CLEAR TECHNICAL REPORT NO. 237

(LAKE ERIE TAT CONTRIBUTION NO. 12)

LAKE ERIE INTENSIVE STUDY: CLUSTER ANALYSIS OF NEARSHORE WATER MASSES

Prepared by

Charles E. Herdendorf
and
John J. Mizera

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

LIST OF FIGURES

INTRODUCTION

In an effort to examine relationships between parameters in a system so complex as western Lake Erie it is often desirable to define areas having similar characteristics. Previous studies (Hartley et al. 1966; Herdendorf 1969) have suggested that because the western basin of Lake Erie receives flow from several sources the nearshore zone of this basin may be characterized by several distinct water masses. To explore relationships between nutrient concentrations and chlorophyll _a_ distributions, an attempt was made to identify these water masses by performing a series of cluster analyses and examining the patterns.

METHODS

A cluster analysis is a technique where objects are partitioned into optimally homogenous groups on the basis of an empirical measure of similarity among the objects (Johnson 1967). The Statistical Analysis System (SAS) was used to perform the analysis (Barr et al. 1976). Data used in the analysis were collected at seventy-seven nearshore stations along the United States portion of the western basin of Lake Erie during the 1978-1979 intensive study (Figure 1). The Lake Erie intensive study included four cruises each year:

Cruise No.	1978	1979
1	April 14-29	March 29-April 15
2	June 26-July 12	July 25-August 5
3	August 23-September 11	September 9-23
4	October 3-17	October 9-23

Two sets of clusters were formed, both using parameters believed to influence primary productivity. One set of clusters used the parameters of temperature conductivity, total phosphorus, soluble reactive phosphorus, nitrate + nitrite, ammonia, silica and total Kjeldahl nitrogen; the second set used conductivity, temperature, transparency as measured by secchi, and turbidity. Temperature and conductivity were included in both groups because earlier investigations (Hartley et al. 1966; Herdendorf 1969) have shown these parameters are the most reliable indicators of water masses regardless of productivity potential.

The algorithm used by SAS was outlined by Stephen C. Johnson (1967). This procedure initially considers each case (in this instance each station) a cluster that consists of one station. The two clusters with the smallest Euclidean distance are joined together, forming one cluster of two stations and one less cluster. This procedure is repeated until only one cluster remains.

The Euclidean distance between clusters I and K is defined by:

$$D(I,K) = (\quad 1 \le J \le N \quad A(I,J) - A(K,J) \ ^{2})^{\frac{1}{2}}$$

where: D(I,K) is the Euclidean distance between cluster I and cluster K.

J is the number of parameters on which the distance is to be calculated.

A(*,J) is the value of parameter J at cluster *.

This is simply the sum of the difference of each parameter between two clusters.

Many parameters were measured in different units; therefore, it is necessary to prescale the variables to make their values comparable (Hartigan 1976; Engleman 1979). Hartigan recommended dividing each observation by the variance of that parameter. Using this weighting factor the Euclidean distances are invariant under changes in units of measurement, and all variables make the same average contribution to the distance (Sokal and Sneath 1963).

A weakness of cluster analysis is that there are no established criteria for determining the proper number of clusters. The authors of this paper chose the number of groupings presented here by combining information about the average distance between clusters and the number of clusters, graphic representations of the groups, and previous experience with water flow patterns in the study area.

RESULTS AND DISCUSSION

The first set of parameters were used to form what will be referred to as "nutrient clusters" and the second set to form "turbidity clusters." The results of the analysis were used to prepare maps of the study area; separate maps for six through ten clusters were generated. Eight clusters appeared to be the most appropriate number to discriminate water masses, based on previous experience with water flow patterns in western Lake Erie. This number is used for the cluster maps (Figures 2-7). Figure 2 depicts the nutrient clusters from 1978 for all cruises, and Figure 3 depicts the nutrient clusters for all cruises except the period of spring runoff (Cruise 1). In a similar way, Figures 4 and 5 show nutrient clusters for 1979. Figures 6 and 7 present turbidity clusters for all cruises in 1978 and 1979, respectively.

Nutrient Clusters

One of the most obvious features of the nutrient clusters for both 1978 and 1979 is a concentration of clusters at major tributary mouths which empty into restricted or semi-restricted bays, such as Sandusky Bay and Maumee Bay. From these bays a series of clusters extend to open portions of the lake, generally increasing in surface area in a lakeward direction. This apparent gradation pattern suggests sources of water masses and modifications of these masses as they disperse and mix with open lake water. The nearshore water beyond the areas impacted by major tributaries is relatively consistent as evidenced by small number of clusters both offshore and inshore along reaches without tributaries.

When Cruise 1 data (March and/or April) for both years is eliminated from the data set, the influence of the major tributaries on the formation of clusters is greatly reduced (Figures 3 and 5). All of the major western Lake Erie tributaries experience their highest flow in the spring and decline dramatically in the summer and fall. For example, the flow rates of the Maumee River for each cruise period are listed below:

Cruise No.	Flow Rate (Cubic Meters/Sec) 1978	1979
1	376.9	442.6
2	171.2	117.3
3	10.5	33.4
4	7.1	13.4

Data Source: U.S. Geological Survey (1979, 1980, 1981)

The high tributary flows during Cruise 1, for each year, are apparently responsible for carrying large volumes of water to the lake with distinct nutrient characteristics resulting in the formation of the cluster series described earlier. The mean concentration of each water quality parameter used in a particular cluster analysis is shown on each cluster figure. For example, for all cruises in 1978 (Figure 2), the mean phosphorus concentration for the cluster at the mouth of the Maumee River (no. 5) was 301 ug/l, while the furthest offshore cluster (no. 1) was only 113 µg/l. When Cruise 1 was eliminated from the data set (Figure 3), the Maumee River mouth cluster (no. 5) concentration for total phosphorus was 252 µg/l and the offshore cluster (no. 1) was 103 µg/l. Water masses do not appear to flow as strongly lakeward after the spring cruises as evidenced by clusters for Cruises 2-4 being packed close to the shore. After the high spring flow, the nearshore waters of the study area appear to be dominated by water masses emanating from the Detroit River (cluster no. 1 in 1978 and cluster no. 3 in 1979).

