C) is sufficient for maintaining measurements accurate to 0.01 pH. Addition of neutral salts ($NANO_3$ or KCl) can not be tolerated in accurate pH measurements, as these salts change not only the mean activity coefficients of solutions, but also unpredictably change the electrode behavior and hence the measured pH. The ruggedness tests have revealed, in addition to the above main effects, three 2FI. These 2FI could not be determined from an experiment which changed only one factor at a time.

References

[1] Paule, R. C.; G. Marinenko, M. Knoerdel, and W. F. Koch, Ruggedness Testing—Part I: Ignoring Interactions, J. Res. Natl. Bur. Stand. **91-1**, 3–8 (1986).

[2] Paule, R. C.; G. Marinenko, M. Knoerdel, and W. F. Koch, Ruggedness Testing—Part II: Recognizing Interactions, J. Res. Natl. Bur. Stand. **91-1**, 9–16 (1986).

[3] Hamer, W. J., and Y. C. Wu, Osmotic coefficients and mean activity coefficients of uni-univalent electrolytes in water at 25 °C, J. Phys. Chem. Ref. Data, **1**(4), 1047–1099 (1972).

[4] Koch, W. F.; G. Marinenko and J. W. Stolz, Simulated Precipitation Reference Materials IV, NBSIR 82-2581 (1982).

[5] Koch, W. F., and G. Marinenko, Simulated Precipitation Reference Materials: Measurement of pH and Acidity, *Sampling and Analysis of Rain*, S. A. Campbell, ed., ASTM STP 823, 10–17 (1983).

[6] Covington, A. K., and P. D. Whalley, Improvements in the precision of pH measurements, Anal. Chim. Acta **169**, 221–229 (1985).

[7] Davison, W., and C. Woof, Performance Tests for the Measurement of pH with Glass Electrodes in Low Ionic Strength Solutions Including Natural Waters, Anal. Chem. **57**, No. 13, 2567–2570 (Nov. 1985).

[8] Koch, W. F.; G. Marinenko and R. C. Paule, An Interlaboratory Test of pH Measurements in Rainwater, J. Res. Natl. Bur. Stand. **91-1**, 23–32 (1986).

Journal of Research of the National Bureau of Standards
Volume 91, Number 1, January-February 1986

An Interlaboratory Test of pH Measurements in Rainwater

William F. Koch, George Marinenko, and Robert C. Paule
National Bureau of Standards, Gaithersburg, MD 20899

Accepted: October 23, 1985

An interlaboratory test of pH measurements in rainwater has been conducted. Various types of electrodes and junction materials were used in this test. The results of this exercise verify that there are significant differences in the pH values of low ionic strength solutions reported by various laboratories. Other work suggests that these differences are due to residual liquid junction potentials. Furthermore, this test confirms the efficacy of using dilute solutions of a strong acid as working standards for pH measurements in acid deposition studies.

Key words: acid rain; electrodes; interlaboratory; liquid junction potential; measurements; pH; rainwater; round-robin; standards.

Introduction

The accurate measurement of pH in rainwater samples is an essential and critically important component in the monitoring of wet deposition for trends in the acidity of rainfall. Our laboratory has shown that serious biases may affect the accuracy, intercomparability, and reliability of these measurements [1][1]. To verify this situation and to seek a remedy, an interlaboratory test was conducted.

In November 1983, sets of samples were sent to 12 laboratories which voluntarily agreed to participate in this interlaboratory test of pH measurements in rainwater. (Appendix A lists the participants.) In addition, five sets of samples were tested independently by three scientists at the National Bureau of Standards (NBS). The purpose of this exercise was to gain information as to the utility of using dilute solutions of a strong acid as working standards for pH measurements in acid deposition studies.

Background

It has been proven that residual liquid junction potentials at the reference electrode can seriously bias pH readings, especially if the ionic strengths of the buffer standards and the test solutions are vastly different [1,2]. In an effort to correct for these biases, the Electroanalytical Chemistry Group (within the Inorganic Analytical Research Division in the NBS Center for Analytical Chemistry), has been conducting research to test the feasibility of providing matrix specific pH standards

About the Authors: William F. Koch and George Marinenko are chemists in NBS' Inorganic Analytical Research Division, while Robert C. Paule is a physical scientist with the Bureau's National Measurement Laboratory. The work they describe was funded in part by the U.S. Environmental Protection Agency (National Acid Precipitation Assessment Program).

---

[1] Figures in brackets indicate literature references.

for selected applications. It is imperative that such working standards be consistent with the National pH scale, as maintained through the NBS Standard Reference Materials pH buffers, in order that measurements of pH be as thermodynamically meaningful as possible and to avoid confusion caused by multiple pH scales. This interlaboratory test is a critical step in the evaluation and implementation of this approach.

Experimental

Each set of samples consisted of eight solutions in 60 mL polyethylene bottles labeled A through H. Participants were asked to measure the pH of each sample according to explicit instructions (Appendix B) and to record the values on a data sheet (Appendix C). Additional information was also requested (Appendix D). The composition of the eight solutions is shown in table 1. The participants were not informed of the identity of these solutions.

Table 1. Composition of test solutions.

Solution	Composition
A	0.05 molal potassium acid phthalate (SRM 185f, pH(S) 4.006 at 25 °C)
B	Hydrochloric acid, pH 3.68 at 25 °C
C	Hydrochloric acid, approx. pH 4.3
D	Simulated acid precipitation, approx. pH 4.1
E	Simulated acid precipitation, approx. pH 3.7
F	Identical to B
G	Identical to C
H	Identical to A

The rationale behind the sample selection and the sequence of measurements was as follows: Solution A was intended as a check on the calibration of each participant's pH measurement system, since the composition of this solution was nearly identical to the one provided for calibration.

Solution B was prepared by dilution of high-purity hydrochloric acid. The pH of the solution was determined using hydrogen gas electrodes in cells without liquid junction. The apparatus and calculations were identical to those used in the certification of NBS buffers [3,4]. The pH value was confirmed through calculations based on independent measurements using high-precision coulometry and ion chromatography. Solution B served as the reference or normalizing solution for this test. Solution C also was prepared by dilution of hydrochloric acid and was intended as a low ionic strength acidic solution in a simple matrix. Solutions D and E were more complex acidic matrices composed of several anions and cations simulating the composition of rainwater. Solution F was a repeat of Solution B to check for instrument drift, and possible hysteresis of the liquid junction. Solution G was a repeat of C to check instrument drift. Solution H was a repeat of A to check calibration drift and hysteresis. Participants were asked to repeat the sequence in order to establish the precision of the measurements. Participants were also asked to supply information as to types of electrodes and standards used.

Results and Discussion

Because the samples were to be sent to participants via the U.S. Postal System during the winter months, it seemed advisable to test whether freezing and thawing the solutions would affect the pH values. A set of samples was frozen in a laboratory freezer for 24 hours and then thawed. No significant differences in pH values were observed with this set versus a control set which had not been frozen. Even after several freeze-thaw cycles, there were no significant differences noted.

Full cooperation was obtained from all participants with regard to quality, completeness, and timeliness of response. All of the measured pH values are shown in table 2, including the buffer standards used by the participants (STD 1 and STD 2). The average pH values for measurement trials 1 and 2 for each solution by each participant are shown in table 3. The laboratory number does not correspond to the alphabetical listing of laboratories in Appendix A. The data and general results will first be presented graphically since this is easier to assimilate, and will then be described in a more quantitative fashion through the use of statistical analyses of variance. As a general rule in routine pH measurements using combination electrodes, an uncertainty of ±0.02 pH units is to be expected. This permissable variability will be applied in the following discussion of the data.

With the exception of laboratory 3, all participants reproduced the value for solution A, the standard buffer solution of potassium acid phthalate. Subsequent to the test, laboratory 3 discovered that the commercial buffer that they were using was biased. They have since corrected this problem. This indicates that good calibration practices were in effect and that the instruments were in a state of control. It also suggests that if accurate results are required, then standardization should be done using quality reference buffers, such as Standard Reference Materials supplied by the National Bureau of Standards. The values for solution H (which is identical to A) scatter a little more, with three labs (7, 8, and 17) being out of compliance. However, this is most likely due to hysteresis at the liquid junction. When comparing the values for A and H, and C with G, no trends in instrument drift are apparent. Furthermore, hysteresis of the electrodes, when subjected to buffers and low ionic strength solutions, should not be a major problem if, as for this

Table 2. pH readings for both measurement trials 1 and 2.

