NYSGI-T-79-011 c. 3

AN ANALYSIS OF SHELLFISH SANITATION DATA

P.K. WEYL

marine sciences research center
state university of new york

SPECIAL REPORT 30

REFERENCE 79-13

MARINE SCIENCES RESEARCH CENTER
STATE UNIVERSITY OF NEW YORK
STONY BROOK, NEW YORK 11794

CIRCULATING COPY
Sea Grant Depository

AN ANALYSIS OF SHELLFISH SANITATION DATA

Peter K. Weyl

August 1979

Sponsored by the New York Sea Grant Institute and the New York State Department of Environmental Conservation

Special Report 30

Reference 79-13

Approved for Distribution

J.R. Schubel

J. R. Schubel, Director

TABLE OF CONTENTS

	Page
List of Figures	ii
List of Tables	ii
Acknowledgements	iii
Maps of Great South Bay	iv
Introduction	1
Statistical characteristics of the coliform data	1
Comparison of total coliform and fecal coliform criteria	3
Tidal variation of bacterial contamination	9
Discussion	11
Area 3, Great South Bay	11
Area 8, Moriches Bay	11
Area 29, Flanders Bay	11
Areas 42, 43, 53 and 46, Huntington Bay	11
The effect of rainfall on coliform contamination	14
Conclusions	17
Statistical aspects	17
Comparison of total and fecal coliform data	23
The effect of storm runoff on the coliform distribution	25
General	25

LIST OF FIGURES

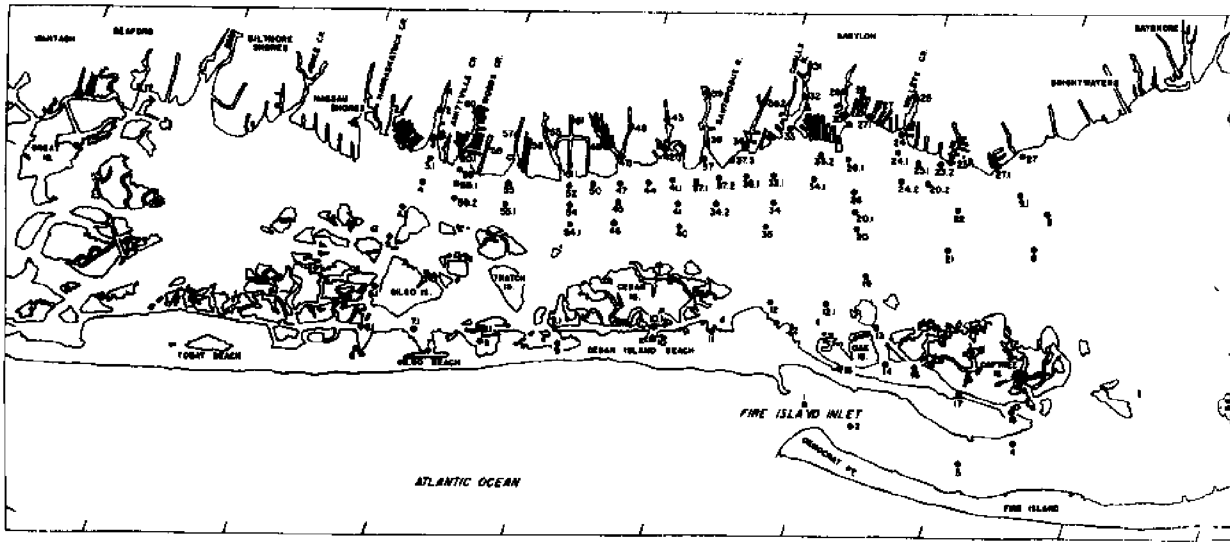
Figure		Page
1	Area 3, Great South Bay	5
2	Area 8, Moriches Bay	6
3	Areas 42 and 43, Northport Harbor and Centerport Harbor	7
4	Areas 29 and 46, Flanders Bay and Huntington Harbor	8
5	Coliform distributions for Area 3, Great South Bay	18
6	Coliform distributions for Area 8, Moriches Bay	19
7	Coliform distributions for Area 29, Flanders Bay	20
8	Coliform distributions for Areas 40-46, Huntington Bay Complex	21
9	Relationship between the median ratio of wet to dry coliform concentrations and the estimate of the fraction of WET observations. T = total coliform, F = fecal coliform	22
10	Comparison of the cumulative distribution of all five tube tests using the transformation matrix (the steps) and using the clustered data (the circles)	24

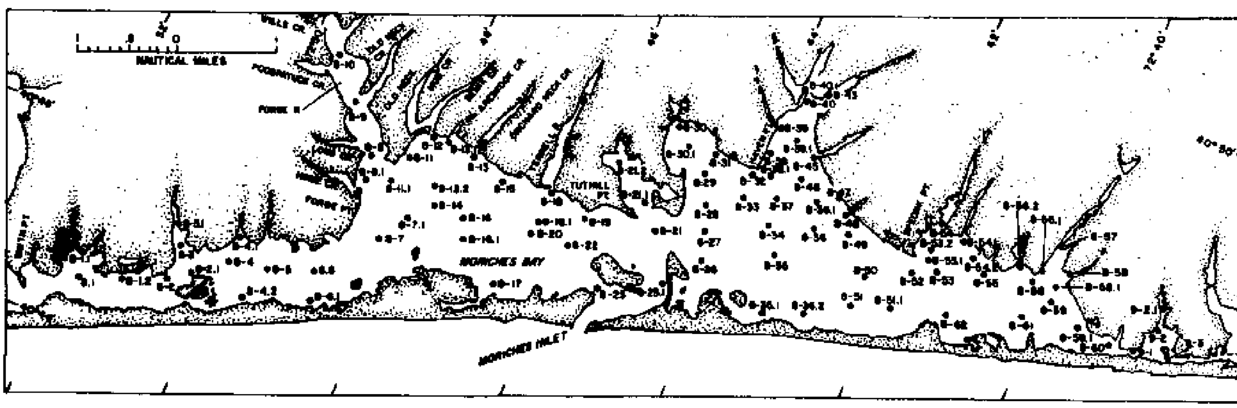
LIST OF TABLES

Table		Page
1	Summary of coliform data	2
2	Number of stations that pass or fail the four criteria for closing shellfish grounds. (Based on all data for five years. To qualify, a station must have at least 10 total coliform observations during the period)	4
3	Fecal to total coliform concentration ratios	9
4	Tidal analysis for station #4 area 3	11
5	Tidal variation of total coliform concentration	12
6	Number of days during the period 1974-1977 when N stations simultaneously fulfilled the rain impact criterion (0.25" or more rain on day or two previous days)	15
7	Rainfall analysis for area 46 stations 5 and 6	15
8	Approximate percentage of the coliform data for all observations that have concentrations greater than any dry day data. Total coliform TC, Fecal coliform FC	16
9	Characteristics of the effect of rainfall on the coliform distributions	22
10	Matrix transformation of five tube coliform data	23

ACKNOWLEDGEMENTS

The work herein reported could not have been carried out without the active support of the Shellfish Sanitation Group of the New York State Department of Environmental Conservation. I am particularly grateful for the assistance given by Bruce MacMillan and James Redman. I also thank A. S. Taormina, Director of Marine Resources, for assistance in coordinating the project.





Maps of Great South Bay and Moriches Bay showing stations used in this study.

INTRODUCTION

This report summarizes the results of a study of the shellfish sanitation data of the New York State Department of Environmental Conservation, carried out under funding from the New York Sea Grant Institute through the Research Foundation of State University of New York. The study covers data for the period 1973-1977 for the following shellfish growing areas:

Area 3, Great South Bay
Area 8, Moriches Bay
Area 29, Flanders Bay
Areas 40-46, Huntington Bay Complex

The study was under the direction of Prof. P. K. Weyl. Computer programming was carried out by Mr. George Carroll, assisted by Mr. Michael Carlin during the summer of 1978. The work on the project was seriously delayed by the delay in the receipt of a Datapoint 1500 terminal for entering the data into the computer. This equipment was ordered on April 28, 1978 but not received until November 8, 1978.

After the input terminal was received, a data entry program was developed and the data were entered by Mr. Paul Giroux, whose services were supplied by DEC. The data for the five years and four areas amounted to almost 30,000 coliform analyses. Professor Marco Retamal, Ms. Leslie Clarke and Mr. Alan Robbins assisted with the data analysis.

STATISTICAL CHARACTERISTICS OF THE COLIFORM DATA

The data on coliform concentrations are obtained by carrying out multiple tube fermentation tests. From 1973-1975, DEC used the five tube, decimal dilution technique using five samples each of 10, 1 and 0.1 ml. During 1976-1977, the same sample sizes were used but the number of tubes used for each decimal dilution was reduced to three.

The coliform data, both total and fecal, are summarized in Table 1. There were 12,672 five tube data for the period 1973-1975 and 16,603 three tube data for 1976-1977, giving a total of 29,275 observations. The distribution by most probable number (MPN) reveals the statistical nature of the multiple tube fermentation test. Some MPN values appear with great frequency while others are relatively rare. (The tabulation of the five tube data originally did not include the rare MPN values of 6 and 63 for the five tube test and these therefore were listed under adjacent numbers by the computer.) Column n in Table 1 lists the number of times that a given MPN occurred in the data. The column labeled P% lists the probability in percent that the particular MPN result would be obtained if the actual number of bacteria per 100 ml is equal to the MPN value. It therefore gives the highest possible probability of obtaining the specific MPN value.

The column n/P is the ratio of the number of occurrences of a given MPN value, divided by the maximum probability of its occurrence. As expected, this ratio varies relatively little compared to the large variability in individual MPN results. For example, for the five tube tests for MPN values between 23 and 34, the number of test results range from 14 to 887. The probabilities range from 0.3 to 17 percent but the ratios range only from 20 to 48, with a mean value of 32.5.

