THE

ONTARIO WATER RESOURCES

COMMISSION

WATER QUALITY PREDICTION EQUATIONS FOR A RIVER INPUT INTO LAKE ERIE

OCTOBER, 1968

WATER QUALITY PREDICTION EQUATIONS FOR A RIVER INPUT INTO LAKE ERIE

October, 1968

M. D. Palmer,
Water Quality Surveys Branch,
Division of Sanitary Engineering.

WQTS-2

WATER QUALITY PREDICTION EQUATIONS
FOR A RIVER INPUT INTO
LAKE ERIE

ABSTRACT

Prediction equations were developed from the 1967 water quality data at 16 lake stations in the Grand River outlet area of Lake Erie. These equations were determined by multiple regression techniques (with transforms) on a trial-and-error basis using the following independent variables:

1. wind force and direction
2. loadings of the Grand River at the outlet
3. geographical location of the lake station (up to 5 miles from the outlet)
4. depth and temperature gradient with depth at the lake station.

Individual equations were developed for each of the following parameters:

1. maximum turbidity
2. average turbidity
3. maximum conductivity
4. maximum total phosphate
5. maximum soluble phosphate maximum dissolved oxygen
6. minimum dissolved oxygen
7. maximum nitrate

As a check on the equations, they were tested on the 1966 data and found to predict the measured water quality parameters.

WATER QUALITY PREDICTION EQUATIONS
FOR A RIVER INPUT INTO LAKE ERIE

INTRODUCTION

Periodic water quality surveys have been carried out for several years in lake stations adjacent to major tributaries. Since this data represents conditions at a specific time only, the need arises to establish a relationship between the major factors influencing the water quality such as weather, material input conditions, and the measured water quality. One approach would be to propose the variables which are operative in a cause and effect relationship, then assign the variables numerical values. Once this is done, it is possible to carry out history matching, in this case multiple regression, on one year's data to develop the best prediction equation. The prediction equation can then be tested on another year's data.

These techniques were applied to 1966 and 1967 data collected on the Grand River and at the adjacent Lake Erie sampling stations in the Port Maitland area.

The variables selected as being operative were:

1. Loadings - represented by the concentration of a substance and the flow in the Grand River at the outlet.

2. Dispersion characteristics:

   a) time factor - represented by distance from the outlet.

   b) current - represented by the wind direction and magnitude in relation to the point considered and the outlet of the Grand River.

   c) dilution volume - represented by the water depth

d) bottom configuration - represented by a depth and distance term.

E) temperature stratification - represented by a vertical temperature gradient at the station.

Using each of these variables as measured at the time of sampling and the actual measured water quality values at a point for many different stations and dates, it was possible to develop cause and effect relationships. In the case of this study, of results at 16 different stations could only be used since it was necessary to co-ordinate river and lake surveys data collected independently. The variables were not used in a linear form as some guidance as to the functional relationships are available from work carried out by others which dealt with one variable at a time. Nevertheless, trial-and-error methods for the best relationships were still required.

It must be appreciated that the analysis has limitations and can only be expected to produce good approximations since no direct time considerations could be made. The conditions at the outlet were assumed to immediately affect a station at a distance of four miles. Similarly, wind conditions existing at the time of the survey were assumed to be directly responsible for existing lake currents, although prevalent currents may have resulted from earlier meteorological conditions. The limited data available necessitated the acceptance of these limitations. It was therefore necessary to assume that the time variations are slow and continuous with resultant minor errors. As will *be* seen, the errors arising from this assumption are not unreasonable, and workable equations commensurate with sampling and analytical accuracies can be formulated.

OUTLINE OF ANALYTICAL PROCEDURES

A multiple regression analysis (U.S. Department Commerce, 1963) employing various transforms of 1967 water quality data collected during surveys in the Port Maitland area was carried out for 22 sets of results at 16 different stations (see Figure 1). Lake data not coinciding with river data was rejected. The prediction equations were developed by trial-and-error employing various transforms and combinations of variables until good prediction equations were arrived at. On the average, fourteen different trials were made to determine each equation.