Lab. No.	STD 1	STD 2	A	B	C	D	E	F	G	H	STD 1	STD 2
1	3.997	6.860	4.016	3.716	4.401	4.265	3.787	3.727	4.415	3.998	4.014	6.860
	4.014	6.863	4.014	3.719	4.402	4.270	3.786	3.727	4.414	3.996	4.016	6.862
2	4.000	7.020	4.000	3.710	4.380	4.160	3.770	3.720	4.370	4.010	4.000	7.020
	4.000	7.010	4.010	3.700	4.370	4.150	3.760	3.710	4.370	4.010	3.990	7.010
3	4.010	6.870	3.960	3.530	4.190	4.010	3.660	3.590	4.210	3.990	4.030	6.880
	4.010	6.850	3.960	3.540	4.210	4.030	3.680	3.620	4.230	4.000	4.020	6.860
4	4.005	6.851	4.006	3.720	4.393	4.203	3.772	3.716	4.385	4.002	4.005	6.849
	4.008	6.859	4.007	3.728	4.392	4.208	3.780	3.730	4.384	4.009	4.010	6.859
5	4.000	6.830	4.000	3.720	4.370	4.140	3.760	3.700	4.360	3.990	3.990	6.820
	4.000	6.840	4.000	3.720	4.370	4.160	3.770	3.710	4.360	4.000	4.000	6.840
6	4.010	7.000	4.006	3.734	4.417	4.187	3.787	3.727	4.404	4.020	4.010	7.003
	4.010	6.973	4.005	3.711	4.365	4.160	3.770	3.710	4.370	4.006	4.022	6.973
7	4.004	6.871	4.024	3.583	4.218	4.047	3.860	3.602	4.088	4.082	4.063	6.914
	4.004	6.863	3.995	3.545	4.171	4.049	3.848	3.630	4.159	4.007	3.991	6.856
8	4.000	7.010	4.010	3.460	4.130	3.970	3.610	3.570	4.210	4.060	4.030	7.030
	4.000	7.000	4.000	3.430	4.070	3.910	3.560	3.520	4.150	4.040	4.030	7.020
9	4.000	7.410	4.020	3.720	4.340	4.140	3.780	3.730	4.320	4.020	4.000	7.390
	4.000	7.410	4.010	3.720	4.320	4.140	3.770	3.720	4.320	4.010	4.010	7.400
10	4.000	7.000	4.000	3.710	4.390	4.170	3.770	3.700	4.380	4.000	4.000	7.000
	4.000	7.000	3.990	3.710	4.390	4.160	3.760	3.700	4.380	4.000	3.990	6.990
11	4.010	6.990	4.010	3.740	4.400	4.170	3.800	3.740	4.400	4.020	4.010	6.970
	4.000	6.990	4.020	3.760	4.450	4.200	3.810	3.760	4.410	4.030	4.030	7.000
12	4.000	6.863	4.000	3.668	4.324	4.121	3.732	3.682	4.344	4.010	4.003	6.865
	3.997	6.862	4.001	3.667	4.331	4.124	3.734	3.676	4.338	4.008	4.003	6.859
13	4.004	6.885	4.006	3.705	4.394	4.165	3.751	3.692	4.390	3.990	3.992	6.875
	4.002	6.866	4.003	3.709	4.399	4.171	3.762	3.704	4.397	4.000	4.001	6.865
14	4.003	6.878	4.005	3.735	4.393	4.181	3.790	3.735	4.400	4.005	4.007	6.866
	4.003	6.873	4.006	3.720	4.392	4.178	3.783	3.720	4.394	4.007	4.005	6.856
15	4.002	6.830	4.002	3.743	4.371	4.161	3.775	3.748	4.383	4.002	4.000	6.801
	4.003	6.879	3.997	3.704	4.368	4.158	3.761	3.704	4.383	3.996	4.004	6.880
16	4.010	6.840	4.000	3.610	4.300	4.090	3.700	3.640	4.290	4.010	4.000	6.840
	4.000	6.840	4.030	3.630	4.290	4.070	3.700	3.640	4.280	4.000	4.010	6.850
17	4.000	6.980	3.990	3.720	4.410	4.190	3.820	3.740	4.440	3.980	4.000	7.000
	4.000	6.990	3.990	3.730	4.450	4.190	3.780	3.720	4.430	3.980	3.990	6.990

Table 3. pH readings—averages of trials 1 and 2.

Lab. No. \ Solution	A	B	C	D	E	F	G	H
1	4.02	3.72	4.40	4.27	3.79	3.73	4.42	4.00
2	4.00	3.70	4.38	4.16	3.77	3.72	4.37	4.01
3	3.96	3.54	4.20	4.02	3.67	3.60	4.22	4.00
4	4.01	3.72	4.39	4.21	3.78	3.72	4.38	4.00
5	4.00	3.72	4.37	4.15	3.77	3.70	4.36	4.00
6	4.01	3.72	4.39	4.17	3.78	3.72	4.39	4.01
7	4.01	3.56	4.19	4.05	3.85	3.62	4.12	4.04
8	4.00	3.44	4.10	3.94	3.58	3.54	4.18	4.05
9	4.02	3.72	4.33	4.14	3.78	3.72	4.32	4.02
10	4.00	3.71	4.39	4.16	3.76	3.70	4.38	4.00
11	4.02	3.75	4.42	4.18	3.80	3.75	4.40	4.02
12	4.00	3.67	4.33	4.12	3.73	3.68	4.34	4.01
13	4.00	3.71	4.40	4.17	3.76	3.70	4.39	4.00
14	4.01	3.73	4.39	4.18	3.79	3.73	4.40	4.01
15	4.00	3.72	4.37	4.16	3.77	3.73	4.38	4.00
16	4.02	3.62	4.30	4.08	3.70	3.64	4.28	4.00
17	3.99	3.72	4.43	4.19	3.80	3.73	4.44	3.98

exercise, a strict measurement protocol is established and followed. In fact, with the exceptions of laboratories 3, 7, and 8, excellent repeatability was observed for the duplicate solutions (A/H, B/F, and D/G). It can be concluded that, with few exceptions, the precision of pH measurements made by a single operator, using a single set of electrodes, is within the permissable variability of ±0.02 pH, and is not a matter of great concern. The data reported in table 2 for the duplicate measurements also shows good within-laboratory precision. However, this conclusion can be misleading because of the constraints of a single laboratory, a single operator, and a single set of electrodes, and because of the inattention to the matter of accuracy. Figure 1 shows the wide range of pH values obtained by the participants for solution B and their relationship to the true pH of this solution. The lower case letters, s through z, above each data point serve to categorize the electrodes used by the participants according to type and manufacturer. The range of values spans more than 0.3 pH units. Only two laboratories are within the permissable variability of 0.02 pH. Most laboratories are biased several hundredths of a unit high. A few laboratories are biased low and once again laboratories 3, 7, and 8 stand out in this regard. It is clear, that in spite of excellent within-laboratory precision, between-laboratory precision is very poor, and the measurements show a great deal of systematic bias. The high degree of scatter is confirmed by the values reported for solutions C, D, and E, as shown in figure 2. The solid line is the mean value of the 17 measurements and should be viewed as a point of reference only, not as the true value. Unlike B, the true pH values for these solutions have not been determined.

Exhibited in figure 3 are the results of normalizing the pH values of C, D, and E, with respect to B. In essence, solution B is used as the calibration standard rather than the customary buffer solution. The vast improvement in terms of decreased scatter is obvious. Most values are now within ±0.02 pH units of the mean. (The extremely high value for E by laboratory 7 must be considered an outlier and is discussed in more detail below.)

As noted, the participants included in their reports the make, model, and type of electrodes used in their measurements. Eight distinct types of electrodes were used in the study. These have been identified in code on figures 1–3 with the lower case letters s through z. An association of the type of electrode with deviations from either the true value or the mean is exhibited by only two of the eight types, namely, "s" and "w." Type "s" is a combination electrode in which the liquid junction of the reference electrode is somewhat larger in area than the other electrodes used in the test. This apparently results in a larger variability in residual liquid junction potential. However, when normalized with solution B, the values obtained by electrode "s" conform quite well. Type "w", used by laboratory 7, is a combination electrode which incorporates a gel-filled reference electrode. During this test, this electrode behaved quite erratically in low ionic strength solutions, as evidenced by its non-conformity even after the normalization process.

Electrode types "y" and "z" had open-junction reference electrodes; that is, the junction between the filling solution of the reference electrode and the sample solution was formed at a capillary tip, rather than across a ceramic or fiber frit as is customary in combination electrodes. In the low ionic strength solutions tested in this exercise there was no apparent advantage to this type of junction, although it has proven its worth in other types of solutions. Correlations with the other requested information (Appendix D) were not readily apparent.

