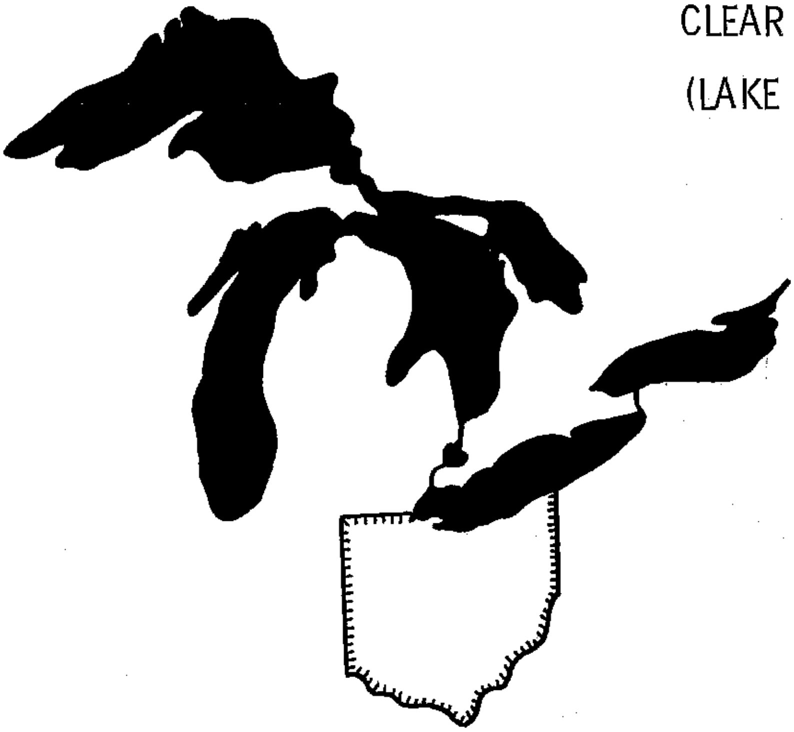
CLEAR TECHNICAL REPORT NO. 247
(LAKE ERIE TAT CONTRIBUTION NO. 22)



LAKE ERIE INTENSIVE STUDY: HISTORICAL WATER QUALITY TRENDS--CLEVELAND, OHIO

Prepared by

R. Peter Richards
Heidelberg College, Tiffin, Ohio

Prepared for

U.S. Environmental Protection Agency
Great Lakes National Program Office
Region V - Chicago, Illinois
Grant No. R005516001

THE OHIO STATE UNIVERSITY
CENTER FOR LAKE ERIE AREA RESEARCH
COLUMBUS, OHIO

DECEMBER 1981

TABLE OF CONTENTS

ABSTRACT

Comparison of recent data from the Central Basin nearshore zone with the historical data of Beeton indicates that the concentrations of most parameters in 1978-1979 were lower than would be expected from the long-term trend based on the historical data. Analysis of a more detailed data set from the Cleveland Division water intake shows a statistically significant decline in chloride, but significant increases in nitrate plus nitrite and sulfate in the period 1968-1979. Total phosphorus shows a significant decline if the data is first filtered to remove seasonal fluctuations. Residuals analysis failed to show a change in the seasonal pattern for any parameter during the period of record. Further analysis suggests that this data set may contain biases, and shows that portions lack the sensitivity necessary to completely assess the historical trends for some parameters. These findings indicate the significant detrimental effect which lab-to-lab variability can have in any attempt to study subtle limnological phenomena, and point to the importance of using data sets from a single lab, if at all possible.

INTRODUCTION

During 1978 and 1979, the nearshore zone of Lake Erie was sampled intensively as part of the monitoring and surveillance program for the Great Lakes (Herdendorf, 1978). The Heidelberg College Water Quality Laboratory (HCWQL) was responsible for sampling in the nearshore zone of the Central Basin between Vermilion, Ohio and Ashtabula, Ohio. Sampling was carried out at 89 stations (see figure 1). Each station was sampled on three successive days during each of four cruises each year. At most stations, samples were collected one meter below the surface and one meter above the bottom.

One major purpose of this nearshore study was to seek to identify historical trends, especially among parameters that may have changed in concentration due to human impact on Lake Erie. This paper addresses that purpose at two levels: comparisons with a long term, but often sketchy database that extends back to 1900, and comparisons with a much more detailed, but localized database for one station from 1968 to 1979.

The attempt to identify historical trends is often frustrated by the scarcity and inadequate quality of historical data. Changes in methods of analysis affect the data in ways which are hard to identify. The methods of analysis themselves are often not specified. Even when

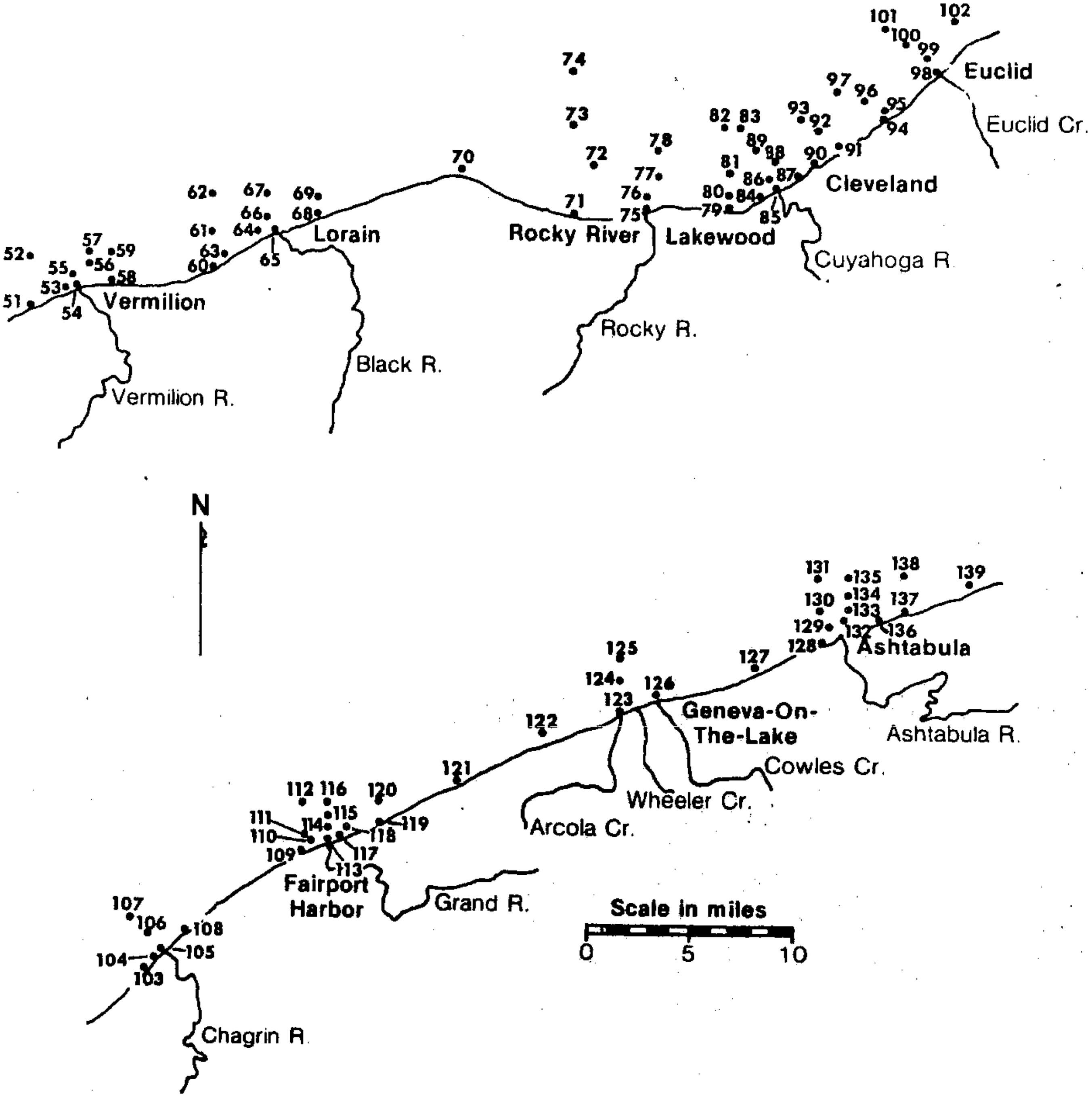


Figure 1-- Map of station locations sampled by the Heidelberg College Water Quality Laboratory in 1978 and 1979. Station 82 is the Division Water Intake of the City of Cleveland. The fifteen stations used for comparison with Beeton's data are 52,62,73,74,82,83,97,101,107,112, 116,125,131,135, and 138.

the methods are specified, and are known to be bias-free as methods, the possibility of biases due to different working ranges and other laboratory-level differences is very real, but usually difficult to evaluate. For the nearshore zone, data is scarce even in comparison to the sketchy data set for the open lake. Much of the data found in the literature before 1950 consists of average values, and often the locations where the data was obtained are not adequately specified. Also, the nearshore zone is much more variable spatially and temporally than is the open lake, making historical trends more difficult to detect.

For all of these reasons, historical analysis is a hazardous exercise, especially in nearshore waters. Even statistically significant changes must be carefully scrutinized to see if they are reasonable in limnological context, or if they are better interpreted as artifacts of problems in the data set. The later sections of this paper suggest some ways this can be accomplished. The point must be made at the outset, however, that statistical trend analysis can show only a significant change in the numbers in a data set as a function of time. Identification of a historical trend in Lake Erie involves a substantial _interpretation_ of the results of the statistical procedure.

LONG TERM HISTORICAL TRENDS

One of the most important historical trend studies is that of Beeton (1961,1965; Beeton and Chandler, 1963). The importance of this study lies in its time span (1902 to 1960) and its concern with chemical parameters of general importance: total dissolved solids, calcium, sodium, potassium, sulfate, and chloride. The concentrations of these parameters in the water are great enough that one can have at least cautious confidence that measurements in the early 1900's are not drastically inaccurate.

The data used by Beeton in his study came from a variety of sources: public water intakes, fisheries studies, early research efforts, and a few early studies of pollution in Lake Erie. The data set includes data from all three basins of Lake Erie, and from nearshore regions and open lake waters. Since gradients in concentration are known to exist from onshore to offshore, and from basin to basin, the data set contains sources of systematic difference other than the historical trends. However, the changes in most of the parameters over time are large compared to the magnitude of these spatial gradients, and the mix of data from different areas is reasonably random with respect to time. Thus, spatial factors may serve to obscure historical trends by increasing overall variance, but they probably do not bias the trends in an important way.