RESULTS

The variables considered and the derived prediction equations together with relative statistical characteristics are presented in Tables 1, 2 and 3 respectively. Accuracies based on 1968 analysis of multiple samples collected in the field employing standard techniques are presented in Table 4. Accuracies of 1967 data were reasonably close to the 1968 values, whereas 1966 values were not as accurate for various operational reasons. Unfortunately, detailed checks of navigation and wind velocity and direction readings are presently not available. A review of the techniques employed would, however, indicate accuracies of approximately 0.25 miles, 0.5 knots/hour and 10° respectively for positioning and wind velocity and direction readings.

The prediction equations developed from the 1967 data were applied to 1966 survey data and are presented in Table 5.A comparison was only possible in a small number of cases with coincident lake and stream data. A further reservation is necessary when considering the comparison, since in 1966, wind direction was determined only to the nearest 45° on the compass and navigation techniques were less accurate than in 1967.

DISCUSSION

The predictions are considered reasonable (except for soluble PO_4) in view of the accuracies of the determinations (Table 4), navigation and wind measurements. One must also consider three major shortcomings when applying a regression analysis. Firstly, there is no direct time factor. It was assumed that the loadings at the outlet of the Grand River are immediately affecting the concentrations at lake stations up to five miles away. Secondly, the water movements were correlated to the wind conditions at the time of the survey, whereas, it is known that longshore lake currents generally reflect winds occurring between 12 and 24 hours earlier (Hamblin, 1967 and Csanady, 1967). Unfortunately, the nearest recording wind station was at Grass Island, Niagara Falls, an inland station over 80 miles away. It does not, however, seem unreasonable to assume that the changes in the Grand River loadings are gradual and that wind conditions at the time of the survey compare reasonably with winds occurring during the proceeding 24 hours. Thirdly, temperature gradient between the lake station and the outlet of the Grand River were not considered due to lack of no data. In addition to this, there are a number of objections and statistical problems in dealing with maximum values.

In examining the prediction equations in Table 2, some scepticism is bound to be engendered by some of the variables appearing in the equations. It should, however, be recognized that all variables appearing in the equations are significant at the 0.01 level (except minimum DO at 0.05). Admittedly, it is difficult to imagine conductivity dependence on distance only. However, despite nearly twenty-five different trials on the data, distance remained as the significant factor. Although no allowance was made in the turbidity equation for turbidity originating in the lake, the equations still produce reasonable results even with winds over 10 knots.

The appearance of a negative predicted turbidity at a station and the negative results for soluble PO_4 cannot be explained at this time. It is, therefore, recommended that the soluble PO_4 equation be used with reservation putting more faith in the prediction equations than the 1966 check. In addition, the lack of a flow term in the equations for total PO_4 is unexpected. While these and other shortcomings might be dismissed with the statement "the equations work are valid within reasonable accuracy", it does not seem to be a very strong position to assume when they are developed on a trial-and-error basis. Consequently, the equations are put forth as workable computational tools only and will require modification to involve more operative physical variables as additional data becomes available. The approach to the complete cause and effect equations will necessarily have to be an evolutionary one.

CONCLUSIONS

Despite the shortcomings mentioned above, the prediction equations, based on actual measurements, are capable of predicting another set of results with reasonable accuracy (see Table 3). This is strong justification for their application appreciating that they will be modified to produce better predictions as more data becomes available.

REFERENCES

Csanady, G. T. *et al*, Douglas Point Saturation Run, University of Waterloo, Great Lakes Institute Report No. PR30, 1967.

Hamblin, P. *F.* and G. K. Rodgers, The Currents in the Toronto Region of Lake Ontario, University of Toronto, Great Lakes Institute, PR29, 1967.

U.S. Department of Commerce, Experimental Statistics, Handbook 91, National Bureau of Standards, Chapter 6, 1963.

TABLE 1: Variables Used

Symbol	Variables
S	average concentration of parameter at the outlet of the Grand River at the same time as the parameter determination at the lake station in the same units as lake station values.
T	difference in temperature between the surface and bottom at the lake station in degrees centigrade +1.0.
Y	depth at the lake station in metres.
W	wind vector determined by finding the resultant wind acting on a line joining the lake station and the outlet of the Grand River at the same time (use approximately 4 hour average wind) as the parameter determination at the lake station measured in knots per hour with the positive direction acting towards the outlet of the Grand River.
X	distance measured from the lake station to the outlet of the Grand River in miles.
Z	Grand River flow measured at the gauge at Brantford in cfs occurring four days earlier than the time of determination of the parameter at the lake station.