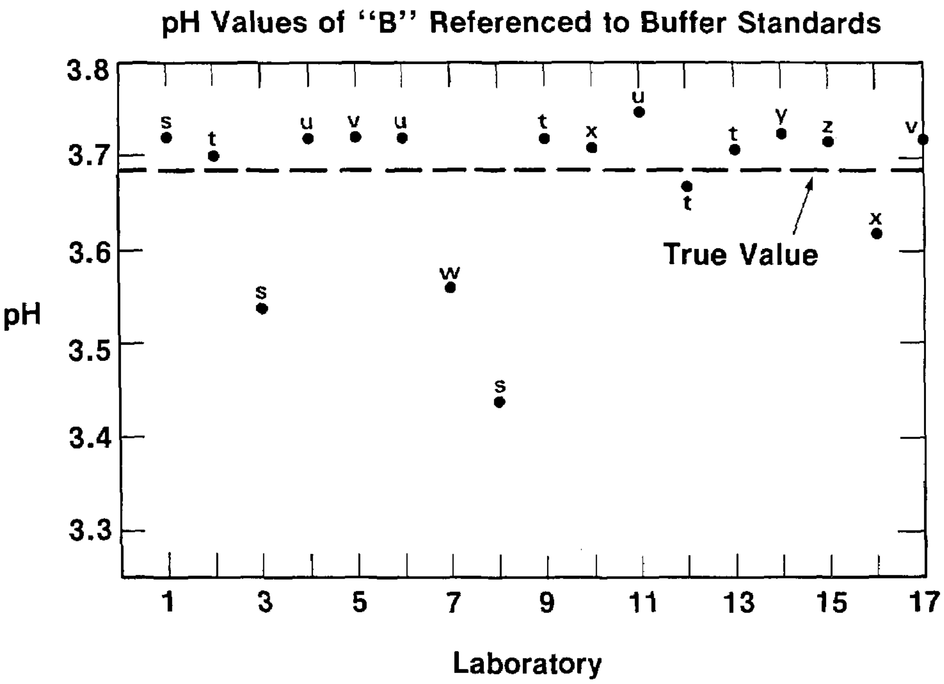

Figure 1-pH values of "B" referenced to buffer standards.

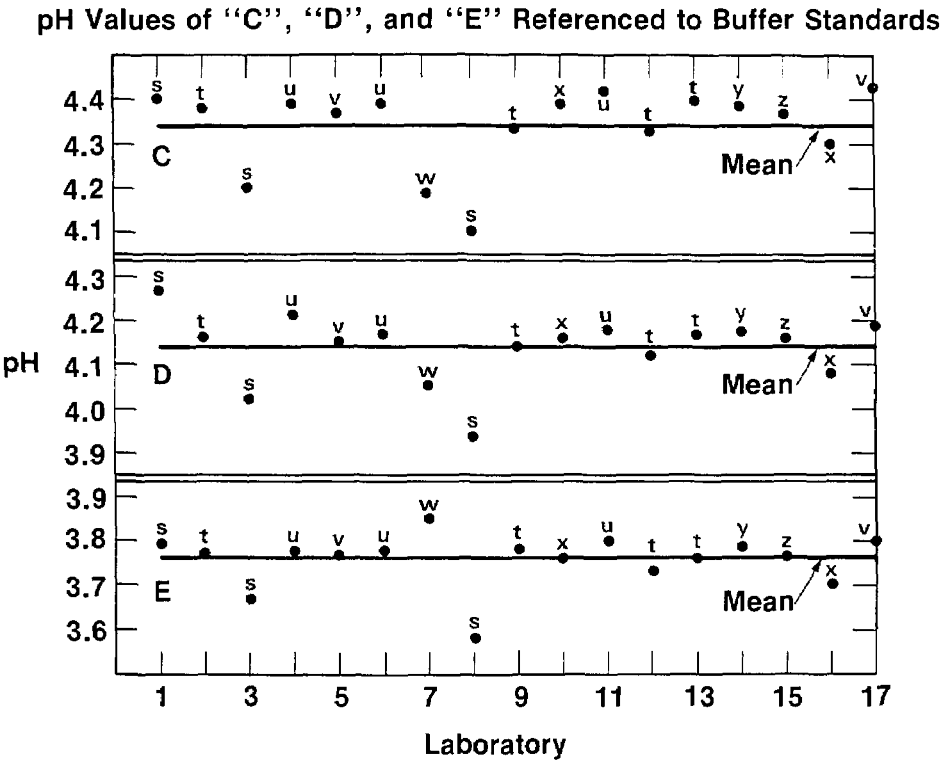

Figure 2-pH values of "C," "D," and "E" referenced to buffer standards.

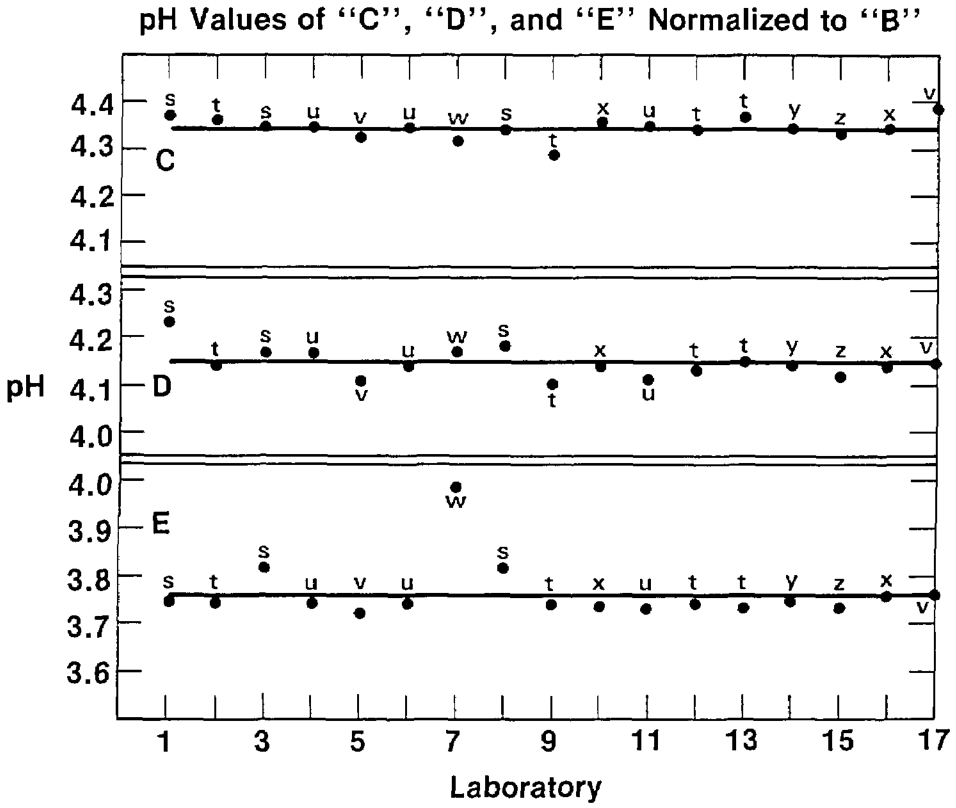

Figure 3-pH values of "C," "D," and "E" normalized to "B."

Statistical Analysis

The original pH data from the laboratories, as well as the result obtained by mathematically recalibrating the original data with the individual laboratory's solution B measurements, have been quantitatively evaluated. Table 2 gives the original data and table 4 the recalibrated data. One-way analyses of variance were run on the original and the recalibrated data for non-buffer solutions C through G. The analyses of variance calculations gives within- and between-laboratory components of variance. The square roots of these quantities are called the within- and between-laboratory components of standard deviation, and are reported in table 5.

Examination of table 5 results shows the within-laboratory components of standard deviation to be about 0.015 pH units, and to be approximately equal for the original and the recalibrated data sets. The between-laboratory components of standard deviation, however, are markedly different between the original and the recalibrated data sets. As expected, the recalibrations with solution B have resulted in greatly improved between-laboratory precisions. The component of between-laboratory standard deviation is about 0.07 pH units for the original data, and is about 0.03 for the recalibrated data. The reader may evaluate the stability of the precision estimates by comparing the results from the identical solutions C and G.

Conclusion

The results of this interlaboratory test show that the use of dilute strong acid solutions of known and independently verifiable pH (e.g., solution B) to standardize pH instrumentation greatly improves between-laboratory precision in the measurement of pH in low ionic strength solutions, such as acidic rainwater. Thus, the use of such solutions by all those involved in network monitoring of wet deposition is recommended to improve the intercomparability of the measurements as a function of time and location. To aid in this endeavor, Standard Reference Material, SRM 2694, "Simulated Rainwater" has been prepared and analyzed by our laboratory and is available through the Office of Standard Reference Materials of NBS. SRM 2694 consists of a set of two 50 mL solutions in polyethylene bottles. The nominal pH of Level I is 4.3 and that of Level II is 3.6. The acidity and the specific conductance have also been determined, as well as several of the major cations and anions commonly found in rainwater. Guidelines for the measurement of pH in rainwater with the use of these materials are included with the Certificate of Analysis.

Table 4. pH values after normalization.