Beeton does not list the data shown in his figures, and much of it is from sources that are not readily available. I have read the data from Beeton's figures as precisely as possible; his data is reproduced graphically in figures 2 to 6. Since many of the data "points" in Beeton's graphs are actually averages, the distortion of the data due to reading the graphs is probably small compared to the distortion (loss of variance) introduced initially by the averaging process. It would be preferrable to begin with the raw data, however to date it has not been possible to reassemble the data set from the sources that Beeton used.

Also shown in the figures are the 1978 and 1979 means and standard deviations for data from the 15 HCWQL stations which are farthest from shore. These stations were chosen as most compatable with Beeton's sources.

In general, the HCWQL values are quite comparable to the values of Beeton (1965) from the late 1950's; indeed they seem to be lower than Beeton's 1950's data in the case of calcium and chloride. Within the limitations of the data, it appears that, for most parameters, the lake is not deteriorating at the rate which typified the first half of the century.

In order to test this conclusion, regression lines were fitted to Beeton's data for each parameter. The slopes of the regression

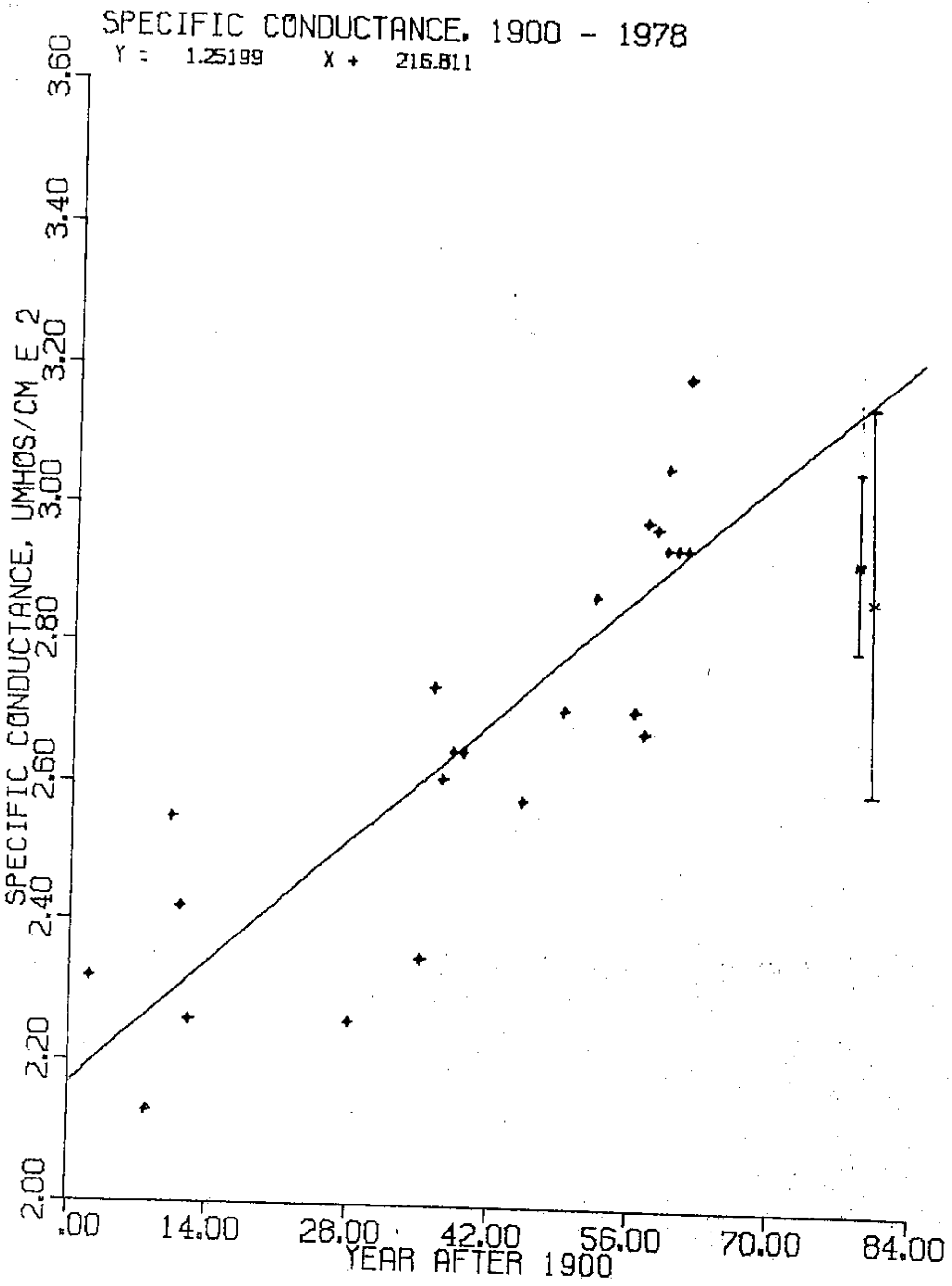


Figure 2-- Long term trend in specific conductance . The points are data from Beeton, and the line is the regression line through the data, extrapolated to 1980. The vertical lines indicate the 1978 and 1979 means plus and minus one standard deviation, as measured by the HCWQL at 15 stations at the outer edge of the Central Basin nearshore zone.

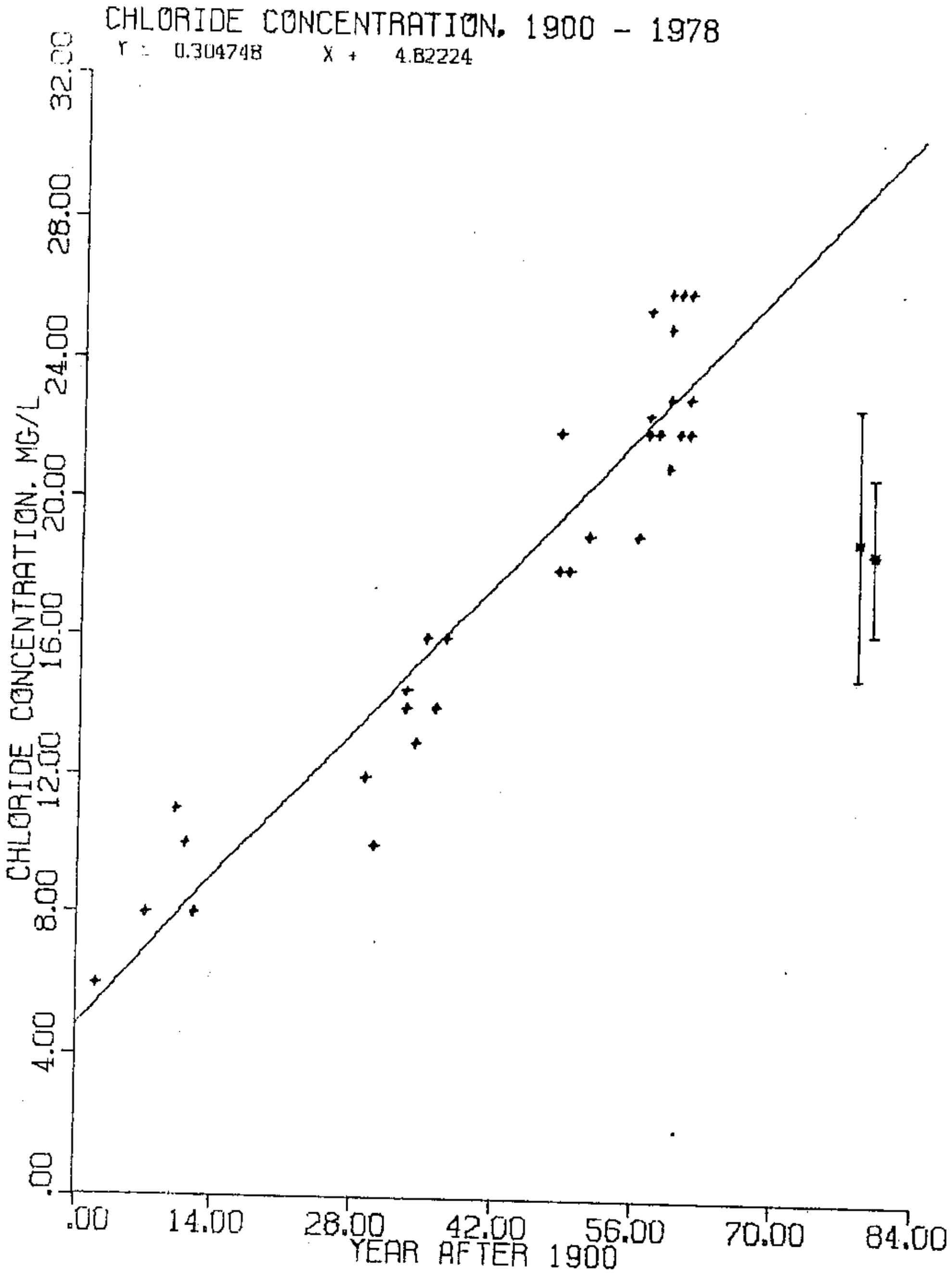


Figure 3-- Long term trend in chloride ion. The points are data from Beeton, and the line is the regression line through the data, extrapolated to 1980. The vertical lines indicate the 1978 and 1979 means plus and minus one standard deviation, as measured by the HCWQL at 15 stations at the outer edge of the Central Basin nearshore zone.