Turbidity Clusters

Clusters for turbidity parameters (Figures 6 and 7) showed patterns similar to those for nutrient parameters, except the nutrient clusters generally graded further offshore before an open lake or "background" cluster was formed. This can be interpreted as consistent movements of distinct water masses and is indicative of particulate settling at a more rapid rate than the dispersion or mixing of dissolved material as the water masses progress in an offshore direction.

Consistent patterns of turbidity clusters are formed for both 1978 and 1979. The most turbid clusters are located in Sandusky and Maumee Bays and at the mouths of the Raisin and Huron Rivers, indicating the tributaries as source areas. The least turbid cluster each year is located offshore from Stony Point, Michigan. It is likely that this water mass originated in the clearer, upper lakes and entered Lake Erie through the mid-channel of the Detroit River. Inshore turbidity clusters, exclusive of the bays, are generally twice as turbid as the offshore clusters. For example, in 1978, the clusters adjacent to the Michigan and Ohio shorelines range from 24-31 NTU (cluster nos. 1, 5 and 6) while the offshore cluster (nos. 3 and 4) range from 9-14 NTV. Although the patterns are similar for both years, the 1979 turbidity values are considerably higher than those for 1978. Presumably this is due to an earlier cruise schedule in 1979 which included measurements during the a spring runoff event in March and April when the Maumee River and other tributaries were extremely turbid.

Chlorophyll Patterns

Figures 7 and 8 depict the annual (March/April-October) distribution of corrected chlorophyll *a* in the United States nearshore water of western Lake Erie for 1978 and 1979, respectively. The distribution patterns indicate high productivity in the bays and adjacent to the shoreline, particularly along the Michigan shoreline. The area of lowest concentration of chlorophyll *a* is offshore near Stony Point, Michigan. The regions of high chlorophyll *a* concentrations correspond to clusters with high nutrient and turbidity values, particularly those associated with the major tributaries. Conversely, low chlorophyll *a* concentrations are found in areas with clusters low in nutrient and suspended solids, especially well offshore of the Michigan shore in the region of the lake strongly influenced by mid-channel Detroit River flow.

Sandusky Bay consistently shows a series of clusters, indicating three distinct, but intergrading water masses. This sequence is verified by a gradation of values for total phosphorus, turbidity, chlorophyll *a* and most other parameters. This bay contains the highest annual concentrations of chlorophyll *a* (in excess of 100 µg/l) in western Lake Erie. Maumee Bay is the area of next highest concentration (over 170 µg/l). Nutrient conditions are more favorable and inorganic turbidity is slightly less in Maumee Bay, but the algal populations are not as abundant as in Sandusky Bay. Herdendorf et al. (1977) speculated that high ammonia concentration in the Maumee River can inhibit algal production. Figures 2-5 show Maumee River clusters with ammonia concentrations between 400-500 µg/l, the approximate level of algal inhibition determined by Abeliovich and Azov (1976).

CONCLUSIONS

Cluster analysis provides a useful technique for assessing the homogeniety water masses, their sources and their movements. The impact of major tributaries on the formation of water masses in the nearshore region can be readily demonstrated by clustering techniques. In western Lake Erie nutrient- and turbidity-based clusters showed a positive relationship to algal productivity as measured by chlorophyll _a_ concentrations. The similarity of cluster patterns over a two-year period indicates a consistency within the lake and provides some verification of the reliability of this approach.

REFERENCES

Abeliovich, A. and Y. Azov. 1976. Toxicity of ammonia to algae in sewage oxidation ponds. Applied and Environ. Microbiol. 31(6):801-806.

Barr, A.J., J.H. Goodnight, J.P. Sall and J.T. Helwing. 1976. A user's guide to SAS. SAS Institute, Raleigh, N.C. 329 p.