An attempt was made to sharpen the data by using a transformation matrix that takes the statistical character of the multiple tube fermentation tests explicitly into account. The transformation was applied to the five tube data from area 29. The results were disappointing, indicating that the actual distribution of bacterial concentrations is broad. If the actual distribution had a narrow spread about one or two well-separated peaks, the transformation would have indicated this. To resolve more complicated distributions by the transformation-matrix technique would have

Table 1. Summary of coliform data

Five tube tests 1973-1975					Three tube tests 1976-1977				
MPN <2	n 1,785	P%	n/P	cum % 100	MPN <3	n 3,792	P%	n/P	cum % 100
2	1,395	40.1	35	85.9	3	235	3.7	63	77.2
4	257	7.7	33		4	2,170	38.9	56	
5	912	26.2	35	72.9					62.7
6	0*	.69							
7	321	8.4	38		7	389	7.0	56	
8	709	21.8	33						
9	65	1.6	41		9	1,703	32	53	
11	336	11.4	30		11	38	.79	48	
12	6	.14	43						
13	508	18.7	27	54.4	14	35	1.2	29	49.8
14	122	3.3	37		15	558	11.9	47	
17	438	16.7	26		20	29	.62	47	
21	31	1.4	22		21	121	2.3	53	
22	145	6.2	23						
23	664	16.8	40	44.5	23	1,761	34.4	51	45.4
26	25	.8	31						
27	39	1.7	23		28	25	.16	156	
31	45	2.3	20						
33	887	26.9	33						
34	14	.29	48		39	118	3.1	38	
43	6	.17	35	31.3	43	1,734	37.6	46	33.9
46	109	5.2	21						
49	762	26.1	29						
63	1*	.58			64	4	.16	25	
70	176	7.9	22						
79	563	22.6	25		75	208	6.6	32	
94	20	1.3	15	18.6	93	1,173	32.8	36	22.2
109	151	11.0	14		120	14	.64	22	
130	343	20.0	17						
141	33	2.9	11		150	278	12.5	22	
172	183	16.1	11						
175	7	.48	15						
221	64	6.9	9		210	48	2.4	20	
240	334	19.0	18	12.3	240	921	36.8	25	13.1
278	20	1.9	11						
345	0	.34							
348	356	31.5	11						
542	260	33.7	8	6.6	460	616	42.9	14	7.5
918	196	34.6	6						
1,609	141	41.0	3	3.0	1,100	306	44.4	7	3.8
>1,609	243			1.9	>1,100	327			2.0
Total	12,672					16,603			

TOTAL 29,275

*These MPN numbers were originally left out and any tests falling into these unlikely categories were listed as adjacent MPN values.

required orders of magnitude more data from single stations.

Instead of using the transformation matrix, the statistical analysis of the data is based on cumulative percentages over intervals that include at least one highly probable MPN value for both the five and three tube tests. The intervals chosen are shown on Table 1. They are: 1)greater than 1,610.This class covers test results in which all tubes give a positive result. It thus combines the category of greater than 1,100 for the three tube test and greater than 1,609 for the five tube test; 2)greater than 1,099. This adds MPN values of 1,609 for the five and 1,100 for the three tube test; 3)greater than 459; 4)greater than 239; 5)greater than 92; 6)greater than 42; 7)greater than 22; 8)greater than 12; 9)greater than 4; and 10)greater than 1. The last category covers all results other than all tubes being negative. It excludes MPN values of less than two for the five tube test and MPN values less than three for the three tube test. The cumulative results for these categories are also shown in Table 1.

COMPARISON OF TOTAL COLIFORM AND FECAL COLIFORM CRITERIA

There are a total of four criteria for closing a shellfish harvesting area, two each based on total or fecal coliform tests. These are:

criterion 1 - the median total coliform MPN value is greater than 70

criterion 2 - more than 10% of the total MPN values exceed

230 for a five tube test

330 for a three tube test

criterion 3 - the median fecal coliform MPN value is greater than 14

criterion 4 - more than 10% of the fecal MPN values exceed

43 for the five tube test

49 for the three tube test

For each station for which there were at least 10 total coliform observations, these criteria were applied to the data. The results are coded as 0 for passing the test and 1 for exceeding the standard. The four criteria are coded in the order 1234. Beneath the four digits are listed the number of total and fecal coliform observations to which the criteria were applied. If the number of observations exceeded 99, this is indicated by **. Thus an entry

0111
**78

means that the results exceeded all but the median total coliform criterion and that there were more than 99 total coliform observations and 78 fecal coliform observations. Maps for each shellfish area for each year from 1973-1977 and for the five-year period have been prepared and are labeled FOUR WAY TC-FC TEST.

There are $2^4 = 16$ possible outcomes for combinations of the four criteria. Of these only nine occurred for the five-year interval taken as the data base. The results for the various areas are summarized in Table 2. A total of 261 stations had at least 10 total coliform observations during the period. Of these, 120 passed all criteria and 39 failed all four. Ninety-nine stations failed at least one of each of the total and fecal coliform criteria. That leaves 37 stations that failed one or both fecal coliform criteria and passed all total coliform criteria. Only five stations from areas 3 and 46 passed all fecal coliform criteria and failed the total coliform test. Thus a conversion from the total to the fecal coliform criteria would generally lead to a greater closure of shellfish areas. This applies to all areas except for area 46. Areas 40, 41, 44 and 45 of the Huntington complex passed all criteria for all stations.

Before discussing the other areas, it is useful to compare the failure rate for the median and 10% criteria. A total of 141 out of 261 stations failed at least one criterion. Of these, 78 failed at

Table 2. Number of stations that pass or fail the four criteria for closing shellfish grounds. (Based on all data for five years. To qualify, a station must have at least 10 total coliform observations during the period.)

Test Result* Criterion					SHELLFISH GROWING AREA										
#	1	2	3	4	3	8	29	40	41	42	43	44	45	46	ALL
0	0	0	0	0	16	36	16	20	17	3	4	3	4	1	120
1	0	0	0	1	5	19	6			1	2				33
2	0	0	1	0	1										1
3	0	0	1	1		1				1	1				3
4	0	1	0	0	2									3	5
5	0	1	0	1	8	5	8							3	24
6	0	1	1	0											0
7	0	1	1	1	5	18	5			2	4				34
8	1	0	0	0											0
9	1	0	0	1											0
10	1	0	1	0											0
11	1	0	1	1											0
12	1	1	0	0											0
13	1	1	0	1	1									1	2
14	1	1	1	0											0
15	1	1	1	1	12	6	10			3	2			6	39
Total					50	85	45	20	17	10	13	3	4	14	261
fail T & F test					26	29	23	0	0	5	6	0	0	10	99
fail T only					2	0	0	0	0	0	0	0	0	3	5
fail F only					6	20	6	0	0	2	3	0	0	0	37
fail 10% only					15	24	14	0	0	1	2	0	0	6	62
fail median only					1	0	0	0	0	0	0	0	0	0	1

* 0 = station passes criterion

1 = station fails criterion

least one median and one 10% criterion. A total of 62 failed one or both 10% criterion only, but only one station, number 17 in area 3, failed the median criterion only. The result for that station was based on only 10 observations.

Maps for the four criteria for the five-year period are shown in Figures 1-4. Fig. 1 shows the results for shellfish growing area 3 in Great South Bay. Two stations, #10 and 46 passed all fecal criteria but failed the total 10 percent criterion. Changing to the fecal criteria would increase the area of closure.

Fig. 2 shows the four criteria for shellfish growing area 8, Moriches Bay. Changing from the total to the fecal coliform criteria would significantly increase the area of closure. The same holds for areas 42 and 43 displayed in Fig. 3. Fig. 4 gives the results for areas 29 and 46. In area 29, Flanders Bay, changing to a fecal standard would significantly increase the area of closure. Three stations in that area, stations 9, 9.1 and 10 failed the 10% total criterion but passed all fecal criteria. These results were based on more than 100 separate determinations. On a single year basis, a result of 0100 was only obtained in 1974 and 1973.

Another way to examine the relationship between total and fecal coliform results is to construct a total coliform versus fecal coliform matrix. For each area (areas 40-46 are combined), we consider all water samples for which both a total and a fecal coliform test was run. We then tabulate the number of times that each total coliform MPN value coincided with each fecal coliform MPN value. Separate matrices are constructed for 1973-1975 when the five tube tests were run and 1976-1977 when three tube tests were

