Analyzing and Presenting pH Data

M. R. MURPHY
Department of Dairy Science
University of Illinois
Urbana 61801

ABSTRACT

Both the practical effects and theoretical basis of error in pH data analysis and presentation were examined. Since the "pH" transformation has been applied for convenience and not for statistical reasons, added care must be exercised in analyzing these data. Calculating a mean of observed pH's introduced a significant and predictable bias as compared to −log (mean [H^+]), henceforth designated mH. Rumen pH data from Holstein steers fed a high concentrate diet twice daily showed that a relative error of from .35 to 2.22%, with an average of 1.25%, resulted when the former method was used. Wide ranges in observed pH or low pH greatly increased the possible relative error. Statistical analyses of pH data should use the untransformed variable, i.e., hydrogen ion concentration. Mean hydrogen ion concentration, the variable in which researchers are interested, still can be transformed for convenience and reported as mH together with the asymmetric confidence intervals that result from transformation.

INTRODUCTION

The pH, or "hydrogen ion exponent," was introduced by Sorensen in 1909 as a convenient way of expressing small hydrogen ion normalities (1, 2). Widespread and longtime usage of pH has caused it to be considered mathematically equivalent to other biological variables. This misconception, coupled with expanded use of computer programs to statistically analyze experimental data, has increased the frequency with which pH is misused. The objectives of this paper are to examine both the practical effects and theoretical basis of error in pH data analysis and presentation.

EXPERIMENTAL PROCEDURES

Rumen pH data from rumen-fistulated Holstein steers fed a high concentrate diet twice daily were used (unpublished data). The pH had been measured at 2-h intervals over a 12-h feeding period and was used to compare effects of various treatments on rumen fermentation.

RESULTS AND DISCUSSION

Each rumen pH observation represented a particular effective hydrogen ion concentration of research interest. A negative logarithmic transformation, from concentration to pH, was applied merely as a customary convenience. Transformation is required for the valid application of tests of significance in the analysis of variance when non-normality exists in the data (3, 4). The logarithmic transformation, in particular, is usually applied to data in which the mean is positively correlated with the variance. It also may be used to convert multiplicative effects on the original scale of measurement to additive effects on the logarithmic scale. Neither of these conditions would be expected to apply to hydrogen ion concentration data. In fact, mixing solutions of different pH results in a hydrogen ion concentration nearly the sum of their unneutralized hydrogen ion equivalents divided by their combined volume. Small deviations from equilibrium shifts are expected.

Since the "pH" transformation has been applied for convenience and not for statistical reasons, added care must be exercised in analyzing this type of data. A standard data analysis would report an arithmetic mean of the observed pH's, i.e., mean (−log[H^+]). Alternatively, the hydrogen ion concentrations implied by these pH measurements can be averaged and reported as the negative logarithm of the mean hydrogen ion concentration, i.e., −log (mean [H^+]), henceforth designated mH.

Received January 29, 1981.

TABLE 1. Analysis of rumen pH data.

Animal	Observed pH Minimum	Observed pH Maximum	mH −log (mean [H+])	Mean pH mean (−log [H+])	Relative error %
1	5.56	6.63	5.86	5.99	2.22
2	5.98	6.66	6.28	6.34	.96
3	5.60	5.99	5.72	5.74	.35
4	5.68	6.55	6.11	6.20	1.47
Mean	....	....	....	....	1.25

Results for both methods are in Table 1. Differences are attributable to the error committed when mean pH is used instead of mH.

Calculating an arithmetic mean of observed pH's introduces a predictable bias in that it will exceed mH except for the case when all samples are identical. The magnitude of this error depends on the range of pH's or hydrogen ion concentrations as well as the mean hydrogen ion concentration. For the data in Table 1, the average relative error was 1.25% with a range of .35 to 2.22% (relative error was defined as |mH − mean pH|/mH).

