Statistical Analysis of Combined Sewer Overflow Receiving Water Data, 1989-1999

Massachusetts Water Resources Authority

Environmental Quality Department
Report ENQUAD 2001-06

Citation:

Ellis B, Rosen J.2001. **Statistical Analysis of Combined Sewer Overflow Receiving Water Data, 1989-1999.** Boston: Massachusetts Water Resources Authority. Report ENQUAD 2001-06. 22 p.

Statistical Analysis of Combined Sewer Overflow Receiving Water Data, 1989-1999

for

MWRA Harbor and Outfall Monitoring Project

Submitted to:

Battelle
397 Washington Street
Duxbury, MA 02332

Prepared by:

Brian Ellis and Jeffrey Rosen
Technology Planning & Management Corporation
Mill Wharf Plaza, Suite 208
Scituate, MA 02066
(781) 544-0026

October 6, 2000

TABLE OF CONTENTS

LISTOF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

1.1 Background

Due to the deleterious influence of combined sewer overflows (CSOs) on the bacteriological quality of receiving waters in rivers and harbors in Greater Boston, the Massachusetts Water Resources Authority (MWRA) implemented a water quality sampling program in 1989. MWRA initiated a series of changes in the system of CSOs that discharge into Boston Harbor and its tributary rivers. Some of these changes include: (1) complete shutdown and/or re-piping of existing CSOs, (2) addition of screening and chlorination treatment facilities to a number of CSOs, (3) increasing the height of weirs within the CSOs to improve their stormwater/sewage withholding capacity and prevent dry weather discharge, and (4) strategic placement of tide gates to prevent seawater from entering the sewer system and increasing the frequency of CSO discharge. The addition of the new Deer Island Secondary Treatment Plant in conjunction with the shutdown of the Nut Island Treatment Plant increased the capacity of the sewer system, further decreasing discharges from CSOs due to the higher pumping capacity of the new treatment plant. Through all of these improvements, MWRA has continued a water quality monitoring program aimed at tracking bacteria indicative of fecal contamination as well as physical and chemical measurements of the water. These organisms (fecal coliforms and *Enterococcus*) in addition to the physicochemical environmental characteristics such as salinity, temperature, and dissolved oxygen, are contained in a database maintained by MWRA.

1.2 Previous Work

MWRA has collected a large dataset of bacteriological water quality measurements in Greater Boston. Based on these data a number of reports have been produced (Rex, 1991 &1993, Gong et al. 1997& 1998, and Solow, 1993) that have attempted to establish a relationship between decreased concentrations of indicator bacteria and improvements within the CSO system. Gong, et al. (1997, 1998) had an increasingly larger number of observations on which to perform analyses. Using a randomized block Analysis of Variance (ANOVA) design, they concluded that improvements were resulting in decreased fecal coliform counts. This same analysis by Gong, et al. was inconclusive with respect to the occurrence of *Enterococci*. The last report by Gong, et al. (1998) covering the years 1989 to 1997 indicated that the available database included 8646 fecal coliform and 8272 *enterococci* observations.

All reports cited above have attempted, with varying degrees of success to relate both anthropogenic and environmental factors to the concentration of sewage indicator bacteria in receiving waters. Both reports by Gong, et al. (1996, 1997) utilized a randomized block ANOVA to assess the influence of environmental factors such as tide, season, location, and rainfall on the occurrence of indicator bacteria in the CSO receiving water system of the Boston Area. All of the reports previously cited have stratified the data according to a number of factors. Temporal stratification into before and after improvement periods was the primary partition used for all analyses in this report as well as all previous reports. Observations on bacterial occurrence prior to and including 1991 were used as baseline measurements on before improvement conditions within receiving waters. Conversely, measurements obtained since the beginning of 1992 were compared to the baseline measurements and were considered as after improvement measurements. For example, see Leo, et al. (1994) and Rex, et al. (1998) for a description of the documented changes that have occurred within the CSO system in the Boston Area receiving waters over the last eleven years.