Lab. No.	STD 1	STD 2	A	B	C	D	E	F	G	H	STD 1	STD 2
1	3.961	6.824	3.980	3.680	4.365	4.229	3.751	3.691	4.379	3.962	3.978	6.824
	3.975	6.824	3.975	3.680	4.363	4.231	3.747	3.688	4.375	3.957	3.977	6.823
2	3.970	6.990	3.970	3.680	4.350	4.130	3.740	3.690	4.340	3.980	3.970	6.990
	3.980	6.990	3.990	3.680	4.350	4.130	3.740	3.690	4.350	3.990	3.970	6.990
3	4.160	7.020	4.110	3.680	4.340	4.160	3.810	3.740	4.360	4.140	4.180	7.030
	4.150	6.990	4.100	3.680	4.350	4.170	3.820	3.760	4.370	4.140	4.160	7.000
4	3.965	6.811	3.966	3.680	4.353	4.163	3.732	3.676	4.345	3.962	3.965	6.809
	3.960	6.811	3.959	3.680	4.344	4.160	3.732	3.682	4.336	3.961	3.962	6.811
5	3.960	6.790	3.960	3.680	4.330	4.100	3.720	3.660	4.320	3.950	3.950	6.780
	3.960	6.800	3.960	3.680	4.330	4.120	3.730	3.670	4.320	3.960	3.960	6.800
6	3.956	6.946	3.952	3.680	4.363	4.133	3.733	3.673	4.350	3.966	3.956	6.949
	3.979	6.942	3.974	3.680	4.334	4.129	3.739	3.679	4.339	3.975	3.991	6.942
7	4.101	6.986	4.121	3.680	4.315	4.144	3.957	3.699	4.185	4.179	4.160	7.011
	4.139	6.998	4.130	3.680	4.306	4.184	3.983	3.765	4.294	4.142	4.126	6.991
8	4.220	7.230	4.230	3.680	4.350	4.190	3.830	3.790	4.430	4.280	4.250	7.250
	4.250	7.250	4.250	3.680	4.320	4.160	3.810	3.770	4.400	4.290	4.280	7.270
9	3.960	7.370	3.980	3.680	4.300	4.100	3.740	3.690	4.280	3.980	3.960	7.350
	3.960	7.370	3.970	3.680	4.280	4.100	3.730	3.680	4.280	3.970	3.970	7.360
10	3.970	6.970	3.970	3.680	4.360	4.140	3.740	3.670	4.350	3.970	3.970	6.970
	3.970	6.970	3.960	3.680	4.360	4.130	3.730	3.670	4.350	3.970	3.960	6.960
11	3.950	6.930	3.950	3.680	4.340	4.110	3.740	3.680	4.340	3.960	3.950	6.910
	3.920	6.910	3.940	3.680	4.370	4.120	3.730	3.680	4.330	3.950	3.950	6.920
12	4.012	6.875	4.012	3.680	4.336	4.133	3.744	3.694	4.356	4.022	4.015	6.877
	4.010	6.875	4.014	3.680	4.344	4.137	3.747	3.689	4.351	4.021	4.016	6.872
13	3.979	6.860	3.981	3.680	4.369	4.140	3.726	3.667	4.365	3.965	3.967	6.850
	3.973	6.837	3.974	3.680	4.370	4.142	3.733	3.675	4.368	3.971	3.972	6.836
14	3.948	6.823	3.950	3.680	4.338	4.126	3.735	3.680	4.345	3.950	3.952	6.811
	3.963	6.833	3.966	3.680	4.352	4.138	3.743	3.680	4.354	3.967	3.965	6.816
15	3.939	6.767	3.939	3.680	4.308	4.098	3.712	3.685	4.320	3.939	3.937	6.738
	3.979	6.855	3.973	3.680	4.344	4.134	3.737	3.680	4.359	3.972	3.980	6.856
16	4.080	6.910	4.070	3.680	4.370	4.160	3.770	3.710	4.360	4.080	4.070	6.910
	4.050	6.890	4.080	3.680	4.340	4.120	3.750	3.690	4.330	4.050	4.060	6.900
17	3.960	6.940	3.950	3.680	4.370	4.150	3.780	3.700	4.400	3.940	3.960	6.960
	3.950	6.940	3.940	3.680	4.400	4.140	3.730	3.670	4.380	3.930	3.940	6.940

Table 5. Summary of results from the analyses of variance.

Components of Standard Deviation (S)	C	G	D	E[a]	F[b]
	Using Original pH Readings (table 2)				
S, *within-lab*	*0.020*	*0.018*	0.014	0.014	0.015
S, between-lab	0.091	0.087	0.077	0.056	0.055
	Using Normalized pH Reading (table 4)				
S, within-lab	0.014	0.022	0.014	0.012	0.014
S, between-lab	0.021	0.036	0.029	0.027	0.029

[a] Laboratory 7 data for solution E were judged to be outliers and were not used in the analysis of variance.

[b] Results for solution B are essentially the same as F.

References

[1] Koch, W. F., and G. Marinenko, Simulated Precipitation Reference Materials: Measurement of pH and Acidity, Special Technical Publication 823, American Society for Testing and Materials, Philadelphia, PA, pp. 10–17 (1983).

[2] Marinenko, G., and W. F. Koch, A Critical Review of Measurement Practices for the Determination of pH and Acidity of Atmospheric Precipitation, Environment International **10**, 315–319 (1984).

[3] Durst, R. A., Standard Reference Materials: Standardization of pH Measurement, Special Publication 260-53, 48 pages (Dec. 1985).

[4] Wu, Y. C.; W. F. Koch, and G. Marinenko, A Report on the National Bureau of Standards pH Standards, J. Res. Natl. Bur. Stand. (U.S.), **89** (5), 395–400 (1984).

APPENDICES

Appendix A

Participants in alphabetical order according to institution:

- Brookhaven National Laboratory, Long Island, NY
- Department of Energy, Environmental Measurements Laboratory, New York, NY
- Environmental Science and Engineering, Inc., Gainesville, FL
- Florida Southern College, Lakeland, FL
- Global Geochemistry Corp., Canoga Park, CA
- Illinois State Water Survey, Champaign, IL
- Martin Marietta Corporation, Baltimore, MD
- National Bureau of Standards, Gaithersburg, MD
- Research Triangle Institute, Research Triangle Park, NC
- Rockwell International, Newbury Park, CA
- Russell Sage College, Troy, NY
- South Coast Air Quality Management District
- United States Geological Survey, Denver, CO

Appendix B

Directions for Measurement of pH for the Interlaboratory Test

1) Make all pH measurements between 20 and 26 °C. Report the solution temperature to within 1 °C. Report all pH values to *at least* 0.01 pH unit. Make all measurements in a quiescent solution.
2) Standardize the pH electrodes and meter using SRM 185e, potassium hydrogen phthalate [pH(S) 4.004 at 25.0 °C, 0.05 molal], or equivalent. Refer to ASTM D1293, "Standard Test Methods for pH of Water" for guidance. Record the value on the enclosed data sheet.
3) With the slope adjustment of the meter set at 100%, check the Nerstian response of the pH measurement system with a second buffer, SRM 186Ic/186IIc, potassium dihydrogen phosphate/disodium hydrogen phosphate [pH(S) 6.863 at 25.0 °C] or equivalent. Refer to ASTM D1293 for guidance. If the reading for the second buffer is *not* within 0.03 pH units of the prescribed value, recheck the calibration of the system. *DO NOT CONTINUE* with the interlaboratory test until the conditions for calibration and Nernstian response have been satisfied. *See Note A*. If the reading for the second buffer is within 0.03 pH units of the prescribed value, record the value and continue.
4) Rinse the electrodes thoroughly with distilled water (ASTM Type II or better). Remove drops of water on the electrode by blotting gently (Do Not Rub!) with a clean lab tissue.
5) Insert the electrodes into a clean beaker (10–20 mL capacity) containing a portion (10–20 mL) of solution A (or subsequent solution). Be certain that the reference junction and glass bulb are completely immersed. Do not insert the electrodes directly into the polyethylene bottles.
6) Stir or swirl the solution to ensure homogeneity and contact with the electrodes.
7) Allow the solution to settle to a quiescent state (approx. 30 seconds). Record the pH value after the reading has stabilized. (Some systems may require five minutes or more to stabilize.) If drifting persists, record the value after 10 minutes in the quiescent state and note this fact in the

"Comments" column. Report only a single value for each solution per trial. Do not average several readings of the same solution.

8) Repeat steps 4 through 7 for solutions B, C, D, E, F, G, and H in this assigned sequence. Do not change the order, and do not recalibrate with buffer standards within this sequence.
9) Remeasure buffer standards #1 and #2 (without adjusting the settings of the pH meter) and record the values.
10) Repeat steps 1–9 for Trial #2.

Note A: Possible causes for deviation from Nernstian response include:

- i) improper calibration
- ii) old or contaminated buffer solutions
- iii) insufficient rinsing of the electrodes between solutions
- iv) plugged reference junction
- v) defective electrodes
- vi) defective meter

It may be necessary to replace the electrodes with a new pair if the non-Nernstian behavior persists even with accurately and freshly prepared buffer standards.