FIGURES

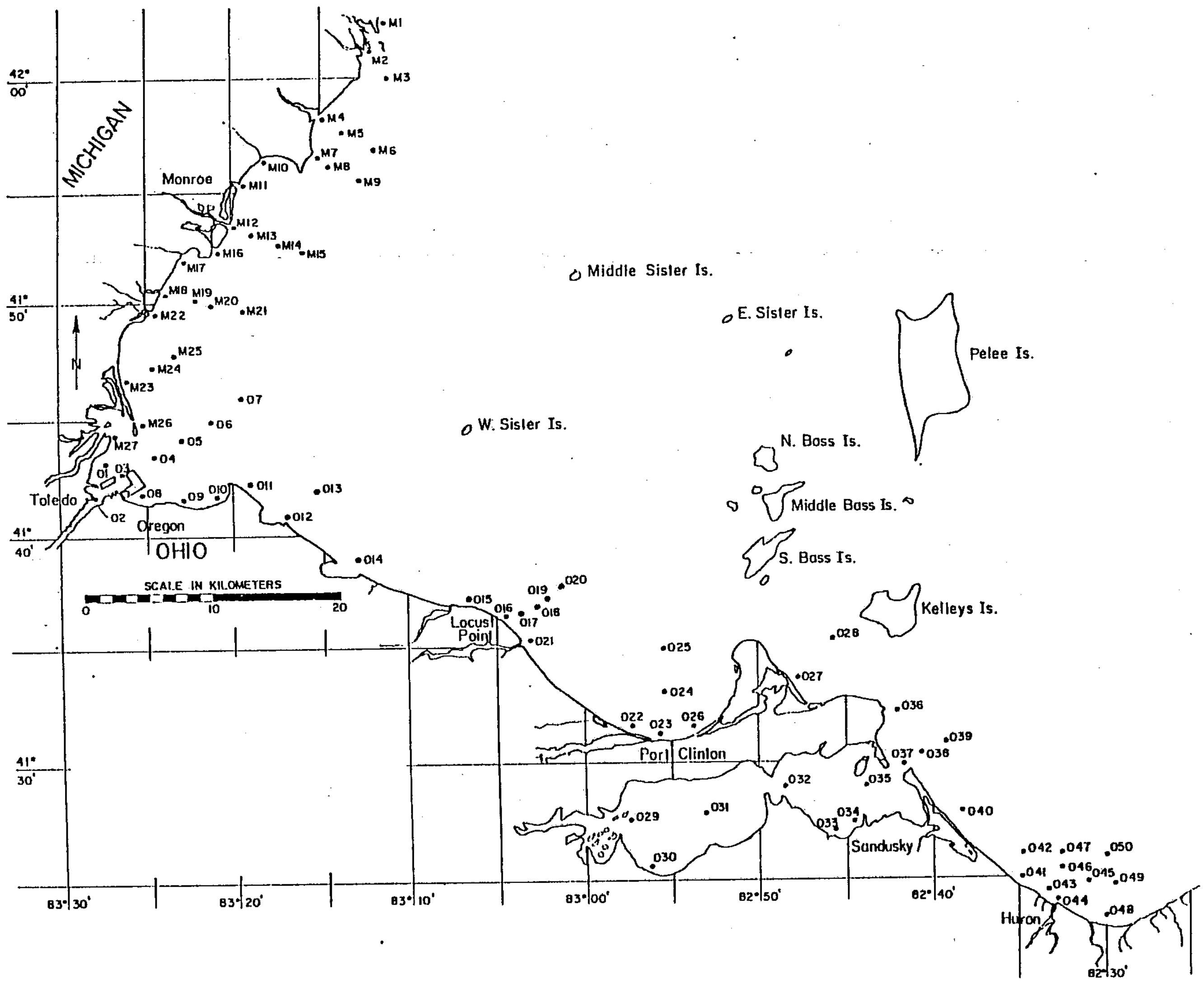

Figure 1. Western Lake Erie Nearshore Stations Used for Cluster Analysis.

MEAN PARAMETER VALUES FOR NUTRIENT CLUSTERS

CRUISES 1-4 (1978)

CLUSTER NO	TEMP (^oC)	COND (umhos/cm)	TP (ug/l)	SRP (ug/l)	NO_2+NO_3 (ug/l)	NH_3 (ug/l)	DRS (ug/l)	TKN (ug/l)	TOTAL STATIONS
1	18.5	245	113	7	316	67	761	317	38
2	19.3	306	163	20	1218	93	1573	458	9
3	17.7	247	111	19	761	73	1298	388	19
4	18.0	363	185	9	789	21	2135	397	4
5	20.6	421	301	58	3303	410	3066	928	1
6	21.3	370	290	46	2029	243	2198	681	2
7	18.4	463	225	26	650	19	5561	439	2
8	17.8	428	226	16	708	16	4161	404	2

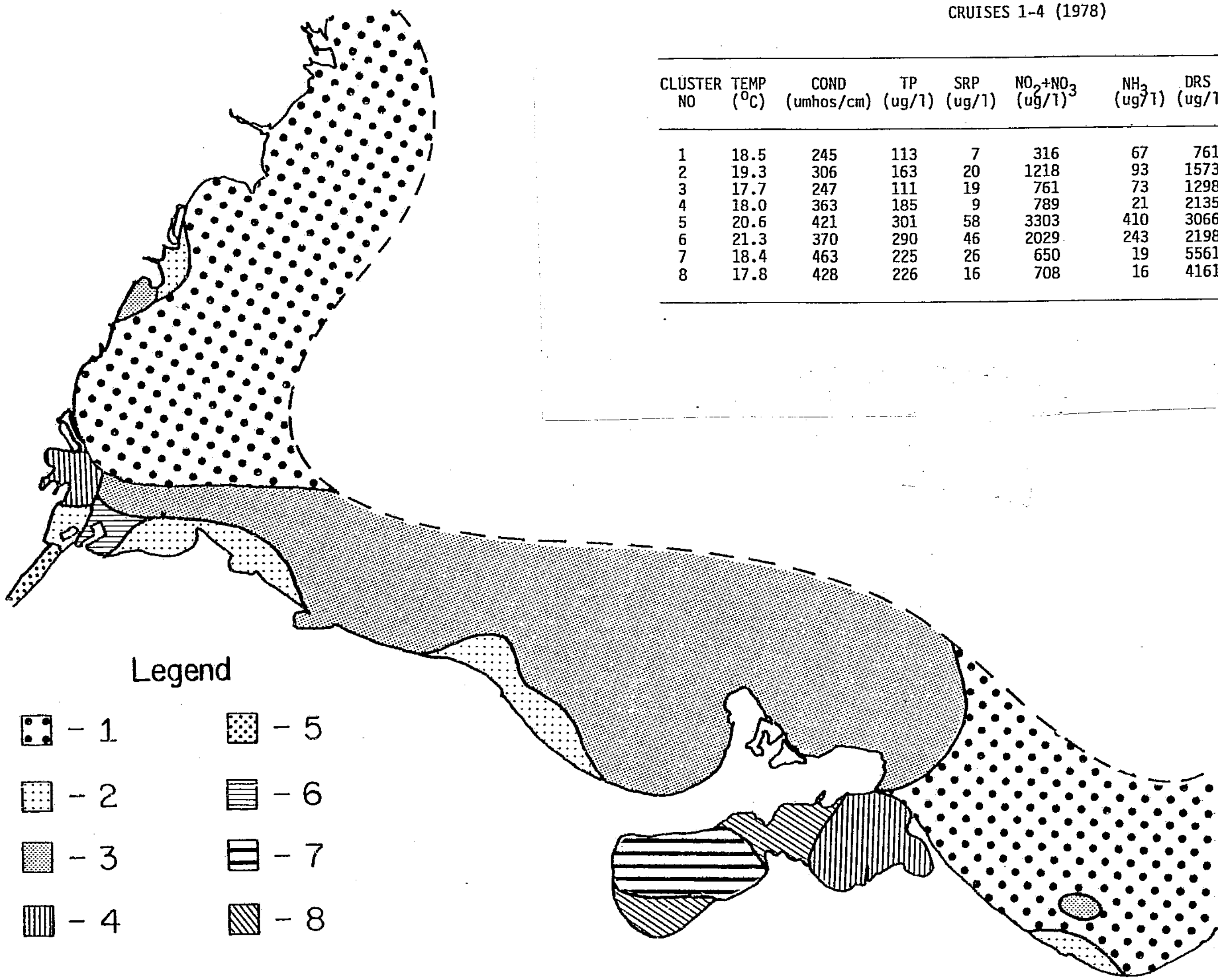

Figure 2. Western Lake Erie Nearshore Nutrient Clusters, Cruses 1-4, (1978).