1.3 Data Characteristics

Table A1 (Appendix A) provides a summary of the data available for these analyses segregated by geographic region. This table reveals that there are a number of regions (Deer Island, Nut Island, Winthrop Bay, Outer Harbor, and Quincy Bay) for which data does not exist prior to the 1991 period. The total number of observations in the data file received from MWRA was 16,033 fecal coliform observations (range = 0 to 1.34 x 10^6 organisms/100 mL) and 15,669 *enterococci* observations (range = 0 to 2.57 x 10^5 organism/100 mL). The data were from the years 1989 to 1999 (first 3 months) and also contained the one day, two day, and three day rainfall sums. In addition, the root mean square ($RMS = ((R^2{}_1 + R^2{}_2 + R^2{}_3)^{1/2})$ of the 3 day summed rainfall was provided to place more emphasis on one day heavy rainfall events. This summed rainfall value was originally devised for placing greater weight on high intensity events, which are more likely to result in a CSO discharge. For several reasons, the root mean square 3 day summed rainfall was not used for correlation analyses in this study. This is because the RMS almost always had a lower correlation with bacterial counts than did the 3-day summed rainfall and there was a very close correlation between 3-day summed rainfall and the RMS 3-day rainfall (i.e. the correlation between these two events was always ≥ 0.95). While other parameters such as salinity, temperature, and conductivity were also supplied, these values were not used with respect this investigation of bacteriological water quality.

The analyses performed in this study were intended to compare the data from the 1989 through 1991 period to the data from the years after 1991 and therefore the actual n-value for the number of bacterial observations is lower than indicated above. This arises due to the lack of data for the after improvement period in the regions indicated above and detailed in Table A.1. The observations in the after improvement period in Table A.1, where no bacterial observations were recorded in the before improvement period were not used for comparison purposes in the System Wide subset of data subsequently mentioned in these analyses. This causes the number of observations (i.e. n-value) for fecal coliforms to drop to 12, 228 when we exclude sampling locations where data is unavailable for the before and after improvement period while enterococci observations drop to 11,919. Regions or watersheds for which data are available for before and after improvement years include Dorchester Bay, Neponset River, Mystic River, Charles River and the Inner Harbor. Given this lack of data for some regions, System Wide analyses were performed with and without the information contained within these regions. Division of the data into geographic regions was performed in order to provide continuity with previous studies in addition to addressing the goals of MWRA in treating the watersheds as individual and separate ecosystems that may require different pollution abatement strategies.

Station names and regional designations used in this study were taken directly from Gong, et al. (1996, 1998) and Rex and Taylor (1998).

1.4 Objectives

MWRA is interested in assessing the impact that CSO improvements have had on bacterial occurrence in receiving waters. With this in mind, we attempt to answer several questions with the analyses performed in this report.

1) Has there been a measurable, detectable, and statistically significant decline in bacterial occurrence concentrations within the Boston area CSO receiving waters?

    a) Are areas that have undergone major changes in CSO discharge characteristics significantly different from before those changes or from areas where minor (or no) changes have occurred?

2) Has the association between rainfall and bacterial pollution in the Harbor and rivers changed due to the pollution abatement projects completed over the past 10 years (i.e. has the correlation between bacteria and rainfall decreased over the years)?

1.5 Analytical Methods

The Statistical Analysis System (SAS Institute) was utilized for all analyses performed for this study. Environmental data (especially bacteriological data), can often be very "noisy", exhibiting extremes in ranges. Data transformation by calculating the logarithm is often utilized for analyzing bacterial counts as well as being indicated as the method of choice for the analysis of microbiological counts (Standard Methods, 1995). This transformation tends to smooth out the effects of large outliers. Due to this property, almost all of the analyses done for this report were performed after the data had been $\log_{10}$ transformed and reported values are the back transformed values (geometric means). All of the bacterial counts reported to be zero were included in the analyses. Due to the technical problem encountered when the data to be transformed includes zeros the transformation $\log_{10}(x+1)$ was employed. In addition, the method of Box and Cox (Sokal and Rohlf, 1995), involving a series of power transformations indicated that in most cases, a logarithmic transformation was sufficient to induce normality in the data for yearly means. This transformation also quite often results in homogeneity of variance in the data, another basic assumption for parametric statistical analyses.