Appendix C

DATA SHEET

Interlaboratory Test NBS-AD-83

Laboratory/Scientists: ____________________

Date of Test: ____________

Temperature of Test Solutions: ______

Sample	pH Reading Trial #1	 Trial #2	Comments
Standard Buffer #1 (Calibration)	________	________	
Standard Buffer #2	________	________	
Solution A	________	________	
Solution B	________	________	
Solution C	________	________	
Solution D	________	________	
Solution E	________	________	
Solution F	________	________	
Solution G	________	________	
Solution H	________	________	
Standard Buffer #1	________	________	
Standard Buffer #2	________	________	

Appendix D

INFORMATION SHEET

Interlaboratory Test NBS-AD-83

-pH Meter

- Make/Model: ____________________________________
- Readout: Digital __________ ; Analog __________
- Sensitivity: 0.01 pH ________ ; 0.001 pH ________

-pH Electrodes

- Make/Model
  - Combination Electrode: ____________________
  - Glass Electrode: ________________________
  - Reference Electrode: ______________________
- Size: micro ______________ ; semi-micro ____________
  other (specify) __________
- Range (in pH units): 0-10 ____ ; 0-11 ____ ; 0-14 ____ ;
  other (specify) __________
- Reference Electrode Filling Solution (specify, e.g., 4 M KCl)
  ____
- Reference Electrode Junction: ceramic ____ ; fiber ____ ; sleeve ____ ;
  double ____ ; other (specify) ______________

-Standard Buffers

- Nominal pH 4: SRM 185e ________ ; Commercial (specify brand and pH values)
  ____
- Nominal pH 7: SRM 18Ic/186IIc ________ ; Commercial (specify brand and
  pH value) ____
- Other (specify brand and pH value) ________________________________

Journal of Research of the National Bureau of Standards
Volume 91, Number 1, January–February 1986

Development of a Standard Reference Material for Rainwater Analysis

William F. Koch, George Marinenko, and Robert C. Paule
National Bureau of Standards, Gaithersburg, MD 20899

Accepted: October 9, 1985

This paper describes the development of Standard Reference Material, SRM 2694, "Simulated Rainwater," intended to aid in the analysis of acidic rainfall. Details of the formulation and preparation of the two levels of solutions (2694-I and 2694-II) are given. The 10 analytical techniques used to measure the 12 components in the solutions are described in brief. The data used in the statistical evaluation of the results are summarized and the recommended values for pH, specific conductance, acidity, fluoride, chloride, nitrate, sulfate, sodium, potassium, ammonium, calcium, and magnesium are tabulated. The instability of ammonium ion in acidic solutions is discussed. Recommendations for the use of SRM 2694, particularly with regard to the measurement of pH, are given.

Key words: acid rain; acidity; ammonium; analytical chemistry; conductivity; measurement; pH; precipitation; rainwater; Standard Reference Material; statistics; sulfate.

1. Introduction

Wet deposition is monitored by various laboratories and agencies as part of national and international networks to record accurately the composition of rainfall. These efforts are intended to determine the extent of the problem of "acid rain," and to establish spatial and temporal trends. Discrepancies in data often occur due to differences in instruments and techniques. These discrepancies limit the conclusions which may be drawn from the data. To establish a common basis for chemical measurements in rainwater, a multi-year research effort has been established in the Inorganic Analytical Research Division of NBS; Center for Analytical Chemistry. This effort has resulted in the issuance of Standard Reference Material (SRM) 2694, Simulated Rainwater.

The initial stages of the evolution of this SRM are described in detail in a previous article [1].[1] In that article, the early formulations of multicomponent solutions are given, and the problems with their stability are discussed. The stability problems inherent in these early formulations stem from both the complexity of the solutions and the use of glass ampoules as storage containers. The progression to polyethylene bottles and to simplified solutions, from which the transition and heavy metals were eliminated, resulted in solutions of greater stability and overall applicability to the measurements of pH, acidity, conductance, nitrate, and sulfate. These components are most critical to acid rain studies.

About the Authors: William F. Koch and George Marinenko are chemists in NBS' Inorganic Analytical Research Division, part of the Bureau's National Measurement Laboratory in which Robert C. Paule, a physical scientist, serves. The work they describe was sponsored in part by the U.S. Environmental Protection Agency (National Acid Precipitation Assessment Program).

---

[1] Figures in brackets indicate literature references.

Associated difficulties with the measurement of pH and acidity are treated in other articles [2-4].

The pressing need of the scientific community for a common reference material for rainwater and the desire by us to field test such a material led to the production and distribution of a Research Material, RM 8409, Simulated Rainwater. RMs are distinguished from SRMs according to the definitions of NBS' Office of Standard Reference Materials [5].

In brief, RMs are high quality materials whose composition has been established by a single technique for each component. The composition of SRMs is certified after much more extensive testing involving at least two independent techniques for each component or analysis by a definitive method. Often an RM uncovers unforeseen problems, as was the situation in this case. This problem involved the long-term stability of the ammonium ion. The extent of the instability and its ramifications will be addressed below. The primary focus of this paper will be the preparation, analysis, and certification of SRM 2694.

2. The Preparation of SRM 2694

The decision to prepare simulated rainwater, rather than collecting natural rainfall, was based on the need to minimize contamination and unwanted components that would compromise the overall stability of the solutions. We also wanted to prepare two solutions of different concentrations of the various components, and we wanted to control the levels closely. Thus, simulated rainwater prepared by the dissolution of salts and acids in water was the best recourse.

The production of this simulated rainwater involved careful coordination of several operations, including bottle cleaning, formulations, dilution, mixing, and bottling. Approximately 2000 bottles of each level were prepared.

2.1 Bottle Cleaning

Based on earlier research [1], low density polyethylene (LDPE) bottles (60-mL capacity) were chosen for this project. Over 4000 bottles were cleaned using a rigorous procedure to minimize contamination. In brief, this cleaning procedure consisted of rinsing and soaking the bottles and caps for extended periods with filtered, distilled/deionized water. The bottles were then dried at 40 °C in a clean oven, and recapped until the filling operation was started. At all times, the bottles were kept away from areas with acid fumes.

2.2 Formulation and Target Values

Two levels of simulated rainwater containing the cations and anions commonly found in acid rain were formulated in such a way so as to span a useful analytical range of concentrations of all components. The target values for the two levels are shown in table 1. It should be noted that the measurement of pH was the primary driving force behind the development of this SRM. Hence, the stability of the solutions with respect to pH was an overriding constraint.

It has been found that unbuffered solutions at about pH 4.5 or above are extremely susceptible to fluctuations in pH and acidity due to absorption and desorption of atmospheric carbon dioxide. These processes occur even through the walls of the polyethylene bottles. For this reason, although it would have been desirable to issue a solution of pH 5.0, SRM 2694-I was targeted at pH 4.3 as a precaution.

Table 1. Target values of SRM 2694, simulated rainwater.

	2694-I	2694-II
pH	4.3	3.6
Specific Conductance, μS/cm	25	130
Acidity, meq/L	0.05	0.28
Fluoride, mg/L	0.05	0.1
Chloride, mg/L	0.25	1.0
Nitrate, mg/L	0.5	7
Sulfate, mg/L	2.7	11
Sodium, mg/L	0.2	0.4
Potassium, mg/L	0.05	0.1
Ammonium, mg/L	0.1	1.0
Calcium, mg/L	0.01	0.05
Magnesium, mg/L	0.025	0.05

2.3 Mixing of Simulated Rainwater

Seven ACS-reagent grade salts and three high-purity acids [6] were used in the preparation of the two levels of simulated rainwater. For convenience in the mixing process, stock solutions of the salts and acids were prepared. Table 2 lists the chemicals and the concentrations of the stock solutions. Table 3 lists the weights of each of the stock solutions used in the final dilutions of the solutions.

Table 2. Stock solutions used in the preparation of SRM 2694.

Stock Solution	Chemical	Concentration
1	$NaNO_3$	1 mg salt/g solution
2	KNO_3	1 mg salt/g solution
3	$CaCl_2 \cdot 2H_2O$	1 mg salt/g solution
4	$MgSO_4 \cdot 7H_2O$	1 mg salt/g solution
5	NH_4Cl	1 mg salt/g solution
6	$(NH_4)_2SO_4$	1 mg salt/g solution
7	H_2SO_4	0.100 N (0.050 mol/L)
8	HNO_3	0.100 N (0.100 mol/L)
9	HCl	0.050 N (0.050 mol/L)
10	NaF	1 mg salt/g solution

Table 3. The weights of the stock solutions used in the preparation of SRM 2694.