TABLE 2: Prediction Equations.

	Parameter at Lake Station	Prediction Equation Parameter = expressions below
1.	maximum turbidity(units)	$+ 0.71 - 0.21W + 4.02 (X)^{-1} + 0.000019 SZ$
2.	average turbidity(units)	$+1.78 - 0.18W - 2.24 (X)^{-1} + 0.000015 SZ$
3.	maximum conductivity (µmhos)	$+ 311.4 + 73.8 (X)^{-1}$
4.	maximum total PO_4 - PO_4 (ppm)	$+ 0.044 - 0.0052 SY (X)^{-1} - 0.0047W + 0.46 (X)^{-1}$
5.	maximum soluble PO_4 - PO_4 (Ppm)	$+ 0.032 - 0.0096 SY (X)^{-1} - 0.0083W + 0.37 (X)^{-1}$
6.	maximum dissolved oxygen (CO) (ppm)	$- 0.76 + 1.44S - 0.13T + 0.15Y$
7.	minimum dissolved oxygen (DO) % saturation	$+ 110 + 0.055S (X)^{-1} - 0.74TY (X)^{-1} + 0.094 X (W)^{-1}$
8.	maximum nitrate - N (ppm)	$- 0.059 - 0.0575S + 0.0029Y + 0.0013W + 0.047 (X)^{-1} + 0.00003Z$

TABLE 3: Statistical Characteristics Of Multiple Regression Analysis.

Parameter	Multiple Correlation Coefficient	Standard Error of Estimates	F. Test
Max. Turbidity Units	0.87	0.96	$F_{3;21} = 22.5 > F_{0.01;3;21} = 4.87$
Avg. Turbidity Units	0.73	1.08	$F_{3;19} = 6.96 > F_{0.01;3;19} = 4.94$
Max. Conductivity µmhos	0.80	8.87	$F_{1;20} = 79.8 > F_{0.01;1;20} = 8.10$
Max. Total PO_4 - (PO_4) ppm	0.90	0.044	$F_{3;18} = 26.7 > F_{0.01;3;18} = 5.09$
Max. Soluble PO_4 (PO_4) ppm	0.95	0.027	$F_{3;18} = 51.1 > F_{0.01;3;18} = 5.09$
Max. DO ppm	0.91	0.57	$F_{3;18} = 27.2 > F_{0.01;3;18} = 5.09$
Min. DO % Saturation	0.86	5.2	$F_{3;18} = 16.1 > F_{0.01;3;18} = 5.09$
Nitrate - (N) ppm	0.90	0.009	$F_{5;16} = 14.4 > F_{0.01;5;16} = 4.44$

TABLE 4: Accuracies Of Determinations - (1968).

Parameter	Range	No. of Readings	Standard Deviation
Turbidity			
Conductivity µmhos	300 - 350	estimate	7
pH	(8.2 - 8.6)	(63)	(0.057)
Total PO_4 - PO_4 ppm	0.02 - 0.31	54	0.03
Soluble PO_4 - PO_4 ppm	0.02 - 0.11	54	0.008
Dissolved Oxygen (DO) % Saturation	94 - 130	135	2.3
Nitrate - N ppm	0.01 - 1.62	48	0.022

TABLE 5: Prediction Equations Applied To 1966 Survey Data.

Stn. No.	Date	Maximum Turbidity Units		Mean Turbidity Units		Maximum Conductivity (μmho)	
		PARAMETER					
		Predicted	Observed	Predicted	Observed	Predicted	Observed
74	June 17	-0.19	2.3	0.68	2.3 (only one value)	329	324
	July 17	1.8	1.1	2.4	1.1	329	322
78	June 17	2.6	3.2	2.7	3.2	345	310
	July 17	3.0	2.5	3.1	2.5	345	323
87	June 18	2.5	2.9	3.0	2.9	326	331
	July 17	2.0	0.6	2.7	0.6	326	320

TABLE 5: Prediction Equations Applied To 1966 Survey Data.