MEAN PARAMETER VALUES FOR NUTRIENT CLUSTERS

CRUISES 2-4 (1978)

CLUSTER NO	TEMP (^oC)	COND (umhos/cm)	TP (ug/l)	SRP (ug/l)	NO_2+NO_3 (ug/l)	NH_3 (ug/l)	DRS (ug/l)	TKN (ug/l)	TOTAL STATIONS
1	20.2	253	103	7	184	46	594	289	63
2	19.5	352	124	3	25	6	1563	310	3
3	24.1	397	254	32	1066	156	1677	561	3
4	20.2	330	263	24	951	186	878	508	2
5	21.6	410	252	52	2720	429	2052	938	1
6	19.1	486	235	26	278	9	5462	420	2
7	18.9	463	263	11	336	6	4640	407	1
8	19.3	425	210	14	11	7	2994	337	1

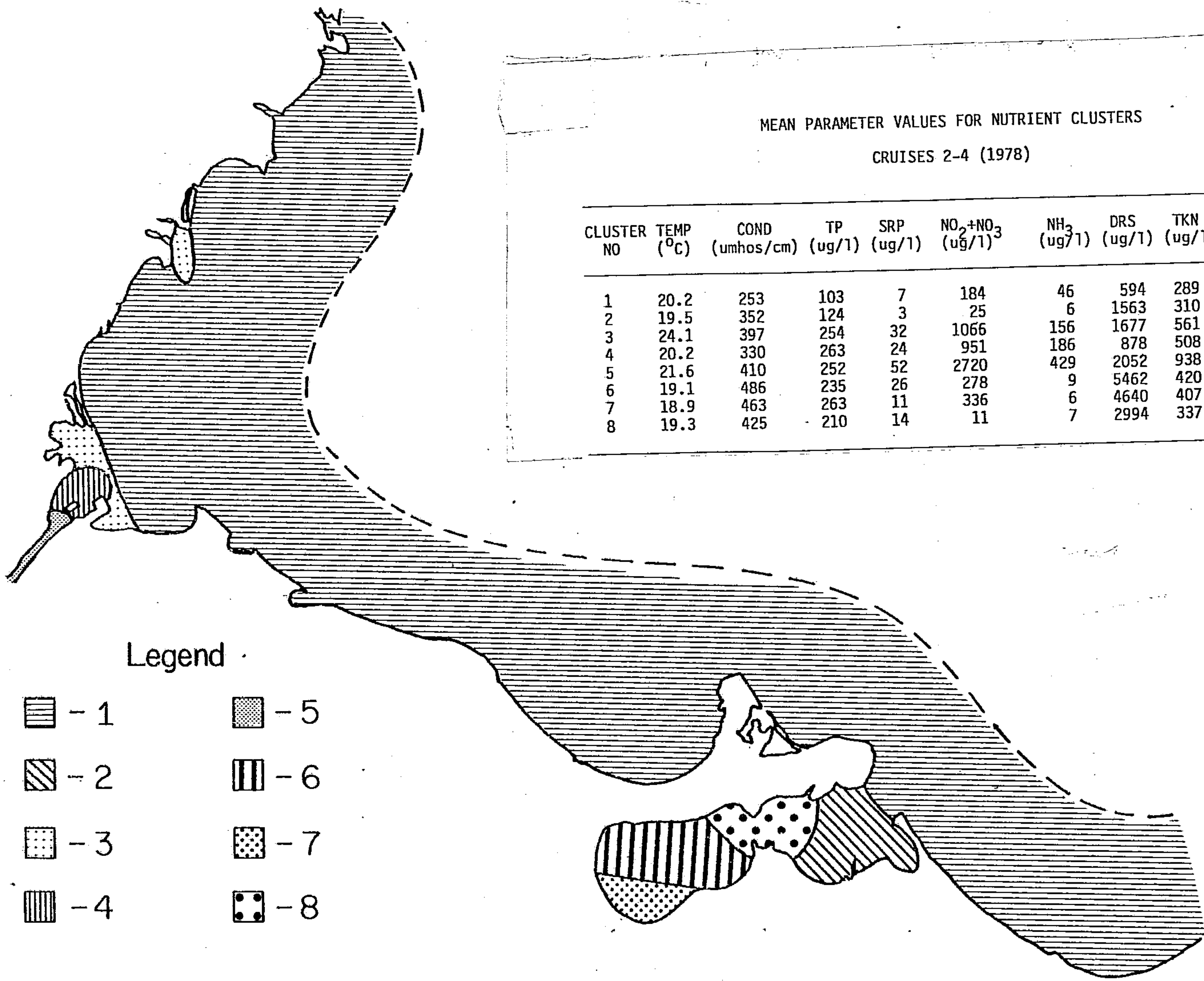

Figure 3. Western Lake Erie Nearshore Nutrient Clusters, Cruises 2-4, (1978).

MEAN PARAMETER VALUES FOR NUTRIENT CLUSTERS

CRUISES 1-4 (1979)

CLUSTER NO	TEMP (^{o}C)	COND (umhos/cm)	TP (ug/1)	SRP (ug/1)	NO_2+NO_3 (ug/1)	NH_3 (ug/1)	DRS (ug/1)	TKN (ug/1)	TOTAL STATIONS
1	15.5	238	66	5	514	53	783	2212	34
2	15.1	272	99	5	629	96	1134	1395	10
3	16.7	340	123	13	1808	68	1863	2276	8
4	15.7	305	106	7	1760	72	972	2132	13
5	15.2	349	118	10	2111	129	1235	3019	5
6	16.7	461	184	31	3448	215	2060	2441	2
7	16.5	459	201	71	2756	468	4493	2572	1
8	18.1	485	174	18	1125	26	3825	2241	3