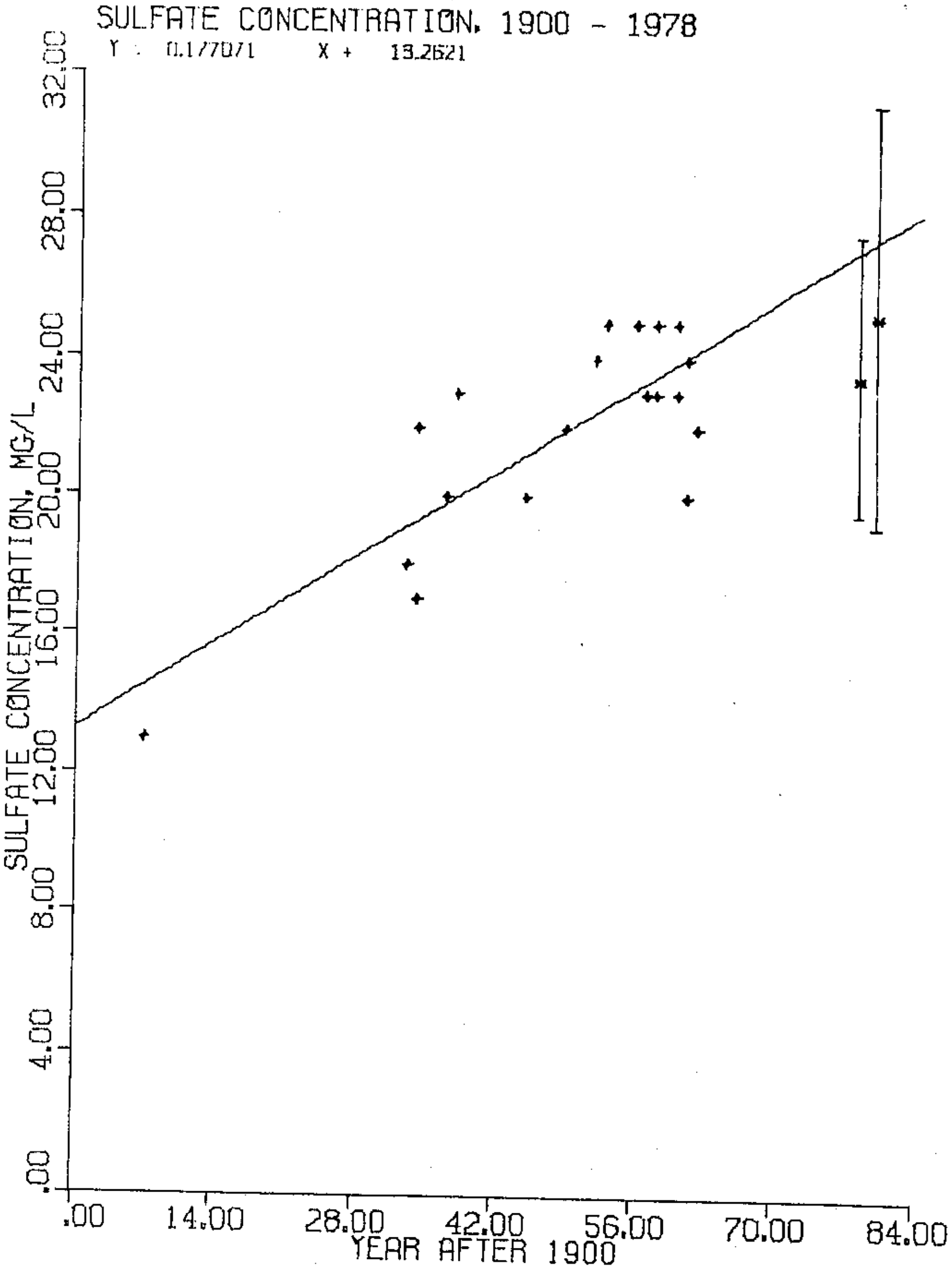


Figure 4-- Long term trend in sulfate ion. The points are data from Beeton, and the line is the regression line through the data, extrapolated to 1980. The vertical lines indicate the 1978 and 1979 means plus and minus one standard deviation, as measured by the HCWQL at 15 stations at the outer edge of the Central Basin nearshore zone.

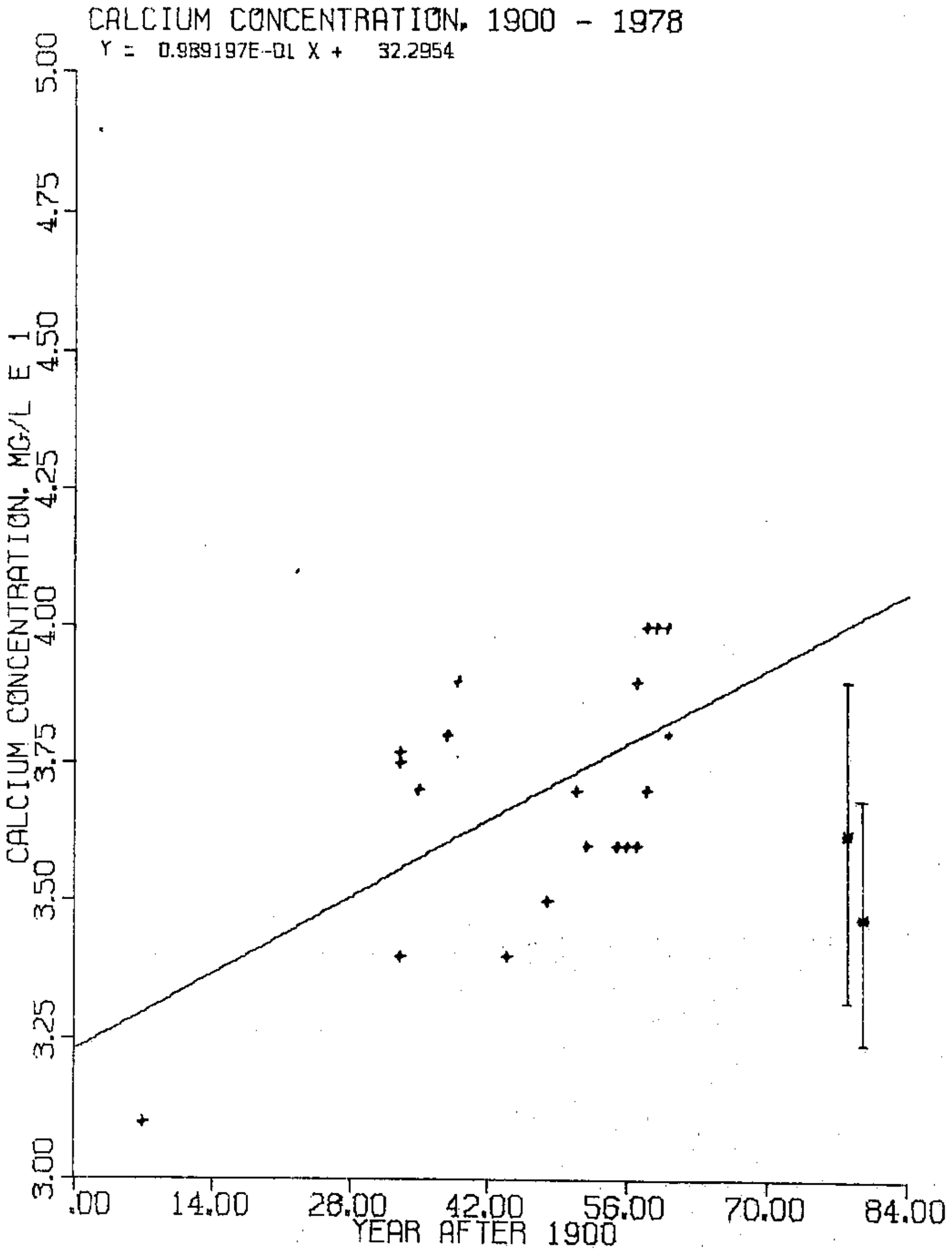


Figure 5-- Long term trend in calcium ion. The points are data from Beeton, and the line is the regression line through the data, extrapolated to 1980. The vertical lines indicate the 1978 and 1979 means plus and minus one standard deviation, as measured by the HCWQL at 15 stations at the outer edge of the Central Basin nearshore zone.

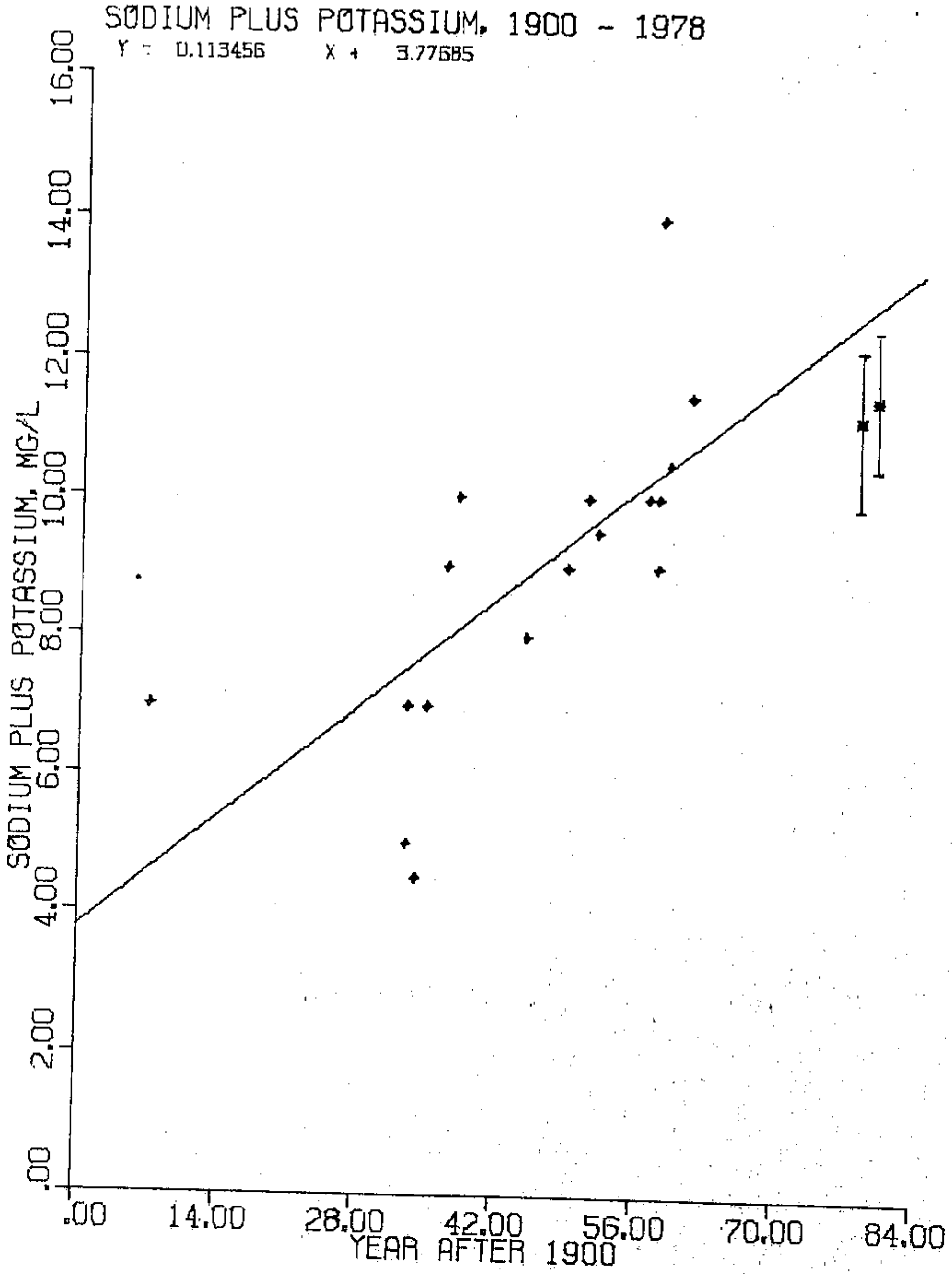


Figure 6-- Long term trend in sodium plus potassium. The points are data from Beeton, and the line is the regression line through the data, extrapolated to 1980. The vertical lines indicate the 1978 and 1979 means plus and minus one standard deviation, as measured by the HCWQL at 15 stations at the outer edge of the Central Basin nearshore zone.