Previous studies have tended to rely on one major strategy such as linear regression (Solow, 1996) or ANOVA (Gong, et al. 1996, 1998) to investigate the effects of improvements on bacterial discharge characteristics for the Boston Area CSO receiving waters. As stated above, we wished to maintain continuity with previous work, such that general classifications were maintained. Two major divisions of the data files could be made and were based on (1) geographic region and (2) before/after improvement. The number of samples available for analysis was considerably higher than was previously used by Gong, et al. (1998) where the analysis considered 1989 through 1996. The ensuing years (1997 through 1999 and part of 2000) had resulted in the collection of an extra 5,262 observations. Due to this increase we thought that a simple, uncomplicated statistical approach should be able to detect changes in indicator bacteria concentrations if they existed.

1.5.1 Data Reduction and Aggregation Strategies

When comparing two treatments or conditions it is important to control the variability so that the researcher understands as many sources of variability as possible. Optimal evaluations of the effects of treatments or changes in condition are based on the same sampling done on two independent replicate systems or on very similar sampling of the same system before and after some change in the conditions of the system. In this situation we are attempting to determine whether or not infrastructure improvements have resulted in significant reductions in contaminant concentrations and effects on ecosystems. Ideal sampling to address this question would be for the same number of samples to be collected at the same number of stations over a similar temporal extent. When hypotheses are tested of whether or not the concentrations before the infrastructure improvements are significantly different from those observed after the improvements all temporal and spatial components of the sampling domain are equally represented in the pre and post improvement data sets.

In this data set there are major differences in the sampling done before and after the infrastructure improvements. These differences complicate the test of hypotheses discussed above. For example: in the Charles River region, stations 1-4 were not sampled during 1993-1996. Stations 7,9 and 10 were not sampled during 1992-1996. Stations 144 and 145 were added to the sampling plan in 1994. When comparing results of pre and post infrastructure improvements we are comparing different sampling plans. This prevents station by station comparisons for pre and post improvements. This has been handled by pooling all the pre-improvement stations and sampling dates together and all of the post-improvement stations and sampling dates together. In order for this approach to be optimized, each sampling at a station and a date must be equally represented in the database. This approach reduces the amount of data available but ensures that unexplainable variance is not introduced into the analysis.

A number of approaches were used for reducing and aggregating data. Data reduction was accomplished by calculating means (for log transformed bacterial concentrations) for each year region combination and then partitioning the years into before and after improvement. These means followed the assumptions for an ANOVA and were analyzed in this manner. These means were used to establish whether or not gross bacterial concentrations were different between the before and after improvement periods. The non-transformed data did not follow a normal distribution and could not be evaluated by parametric analyses. Therefore, non-parametric analyses were also performed on the non-transformed non-aggregated data in the same manner in addition to t-tests on the log10 transformed aggregated data.

1.5.2 Linear Regression Analyses

Linear regression was used to determine if there were significant decreases (or increases) in the yearly mean bacterial concentrations aggregated as noted above (i.e. yearly means). Regression analyses were performed on the System Wide data using all observations. The System Wide subset of data was analyzed in a similar manner to determine if the low bacterial concentrations noted in the marine regions (Deer Island, Nut Island, Winthrop Bay, Outer Harbor, and Quincy Bay) were introducing a bias in the results. Individual geographic regions where sufficient data was available for the before and after improvement years were also examined on a region by region basis.

1.5.3 Correlation Analyses

Pearson product moment correlations were calculated to ascertain the intensity of association between rainfall and bacterial occurrence System Wide, within geographic regions, and partitioned by the before and after designations. The yearly average rainfall (i.e. one, two, and three-day rainfall) was calculated to establish if the presence of a trend in the amount of rainfall could account for any of the noticeable changes in bacterial concentrations. Rainfall was partitioned by selecting only those samples where the three day summed rainfall was greater than 0.25 inches. This amount of rainfall represents an event that is likely to cause discharge from the CSOs. The three-day rainfall consists of the sampling date plus the two previous days. Partitioning the data in this manner removes information below a threshold value where CSO impacts are unlikely. The equality of correlation coefficients was tested by use of a t-statistic after transformation. Correlation coefficients are not normally distributed and therefore require a transformation prior to analysis. The method utilized is detailed in Sokal and Rohlf (1995) and the test statistic used a t-test where t_s is compared to $t_{\alpha[\infty]}$ with $\alpha = 0.05$. In the case of comparing two correlation coefficients

$$t_s = (z_1 - z_2) / [1/(n_1-3) + 1/(n_2-3)]^{1/2}$$

where z_1 and z_2 are the transformed (i.e., inverse hyperbolic tangent of r) correlation coefficients and n_1 and n_2 have their usual meaning. The null hypothesis in this case is for equal correlation coefficients versus the alternative of unequal correlation coefficients.