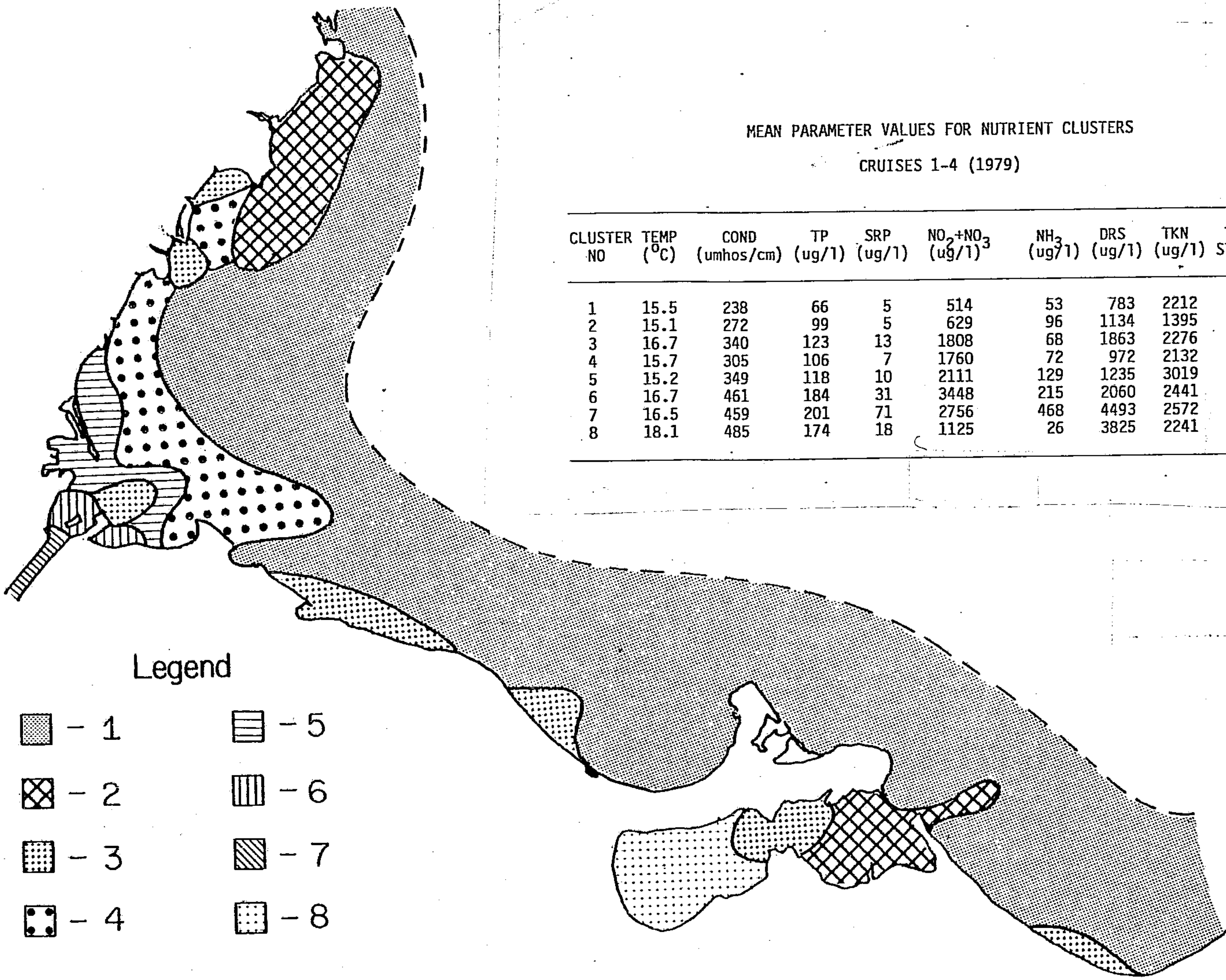

Figure 4. Western Lake Erie Nearshore Nutrient Clusters, Cruises 1-4, (1979).

MEAN PARAMETER VALUES FOR NUTRIENT CLUSTERS

CRUISES 2-4 (1979)

CLUSTER NO	TEMP (^oC)	COND (umhos/cm)	TP (ug/l)	SRP (ug/l)	NO_2+NO_3 (ug/l)	NH_3 (ug/l)	DRS (ug/l)	TKN (ug/l)	TOTAL STATIONS
1	19.2	294	104	5	564	42	522	3597	8
2	18.5	298	110	3	945	22	348	2287	13
3	19.1	247	66	3	228	16	596	596	40
4	19.2	263	91	4	286	46	756	756	10
5	19.4	473	182	68	3470	488	3838	3838	1
6	21.9	466	150	30	2820	196	2101	2101	1
7	20.5	488	188	17	273	3	3296	3296	2
8	20.4	518	166	8	67	2	3478	3477	1