Table 1. Comparison of historical data from Beeton with 1978 Heidelberg College Water Quality Laboratory data.

Parameter	Regression analysis of Beeton's data						Extrapolation to 1979		HCWQL data			Comparison	
	$\bar{Y}$	$\bar{X}$-1900	N	b	t^*_b	p^*_b	$\hat{Y}$	$S_{\hat{y}}$	N	$\bar{Y}$	S_y	t**	p**
Conductivity													
@ $\frac{\text{TDS}}{.62}$	267.04	40.20	24	1.25	7.97	<.001	315.67	6.81	1041	293.09	14.53	3.31	<.001
@ $\frac{\text{TDS}}{.65}$	254.72	40.20	24	1.20	7.97	<.001	300.91	6.45	1041	293.09	14.53	1.21	n.s.
Calcium	36.9	46.6	20	.098	3.09	<.01	40.09	1.13	146	35.79	3.06	3.73	<.001
Sodium plus Potassium	8.91	45	17	.115	3.87	<.01	12.84	1.11	145	11.48	2.19	1.21	n.s.
Chloride	17.6	42.4	30	.305	16.12	<.001	28.91	.767	1031	19.16	2.47	12.65	<.001
Sulfate	22	48.75	20	.177	5.58	<.001	27.25	1.08	1029	24.13	5.56	2.86	<.01

* calculated t value, and associated probability level, for t-test of the null hypothesis, H_0: b = 0. All regression slopes are significant, i.e., significantly different (greater) from 0.

** calculated t value, and associated probability level, for t-test of the null hypothesis, H_0: $Y_{1978} = \bar{Y}_{1978}$. All parameters show highly significant decreases except sodium plus potassium, which is lower than, but not significantly different from the trend of Beeton's data.

lines were tested for significance using a t-test. The regression equations were then used to extrapolate Beeton's data to best estimates for 1979. A standard error of the estimate was also calculated for a sample size comparable to the HCWQL data base for each parameter (Sokal and Rohlf, 1969), and these were compared with the HCWQL data by use of a modified two-tailed t-test, adjusted for unequal variances. Table 1 summarizes the results of this procedure. The statistical procedure reveals that all of Beeton's parameters increased significantly ($p < .01$) from 1900 to 1960. It also shows that, for all parameters except sodium plus potassium, values in 1978-1979 fall significantly below the values extrapolated from the historical data. In some cases, there probably has been an absolute decline in concentration since 1960. In others, especially sulfate and conductivity, there may only be a lessening in the rate of increase. I believe the changes indicated by these analyses are real, with one possible exception discussed below, and are consistant with the findings of other Lake Erie workers. However, had the analysis been done on the original data, the variance would have been greater, and thus the statistical significance of some of the results would have been reduced, perhaps even below the standard acceptable limit of $p < .05$.

The decrease in specific conductance, while highly significant statistically, is strongly dependent on the conversion factor used to convert Beeton's data, expressed as total dissolved solids, to specific conductance. The analysis was done using a conversion factor of

$$\text{Specific Conductance} = \text{TDS}/.62$$

recommended by A. Fraser (personal communication, 1978). The analysis was redone with a conversion factor of .65, which has been used elsewhere on Lake Erie. This second analysis yielded no significant deviation of HCWQL data from the trend of Beeton's data. Analysis of specific conductance and total dissolved solids data from the Division Water Intake of the City of Cleveland, measured between 1968 and 1975 by the U.S. EPA (Westlake) and the City of Cleveland lab at Whiskey Island, produced a ratio of 0.66 . This ratio is not significantly different from the value of 0.65 discussed above, but is higher ($p < .05$) than the value of 0.62 used initially. However, the rather poor correlation between the two parameters (r-square < .15) makes the data of questionable value in establishing the "true" ratio. There is, at present, no adequate way to be certain which ratio is correct. Until a definitive study of this relationship is made, the long term history of specific conductance cannot be assessed with certainty.

SHORT TERM HISTORICAL TRENDS NEAR CLEVELAND

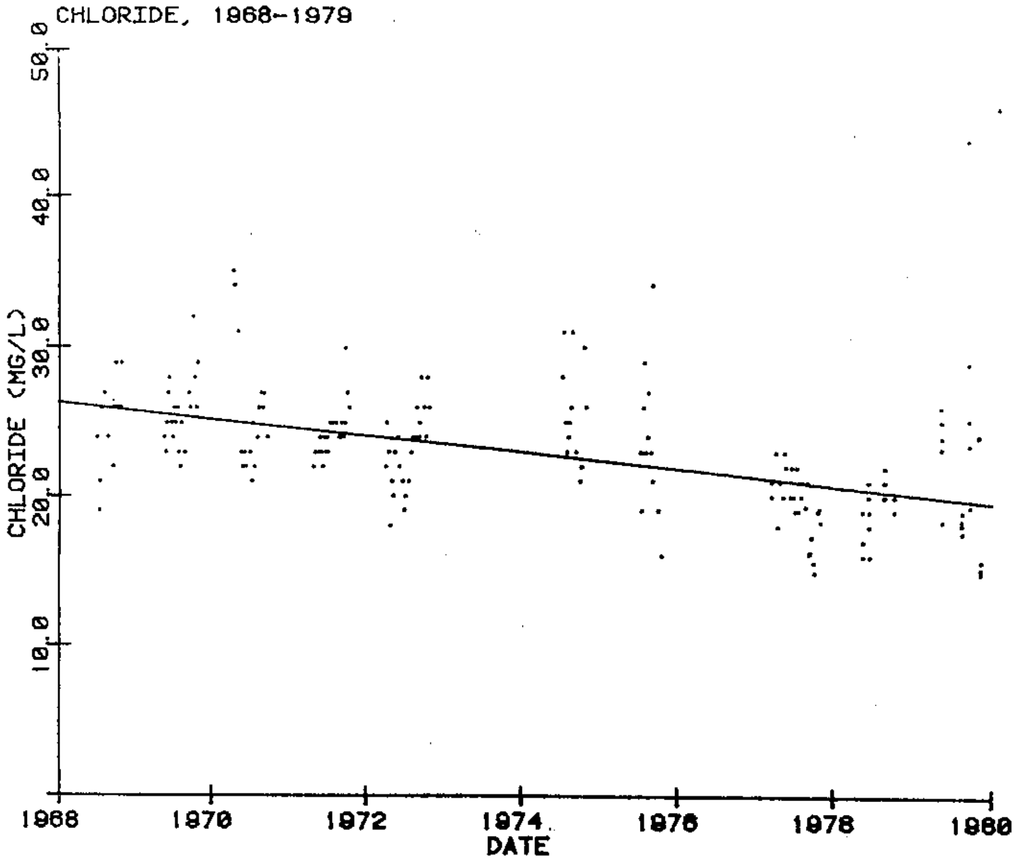
Because Beeton's data set ends about 1960, a more recent data set was sought to help evaluate the changes in the nearly 20 years falling between Beeton's data and the Lake Erie nearshore study. I also hoped to find data on parameters not covered in Beeton's study, particularly the nutrient parameters. The best source of information within the HCWQL study area is the records for the Division Water Intake of the City of Cleveland, which include data on alkalinity, specific conductance, pH, total phosphorus, soluble reactive phosphorus, ammonia, nitrate plus nitrite, chloride and sulfate, obtained between 1968 and 1973 by the U. S. EPA office now in Westlake, and between 1974 and 1977 by the Water Quality Lab of the City of Cleveland, formerly located at Whiskey Island. These data, plus our 1978 and 1979 data for the same location, are plotted in figures 7 through 16.

The data for each parameter were subjected to regression analysis to detect statistically significant linear trends. Inital analysis used all data in raw form, but subsequent analyses involved various modifications of the data, as described below.

Tests of the slopes of the regression lines for significant deviation from 0 (no trend) indicated no trend for alkalinity or

Figure 7-- Concentration of chloride ion as measured at the Division Water Intake of the City of Cleveland between 1968 and 1979. The regression line indicates the presence of a linear trend in the raw data significant at $p < .05$ or better.

Figure 8-- Concentration of sulfate ion as measured at the Division Water Intake of the City of Cleveland between 1968 and 1979. The regression line indicates the presence of a linear trend in the raw data significant at $p < .05$ or better.