Stock Solution	2694-I (final volume, 170 L)	2694-II (final volume, 150 L)
1	83.43 grams	147.27 grams
2	22.06 grams	38.74 grams
3	9.61 grams	27.66 grams
4	34.88 grams	61.60 grams
5	50.97 grams	0 grams
6	0 grams	550.05 grams
7	85.00 grams	225.04 grams
8	1.78 grams	149.98 grams
9	0 grams	75.00 grams
10	19.89 grams	35.24 grams

A 200-liter polyethylene, cylindrical tank, which had been cleaned according to the same procedure used for the bottles, was used in the final dilution of the appropriate amounts of each of the stock solutions. Graduations at 10-liter increments were made on the outside of the tank as it was being filled during the cleaning process. These marks served only as an indication of the volume, and were not intended to calibrate the volume accurately. A stirring motor with a teflon paddle was used to thoroughly mix the solutions. Figure 1 shows the apparatus in operation.

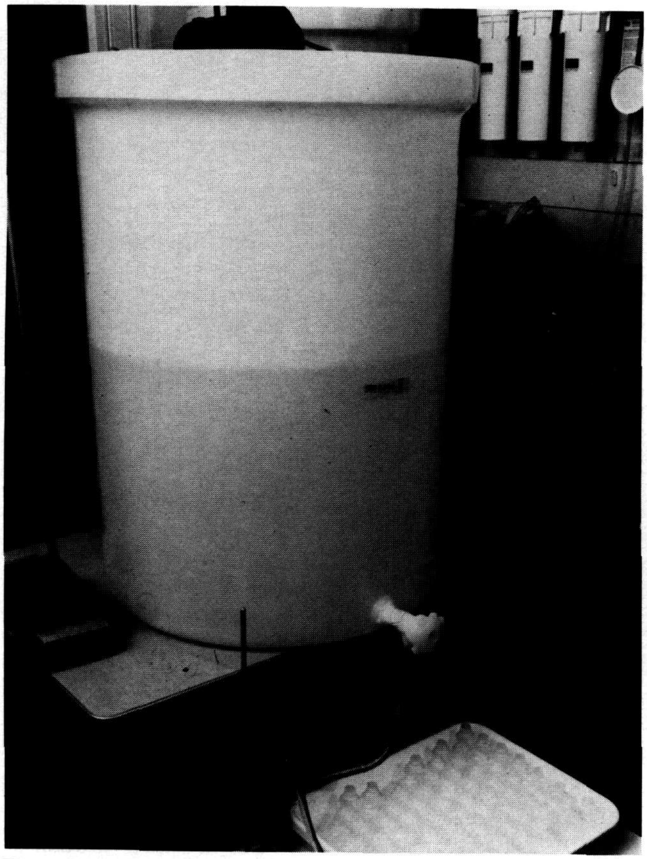

Figure 1–Apparatus used in the preparation and dispensing of SRM 2694.

Approximately 170 liters of 2694-I and approximately 150 liters of 2694-II were prepared in February 1985, according to the following procedure. The calculated amounts of each of stock solutions #1 through #9, (see table 3), were added to 100 liters of filtered, distilled/deionized water in the tank. The resulting solution was mixed thoroughly. Then the calculated amount of solution #10 (sodium fluoride) was slowly added. It was necessary to delay the input of sodium fluoride to prevent the possible precipitation of calcium fluoride, which is extremely difficult to redissolve. Filtered, distilled/deionized water was then added to bring the total volume to the pre-established mark. The solution was thoroughly homogenized by intermittent vigorous stirring over a 24-hour period. Note that although care was taken to add exact amounts of each component and to dilute with the correct amount of water, there was no convenient way to accurately assess the final volume. Hence, the weights and volumes could not be used as an analytical measure of the concentrations of the various components. This was left to the analytical chemists and their myriad of techniques.

2.4 Bottling

Before the bottling commenced, a sample of each level was analyzed by ion chromatography and potentiometry to verify that the target values had been met. The bottling of each level was completed in a single day. The bottles were filled manually, capped immediately, and placed sequentially in numbered cartons. Concurrent with the bottling operation,a homogeneity test was run. One out of every 120 bottles was pulled from the line and tested for specific conductance. The results, shown in tables 4 and 5, demonstrate that the two solutions are homogeneous.

3. Analysis of Simulated Rainwater

After establishing the homogeneity of the solutions by the measurement of conductivity, the statistical design for the sampling and analysis of the other components was configured. Each component was to be determined in triplicate in each of three bottles (selected from the beginning, middle and end of the bottling operation) by each technique. Because of experimental exigencies of some techniques, this analysis design was not strictly adhered to in every case, but was followed whenever possible.

The techniques used in the analysis of SRM 2694 were ion chromatography (IC), conductivity, potentiometry, coulometry, isotope dilution mass spectrometry (IDMS), spectrophotometry, laser enhanced ionization flame spectrometry (LEIS), flame emission spectrometry (FES), inductively coupled plasma (ICP), and

Table 4. Homogeneity test—specific conductance of SRM 2694-I in units of microsiemens per centimeter at 25 °C.

Bottle No.	Run #1	Run #2
101a	25.43	25.52
101b	25.48	25.50
102a	25.45	25.44
102b	25.43	25.47
103a	25.48	25.50
103b	25.47	25.44
104a	25.44	25.47
104b	25.46	25.50
105a	25.45	25.44
105b	25.45	25.47
106a	25.44	25.48
106b	25.50	25.48
107a	25.48	25.46
107b	25.44	25.50
108a	25.46	25.48
108b	25.45	25.45
109a	25.49	25.48
109b	25.45	25.48
110a	25.47	25.44
110b	25.46	25.46
Mean	25.47 μS/cm at 25.0 °C	
Std. Dev. of a Single Meas.		0.02

Table 5. Homogeneity test—specific conductance of SRM 2694-II in units of microsiemens per centimeter at 25 °C.

Bottle No.	Run #1	Run #2
1a	130.1	130.3
1b	129.4	130.2
2a	129.8	130.0
2b	129.8	129.6
3a	130.2	129.8
3b	130.1	129.7
4a	130.1	130.0
4b	130.3	130.2
5a	129.4	129.8
5b	130.3	130.2
6a	129.6	129.9
6b	130.0	130.7
7a	130.2	130.8
7b	130.4	129.9
8a	130.1	129.9
8b	130.4	130.8
9a	129.9	129.7
9b	130.1	130.1
Mean	130.1 μS/cm at 25.0 °C	
Std. Dev. of a Single Meas.		0.3

flame atomic absorption spectrometry (FAAS). Details of the analytical procedures for each technique will be presented in a forthcoming publication. An overview of the methods will be given here.

Ion chromatography was used to determine the concentrations of fluoride, chloride, nitrate, sulfate, sodium, potassium, and ammonium. Dual channel, dual column IC with hollow fiber chemical suppressor systems was employed. For each level and each component, three calibration points (peak height versus concentration) which bracketed closely the concentrations of the individual components were established. Chloride, nitrate, and sulfate were determined sequentially under one set of chromatographic conditions. A different set of conditions, involving an extremely weak carbonate eluent, was required to resolve the fluoride peak from the negative water-dip to allow the accurate determination of the fluoride concentrations. The cations were determined using a hydrochloric acid eluent. A concern about the stability of ammonium ion will be discussed below.

The specific conductance was measured using a dip-type conductance cell of nominal cell constant 0.1 cm^{-1} and an AC conductivity bridge operating at 1 kHz. Measurements for certifications were made at 25.0 °C by thermostating the solutions in a water-jacketed beaker, the outer chamber of which contained circulating constant temperature water to maintain the temperature at 25 °C. The exact cell constant of the cell was determined using 0.001 demal KCl, which has a specific conductance of 146.93 microsiemens per centimeter [μ S/cm]. Measurements were also made at temperatures ranging from 20 °C to 28 °C to establish the temperature coefficient for the specific conductance of these solutions. The temperature coefficient for both levels was determined to be 1.5% per °C at 25.0 °C.

Potentiometry was used to measure pH, fluoride, ammonium, and chloride. Measurements of pH were made with a combination glass electrode according to the procedure established at NBS. All measurements were corrected for residual liquid junction potential bias by normalizing to a dilute solution of sulfuric acid, whose pH had been determined accurately in a hydrogen cell without liquid junction. This cell is of the type used in the certification of the NBS pH buffers. Fluoride was measured with a fluoride ion-selective electrode after addition of a total ionic strength adjustment buffer. Standards which bracketed the concentrations of the rainwater samples were used to calibrate the measurement system. Ammonium ion was determined using an ammonia electrode. The method involves addition of concentrated base to the sample thereby liberating ammonia which diffuses through the semipermeable membrane of the electrode and is sensed by an internal glass pH electrode. The long-term stability of ammonium ion will be discussed below. Chloride was also determined potentiometrically using a chloride ion selective electrode. However, the concentrations of chloride in the simulated rainwater are well below the linear range of response. Thus they are prone to a large uncertainty. The chloride measurements serve only as an indication of concentration and were not used in the final statistical analysis. This method and its values are presented for information only.