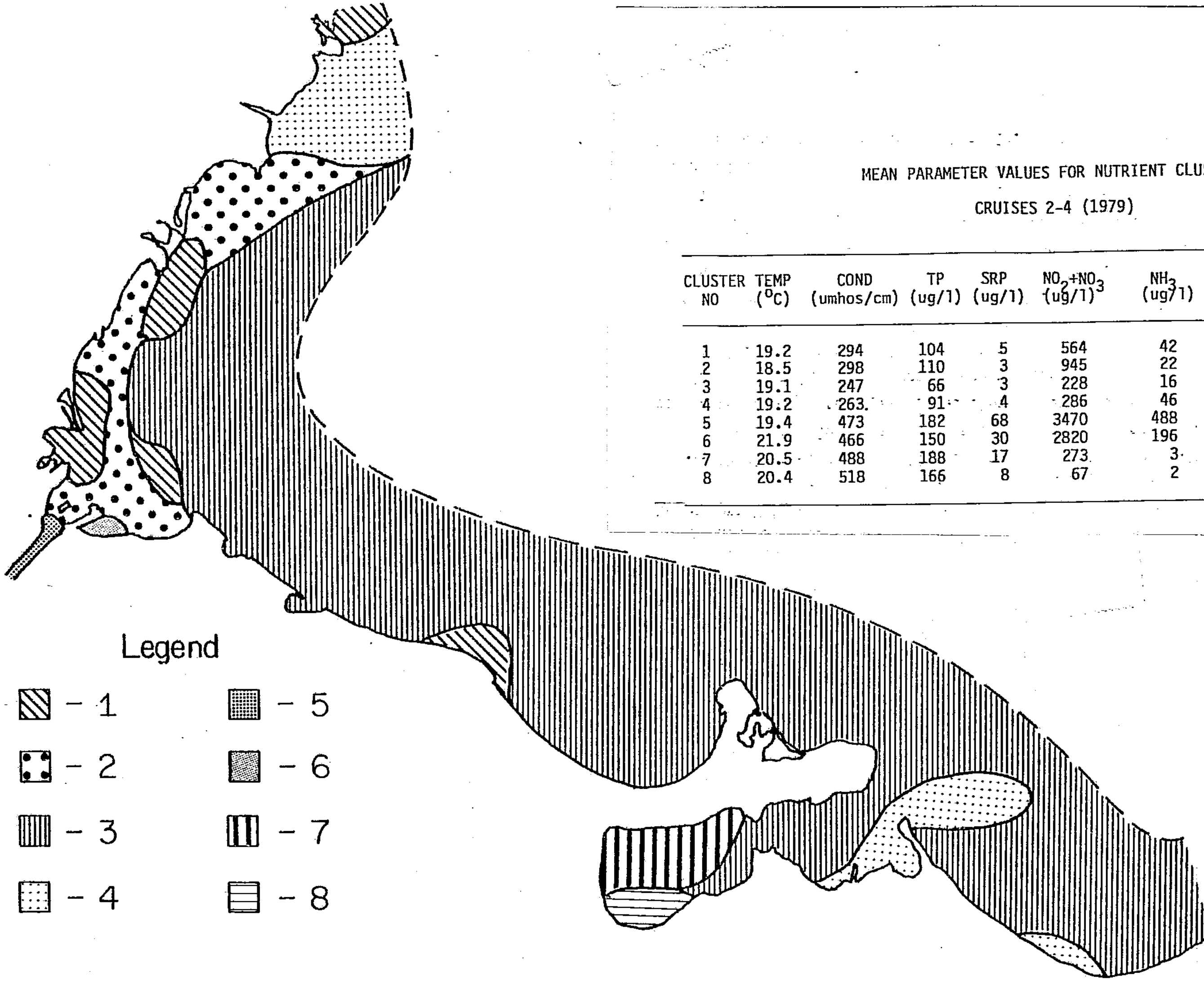

Figure 5. Western Lake Erie Nearshore Nutrient Clusters, Cruises 2-4, (1979).

MEAN PARAMETER VALUES FOR TURBIDITY CLUSTERS

CRUISES 1-4 (1978)

CLUSTER NO	TURB (NTU)	TRANS (m).	COND (umhos/cm)	TEMP (^{o}C)	TOTAL STATIONS
1	27	0.6	263	18.2	21
2	16	0.6	297	18.8	3
3	14	1.0	227	17.9	31
4	9	1.0	189	18.6	3
5	31	0.5	365	19.2	8
6	24	0.4	413	20.0	4
7	35	0.4	449	20.0	4
8	40	0.3	459	18.3	3

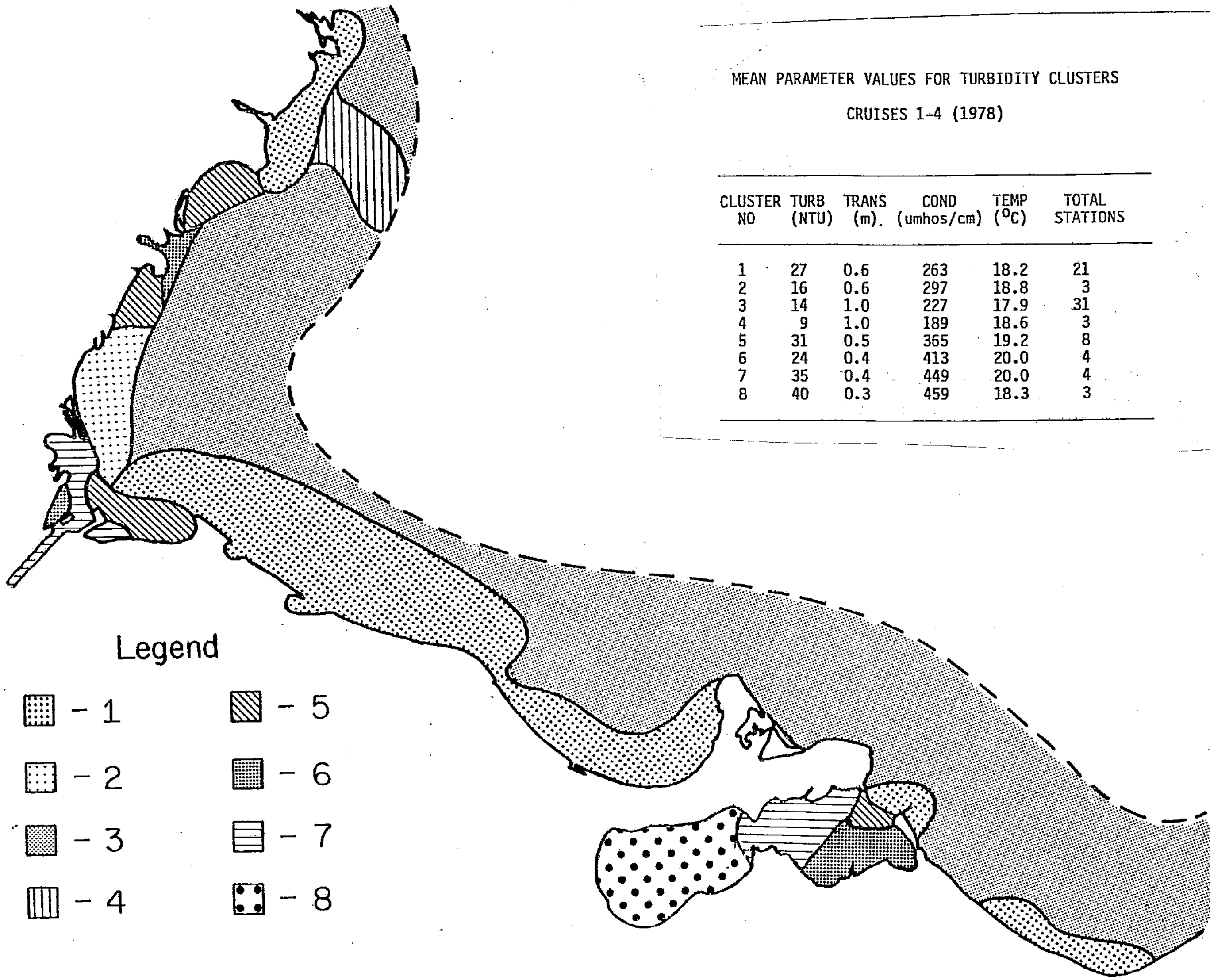

Figure 6. Western Lake Erie Nearshore Turbidity Clusters, Cruise 1-4, (1978).