CHLORIDE, 1968-1979
CHLORIDE (MG/L)
50.0
40.0
30.0
20.0
10.0
1968
1970
1972
1974
1976
1978
1980
DATE

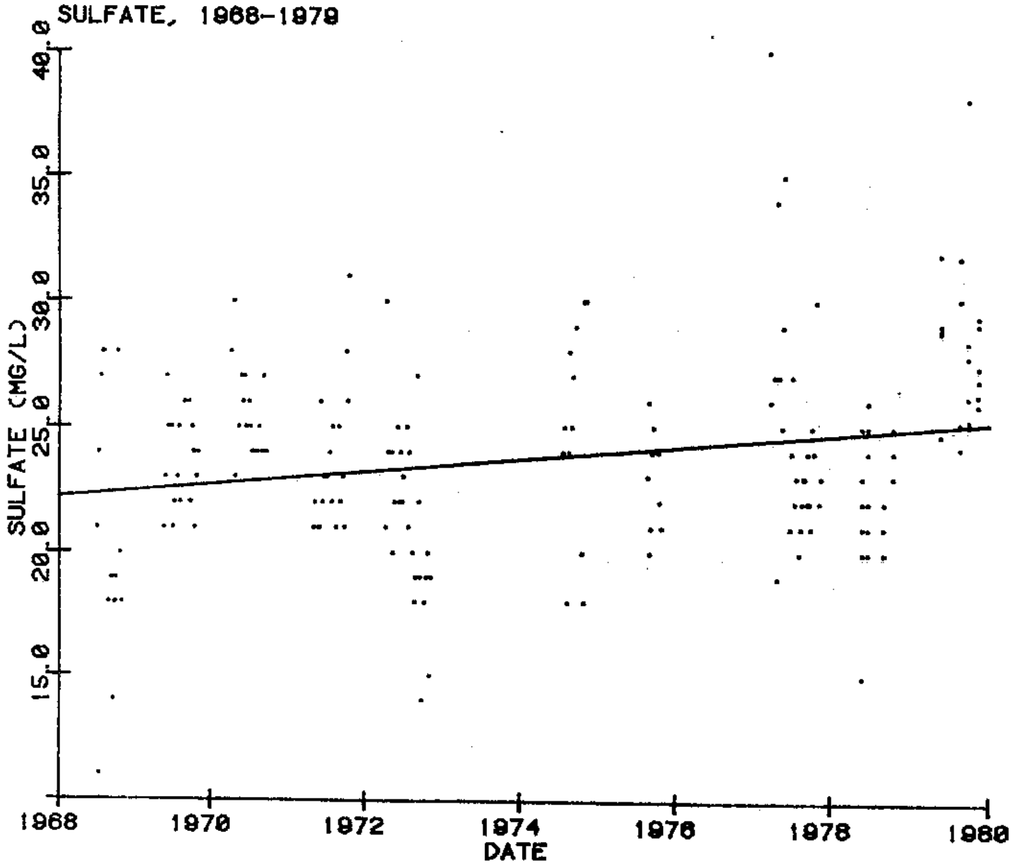
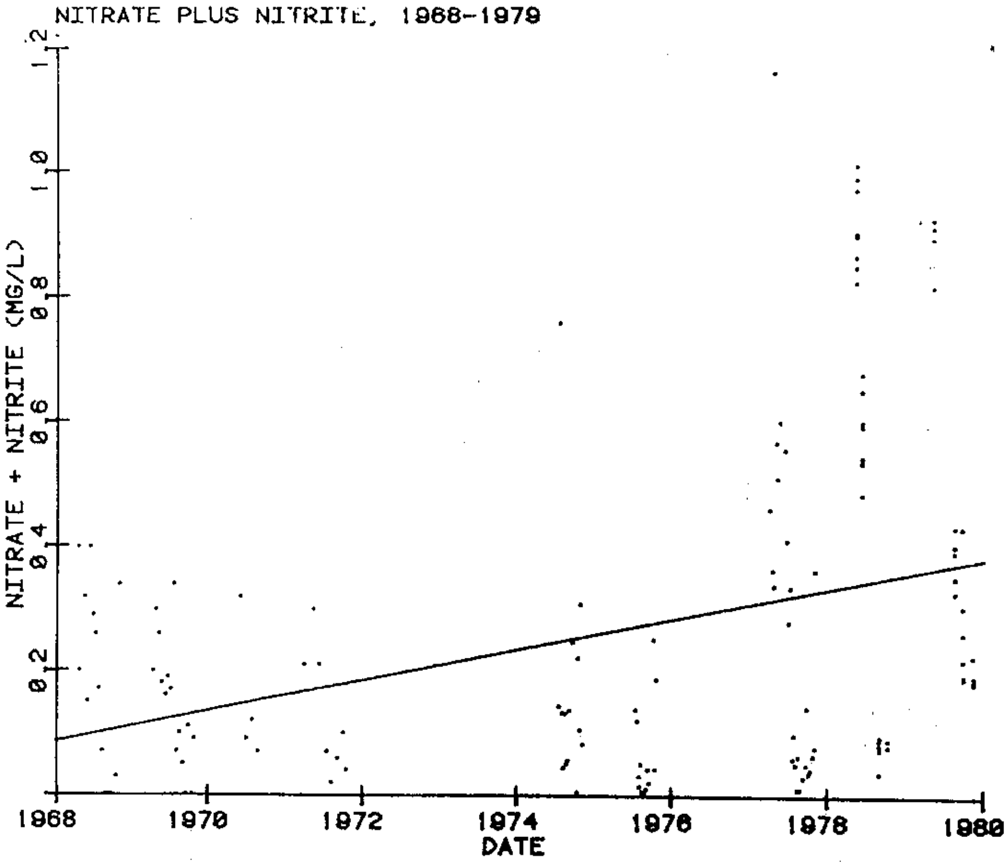

SULFATE, 1968-1979
SULFATE (MG/L)
40.0
35.0
30.0
25.0
20.0
15.0
1968
1970
1972
1974
1976
1978
1980
DATE

Figure 9-- Concentration of nitrate plus nitrite as measured at the Division Water Intake of the City of Cleveland between 1968 and 1979. The regression line indicates the presence of a linear trend in the raw data significant at $p < .05$ or better.

Figure 10-- Concentration of total phosphorus as measured at the Division Water Intake of the City of Cleveland between 1968 and 1979. The lack of a regression line indicates that the raw data failed to show a linear trend significant at $p < .05$.


NITRATE PLUS NITRITE, 1968-1979
NITRATE + NITRITE (MG/L)
1.2
1.0
0.8
0.6
0.4
0.2
1968
1970
1972
1974
1976
1978
1980
DATE

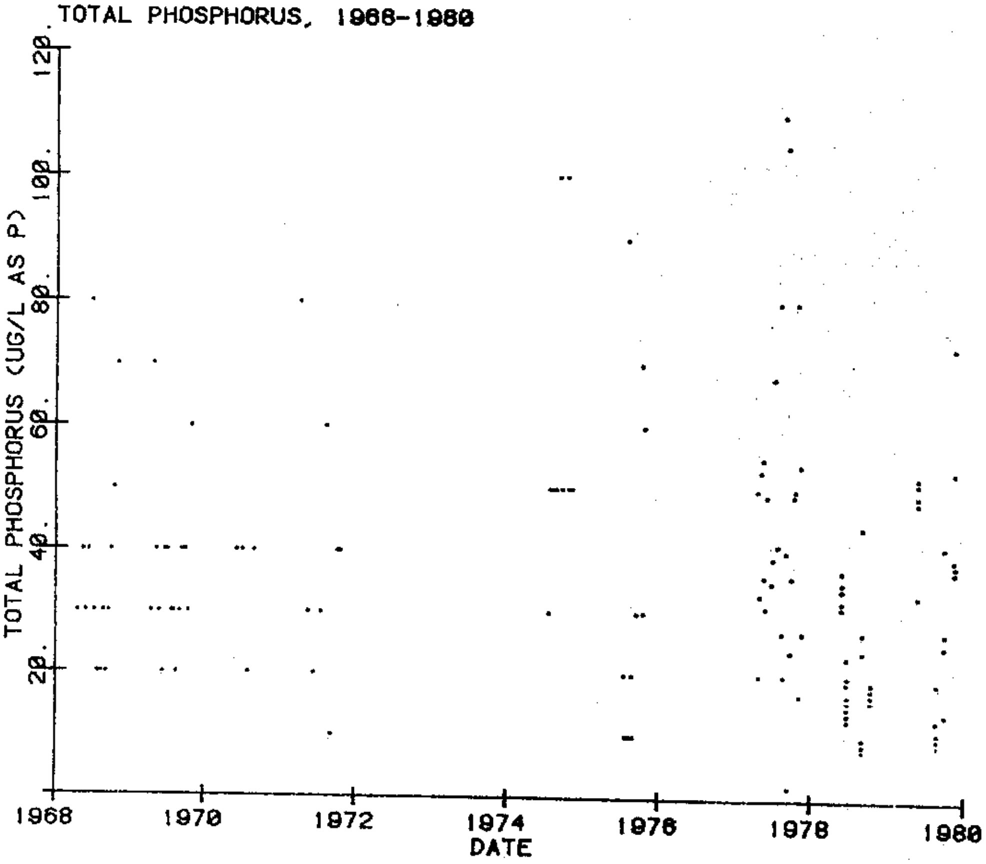
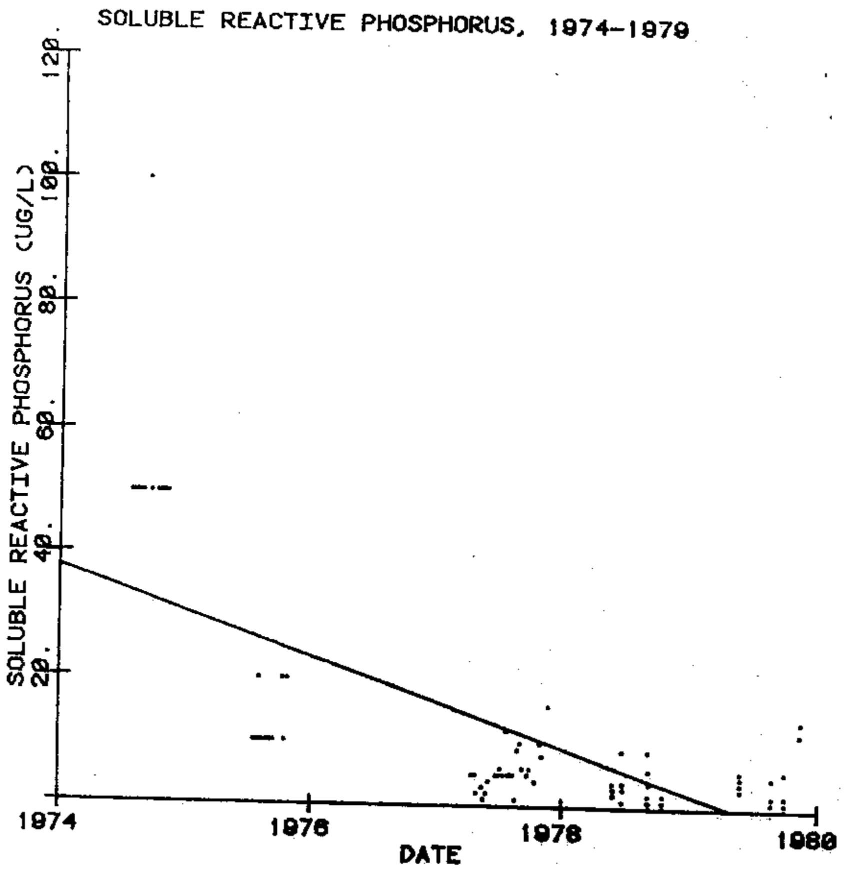
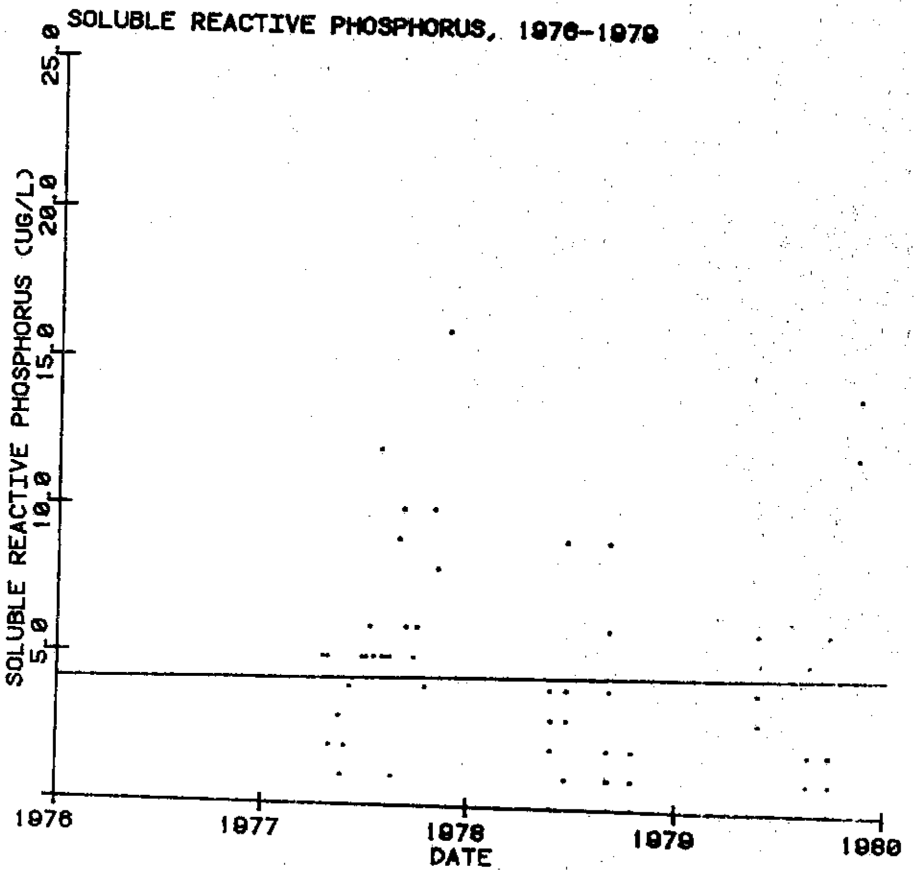

TOTAL PHOSPHORUS, 1968-1980
TOTAL PHOSPHORUS (UG/L AS P)
120.
100.
80.
60.
40.
20.
1968
1970
1972
1974
1976
1978
1980
DATE

Figure 11-- Concentration of soluble reactive phosphorus as measured at the Division Water Intake of the City of Cleveland between 1974 and 1979. The regression line indicates the presence of a linear trend in the raw data significant at $p < .05$ or better.