The acidity of SRM 2694-II was determined by coulemetric reduction of hydrogen ion in a weighed sample, from which CO_2 had been removed by purging with argon [3,4]. Titrations were carried out to the neutrality point, determined potentiometrically.

Thermal ionization isotope dilution mass spectrometry was used to determine the total sulfur in the simulated rainwater. This value was then converted to sulfate concentration. Briefly, the method involves the reduction of sulfate to sulfide, distillation of sulfide and collection in a basic arsenic [III] solution, followed by mass spectrometric identification and quantitation of AsS^+ isotopes [7,8].

Spectrophotometry was used for the determination of nitrate. The procedure was based on the color reaction produced by interaction of nitrate ion and the organic reagent, brucine. Careful attention to experimental procedures and frequent calibration with nitrate standards were required to obtain satisfactory results by this method.

Laser-enhanced ionization flame spectrometry, a new analytical tool in the Inorganic Analytical Research Division, was utilized for the determination of sodium, potassium, calcium, and magnesium. Wavelength scans near the analysis lines were performed to check for spectral background and interfering lines. Minor corrections were applied to the magnesium determination to correct for sodium interference. Magnesium concentration was also determined by flame atomic absorption spectrometry.

Inductively coupled plasma spectrometry was used to measure the calcium concentration in the samples, and flame emission spectrometry was used to determine sodium and potassium. The recovery of each analyte was checked by the single standard addition method.

In addition, the density of the solutions was determined to be 0.997g/ml at 23 °C, essentially identical to pure water at this temperature.

4. Results and Statistical Analysis

The average values from each measurement technique are shown in table 6. Values in parentheses represent the standard deviation of a single measurement. Values in braces are the number of analyses performed. We have used established techniques to calculate weighted averages for the SRM certificate values. The statistical weighting is based on the observed variablities of the various data sets for each analyte. The procedures for deriving the weighted averages and their uncertainties have been described in a previous article [9]. Table 7 contains the certified values and uncertainties for SRM 2694.

The uncertainties associated with fluoride, nitrate, sodium, potassium, calcium, and magnesium are two standard deviations of the certified values. The uncertainties in the certified values for pH, acidity, and specific conductance are based on scientific judgment and experience, rather than on true statistical evaluations because there were no practical second methods of analysis for these components.

The uncertainties tabulated for sulfate are also based on scientific judgment even though there were two independent methods of analysis. For this analyte, the agreements both within and between the methods of analysis were so close that the statistical evaluations of the uncertainties were not believed to be realistic. The uncertainties for these latter four components [pH, acidity, specific conductance, and sulfate] are believed to be roughly equivalent to two standard deviations of the certified values.

The values for chloride and ammonium are listed for information only. Chloride was not certified because the potentiometric method had insufficient precision and accuracy at these levels to corroborate the ion chromatographic data. Until a second independent technique verifies the IC measurements, chloride will not be certified. Ammonium ion was not certified because of very real concerns about the stability of this ion in these solutions.

5. Discussion

The problem with the stability of ammonium was first noticed upon reanalysis of RM 8409 eight months after its preparation. The concentration of ammonium in RM 8409-I had decreased from 0.085 mg/L to 0.025 mg/L. Reanalysis of other samples of simulated rainwater, which were part of a long-term stability study of pH and conductivity, indicated that similar decreases in ammonium ion had occurred. The decrease was significant when the initial concentration of ammonium was below 0.2 mg/L and the pH was above 4.0 (see table 8). The cause of this decrease is not known at this time, but it is suspected to be biological activity. The loss of ammonium does not appear to have significantly affected any

Table 6. Summary of data for SRM 2694.

Component Technique	2694-I[1] Mean (Std. Dev.) {n}	2694-II[1] Mean (Std. Dev.) {n}
1. **pH**		
A. Potentiometry	4.297 (.010)[1] {7}	3.591 (0.007) {8}
2. **Specific Conductance** (μS/cm @ 25 °C)		
A. Conductimetry	25.47 (.02) {40}	130.1 (.3) {36}
3. **Acidity, meg/L**		
A. Coulometry	0.0501 (.0015) {6}	0.2835 (.0045) {6}
4. **Fluoride, mg/L**		
A. IC	0.055 (.001) {12}	0.095 (.001) {11}
B. Potentiometry	0.0532 (.0015) {8}	0.102 (.003) {8}
5. **Chloride, mg/L**		
A. IC	0.242 (.007) {9}	1.003 (.032) {9}
B. Potentiometry	0.29 (.05) {8}	0.87 (.19) {8}
6. **Nitrate,mg/L**		
A. IC	0.513 (.007) {9}	7.141 (.044) {9}
B. Spectrophotometry	0.49 (.02) {9}	6.98 (.05) {9}
7. **Sulfate, mg/L**		
A. IC	2.688 (.018) {9}	10.83 (.06) {12}
B. IDMS	2.697 (.012) {5}	10.81 (.03) {6}
8. **Sodium, mg/L**		
A. IC	0.204 (.004) {12}	0.412 (.002) {12}
B. LEIS	0.1974 (.0030) {22}	0.411 (.006) {22}
C. FES	0.212 (.004) {5}	0.432 (.006) {5}
9. **Potassium, mg/L**		
A. IC	0.053 (.008) {8}	0.107 (.010) {12}
B. LEIS	0.0478 (.0015) {12}	0.1023 (.0031) {16}
C. FES	0.055 (.001) {6}	0.111 (.001) {6}
10. **Ammonium, mg/L**		
A. IC	—	1.064 (.005) {12}
B. Potentiometry	0.1031 (.0039) {6}	1.034 (.010) {7}
11. **Calcium, mg/L**		
A. LEIS	0.0126 (.0004) {12}	0.0436 (.0005) {12}
B. ICP	0.015 (.002) {6}	0.054 (.003) {7}
12. **Magnesium, mg/L**		
A. LEIS	0.0231 (.0003) {14}	0.0492 (.0005) {13}
B. FAAS	0.025 (.001) {6}	0.052 (.006) {6}

[1] Values in parentheses represent the standard deviation of a single measurement. Values in braces are the number of determinations.

other components. There have been isolated incidences of visible fungal growth in a few bottles of RM 8409. In such cases the values for nitrate, pH, acidity, and specific conductance, have changed. A solution to this problem involving the sterilization of the "simulated rainwater" is currently under investigation.

It must be noted that the solutions of SRM 2649 are very dilute, unbuffered solutions, and, as such, are very susceptible to contamination causing gross changes in the certified values. Therefore, the solutions should be used immediately upon opening. No assurance can be made as to the composition or stability of the solutions after being opened and recapped. It has been reported to us that a sample of RM 8409 which had been opened, recapped, and stored in a refrigerator lost virtually all of its nitrate content but gained a significant amount of nitrite.

SRM 2694 should be stored in an area free from acid and/or ammonia vapors. These vapors can permeate the polyethylene bottles and contaminate the samples. A set of samples placed in our laboratory refrigerator, which also contained a polyethylene bottle of concentrated ammonium hydroxide, showed a substantial increase in ammonium ion concentration. Refrigeration of SRM 2694 is not necessary. However, the solutions should not be exposed to extreme heat (i.e., temperatures above

Table 7. Certified values for SRM 2694.

Component	2694-I[1]	2694-II[1]
pH @ 25 °C	4.30±0.02	3.59±0.02
Specific Conductance (μS/cm @ 25 °C)	26±2	130±2
Acidity, meq/L	0.050±0.002	0.284±0.005
Fluoride, mg/L	0.054±0.002	0.098±0.007
Chloride,mg/L	(0.24)[2]	(1.0)[2]
Nitrate, mg/L	0.501±0.026	7.06±0.15
Sulfate, mg/L	2.69±0.03	10.8±0.1
Sodium, mg/L	0.205±0.009	0.419±0.015
Potassium, mg/L	0.052±0.007	0.106±0.008
Ammonium, mg/L	—	(1.0)[2]
Calcium, mg/L	0.014±0.003	0.049±0.011
Magnesium, mg/L	0.024±0.002	0.051±0.003

[1] The estimated uncertainties are two standard deviations of the certified values, except in the cases of pH, acidity, conductance, and sulfate which are based on scientific judgment and are roughly equivalent to two standard deviations of the certified value.

[2] The values for chloride and ammonium are not certified, but are listed for information only.

35 °C) as this will accelerate transpiration of water vapor from the bottles. The bottles, sealed in aluminized bags to retard transpiration, should remain within the sealed bags until just before use. The search for a better container has been initiated.

If conductance and pH are to be measured on the same sample, conductance should be measured first. Otherwise, leakage of concentrated KCl from the pH reference electrode will affect the conductance reading. The measurement of pH should be performed according to the guidelines set forth in an attachment to the Certificate of Analysis of SRM 2694 and appended to this report (appendix 1). Adherence to this procedure will minimize the bias caused by residual liquid junction potentials. Acidimetric titrations should be performed on samples that have been purged of dissolved carbon dioxide to prevent drifting endpoints and high results.