Presented in Figure 1 is a graph of the percent relative error versus pH for the hypothetical case of physiologically common 10-fold differences in hydrogen ion concentration. Observed pH's 6 and 7 would have a mean pH of 6.50 whereas mH indicates 6.26 is correct. A 3.83% error (graphed as the ordinate in Figure 1) would have occurred had 6.50 (graphed as the abscissa in Figure 1) been used. The phenomenon of increasing relative error as pH decreases also is illustrated in Figure 1. In addition, the relative error at any given pH would increase tremendously if 100-fold or 1000-fold ranges in hydrogen ion concentration were encountered. For a "normal" rumen pH range of 5 to 7 the possible error, although unacceptable, is not nearly as great as that which could occur over the "normal" abomasal pH range of 2 to 4.

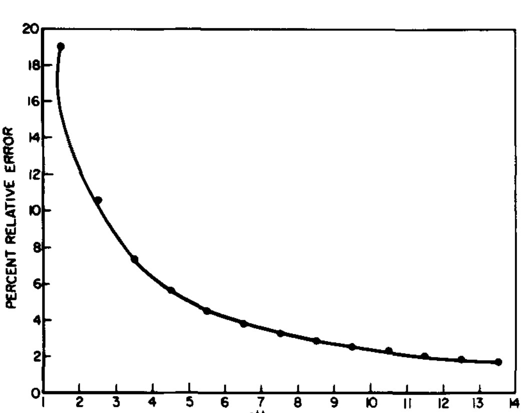

Figure 1. Percent relative error versus pH for a hypothetical 10-fold difference in hydrogen ion concentration (see text).

This type of error can be avoided easily by performing all statistical analyses, including calculation of means, with the untransformed variable, i.e., hydrogen ion concentration. Mean hydrogen ion concentration, the variable researchers actually wish to compare, still can be transformed and reported as mH; however, it must be remembered that confidence limits computed in the original scale will not be symmetrical once transformed. For example, if the mean hydrogen ion concentration was .0100 molar with a 95% confidence interval of ± .0020 molar, the mH would be 2.000 with a 95% confidence interval of 1.921 to 2.097. The transformed 95% confidence interval is 2.000 plus .097 but minus .079. Since standard errors would be misleading under such circumstances, the asymmetric confidence limits should be reported.

ACKNOWLEDGMENTS

The author wishes to thank Neil Forsberg for calling his attention to the peculiarity of pH data.

REFERENCES

1 Bates, R. G. 1954. Pages 18–19 *in* Electrometric pH determinations. John Wiley & Sons, Inc., New York.

2 Jðrgensen, H. 1950. Pages 42–51 *in* Theorie, mesure et applications du pH. Dunod, Paris.
3 Sokal, R. R., and F. J. Rohlf. 1969. Pages 380–384 *in* Biometry. W. H. Freeman and Co., San Francisco.
4 Steel, R.G.D., and J. H. Torrie. 1960. Pages 156–158 *in* Principles and procedures of statistics. McGraw-Hill, Inc., New York.

ERRATUM

Effect of Limestone Buffers on Digestibility of Complete Diets and on Performance by Dairy Cows. W. E. Wheeler. J. Dairy Sci. 63:1848.

Effect of Limestone Buffers on the Digestion of Complete Mixed Rations by Dairy Cows. W. E. Wheeler. J. Dairy Sci. 60 (Suppl. 1):117. (Abstr.)

Limestone Buffers in Complete Mixed Rations for Dairy Cattle. W. E. Wheeler and C. H. Noller. J. Dairy Sci. 59:1788.

Effect of Forage-to-Concentrate Ratio in Complete Feeds and Feed Intake on Digestion of Starch by Dairy Cows. W. E. Wheeler, C. H. Noller, and C. E. Coppock. J. Dairy Sci. 58:1902.

Accounting for Variability in Digestible Energy of Dairy Rations by Multiple Regression Analysis. W. E. Wheeler, C. H. Noller, C. E. Coppock, J. R. Robertson, and P. J. Van Soest. J. Dairy Sci. 57:621. (Abstr.)

"Irregularities have been discovered in data reported in abstracts and manuscripts published in the Journal of Dairy Science on which W. E. (William) Wheeler is senior author. Therefore, the validity of the reported information is questionable."

Signed: William E. Wheeler