MEAN PARAMETER VALUES FOR TURBIDITY CLUSTERS

CRUISES 1-4 (1979)

CLUSTER NO	TURB (NTU)	TRANS (m)	COND (umhos/cm)	TEMP (^oC)	TOTAL STATIONS
1	21	0.9	241	15.3	43
2	105	0.7	245	17.8	1
3	9	1.1	142	14.1	1
4	43	0.5	311	15.8	17
5	45	0.3	394	16.4	8
6	244	0.5	309	17.1	1
7	68	0.3	484	17.1	5
8	192	0.2	477	18.4	1

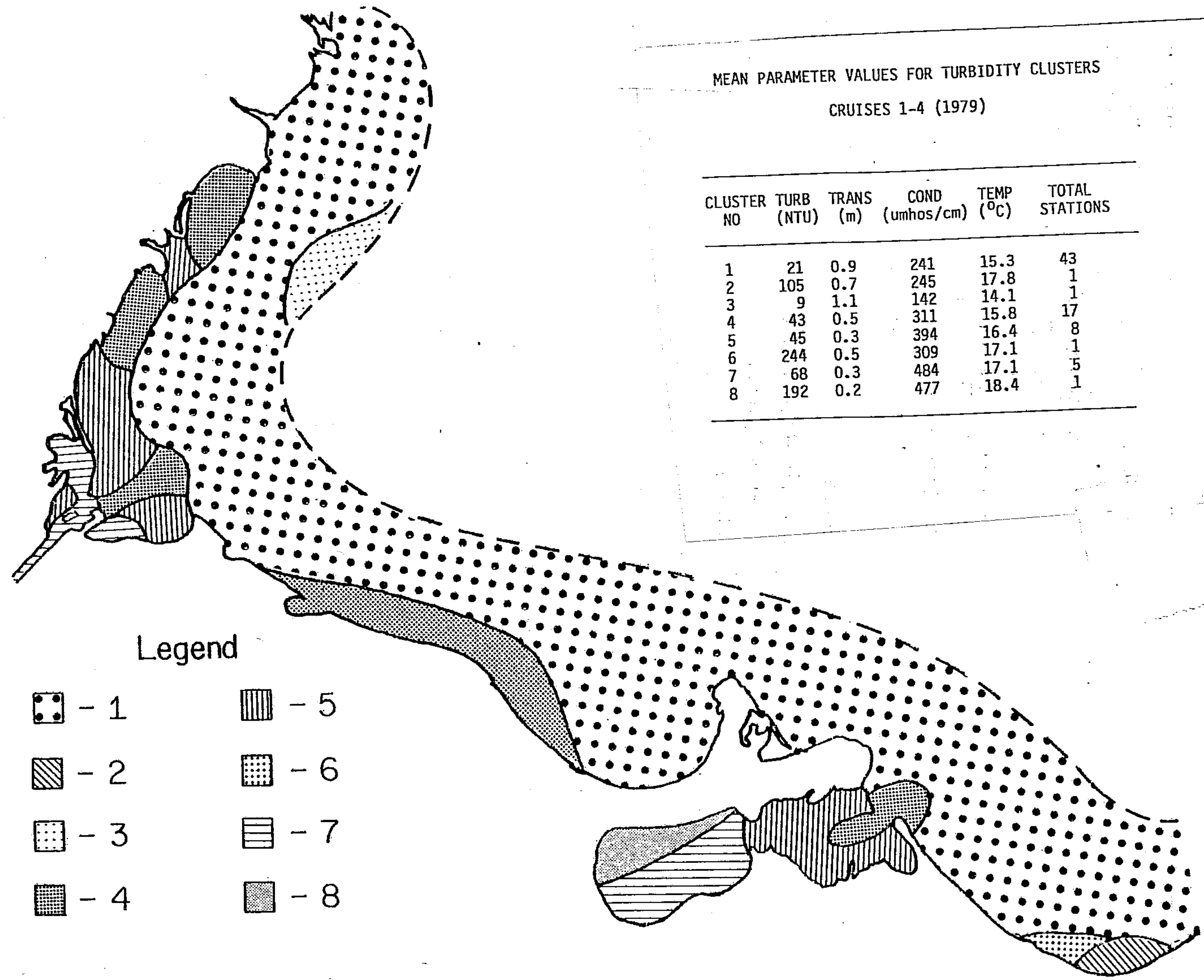

Figure 7. Western Lake Erie Nearshore Turbidity Clusters, Cruises 1-4, (1979).