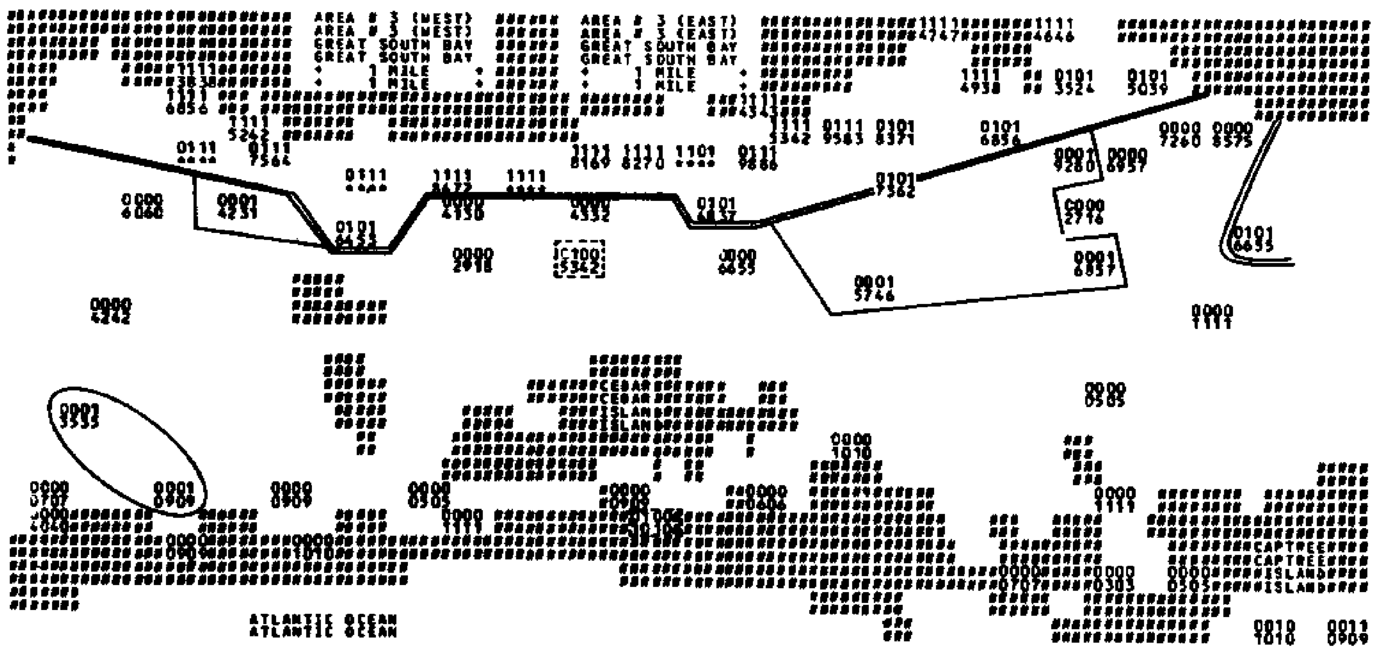


Fig. 1 Area 3, Great South Bay

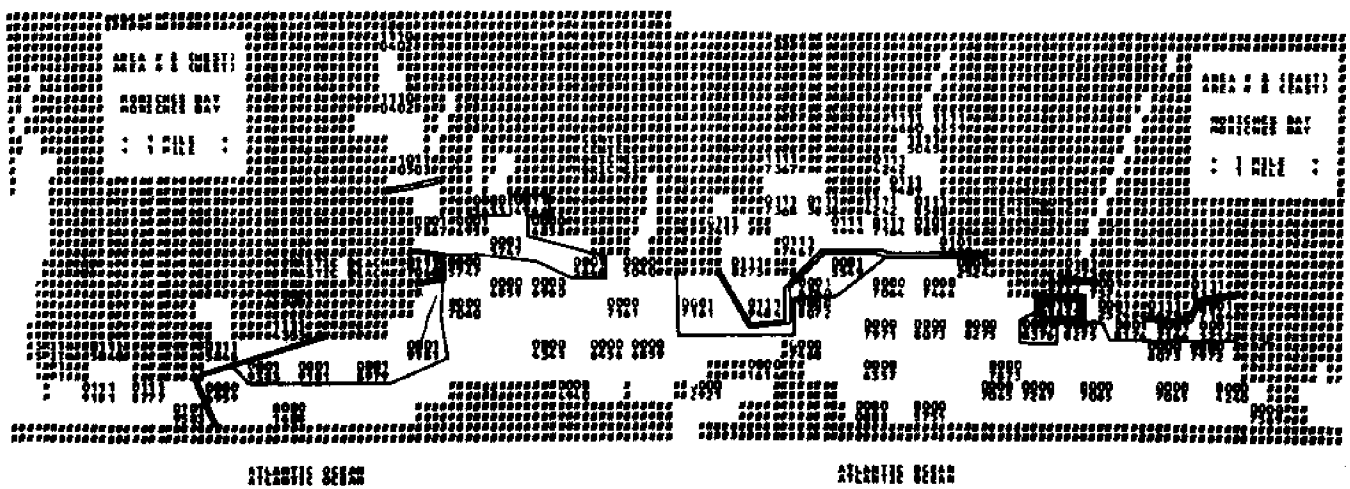


Fig. 2 Area 8, Moriches Bay

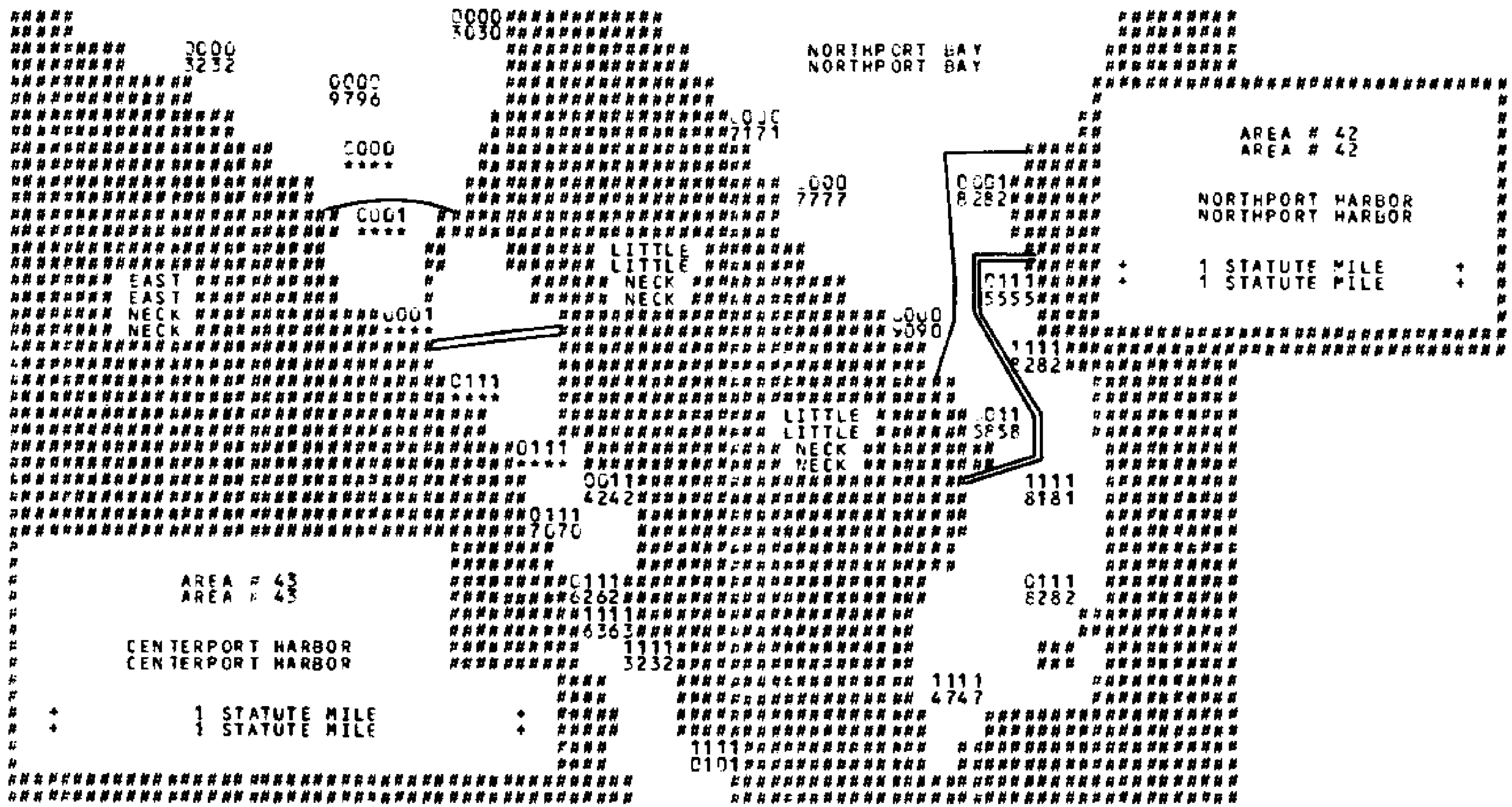


Fig. 3 Areas 42 and 43, Northport Harbor and Centerport Harbor

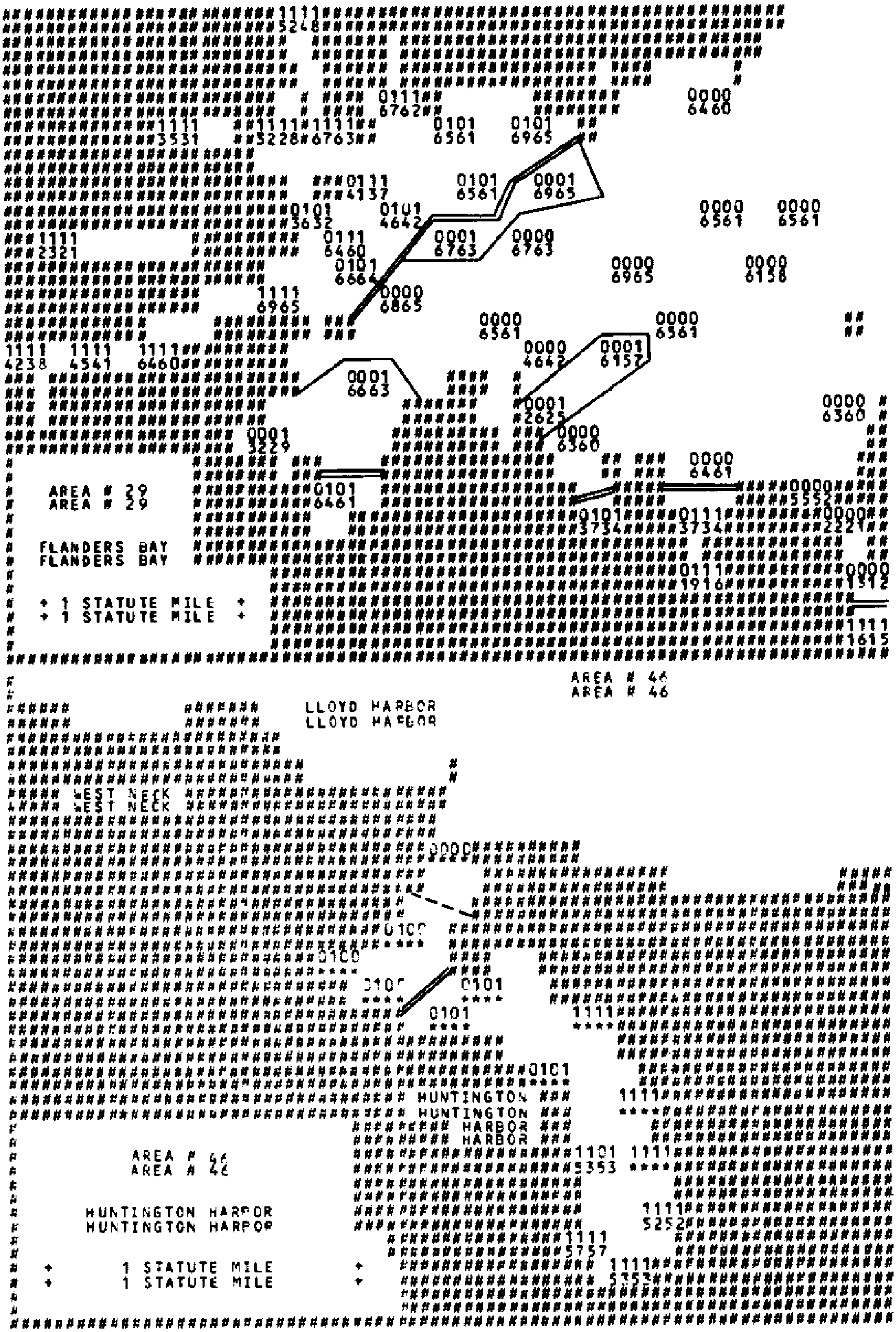


Fig. 4 Areas 29 and 46, Flanders Bay and Huntington Harbor

Table 3. Fecal to total coliform concentration ratios

Total MPN		SHELLFISH GROWING AREA 3	8	29	40 - 46
three tube test	49	.23	.33	.31	.29
	79	.24	.33	.32	.19
	240	.16	.30	.30	.13
five tube test	43	.12	.50	.42	.30
	93	.21	.39	.34	.28
	240	.23	.34	.28	.18
	460	.19	.44	.21	.17
geometric mean		.19	.37	.31	.21

GEOMETRIC MEANS OF RATIO AT DIFFERENT MPN

Total MPN	Ratio
49,43	.29
79,93	.28
240	.23
460	.23

RATIOS FOR CRITERIA

median criteria		0.20
10% criteria	three tube test	0.15
	five tube test	0.19

run. The statistics of the testing process complicate the distributions of data in the matrices.