Figure 12-- Concentration of soluble reactive phosphorus as measured at the Division Water Intake of the City of Cleveland between 1976 and 1979. Same data as figure 11, but with suspect data from 1974 eliminated. The regression line is not significant at $p < .05$.



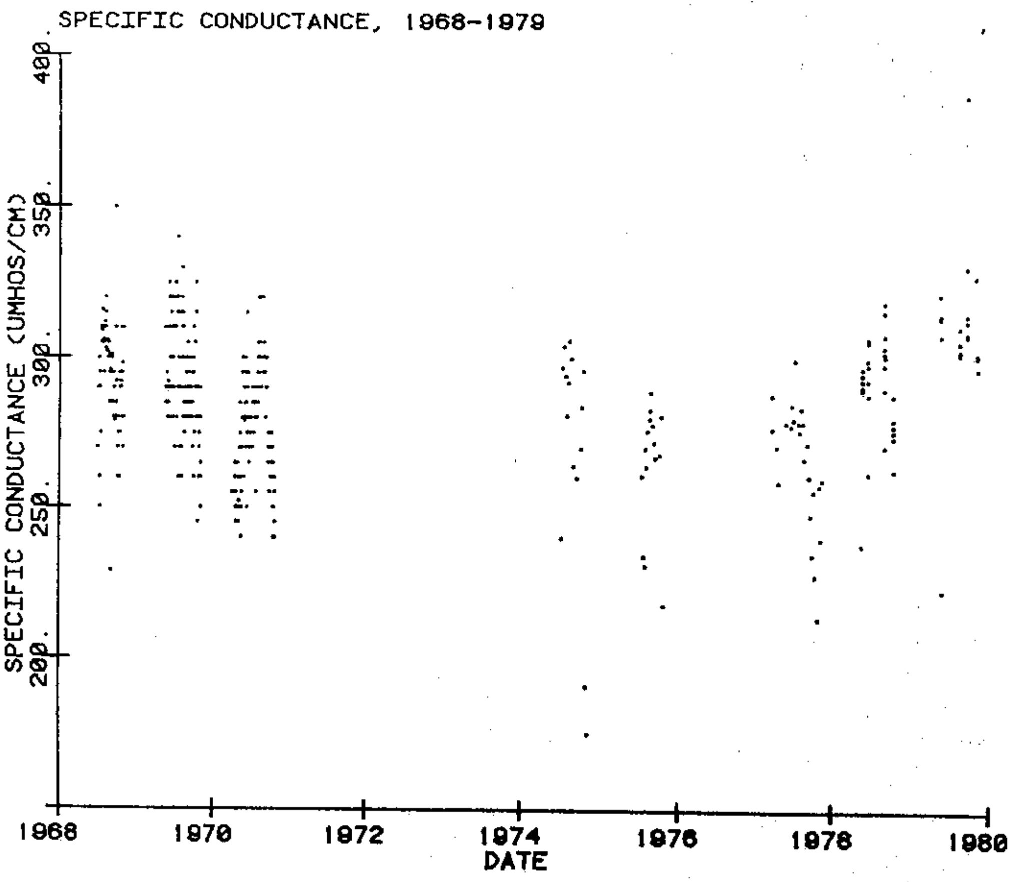
SOLUBLE REACTIVE PHOSPHORUS, 1974-1979



SOLUBLE REACTIVE PHOSPHORUS, 1976-1979

Figure 13-- Specific conductance, as measured at the Division Water Intake of the City of Cleveland between 1968 and 1979. The lack of a regression line indicates that the raw data failed to show a linear trend significant at $p < .05$.

Figure 14-- Alkalinity, as measured at the Division Water Intake of the City of Cleveland between 1968 and 1979. The lack of a regression line indicates that the raw data failed to show a linear trend significant at $p < .05$.


SPECIFIC CONDUCTANCE, 1968-1979
SPECIFIC CONDUCTANCE (UMHOS/CM)
400.
350.
300.
250.
200.
1968
1970
1972
1974
1976
1978
1980
DATE

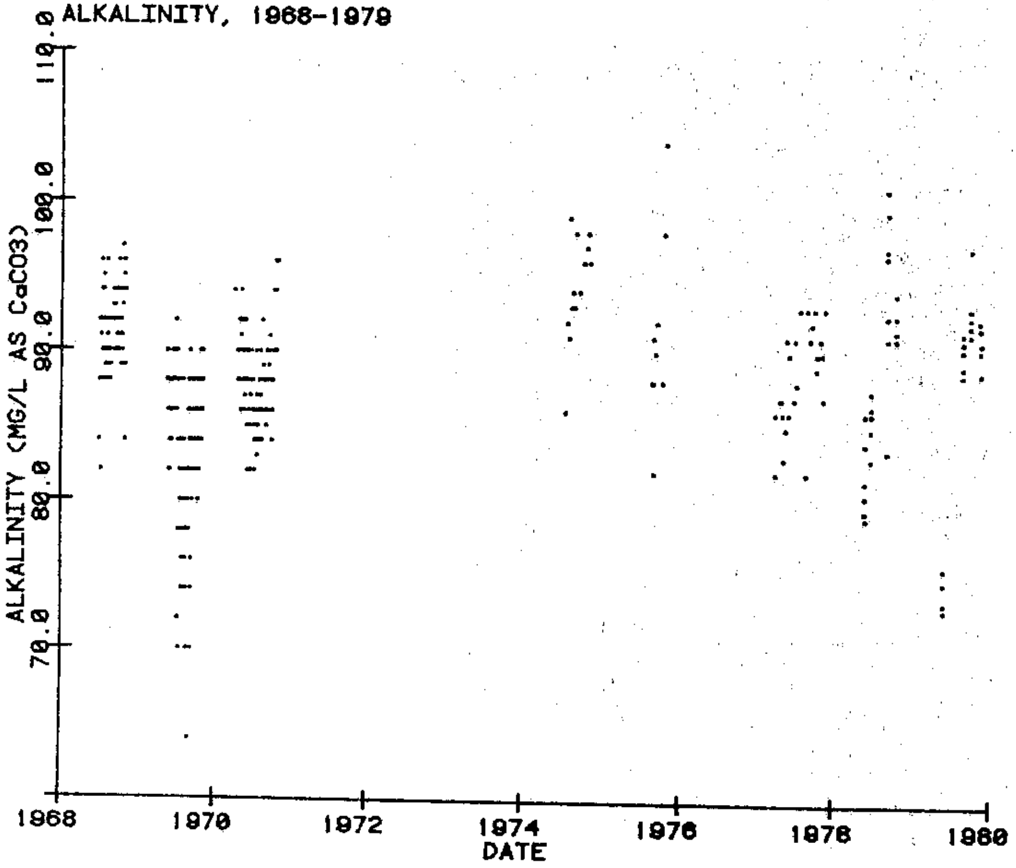
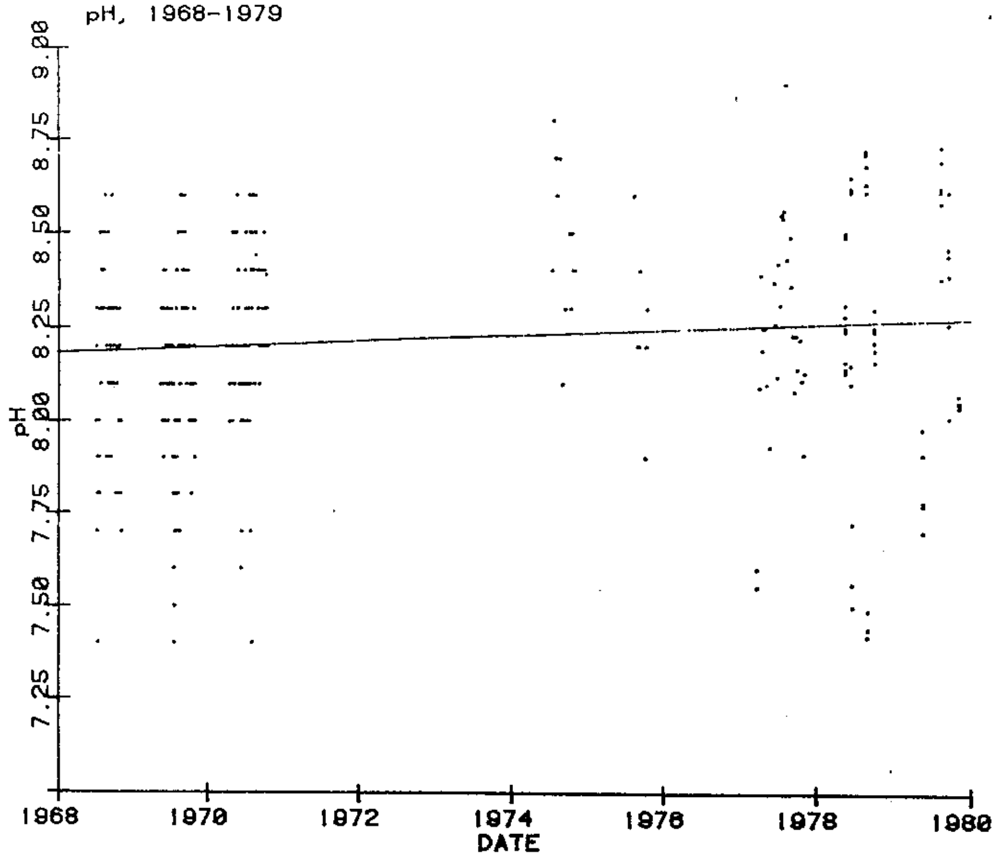
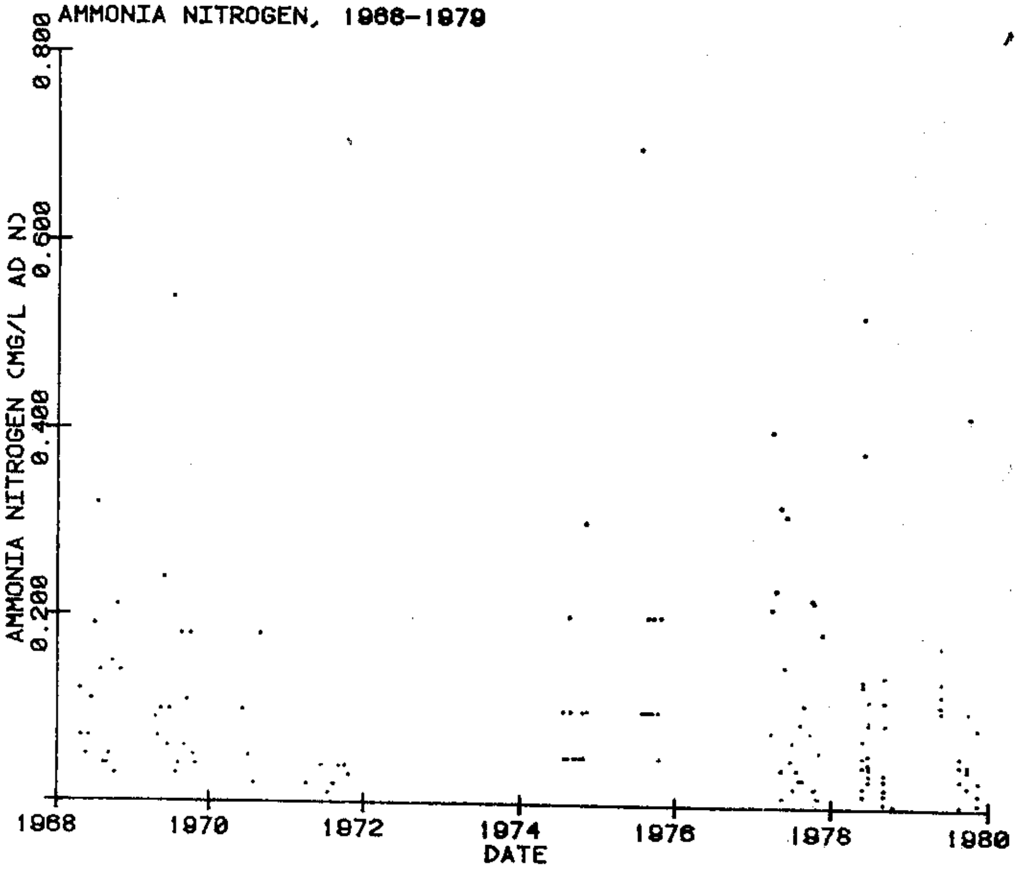