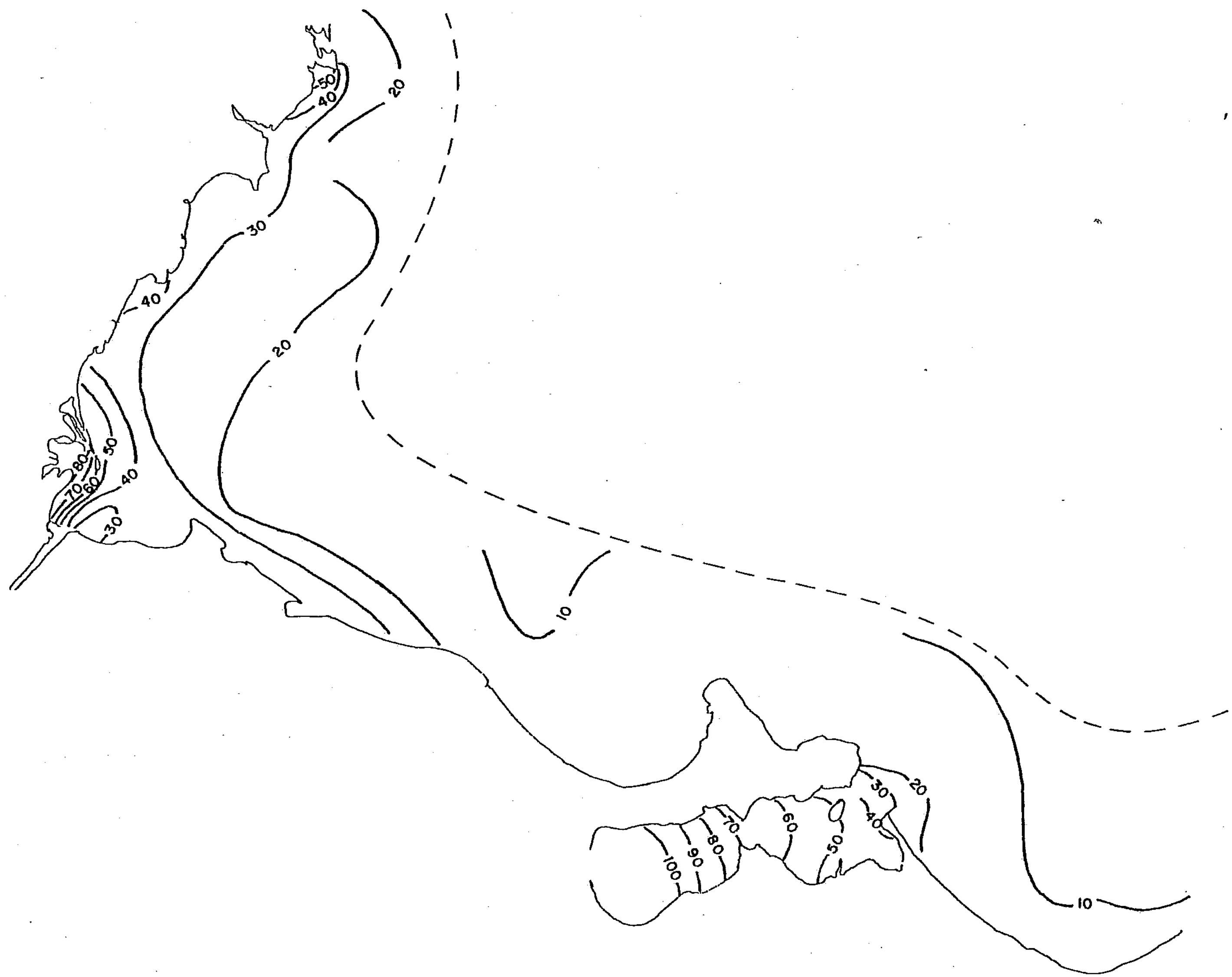

Figure 8. Western Lake Erie Nearshore Corrected Chlorophyll a (µg/l), Mean Annual Concentrations, Cruises 1-4, (1978).

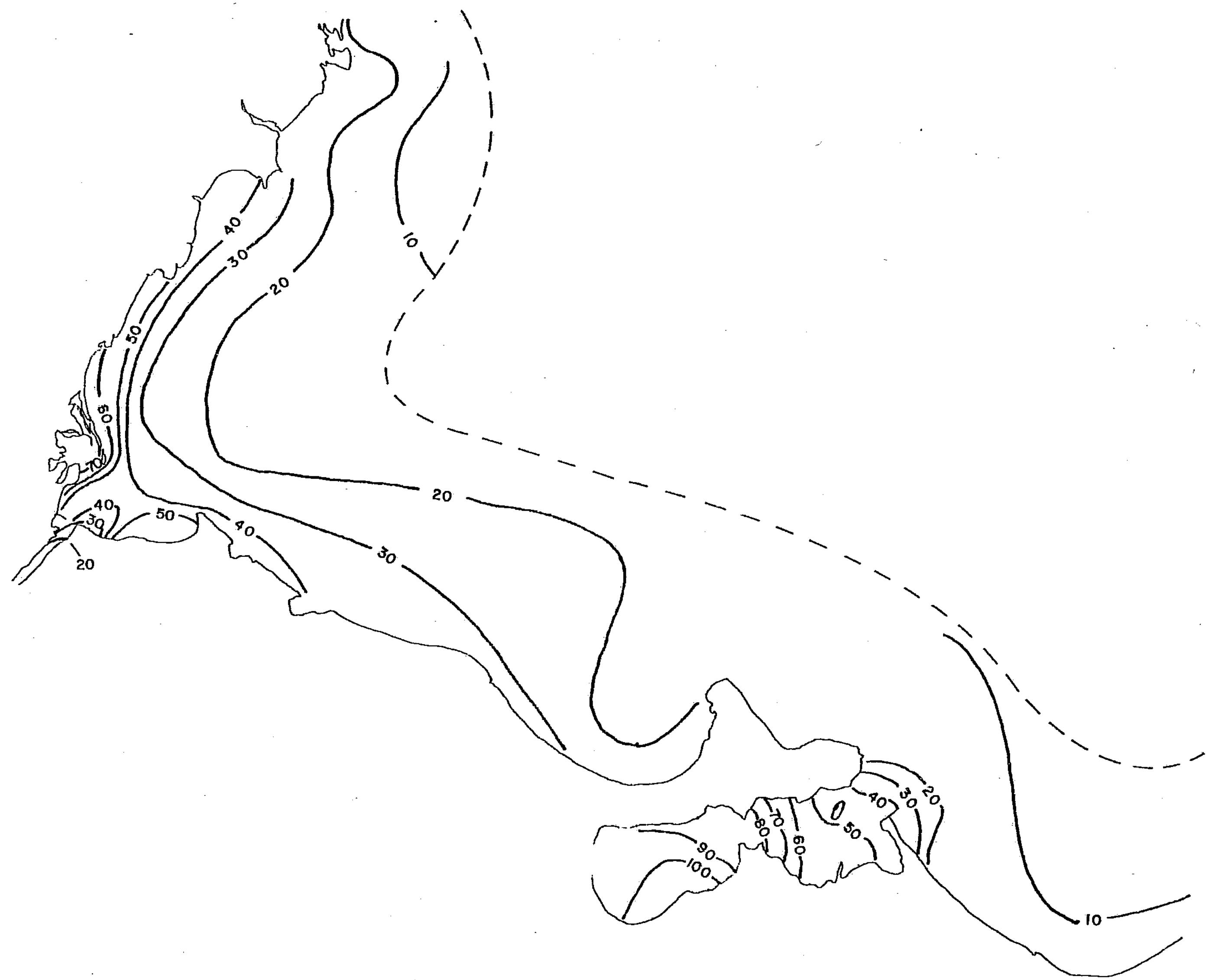

Figure 9. Western Lake Erie Nearshore Corrected Chlorophyll a ($\mu g/l$), Mean Annual Concentration, Cruises 1-4, (1979).