To simplify the analysis, we have considered all tests that gave total coliform values of 49, 79 and 240 for the three tube tests and values of 43, 93, 240 and 460 for the five tube tests. For each area, we then determine the geometric mean of the fecal coliform value that corresponds to a given total coliform MPN value. The ratios of fecal to total coliform MPN are tabulated in Table 3. The ratios are close to 0.2 for areas 3 and 40-46 and larger than 0.30 for areas 8 and 29, indicating that the latter two areas have a relatively higher level of fecal coliform bacteria. Considering all the areas, the ratio of fecal to total coliform concentration seems to decline as the MPN value increases. This suggests a relatively higher ratio of fecal to total coliform bacteria at lower levels of contamination. One would expect this if the mortality of fecal coliform bacteria in the estuary is slightly less than that for total coliform bacteria.

TIDAL VARIATION OF BACTERIAL CONTAMINATION

On the average, one can expect that the coliform concentration at a specific station varies with the tide. This is because the station locations are fixed in space, but the waters that contain the bacteria move back and forth with the tide. By sampling the same location at different phases of the tide, one is effectively sampling waters with different mean tidal positions. The situation is actually more complex, because water in the bays moves not only in response to the tide but also as a result of atmospheric disturbances. Thus the astronomical tide is responsible for only part of the local variability in bacterial concentration.

The phase of the tide, the relative times of high and low water, differs in different bays and within each shellfish growing area. The dominant frequency of the tide in Long Island waters is the semi-diurnal lunar tide with a period of 12.42 hours. To simplify the tidal analysis, the semidiurnal lunar period was divided into 100 segments of 7.45 minute duration. The times when the moon crossed longitudes 0° and 180° was entered into the computer using information in the Nautical Almanacs for the years of interest. Each coliform data set contains the time (in Eastern Standard Time) that the water sample was taken. The computer was programmed to convert that time into a segment of the lunar period. A tidal value of 00 means that the water sample was taken within 3.7 minutes of the time that the moon made an upper or lower crossing of the Greenwich meridian. The tidal code is tabulated for each observation on the station data sheets.

To illustrate how the tidal variation of the coliform concentrations was determined, consider the data for station #4 of area 3. During the five-year period there were 100 total coliform determinations for this station. The statistical nature of the multiple fermentation tube tests makes it necessary for one to analyze grouped data. The cumulative percentage distribution of MPN values was therefore obtained for ranges of 50 units in the tide code (for one half of the semidiurnal lunar period). Overlapping half periods, each shifted by 10 percent of the the tidal period were used, from 00-49; 10-59; 20-69 and so on. The print out for station #4 area 3 is shown in Table 4.

The first task is to determine the phase of the tide that corresponds to a maximum coliform concentration. Using the statistics for all the data, 00-99, one locates the approximate median value. Sixty-one percent of all data had MPN values greater than 42. For the half-tidal intervals, the percentages ranged from 71 to 49. One now selects two exclusive half-tidal intervals that represent the highest and lowest concentrations. These are 20-69 for high and 70-19 for low. Selection involves some subjective judgement. For example, if one had examined the greater than 92 line in the distribution, the highest percentage (45) occurs in the interval 50-99 and the lowest value (31) falls in the 70-19 interval. These cannot be selected as the high and low interval, since they overlap. Station data were only used if exclusive high and low half-cycles can be reliably assigned. In many cases no unambiguous assignment was possible. This can be due to insufficient data or because the actual variation in concentration with tidal phase is very small.

For station #4 area 3, there was no problem. The distributions vary fairly regularly with tidal phase and the high level of the concentration is reached in the interval 20-69, so that the concentration peaks at a tidal code of 45, with an error of ±10.

The next task is to estimate the amplitude of the tidal variation of the coliform concentration. This is done by estimating the median MPN value for the high and low concentration intervals. For the high interval, 20-69, the median falls between MPN values of 42 and 92, the median being a fraction (71-50)/(71-42) = 0.72 towards the higher value 92. A logarithmic interpolation of the MPN value gives 74 for the median. Similarly, the median for the low interval 70-19 falls between MPN values of 22 and 42. A logarithmic interpolation gives a value of 41. The ratio of these values is 1.8. This is the ratio between the average concentrations over the intervals 20-69 and 70-19. If the relative bacterial concentrations vary sinusoidally, then the average ratios must be enhanced by a factor of $\pi/2$. The peak to peak tidal ratio is obtained by the following equation:

$$\text{peak to peak ratio} = (\text{Obs. ratio} - 1) \times \pi/2 + 1$$

For station #4 area 3 we obtain $(1.8 - 1) = 0.8$; $0.8 \times \pi/2 = 1.26$; $1.26 + 1 = 2.26$. Thus on the average, the coliform concentration would be 2.3 times as great when the tide code is 45, relative to the concentration at a tide code of 95.

Table 4. Tidal analysis for Station #4 area 3

Station #4, area #03, 1973-1977, TC vs. TIC

	0 to 49	10 to 59	20 to 69	30 to 79	40 to 89	50 to 99	60 to 9	70 to 19	80 to 29	90 to 39	0 to 99	Total
>1610	3	4	4	4	2	0	0	0	0	2	2	2
>1099	5	5	5	7	2	3	2	2	2	5	4	2
>459	12	13	9	13	9	13	11	16	11	14	12	8
>239	15	16	16	22	20	25	23	22	16	18	19	7
>92	32	41	42	42	43	45	32	31	33	32	37	18
>42	60	70	71	67	64	63	50	49	56	59	61	24
>22	73	80	80	80	77	75	66	67	69	71	74	13
>12	77	84	87	89	89	85	75	71	73	73	80	6
>4	90	91	93	96	95	90	89	87	85	86	90	10
>1	92	95	98	100	100	100	95	91	91	91	95	5
	100	100	100	100	100	100	100	100	100	100	100	5
Total	60	56	55	45	44	40	44	45	55	56	100	100

The data for the tidal variation of total coliform concentrations for all stations where they could be determined are given in Table 5. For each area, the table lists the station number, the phase at which the maximum total coliform concentration occurs and the peak to peak ratio in concentration. The data for each area are then summarized by noting how often each tidal phase occurs for groups of peak to peak ratios. For each area, the approximate time of high and low tide is also indicated and the number of observations for the low and high half-tidal cycle when the ratios were 2 or greater are shown.

Altogether, there are phase and ratio data for 160 stations. For 129 of these, the ratio is 2 or greater.

Discussion

Area 3, Great South Bay

The tide in this area shows relatively little phase shift. High tide occurs at a tide index of 24 and low tide at 74. The phases at which higher coliform values occur vary significantly over the area. Near the mainland shore, high values occur more often at low tide, whereas high values away from the shore generally occur more frequently near high tide. Altogether, twice as many stations show ratios of 2 or greater in the high half-tidal cycle.

Area 8, Moriches Bay

The phase of the tide in this bay is somewhat more variable. Just inside Moriches Inlet, low tide occurs at a tidal index of 65 and then occurs progressively later away from the inlet to reach index values of about 85. An index of 75±10 fairly well represents low tide for most of the bay. The phase when the relative coliform values are high occur mainly at low tide with tidal indexes between 65 and 95.

Area 29, Flanders Bay

The tide in Flanders Bay shows relatively little shift in phase, with high tide occurring at an index of 35 and low tide at an index of 85. The phases of high relative coliform values are fairly well-distributed throughout the tidal cycle but the distribution is not random. High values occur about low tide near Riverhead, in the central part of the Bay and in inlets on the South shore. In the north and to the east, the peaks tend to occur near high tide.

Areas 42,43, 53 and 46, Huntington Bay Complex

In these North Shore bays, the tide is essentially synchronous with the tide

Table 5. Tidal variation of total coliform concentration

AREA #3 GREAT SOUTH BAY

Station Data

Station	Phase	Ratio	Station	Phase	Ratio	Station	Phase	Ratio	Station	Phase	Ratio
3.1	95	1.5	4	45	2.3	4.1	65	3.2	20	45	4.8
22	35	4.0	23.1	05	2.1	23.2	15	2.0	24.1	25	3.8
26	45	3.4	26.1	05	3.0	30.0	75	2.3	33.0	35	3.3
33.1	65	1.6	33.2	25	3.9	34	25	2.6	34.1	45	2.2
34.2	95	11.8	35	65	5.3	36.1	25	2.5	37	55	2.3
37.1	25	1.8	37.2	25	2.4	40	95	3.3	41	05	1.4
41.1	75	2.7	44	85	2.9	45	05	5.4	46	55	1.9
47	05	2.3	50	95	1.5	52	95	2.3	54	25	4.1
55	55	1.4	55.1	25	3.7	59	15	4.8	59.1	75	2.6
59.2	05	6.3									

Summary

RATIO	PHASE 05	15	high 25	35	45	55	65	low 75	85	95	total
1- 1.9	1		1			2	1			2	7
2- 3.9	3	1	6	1	3	1	1	3	1	2	22
4- 7.9	2	1	1	1	1		1			-	7
8-15.9										1	1
16+											0
total	6	2	8	2	4	3	3	3	1	5	37
≥2	5	2	7	2	4	1	2	3	1	3	30
			20					10			

AREA #8 MORICHES BAY

Station Data

Station	Phase	Ratio	Station	Phase	Ratio	Station	Phase	Ratio	Station	Phase	Ratio
1	85	2.9	1.1	55	1.4	1.2	95	2.9	2	85	2.3
2.1	75	8.3	3	95	3.2	4	75	8.4	4.1	55	3.8
5	75	7.9	6	75	6.2	7	75	7.9	7.1	75	15.5
8	95	2.7	8.1	65	1.6	11	75	3.8	11.1	75	2.9
12	95	4.1	13	75	6.1	13.1	25	1.7	13.2	85	1.6
14	65	16	15	65	16	16	75	5.8	16.1	85	3.0
17	65	7.8	18	65	2.5	19	65	14	20	25	7.7
21	85	6.7	21.1	65	3.9	22	55		26	25	1.7
27	65		28	75	4.7	29	05	3.1	30	35	4.8
30.1	85	3.2	32	85	1.6	33	75	12	34	65	
35	75		36	85		36.1	75	2.7	36.2	05	
37	75	8.3	38	25	7.2	38.1	95	2.1	39	55	1.2
39.1	55	2.6	40	55	6.5	40.1	75	2.7	43	85	3.1
45	85	2.3	46	85	3.4	49	95	3.4	50	05	
51	25		51.1	15		52	85	4.5	53	85	7.9
54	75	3.6	54.1	85	2.9	55	85	4.8	56	85	5.7
56.1	75	2.9	56.2	85	3.1	57	95	2.8	58	85	6.3
59	75	7.4	59.1	05	3.7	60	95	2.8	53.2	85	9.0