		PARAMETER					
		Maximum Total PO_4 ppm		Maximum Soluble PO_4 ppm		Maximum DO ppm	
Stn. No.	Date	Predicted	Observed (1 value only)	Predicted	Observed (1 value only)	Predicted	Observed
74	June 17	0.04	0.05	-0.07	0.03	15.7	10.6
	July 17	0.07	0.07	-0.02	0.02	14.3	8.4
78	June 17	0.12	0.14	-0.03	0.12	15.3	10.4
	July 17	0.13	0.06	-0.03	0.04	14.2	8.2
87	June 18	0.20	0.15	+0.002	0.13	16.3	10.3
	July 17	0.043	0.06	-0.058	0.03	14.8	8.2

TABLE 5: Prediction Equations Applied To 1966 Survey Data.

Stn. No.	Date	PARAMETER			
		Minimum DO (% Saturation)		Maximum Nitrate (ppm)	
		Predicted	Observed	Predicted	Observed
74	June 17	99	95	0.025	-
	July 17	105	78	0.016	-
78	June 17	98	92	0.023	-
	July 17	100	87	0.020	-
87	June 18	101	99	0.018	-
	July 17	99	87	-0.023	-

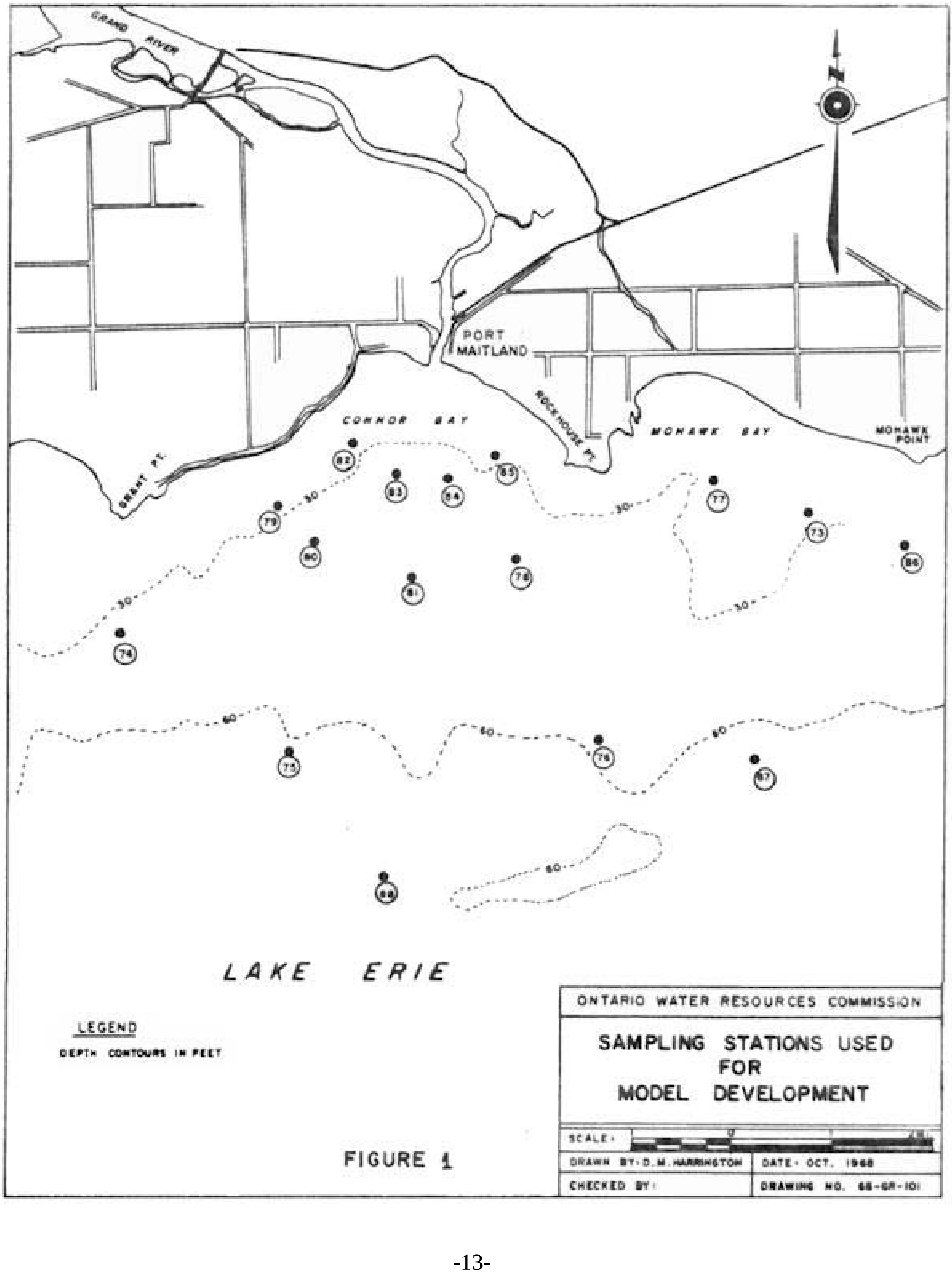
GRAND RIVER
PORT MAITLAND
CONNOR BAY
GRANT PT.
ROCKHOUSE PT.
MOHAWK BAY
MOHAWK POINT
82
83
84
85
77
79
80
81
78
73
86
74
75
76
87
88
LAKE ERIE
LEGEND
DEPTH CONTOURS IN FEET
ONTARIO WATER RESOURCES COMMISSION
SAMPLING STATIONS USED FOR MODEL DEVELOPMENT
SCALE:
DRAWN BY: D. M. HARRINGTON
DATE: OCT. 1968
CHECKED BY:
DRAWING NO. 68-GR-101