ALKALINITY, 1968-1979
ALKALINITY (MG/L AS CaCO3)
110.0
100.0
90.0
80.0
70.0
1968
1970
1972
1974
1976
1978
1980
DATE

Figure 15-- pH, as measured at the Division Water Intake of the City of Cleveland between 1968 and 1979. The regression line indicates the presence of a linear trend in the raw data significant at $p < .05$ or better.

Figure 16-- Concentration of ammonia nitrogen as measured at the Division Water Intake of the City of Cleveland between 1968 and 1979. The lack of a regression line indicates that the raw data failed to show a linear trend significant at $p < .05$.


pH, 1968-1979
9.00
8.75
8.50
8.25
8.00
7.75
7.50
7.25
pH
1968
1970
1972
1974
1976
1978
1980
DATE


AMMONIA NITROGEN, 1968-1979
0.800
0.600
0.400
0.200
AMMONIA NITROGEN (MG/L AD N)
1968
1970
1972
1974
1976
1978
1980
DATE

specific conductance, but a significant increase in nitrate plus nitrite ($p < .001$) and sulfate ($p < .001$), and a highly significant decrease in chloride ($p \ll .001$) in the last decade at this station. Total phosphorus shows no significant change, but soluble reactive phosphorus decreased significantly, even though the period of record is shorter (1974-1979). The SRP data contains a number of suspiciously high points early in the record, however (see the discussion of SRP in the general discussion section below), and when these are removed, the trend disappears.

Filtering the data to remove seasonal patterns

Many if not all of the parameters studied for historical trends also undergo seasonal fluctuations in concentration as a result of seasonal fluctuations in their supply to the water column and, in some cases, as a result of biological activity. The seasonal changes are most pronounced in the nutrients, the extreme case being nitrate plus nitrite, which declines sharply in late summer, often to values less than 20% of those found in early spring.

These seasonal fluctuations tend to mask longer term historical trends, because they increase the overall variance of the data set. Typically, their effect is to decrease the achieved statistical significance in a test of the slope of the regression line. Since seasonal fluctuations are a different phenomenon than the one I am trying to examine, it would be helpful if these fluctuations could be removed from the data. This can be done in the following way.

Assume that the seasonal effect is constant from year to year, and is not linked to the long term historical pattern. If this is so, the data can be fit by a function of the form:

$$y = mx + s(x) + b$$

where s(x) is a periodic function (perhaps a very irregular one) of period one year.

Under this assumption then, the procedure is as follows. A standard regression is performed on the data, in effect ignoring the periodic component. Since it is assumed to repeat exactly each year, it does not change the regression equation except to increase the components of variance associated with it.

The regression equation is then used to calculate the values of y (the concentration) predicted for the given values of x (time), and the predicted y values are subtracted from the actual values. In statistical terms, the residuals are calculated.

The residuals are then grouped together by some sub-interval of the period of the postulated periodic function. The grouping interval should be small enough to capture the essence of the periodic changes, yet large enough to contain enough data to be statistically useful. Some compromise is often necessary. In this study, I grouped data by month, which gave at least 30 data points in a month, with very few exceptions. This grouping gathered January data for all years in one group, February data for all years in another group, etc.

The average of the residuals in each of these groups is calculated. If there is no seasonal effect in a data set, the residuals reflect only random error, and these averages should all be very close to zero. Thus any non-zero group average may be taken as an estimate of the seasonal effect for that interval of time.

These estimates of the seasonal effect are then subtracted from the raw data, and the regression is recalculated. The result should be a regression equation which is very similar to the one calculated from the raw data, but it should have a higher associated r-square, or, alternatively, a t-test comparing the slope with 0 should achieve a higher level of significance.

Results of the procedure and discussion

Application of the above procedure to the Division Water Intake data set produced the results summarized in Table 2. Examination of Table 2 shows that most parameters showed improvement in the resolution

Table 2. R-square and t values for regression analyses of data from the Division Water Intake, before and after filtering the data to remove seasonal fluctuations. n.s. indicates slope not significantly different from 0 at the .05 level of significance. A negative t value indicates a decrease in that parameter's concentration through time, a positive t value indicates an increase.

Parameter	Before			After		
	r^2	t	significance	r^2	t	significance
Total Phosphorus	.021	-1.8	n.s.	.028	-2.07	<.05
Soluble Reactive P	.354	-7.66	<.001	.454	-9.43	<.001
Nitrate + Nitrite	.122	4.53	<.001	.319	8.32	<.001
Ammonia Nitrogen	.010	-1.21	n.s.	.010	-1.25	n.s.
Sol. Reactive Silica	.029	1.61	n.s.	.015	1.17	n.s.
Alkalinity	.005	1.52	n.s.	.004	1.30	n.s.
pH	.012	2.29	<.05	.030	3.72	<.001
Specific Conductance	.001	-0.65	n.s.	.001	-0.66	n.s.
Sulfate	.053	3.43	<.001	.050	3.35	<.001
Chloride	.264	-8.83	<.001	.298	-9.63	<.001

of their historical trends, as measured by increases in the t value. The greatest increases were among the nutrients, where seasonality is typically most pronounced. In a few cases, the t value decreased. This is to be expected in cases where there is little seasonality, due to the error component of the estimate of the seasonal effect. There was in most cases no change in the achieved level of significance of the slope of the regression line. Most of the parameters for which this was true either showed no significant change, and the improvement brought about by removing seasonality was not sufficient to give significance at $p < .05$, or they already showed highly significant historical change ($p < .001$) and the computer program did not give significance levels less than .001. In one very significant case, seasonal filtering produced a significant trend. Total phosphorus did not show a significant trend before filtering, but showed a decrease significant at $p < .05$ after filtering.

Because the procedure assumes that the seasonal patterns were constant during the period of record, monthly averages were plotted for each year to verify the validity of the assumption. While the averages fluctuated considerably from year to year, no systematic change was seen in any of the parameters.

In general, the seasonal filtering process tended to improve resolution of historical changes, but not sufficiently to have a great impact on the conclusions of this study. It appears that other factors than predictable seasonal changes dominate the variance of this data set. These factors may include laboratory accuracy and precision, and the effects of fluctuating currents, which may alternately bring by the sampling site waters of more nearshore or more offshore character.

General Discussion of Short Term Trends

The data set used in this part of the study is less than ideal because it was generated by three different laboratories, using different working ranges and in some cases different analytical techniques. In addition, the U.S. EPA analyses were of samples from the water as it entered the purification plant, while the other samples were of lake water at the site of the water intake. Thus, the EPA samples were of bottom water modified by passage through the intake pipe. By comparison, the City of Cleveland samples were mostly surface water, and the HCWQL samples were both surface and bottom waters. Because of the composite nature of the data set, various techniques were used to try to evaluate possible biases or inadequacies in the data. These included examination of the data plots themselves for sudden changes in typical values between years, especially between years analyzed by different labs. The plots also indicated the degree of resolution in the data for a given year. For some parameters, studies of correlation between parameters on a lab by lab basis also gave some indication of the accuracy with which the parameters were measured, both through the correlation coefficient and through the linear relationship between the parameters as compared to values from the literature. The results of this scrutiny of the data and the results of the earlier analysis are given below on a parameter-by-parameter basis, beginning with parameters which showed the strongest historical trends.

Chloride

The chloride data shows the most visually obvious trend, with the data for most years centered reasonably well on the trend line. Even in this data set, however, the typical yearly spread of values is comparable in magnitude to the total change in the average value over the eleven years of the data. The seasonal filtering procedure improves the trend slightly, but, as is true of all parameters, much of the yearly scatter is not attributable to a uniform yearly concentration cycle. Nothing in the data suggests that the indicated trend is not correct, and it agrees at least qualitatively with the findings of other Lake Erie workers (Weiler and Heathcote, 1979).