Summary

RATIO	PHASE										
			high					low			
	05	15	25	35	45	55	65	75	85	95	total
1- 1.9			2			2	1		2		7
2- 3.9	2					2	2	6	9	7	28
4- 7.9			2	1		1	1	7	6	1	19
8-15.9							1	4	2		7
16+							2				2
total	2	0	4	1	0	5	7	17	19	8	63
≥2	2	0	2	1	0	3	6	17	17	8	56
			5					51			

AREA #29 FLANDERS BAY

Station Data

Station	Phase	Ratio	Station	Phase	Ratio	Station	Phase	Ratio	Station	Phase	Ratio
1	35	5	2	25	4.0	3	55	1.5	4	25	1.3
5	55	3.6	5.1	25	1.6	6	25	2.0	7	75	1.7
9	65	1.8	9.1	85	1.8	10	05	2.8	11	85	1.5
12	35	2.0	13	95	4.8	13.1	75	9.2	15	95	5
17	15	4.2	18	95	6.7	19	15	16	19.1	95	4.5
22	05	3.7	23	55	4.3	24	85	8.3			

Summary

RATIO	PHASE										
			high					low			
	15	25	35	45	55	65	75	85	95	05	total
1- 1.9		2			1	1	1	2			7
2- 3.9		1	1		1			1		2	6
4- 7.9	1	1	1		1				4		8
8-15.9							1	1			2
16+	1										1
total	2	4	2	0	3	1	2	4	4	2	24
≥2	2	2	2	0	2	0	1	2	4	2	17
			8					9			

AREAS #40-46 HUNTINGTON COMPLEX

Station data

Area #42			Area #43			Area #45			Area #46		
Station	Phase	Ratio	Station	Phase	Ratio	Station	Phase	Ratio	Station	Phase	Ratio
1	75	2.6	1	85	2.6	1	05	6.4	1	75	2.3
2	95	2.6	2	75	7.1	2	05	2.1	2	35	4.6
3	15	1.6	2.1	75	2.6	3	45	2.3	3	25	1.3
4	05	2.0	3	75	2.5				4	15	2.8
4.1	85	1.9	4	75	2.9				5	25	1.5
5	95	1.8	5	85	1.7				6	65	1.9
5.1	85	5.1	6	65	2.7				7	65	3.2
6	05	3.5	7	65	1.7				8	85	3.8
7	75	2.2	8	35	3.2				8.1	75	1.6
			9	35	2.2				8.2	75	3.1
									9	75	3.3
									9.1	85	1.3
									10	65	3.4
									11	85	6.3

Table 5. (continued)

Summary

RATIO	PHASE 5	15	high 25	35	45	55	65	low 75	85	95	total
1- 1.9		1	2				2	1	3	1	10
2- 3.9	3	1		2	1		3	8	2	1	21
4- 7.9	1			1				1	2		5
8+											0
total	4	2	2	3	1	0	5	10	7	2	36
≥2	4	1	0	3	1	0	3	9	4	1	26
			9					17			

in Long Island Sound. High tide occurs at a tidal index of 22 and low tide occurs at an index of 72. High relative coliform values tend to occur predominantly at low tide in these areas.

THE EFFECT OF RAINFALL ON COLIFORM CONTAMINATION

A significant fraction of the rainfall that falls in coastal areas is carried as storm runoff into adjacent marine waters. In the process, the runoff water is contaminated by bacteria from the soil, roads and other impervious surfaces. The bacterial concentration in runoff varies with the duration and intensity of the rainfall event and also depends on the length of the dry period that preceeded the rainfall event. Originally, it was planned to analyze the temporal relationship between rainfall events and the coliform bacterial concentrations in detail. An examination of rainfall data and of the statistical variability of the coliform results, however, showed that such an analysis for the available data base would be inconclusive.

We obtained daily precipitation data from eight weather stations in the Long Island area (see Table 6) for the years 1974-1977 on magnetic tape from the National Weather Service. The experience of the shellfish sanitation program of the New York State Department of Environmental Conservation suggests that rainfall events can have an impact on coliform levels for up to two days after a rainfall event, if the total precipitation was 0.25 inches or more. Using these criteria, we analyzed the precipitation records for the eight weather stations and labeled each day for each station as being affected by rain, R, or not being affected by rain, N. For each day, we summed the number of rain-affected stations. (A few gaps in the record led to a designation of X).

The precipitation record indicates that many rain events are patchy. Out of the four-year record, 42.6 percent of the days showed no rainfall impact at any of the stations and 15.4 percent of the days had rainfall impact at all the stations (Table 6). That leaves 42 percent of the days when some but not all stations were impacted by rainfall. Because of this patchiness, the data do not permit a positive identification of all days, when a particular shellfish area was impacted by runoff. Such an impact can be assumed on "universal rain days", days when all eight stations were registered as R. There were 223 such days, but unfortunately, the number of water samples collected on these days were insufficient to warrant a detailed analysis.

Almost half the coliform observations coincided with "universal dry days", days when none of the weather stations indicated a rainfall impact. During these days, it is unlikely that any of the shellfish growing areas were impacted by rainfall.

Table 6. Number of days during the period 1974-1977 when N stations simultaneously fulfilled the rain impact criterion (0.25" or more rain on day or two previous days)

N	no. of days	percent
0	623	42.6
1	124	8.5
2	74	5.1
3	95	6.5
4	71	4.9
5	80	5.5
6	63	4.3
7	108	7.4
8	223	15.4
Total	1461	100

The eight weather stations are:

New York Ave., Brooklyn; N.Y. Central Park; Greenport Power House; N.Y. Kennedy Airport; N.Y. La Guardia Airport; Montauk; Patchogue and Setauket.

Table 7. Rainfall analysis for area 46
Stations 5 and 6

Coliform vs. station, 1974-1977, area 46

	5				6			
	TC ALL	TC DRY	FC ALL	FC DRY	TC ALL	TC DRY	FC ALL	FC DRY
>1610	8	0	0	0	9	0	1	0
>1099	12	0	1	0	15	5	1	0
>459	21	2	5	0	25	9	7	0
>239	41	16	10	0	34	14	13	2
>92	67	44	18	2	58	42	23	5
>42	78	58	36	17	75	63	33	14
>12	88	77	49	21	88	79	47	29
>4	91	81	60	36	91	86	58	45
>1	100	100	78	64	96	91	74	62
	100	100	90	86	99	98	88	83
	100	100	100	100	100	100	100	100
Total	106	43	105	42	105	43	104	42

We therefore generated a statistical analysis for the period 1974-1977, that contrasts ALL data with data limited to universal DRY days. A copy of the print-out for stations 5 and 6 of area 46 is shown in Table 7. Cumulative percentages for ALL and DRY days are shown for total coliform (TC) and fecal coliform(FC) bacteria. The print-outs clearly show a reduction in the coliform concentrations during "universal dry days".

To summarize the impact of rainfall on the coliform concentrations for each station, we developed a simple rainfall index. One locates the lowest MPN value for which the cumulative DRY percentage is 10 or less. One then subtracts the percentage from the percentage of ALL data in the same MPN row. For example, for station 6 area 46 one obtains 25 - 9 = 16 for the total coliform difference. The corresponding value for fecal coliform is 23 - 5 = 18. The index is a measure of the percentage of ALL data that lies above the DRY data in MPN concentration. Values of this index for all stations for which there were at least 10 DRY total coliform observations during the four-year period are given in Table 8. With very few exceptions, all stations show a positive excess

The probability that a particular shellfish growing area is affected by

Table 8. Approximate percentage of the coliform data for all observations that have concentrations greater than any dry day data. Total coliform TC, Fecal coliform FC

Area 3

Stn	TC	FC
3	18	8
3.1	5	11
4	11	1
4.1	9	5
5	9	4
5.1	9	10
6	10	10
20	16	6
22	29	11
23.1	8	13
23.2	10	8
24.1	31	23
24.2	6	18
26	14	4
30	25	25
33	22	18
33.1	26	24
33.2	22	21
34	14	3
34.1	15	12
34.2	32	24
35	6	2
36	9	9
36.1	19	29
37	35	20
37.1	22	13
37.2	9	18
40	12	2
41	21	10
41.1	17	4
44	7	13
46	5	2
47	16	5
50	12	7
52	8	7
55	6	2
55.1	12	17
59	6	1
59.1	9	6
59.2	2	5

Area 8

Stn	TC	FC
1	2	10
1.1	19	8
1.2	8	22
2	5	3
2.1	9	20
3	17	22
4	15	11
4.1	13	27
5	1	0
6	15	5
7	5	3
7.1	9	10
8	7	21
8.1	5	15
11	19	22
11.1	9	0
12	7	4
13	9	9
13.1	3	-1
13.2	10	9
14	10	7
15	4	9
16	6	3
16.1	7	4
17	18	3
18	-2	-4
18.1	4	5
19	6	4
20	12	12
21	9	6
21.1	8	8
21.2	33	53
22	8	5
23	7	6
26	3	5
27	8	7
28	3	3
29	5	6
30	5	6
30.1	8	7
31	-3	-3

Stn	TC	FC
32	6	2
33	13	6
34	7	6
35	8	7
36	5	4
36.1	5	2
36.2	1	1
37	3	0
38	6	13
38.1	9	9
39	8	13
39.1	8	4
40	10	6
40.1	15	12
43	10	7
45	9	7
46	5	9
47	10	13
48	3	7
49	6	3
50	3	4
51	4	1
51.1	8	6
52	13	11
53	14	9
53.1	5	5
53.2	8	-2
54	8	2
54.1	15	12
54.2	10	15
55	11	14
56	15	14
56.1	5	7
56.2	12	10
57	6	5
58	10	13
59	4	5
59.1	19	9
60	13	5
61	16	9
62	9	8