The contributions of the following individuals at the National Bureau of Standards are gratefully acknowledged: D. G. Friend, D. E. Swearingen, and F. Smithers for bottle cleaning, bottling and packaging; T. C. Rains, R. W. Burke, M. Knoerdel, R. W. Kelly, G. C. Turk, Mo De-Ming, M. S. Epstein, T. A. Rush, T. A. Butler, Han Kai, and M. V. Smith for the chemical measurements; and T. E. Gills and L. J. Powell for coordinating the technical and support aspects from the Office of Standard Reference Materials.

References

[1] Koch, W. F.; G. Marinenko and Y. C. Wu, The Development of Reference Materials for Acid Rain Research, Environment International **10**, 117–121 (1984).

[2] Koch, W. F., and G. Marinenko, Simulated Precipitation Reference Materials: Measurement of pH and Acidity, *Sampling and Analysis of Rain*, ASTM STP 823, S. A. Campbell, ed., American Society for Testing and Materials, 10–17 (1983).

[3] Marinenko, G., and W. F. Koch, A Critical Review of Measurement Practices for the Determination of pH and Acidity of Atmospheric Precipitation, Environment International **10**, 315–319 (1984).

[4] Marinenko, G., and W. F. Koch, Evaluation of Methods Used for the Determination of Acidity in Acid Rain Samples, Natl. Bur. Stand. (U.S.), NBSIR 85-3114, 16 pp. (March 1985).

[5] NBS Standard Reference Materials Catalog 1984–1985, C. H. Hudson, ed., Natl. Bur. Stand. (U.S.), Special Publication 260, p. 2, (February 1984).

[6] Moody, J. R., and E. S. Beary, Purified Reagents for Trace Metal Analysis, Talanta **29**, 1003–1010 (1982).

[7] Paulsen, P. J., and W. R. Kelly, Determination of Sulfur by Isotope Dilution Thermal Ionization Mass Spectrometry as AsS^+ Ions, Anal. Chem. **56**, 707–713 (1984).

[8] Kelly, W. R., and P. J. Paulsen, Determination of Sulfur in NBS Coals by Isotope Dilution Thermal Ionization Mass Spectrometry, in *Methods and Procedures used at the National Bureau of Standards to Certify Sulfur in Coal SRM's for Sulfur Content, Calorific Value, Ash Content*, T. E. Gills, ed., Natl. Bur. Stand. (U.S.), Special Publication 260-94, pp. 7–13 (December 1984).

[9] Paule, R. C., and J. Mandel, Consensus Values and Weighting Factors, J. Res. Natl. Bur. Stand. **87**, 377–385 (1982).

Table 8. Stability of ammonium ion in simulated rainwater.

Sample	Preparation Date	pH	Ammonium Ion, mg/L Initial	Ammonium Ion, mg/L Feb. 20, 1985
RM 8409-I	June 1984	4.3	0.085	0.025
RM 8409-II	June 1984	3.6	1.07	1.07
DSPRM V-1	Feb. 1983	4.5	0.14	0.09
DSPRM V-2	Feb. 1983	3.5	1.11	1.08
DSPRM V-3	Feb. 1983	4.0	0.42	0.42

Appendix

Guidelines for the Measurement of pH in Acidic Rainwater

This report presents a recommended procedure for the measurement of pH in acidic rainwater. The intent of this guideline is to improve the accuracy and precision of the pH measurement with special emphasis on reducing the effect of the residual liquid junction potential. It consists of three major parts: Calibration Sequence, Control Sequence, and Rainwater Measurement Sequence. The purposes of the Calibration Sequence are to accurately calibrate the pH measurement system with robust buffer solutions, to accurately set the slope, and to verify that the measurement system is functioning properly. The purpose of the Control Sequence is to quantitatively determine the magnitude of the residual liquid junction potential bias for a particular set of electrodes which must be applied in the rainwater measurement sequence to obtain more reliable and intercomparable results. Each sequence should be executed in stepwise order with strict adherence to detail.

Note: This guideline is applicable only to the measurement of pH in acidic rainwater and acidic low ionic strength aqueous solutions. It should not be used for any other applications as inaccuracies may ensue.

General Directions

Record the solution temperature to within 1 °C. Record all pH values to at least 0.01 pH unit. Make all measurements in a quiescent solution. Fully document all calibration and control standards.

Calibration Sequence

1) Standardize the pH electrodes and meter using SRM 185f, Potassium Hydrogen Phthalate [pH(S) 4.006 at 25.0 °C, 0.05 molal], or equivalent.[1] Refer to ASTM D1293, "Standard Test Methods for pH of Water" for guidance. Record the value. Rinse the electrodes with distilled water (ASTM Type II or better).

2) With the slope adjustment of the meter set at 100 percent, and the temperature adjustment set at the temperature of the buffer solution, check the Nernstian response of the pH measurement system with a second buffer, SRM 186Ic/186IIc, Potassium Dihydrogen Phosphate/Disodium Hydrogen Phosphate pH(S) 6.863 at 25 °C], or equivalent.[1] Refer to ASTM D1293 for guidance. If the reading for the second buffer is not within 0.03 pH units of the prescribed value, recheck the calibration of the system.[2] DO NOT CONTINUE until the conditions for calibration and Nernstian response have been satisfied. If the reading for the second buffer is within 0.03 pH units of the prescribed value, record the value and continue.

Control Sequence

3) Rinse the electrodes thoroughly with distilled water (ASTM Type II or better). Remove drops of water on the electrode by blotting gently (Do Not Rub!) with a clean lab tissue.

4) Insert the electrodes into a clean beaker (10–20 mL capacity) containing a portion (10–20 mL) of the rainwater control standard (e.g., SRM 2694-I).[3] Be certain that the reference junction and glass bulb are completely immersed. Do not insert the electrodes directly into the polyethylene bottles.

5) Stir or swirl the solution to ensure homogeneity and contact with the electrodes.

6) Allow the solution to settle to a quiescent state (approx. 30 seconds). Record the pH after the reading has stabilized.[4]

7) Discard this portion of the control standard. Do not use for subsequent control checks or for other analytical determinations such as specific conductance, anions, cations, and acidity.

8) Repeat steps 3 and 7 with a second rainwater control standard (e.g., SRM 2694-II).[3]

9) Calculate the differences between the true pH values of the rainwater control standard and the values as determined by the pH measurement system.[5] Average the differences and apply this bias correction to subsequent rainwater measurements. (For example, if the pH measurement system displays the pH of the control

0.13 pH units lower than the true value, add 0.13 pH units to the subsequent pH measurements of rainwater.)

Rainwater Measurement Sequence

10) Rinse the electrodes thoroughly with distilled water (ASTM Type II or better). Remove drops of water on the electrode by blotting gently (Do Not Rub!) with a clean lab tissue.

11) Insert the electrodes into a clean beaker containing a portion (10–20 mL) of the rainwater sample. Be certain that the reference junction and glass bulb are completely immersed.

12) Stir or swirl the solution to ensure homogeneity and contact with the electrodes.

13) Allow the solution to settle to a quiescent state (approx. 30 seconds). Record the pH after the reading has stabilized.[4]

14) Apply the bias correction as determined in step 9 and report this corrected value as the pH of the rainwater sample.

15) Discard this portion of the rainwater sample. Do not use it in other analytical tests.

16) Repeat steps 10 through 15 for subsequent rainwater samples.

17) Repeat the Control Sequence at regular intervals, based upon quality control guidelines, performance history of the measurement system, frequency of measurements, and required accuracy.

Storage of Electrodes

18) When not in use, soak the electrodes in a solution which is 0.1 mol/L potassium chloride and 1×10^{-4} mol/L hydrochloric acid. Do not store the electrodes in buffers, concentrated acids, concentrated potassium chloride, basic solutions, or distilled water. Do not allow the electrodes to dry out.

19) Use these electrodes exclusively for rainwater measurements.

---

[1] For the highest accuracy and the most direct traceability to the National Bureau of Standards, it is recommended that freshly prepared solutions of Standard Reference Materials (SRM's) be used.

[2] Possible causes for deviation from Nernstian response include:

i) improper calibration
ii) old or contaminated buffer solutions,
iii) insufficient rinsing of the electrodes between solutions,
iv) plugged reference junction,
v) defective electrodes.
vi) defective meter.

It may be necessary to replace the electrodes with a new pair if the non-Nernstian behavior persists even with accurately and freshly prepared buffer standards.

[3] Standard Reference Materials 2694, Simulated Rainwater, is issued by the National Bureau of Standards, Office of Standard Reference Materials.

[4] Some systems may require five minutes or more to stabilize. If drifting persists, record the reading after 10 minutes and annotate the data accordingly. With such severe drift, it would be advisable to acquire a different type of pH electrodes which do not exhibit this adverse characteristic.

[5] The difference should agree to within 0.05 pH units. If not, repeat the Control Sequence.