Sulfate

Sulfate has a highly significant increasing trend over the period of record. While the pattern is not as uniform as is true for chloride, nothing in the data suggests that the trend is not real. Seasonal filtering decreases the significance of the slope slightly, perhaps because the variance-building effects of some of the more extreme points are spread over more of the data set by the procedure. Analysis of 1978-1979 data for the Cleveland region as a whole (Richards, 1981) shows that sulfate is reasonably constant for these years during the months represented by most of the data (April-October). If this is true of sulfate concentrations over the entire time span of the data, the seasonal filtering approach would be expected to have a small, random effect on the trend analysis statistics.

One might expect a reasonably strong correlation between chloride and sulfate, and indeed the HCWQL data show a correlation significant at $p < .001$, with an r-square of .22. However, the EPA data correlates only at $p < .05$, with an r-square of .033, and the Cleveland data fails to show a correlation, with r-square of .002. There are many factors which could have contributed to these differences, including the possibility that one or both of the parameters may not have been measured with adequate precision by the earlier labs. If one parameter alone is inaccurate, it is probably sulfate, given the form of the data itself, and the reliability of the analytical methodologies.

pH

pH has a statistically significant but not visibly obvious increase over the period of record. Seasonal filtering improves the significance of the slope from $p < .05$ to $p < .001$, and increases the explained variance (r-squared) by a factor of 2.5, but only from .012 to .030. The data seems to show a peak in the years 1975-76, and to be uniformly lower in 1968-70. These latter data were bottom water samples, while the former were mostly surface samples. In the HCWQL data for the Central Basin nearshore zone as a whole (1978-79), bottom pH averaged 0.2 units lower than surface pH. In the data for this station bottom pH averaged .47 units lower than surface pH. Thus the net apparent change over 11 years (about 0.1 pH units) is less than the local or regional difference between annual average surface and bottom pH. It is entirely probable that the apparent historical trend is an artifact of the different depths sampled by the three labs.

Nitrate plus Nitrite

Nitrate plus nitrite shows a highly significant increase in concentration over the period of record. The significance of the trend is increased substantially by seasonal filtering, as indicated by an increase in the explained variance from .122 to .319, and a corresponding increase in the t value for the significance of the slope of the trend line. This result was expected since nitrate plus nitrite is known to undergo substantial seasonal changes in concentration.

While the trend is highly significant statistically, the data do not show a uniform change over time. Most of the increased values are present in the data for the last three years , and indeed the earlier data visually suggest a decrease in concentration. While this may accurately reflect what has happened in the lake, it could also stem from laboratory-related causes. The increasing trend probably reflects what is happening in the lake, especially since other workers have found similar patterns (Weiler and Heathcote, 1979), but the *rate* of increase cannot be established with any degree of confidence from this data. It would be desireable to study the history of nitrate plus nitrite using some other long-term data sets which will have continuity into the future, and will be the product of one lab and one set of methodologies only.

Total Phosphorus

The data for total phosphorus showed no significant change in concentration during the period of record, but displayed a downward trend which did not fall far short of significance ($p = .08$). After seasonal filtering, the slope of the trend line was significant ($p <.05$), indicating a decrease in total phosphorus through time. There are problems with the data, however. The detection limit in 1974 was 100 ug/l, a value far higher than typical lake values. Values below detection limit were entered as half the detection limit, which is still higher than typical values. The analysis was redone with all 1974 data removed, and the result was a reduction of the t value from 1.8 to 1.65, which corresponds to an achieved significance level of about 0.1.

The detection limit from 1968 to 1972 was 20 ug/l. Since many of the values in the later years fell below 20 ug/l, there is the possibility that the "trend" may result only from the failure to adequately measure low-level total phosphorus early in the period of record. To check this, all data equal to 20 ug/l for 1968 to 1972 were recalculated to fit the same frequency distribution as the data less than or equal to 20 ug/l in 1977 to 1979. When the modified data was re-analyzed (with the 1974 data omitted), the result was a further decrease in t from 1.65 to 1.42. While the assumption that the data below 20 ug/l should follow the same distribution may have been too severe, the exercise does show that those few data points have an important effect on the trend analysis. Similarly, if all data less than 20 ug/l are recoded to 20 ug/l, to simulate a constant detection limit throughout the period of record, the t value is decreased from 1.65 to 1.14.

Since the trend for total phosphorus barely achieves statistical significance after seasonal filtering, and given the problems of high detection limits early in the study, it seems appropriate to suggest that total phosphorus is probably declining, but that more years of better data will be required to put that conclusion on sound statistical footing.

The correlation between total phosphorus and nitrate plus nitrite was examined for the data of each laboratory. The HCWQL data showed a significant correlation ($p < .01$, r-square .123), but the relationship was not significant in the data from the EPA (r-square .037) or the Cleveland lab (r-square .000). If these results reflect a problem with analyses, the problem likely lies with total phosphorus, because it was being measured much closer to detection limit, and part of the time with a resolution that would give only three or less values in the typical range of lake water in the area (i.e. values would have been 20, 30, or 40 ug/l).

Soluble Reactive Phosphorus

Soluble reactive phosphorus raw data shows a significant decrease that improves substantially with filtering. The data spans only five years, with no data in the second year of that interval. The first year's data was analized with a detection limit of 100 ug/l, roughly 6 times the typical values in the area. This data consists of one value of 100 ug/l, and the rest at detection limit, which were entered as 50 ug/l. In all liklihood, all or most of these values substantially overestimate the actual concentrations. When this data is eliminated, the "trend" disappears. The remaining three years of data show a slight non-significant increase. More time will be required to show the decrease in soluble reactive phosphorus which is expected to accompany improvements in sewage treatment in the Cleveland area.

Alkalinity

Alkalinity does not show a significant trend over the period of record. However, the data suggest a peak about 1974, with values generally lower before and after. The last five years of data (1974-1979) are suggestive of declining alkalinity values, and within this time the decline is significant (p <.001, r-squared .09).

The alkalinity values are subject to the same sort of sampling bias as was seen for pH: bottom water in 1978-79 averaged nearly 3 mg/l higher than surface water. However, this bias would tend to reduce the peakedness of the alkalinity data, not enhance or create it. The form of the data may be related to lake levels, or may be an artifact of lab bias.

Specific Conductance

Specific conductance shows a non-linear pattern which is the reverse of that for alkalinity, with a minimum in 1975 and higher values before and after. This pattern may also be related to lake level changes, or may be a result of lab bias. Specific conductance should be well correlated with chloride. The correlation coefficients based on the data from each lab are as follows: HCWQL r-squared .407 (t=6.09, sample size n=56), Cleveland lab r-squared .035 (t=1.34, n=52), EPA r-squared .002 (t=0.35, n=51). The low correlations in the data of these last two labs is disturbing, and raises the possibility that the first nine years of data for specific conductance may not be valid.

CONCLUSIONS

1. Analysis of the long-term historical data of Beeton documents statistically significant increases in all parameters. By comparison, data from the Cleveland area in 1978-1979 is comparable to or lower than Beeton's data for 1960, indicating a decline in the rate of increase in concentration for most parameters, and an actual decline in concentration for some. The only exception is specific conductance, for which the analysis is made uncertain because the proper conversion factor between total dissolved solids and specific conductance is not known with sufficient accuracy.

2. Analysis of historical data from the Cleveland Division Water Intake suggests that the concentration of chloride has declined over the period 1968-1979, and that sulfate has increased in the same time. Both trends are reasonably linear, and are highly significant statistically.

3. Other parameters in the short-term data set show no significant linear trends (e.g. specific conductance), show trends that are significant but data that is decidedly non-linear (e.g. nitrate plus nitrite), or show trends that may be at least partly a result of laboratory bias (phosphorus forms) or sampling of different water masses (surface vs. bottom, e.g. pH) rather than historical change. These parameters require further study to establish adequate historical trend information.

4. A technique was presented which can be used to filter out the

effects of seasonal change in concentration of a parameter. The technique can effect substantial improvement in the resolution of long-term trends in parameters that are strongly seasonal in nature (e.g. nitrate plus nitrite). However, much of the annual variation displayed by this data set was not systematic, and thus was not removed by the filtering process.

5. Examination of correlations between parameters which may be expected to show correlated concentrations can help to indicate the quality of the data. This approach indicated that this data set contains some data of suspect quality, and helped in the interpretation of the historical patterns shown by some parameters.

6. Decade-long historical trends are often comparable in magnitude to the annual scatter in the data, or even to the difference between surface and bottom water concentrations. Biases between labs or between years within a lab, or differences in values obtained with different analytical methodoligies, may be sufficient to mask subtle historical trends, or to create "trends" which reflect the history of analytical methodology and bias rather than the history of the body of water under study. The attempt to recognise subtle historical trends, which may none-the-less be of great interest to the public, requires data of the best quality. Wherever possible, the data base should be the work of one lab using one set of methodologies and a carefully designed quality control program to guarantee the comparability of data from day to day, month to month, and year to year. At the very least, frequent participation in "round-robin" exercises (such as that carried out by the International Joint Commission) is necessary by all labs

contributing to a historical data base. The results of these round-robin studies must be made part of the laboratory quality control program, and used to adjust biases, if compatable data are to be generated. These results should also be considered in the historical analysis, since they may suggest biases that were not corrected, and that might not otherwise be apparent.

Finally, the researcher who conducts the historical analysis should seek as much background information as possible related to quality control, and should assume that biases will exist in many data sets. In the matter of historical analysis, scientific scepticism must extend to the data itself.

LITERATURE CITED

Beeton, A. M. 1961. Environmental Changes in Lake Erie. Trans. Amer. Fish. Soc. 90(2): 153-159.

Beeton, A. M., and D. C. Chandler. 1963. The Saint Lawrence Great Lakes. p. 535-558. In D. G. Frey, ed., Limnology in North America. Univ. Wis. Press, Madison.

Beeton, A. M. 1965. Eutrophication of the St. Lawrence Great Lakes. Limnol. and Oceanogr. 10(2): 240-254.

Herdendorf, C.E. 1978. Lake Erie nearshore surveillance station plan for the United States. CLEAR Tech. Report No. 77. Ohio State University. 50 pgs.

Richards, R.P. 1981. Chemical Limnology in the Nearshore Zone of Lake Erie between Vermilion, Ohio and Ashtabula, Ohio, 1978-1979: Data Summary and Preliminary Interpretations. Final Report to U.S. EPA Region V, Robert Bowden, Project Officer. Heidelberg College. 86 p.

Sokal, R. R. and F. J. Rohlf. 1969. Biometry. W. H. Freeman, San Francisco. 776 p.

Weiler, Ronald R. and Isobel W. Heathcote. 1979. Nanticoke Water Chemistry, 1969 to 1978. Ontario Ministry of the Environment, Toronto. 26 pgs.