Area 29

Stn	TC	FC
1	20	9
2	7	9
3	2	22
4	7	8
4.1	5	2
5	14	8
5.1	8	5
5.2	14	14
5.3	7	19
6	17	14
6.1	10	7
7	0	4
7.1	9	14
8	13	9
8.1	16	16
8.2	12	15
9	5	-1
9.1	11	12
10	5	9
10.1	17	12
11	10	22
12	18	12
12.1	4	8
12.2	8	5
13.	11	9
13.1	11	9
13.2	9	9
14	19	8
15.	17	2
16.	20	11
16.1	7	6
17.	5	7
17.1	25	10
17.2	12	4
18.	8	8
19.	12	15
19.1	15	12
19.2	9	1
20	13	12
21	15	13
22	17	22
23	5	4
24	12	11

Areas 40-46

Area	Stn	TC	FC
40	1	7	-1
42	1	5	19
	2	-1	-1
	3	20	11
	4	1	12
	4.1	25	23
	5	19	6
	5.1	5	12
	6	7	5
	7	20	7
	8	9	14
43	1	11	12
	1.1	5	1
	1.2	9	-1
	2	13	10
	2.1	15	10
	3	13	10
	4	13	14
	5	7	2
	6	16	10
	7	9	4
	8	35	20
	9	4	15
45	2	18	29
	3	11	-2
46	1	21	18
	2	17	18
	3	25	26
	4	8	18
	5	19	16
	6	16	18
	6.1	23	33
	7	14	9
	8	20	14
	8.1	8	15
	8.2	10	9
	9	17	23
	9.1	12	8
	10	9	8
	11	17	3

runoff during a "universal dry day" is about 1 percent (1/8 of the probability of only one out of eight weather stations experiencing rainfall). At any one location, however, the total number of dry days greatly exceeds the number of "universal dry days". If the fraction of the total observations at any given station during which dry conditions existed were known, one can estimate the distribution of coliform concentrations during runoff-affected days. If the fraction of wet days is w, then the distribution for ALL days = (1 - w)times the distribution DRY days + w times the distribution for WET days. Since the ALL and the DRY distributions are known, one can calculate the WET distribution by difference.

There are two problems with the above scheme. First, the statistical variability of the observations for any single station is too great to permit a reliable estimate of the WET distribution by difference. Second, we do not know the fraction of WET days, w.

The statistical problem can be overcome by aggregating all the data from each shellfish growing area (areas 40-46 are considered as one area). This gives us at least 1,000 ALL days per area, for stations that have at least 10 DRY total coliform observations. The ALL and DRY distributions for the four areas are shown in Figs. 5-8. To estimate the value of w, the fraction of the ALL observations that correspond to WET days, one can make use of two limits. The number of WET days cannot exceed the difference between ALL days and DRY days. A minimum value for the number of WET days can be obtained by finding the value of w that will give zero percent for one of the low concentration increments in the statistical distribution of the WET data. Any smaller value of w would lead to negative values for the increment.

In all but one case, the lowest possible value of w produced a contribution of zero to the interval, less than 1. The lone exception was the total coliform data for area 3, where the zero occured in the interval from 4-12 MPN. Thus the minimum value of w is valid, as long as rainfall-affected coliform determinations never produce negative results for all tubes of the multiple fermentation tube tests. In order that less than 1 percent of the tests give this result, the actual coliform concentration must be more than 9 bacteria per 100 ml for the five tube test and more than 16 for the three tube test.

The WET distributions of coliform concentrations were calculated using this minimum value for w. This assumption resulted in percentages of WET days ranging from 20 to 40. In contrast, the maximum values of w ranged from 40 to 60. The WET distributions using minimum values of w are also plotted in Figs. 5-8. In all cases, the WET distribution showed significantly higher coliform bacterial concentrations. An index of this higher concentration level is the ratio of the WET to the DRY concentration at the median. These ratios as well as other information about the distributions are given in Table 9. Areas 8 and 29 show the greatest increase in contamination due to runoff, whereas the combined areas 40-46 show the least effect. The significance of these differences is not clear, because of the strong correlation between the median ratios and the percentage of wet days assumed. This is shown in Fig. 9. As would be expected, low values of w lead to large values for the median ratio. The differences between areas at similar values of w are rather small.

CONCLUSIONS

The original objectives of this study were:

1. To determine the statistical significance of coliform data.
2. To make a statistical comparison between coliform and fecal coliform data.
3. To determine the relative importance of point and non-point sources of bacterial contamination on the closing of shellfish areas.

Statistical Aspects

Computerizing the shellfish sanitation data for four areas for a period of five years has made it possible to examine the statistical nature of the multiple tube fermentation data in detail. Of the 12,672 five tube data, 50 percent fall into the six most abundant MPN classes <2, 2, 5, 33, 49 and 8. Of the 16,603 three tube data, 57 percent fall into the four most abundant MPN classes, <3, 4, 23 and 43. The clustering of the data, a result of the statistical nature of the multiple tube fermentation tests, seriously limits the information content. Interpretation is further complicated by the large variability introduced by the tides and by rainfall events. On the average,

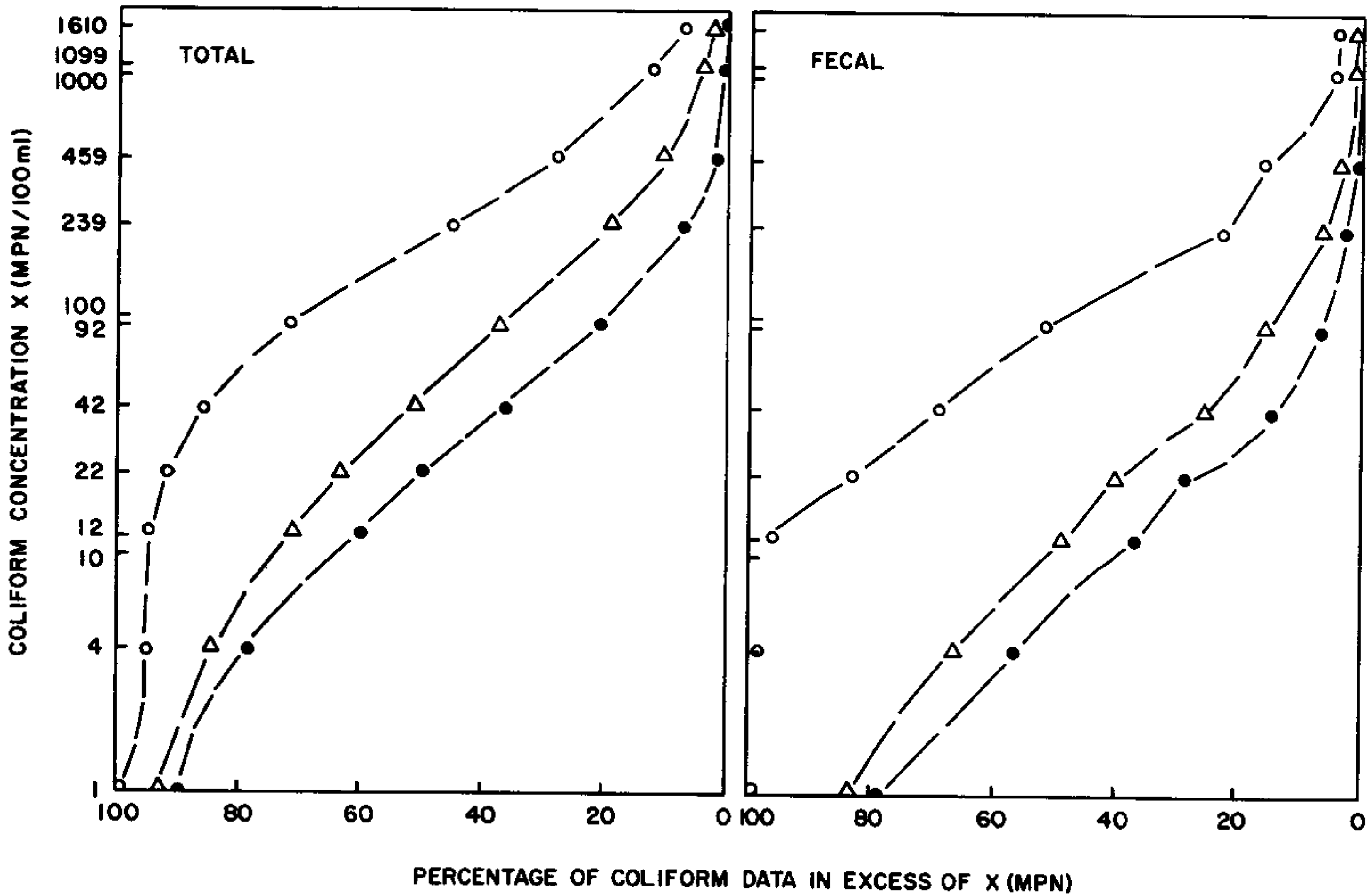


Fig. 5 WET(○) DRY(●) and ALL(△)Coliform distributions for Area 3, Great South Bay