FIGURE 1

APPENDIX 1: WATER QUALITY DATA.

1967

Station	Date	Wind		Depth "Y"	Temperature "T"	Max. Turbidity
		Direction	Force Knots/hr			
82	30 May	237	6	7	1.4	2.3
80	30 May	237	6	13	2.5	2.9
75	31 May	-	calm	19	5.3	2.3
	22 June	225	7		8.6	2.6
74	30 April	90	5	14.3	1.0	8.0
	31 May	-	calm		1.5	4.5
	8 June	130	4		10.4	3.0
	22 June	225	7		2.1	1.8
78	30 May	237	6	12.5	2.7	4.5
	8 June	130	4		3.9	3.5
83	30 May	237	6	11	2.2	3.5
86	29 April	180	9	11.5	1.5	3.2
	29 May	130	5		6.0	2.2
84	30 May	237	6	9	2.8	4.2
85	30 May	237	6	8	2.9	4.8
79	30 May	237	6	5	1.0	3.2
88	31 May	-	calm	23	4.2	2.6
76	30 May	237	6	19	4.8	3.2
77	29 May	130	5	7	1.8	2.6
87	30 May	237	6	20	5.5	1.8
81	30 May	237	6	14	3.5	2.9
73	29 May	130	5	5	1.0	2.2

APPENDIX 1(Cont' d): WATER QUALITY DATA

1967

Station	Max. Conductivity	Max. T PO_4	Max. Sol PO_4	Max. DO ppm	Min. DO % saturation	Max. Nitrate
82	364.	0.3	0.2	11.6	107.	0.04
80	335.	0.1	0.03	12.0	107.	0.03
75	326.	0.08	0.04	14.0	99.	0.05
	326.	0.08	0.03	9.4	83.	0.03
74	322.	0.16	0.1	13.2	110.	0.04
	340.	0.19	0.11	12.5	115.	0.01
	332.	0.12	0.03	10.6	87.	0.01
	328.	0.08	0.03	9.2	97.	0.03
78	334.	0.12	0.06	12.4	106.	0.02
	332.	0.11	0.03	12.5	82.	0.02
83	386.	0.39	0.29	11.6	104.	0.05
86	327.	0.07	0.03	13.3	112.	0.09
	338.	0.11	0.05	12.0	108.	0.02
84	368.	0.28	0.2	11.6	101.	0.03
85	389.	0.33	0.25	12.0	97.	0.04
79	352.	0.22	0.13	12.0	111.	0.02
88	325.	0.08	0.06	13.6	100.	0.05
76	330.	0.08	0.04	13.6	95.	0.03
77	335.	0.11	0.09	11.8	105.	0.01
87	326.	0.06	00.02	13.6	97.	0.05
81	332.	0.11	0.07	12.4	93.	0.05
73	338.	0.12	0.06	11.8	109.	0.01