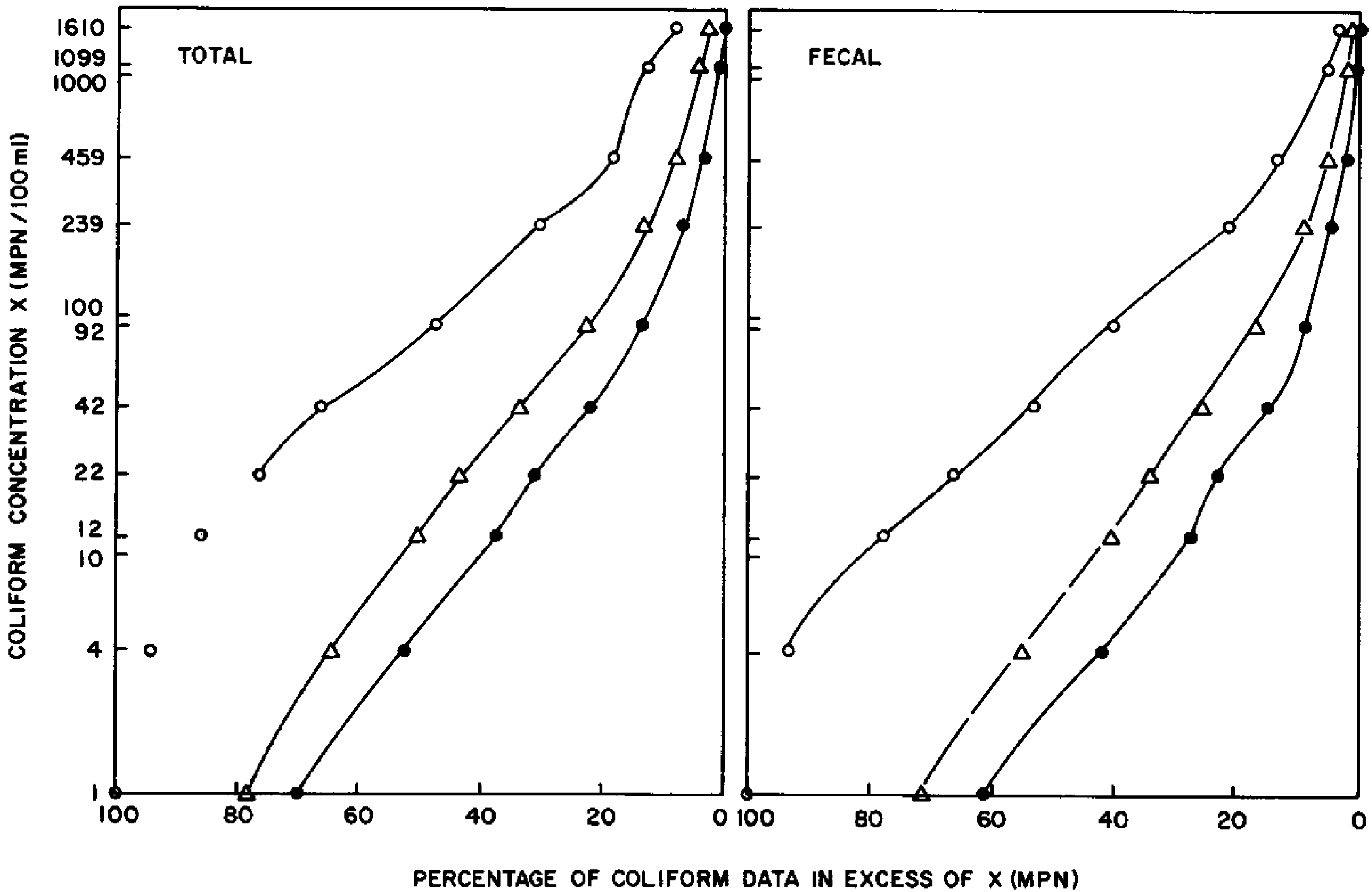


Fig. 6 WET(○)DRY(●)and ALL(△)Coliform distributions for Area 8, Moriches Bay

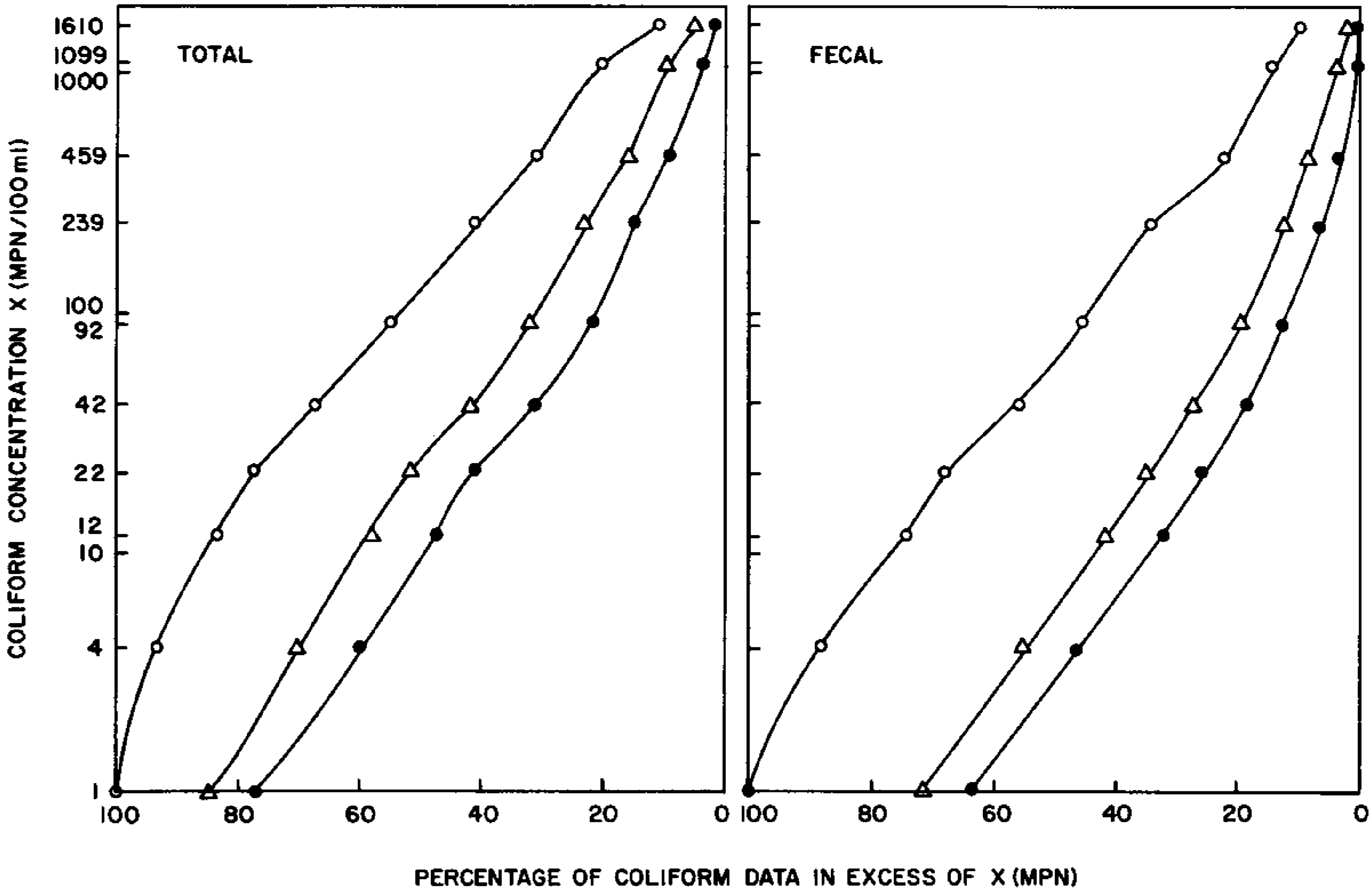


Fig. 7 WET(○)DRY(●)and ALL(△)Coliform distributions for Area 29, Flanders Bay

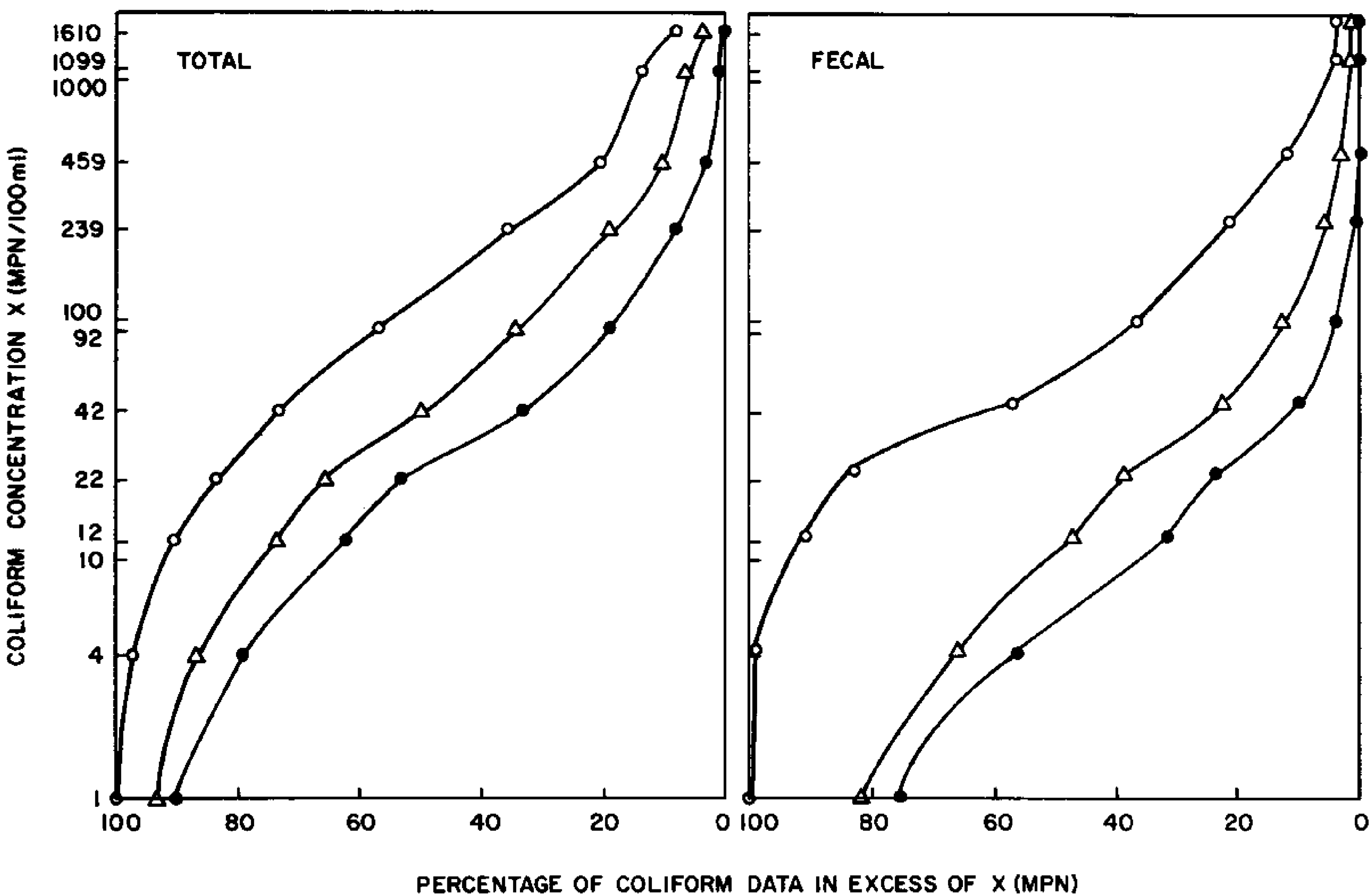


Fig. 8 WET(○)DRY(●)and ALL(△)Coliform distributions for Areas 40-46, Huntington Bay Complex

Table 9. Characteristics of the effect of rainfall on the coliform distributions

Area #	total T or fecal F	number of ALL observations	value of w fraction wet min.	max.	median ratio WET/DRY
3	T	2,100	.316	.595	9
	F	1,747	.200	.575	17
8	T	4,770	.267	.455	16
	F	4,207	.256	.463	21
29	T	1,891	.304	.452	14
	F	1,740	.216	.405	22
40-46	T	2,648	.400	.598	5
	F	2,624	.250	.601	12
All areas combined	T	11,409	.318	.513	12
	F	10,318	.273	.507	12

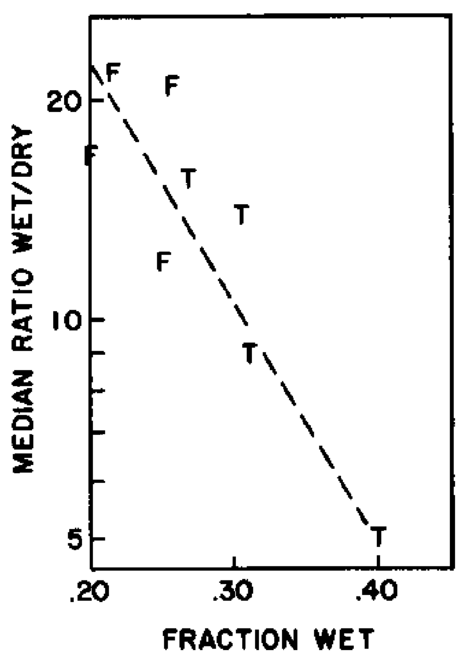


Fig. 9 Relationship between the median ratio of wet to dry coliform concentrations and the estimate of the fraction of WET observations. T = total coliform, F = fecal coliform

these causes introduce variations in the median values by a factor of 3 and a factor of 10 respectively.

The original intention was to utilize a transformation matrix in order to reduce the statistical variability introduced by the multiple tube fermentation test procedure. The transformation maps the MPN results into a set of nine, geometrically spaced discrete concentrations from 2.2 to 1,000. Two additional categories, 0 and ∞, complete the series.

Table 10. Matrix transformation
of five tube coliform data

#	concentration	Area 29 Stn. 3	Area 29 Sta. 4.1	all five tube data number	all five tube data %
1	0	-2.4	-1.1	932	7.4
2	2.2	8.3	11.4	2,713	21.4
3	4.6	.8	-2.2	1,083	8.5
4	10	3.0	1.1	1,902	15.0
5	22	-1.6	12.7	1,676	13.2
6	46	23.4	3.8	1,908	15.1
7	100	-8.5	3.9	619	4.9
8	220	7.8	.1	972	7.7
9	460	-2.3	1.5	329	2.6
10	1,000	7.3	1.1	349	2.8
11	∞	.3	-.1	205	1.6
Total count		36	32	12,672	

Examples of the application of the transformation to two typical stations from area 29 are shown in Table 10. The large fluctuations in the resultant distribution and the presence of large negative values indicated that the number of data available for each station are grossly inadequate to yield useful results.

The transformation was also applied to all five tube data, a total of 12,672 observations. Figure 10 shows the cumulative distribution one obtains. Since the transformation clusters the concentrations at discrete values of 0, 2.2, 4.6, etc., the cumulative distribution curve consists of a series of steps. The statistical distribution obtained by using the standard MPN intervals of Table 1 is also indicated in the figure. The two distributions agree quite well. The main discrepancy is that the transformation reduces the zero category (less than 2 in MPN value) from 14.1 percent to 7.4 percent. This is because a bacterial concentration of 2.2 per 100 ml would give an all negative result for 30 percent of the tests.

Comparision of Total and Fecal Coliform Data

The study showed that a conversion from a total to a fecal coliform standard would generally lead to a greater closure of shellfish growing areas. The ratio of the fecal to total coliform concentrations was greatest in area 8 (0.37), somewhat less in area 29 (0.31) and least in areas 3 and 40-46 (0.19, 0.21). As the bacterial concentrations decrease, the ratio of fecal to total coliform increases slightly, suggesting a slightly slower die-off rate for fecal coliforms. The correlation between total and fecal coliform determinations appears to be as good as could be expected given the statistical nature of the fermentation tube tests. Of the 37 stations that passed the total coliform criteria but failed one or both fecal coliform criteria, four were based on too few observations to be meaningful (12 or less). The other 33 stations had an average of 10.1 percent of the observations in excess of 239 bacteria per 100 ml with a standard deviation of 3.5.

The station showing the greatest discrepancy between fecal and total coliform concentrations was station 15 of area 29. During the four years when there were more than two observations, the station failed the fecal criteria twice and never failed the total coliform criteria. Overall, only three percent of the total coliform determinations exceeded 239 MPN. Station 12.1 close to station 15 was similar but did not have as low total coliform values.

With the possible exception of the above two stations from area 29, Flanders Bay, the total and fecal coliform distributions show similar areal distributions

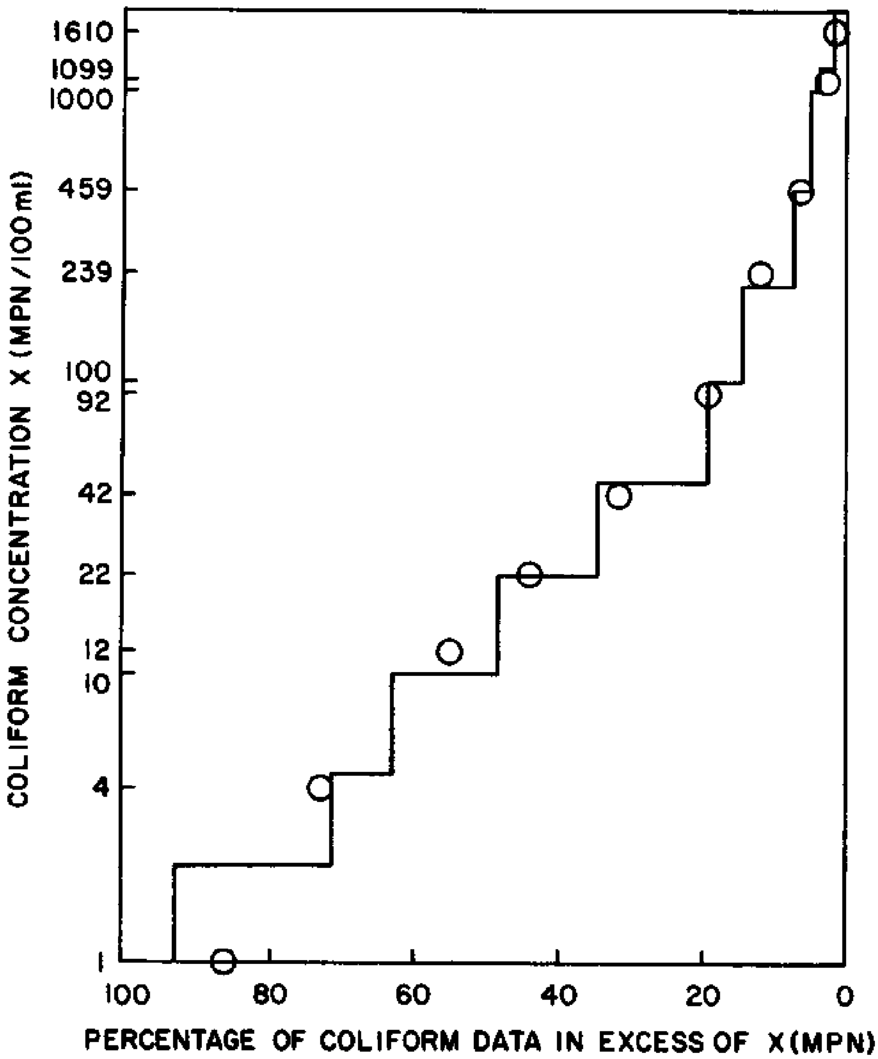


Fig. 10 Comparision of the cumulative distribution of all five tube tests using the transformation matrix (the steps) and using the clustered data (the circles)

of bacterial contamination. A change from the total to the fecal criteria would lead to an increase in the area closed to shellfishing. The additional closures would generally be adjacent to areas closed under the total coliform criteria.

The Effect of Storm Runoff on The Coliform Distribution

It was not possible to make detailed spatial and temporal studies of the effect of storm runoff on the coliform concentration. A statistical study comparing all data with data obtained on days when none of the weather stations in the area indicated a rainfall impact demonstrates that rainfall events significantly increase the coliform concentration. The results suggest that the median coliform concentration on rain impact days is about ten times that on days having less than 0.25 inches of rainfall and that are preceeded by at least two other dry days.

To estimate the statistical impact of storm runoff, we can consider the total coliform data for area 3, which gave intermediate results. Of all total coliform determinations, 42 percent exceeded 70 MPN but only 25 percent exceeded that value on "universal dry days". In contrast, on days that were impacted by rain, the data indicates that as many as 77 percent of the determinations exceeded that value. This suggests that if bacterial contamination by storm runoff could be eliminated, the number of stations failing the median criterion could be reduced approximately to 60 percent of the present number. A more accurate assessment will require a detailed study of individual marginal stations. Time limitations prevented such a study under the present grant; however, the analysis here presented and the various detailed print-outs will greatly facilitate such an effort. Supplementing the data set with observations from 1978 and 1979 would improve the reliability of the conclusions. For such a study, one should simultaneously consider the effect of tides and of storm runoff for each station.

That the abnormally high concentration during storm events plays an important role in closing shellfish areas is also suggested by the great significance of the 10 percent criteria. For the five-year period, 120 out of 261 stations failed one or more of the closure criteria (Table 2). Of these 39 failed all criteria leaving 102 that failed some. Of these 62 only failed one or both 10 percent criteria, but the median criteria alone indicated closure for only one station. Considering only the total coliform criteria, 41 stations required closure because both the median and the 10 percent concentrations were excessive. An additional 63 stations indicated closure because they exceeded the 10 percent criterion and no station passed the 10 percent criterion that failed the median test.

General

This study has examined the statistical aspects of the shellfish sanitation data. The nature of the multi-tube fermentation tests greatly limit the information content of the shellfish sanitation data. This becomes a problem if one wants to use the data to understand the behavior of the system. The statistical fluctuations make it difficult to untangle the relative importance of storm runoff and the tides and to compare the fecal and total coliform results. The statistical variability, however, does not reduce the utility of the multiple-tube fermentation tests as a management tool for protecting public health. The criteria used are statistical and the tests are efficient in that they probably provide a maximum amount of information per unit expenditure.

The problem arises when one wants to use the data for purposes for which they were not intended. Nevertheless, the study has shown that it is possible to obtain some answers by analyzing many years of data. Such analysis, however, is only feasible if the data can be processed by computer. To produce the present analysis by manual methods would have required between 10 and 100 man years of effort.

DEMCO