If the contaminant concentrations in the receiving waters are strongly influenced by the CSOs then we would expect to see strong correlations between the contaminants measured and the amount of rainfall over the threshold that would cause the CSOs to be active. If the infrastructure improvements have been successful at either raising the threshold for the CSOs to be active or have decreased the types and concentrations of contaminants introduced by the CSOs (when they are active), we would expect to see a decrease in the significance of the correlations between rainfall and contaminant concentrations in the receiving waters. At least we would expect to see the magnitude of the correlation between the rainfall and the concentration of contaminants in the receiving waters. Therefore, if we calculate correlations prior to the infrastructure improvements and after we would expect these correlation coefficients to be different from one another. If they are the same then either the improvements had no effect on the amount of the contaminant entering the receiving water or there was little to no correlation between rainfall and the contaminant concentration in the first place. Either way this result would suggest that whatever was done during the improvement process did not change the relationship between the amount of rainfall and the concentration of the contaminant in question. If on the other hand the correlation coefficients do change and can be detected statistically then there is a suggestion that the improvement process has had an influence on the relationship between the contaminant and the amount of rainfall. We will be looking for decreases in the correlation suggesting that following the infrastructure improvement process the amount of rainfall and the concentration of the contaminant were not correlated or the degree of correlation had decreased. This result would suggest (but not prove) that the CSO infrastructure improvements resulted in decoupling CSO inputs from contaminant concentrations.

Thus, geographic separation of data and comparison of the later data (i.e. ≥1992) to the earlier data were used to detect gross changes in the bacterial counts with respect to spatial and temporal scales. Linear regression was used to analyze the data with respect to time, based on yearly means and scaled over the entire CSO receiving water system as well as within specific geographic regions. Correlation analyses were used to monitor the correlation between rainfall and bacterial concentrations through the before and after improvement stages of the surveys. These analyses were used to assess the strength of association (if any) between rainfall and bacterial counts overall and comparatively between the before and after improvement period. In addition, the data were analyzed on an individual region by region and station by station manner in order attempt to capture anomalous behavior within regions or at specific stations. Analysis of variance was attempted on a number of the regions, but this method was contraindicated without data reduction (yearly means) due to the lack of normality and homogeneity of variance in the data.

2.0 RESULTS

2.1 Regional and System Wide Annual Means

Simple comparisons between the before and after means in each geographic region were performed as a prelude to a more in depth data analysis. As shown in Appendix A (Table A.1.) and Figure 2.1- Panels A and B, there has been a significant reduction in the number of indicator bacteria, both system wide and within the individual CSO receiving water regions. The system wide data utilized all values populating the database. This introduced a significant bias toward lower bacterial occurrence values due to the regions, which are indicated in Table A.1 by the entry nd. For this reason, the system wide analysis was repeated on the subset data after removing these stations/regions. These observations were not included due to lack of before improvement data and the tendency towards significantly lower bacterial occurrence counts (See Tables A.1. through A.3.). This sub categorization of the data does not invalidate the conclusions that can be drawn from Figure 2.1, Panels A and B. However, the incorporation of the data from the marine stations results in a decrease in the mean concentrations for both indicator bacteria. This is due to the significantly lower bacterial concentrations measured at the marine stations as compared to the freshwater and estuarine stations (see Tables A.1. through A.3.). All values within specific regions (and system wide) are significantly different between the before and after period for the data shown in Figure 2.1 (see Table A.1 for n-values, p-values and averages). This raw data could not be normalized by transformation and therefore, a nonparametric test for independent groups was used (Wilcoxon Rank Sum Test) that compared the ranked values of both groups (ranked untransformed values). All associated p-values were less than 0.0074. While the raw data could not be normalized, the yearly regional averages could be normalized and a nested ANOVA was performed with regions nested within the before and after improvement time periods. This analysis gave essentially the same results as the above non-parametric test and the results are reported in Table A.2 Appendix A.

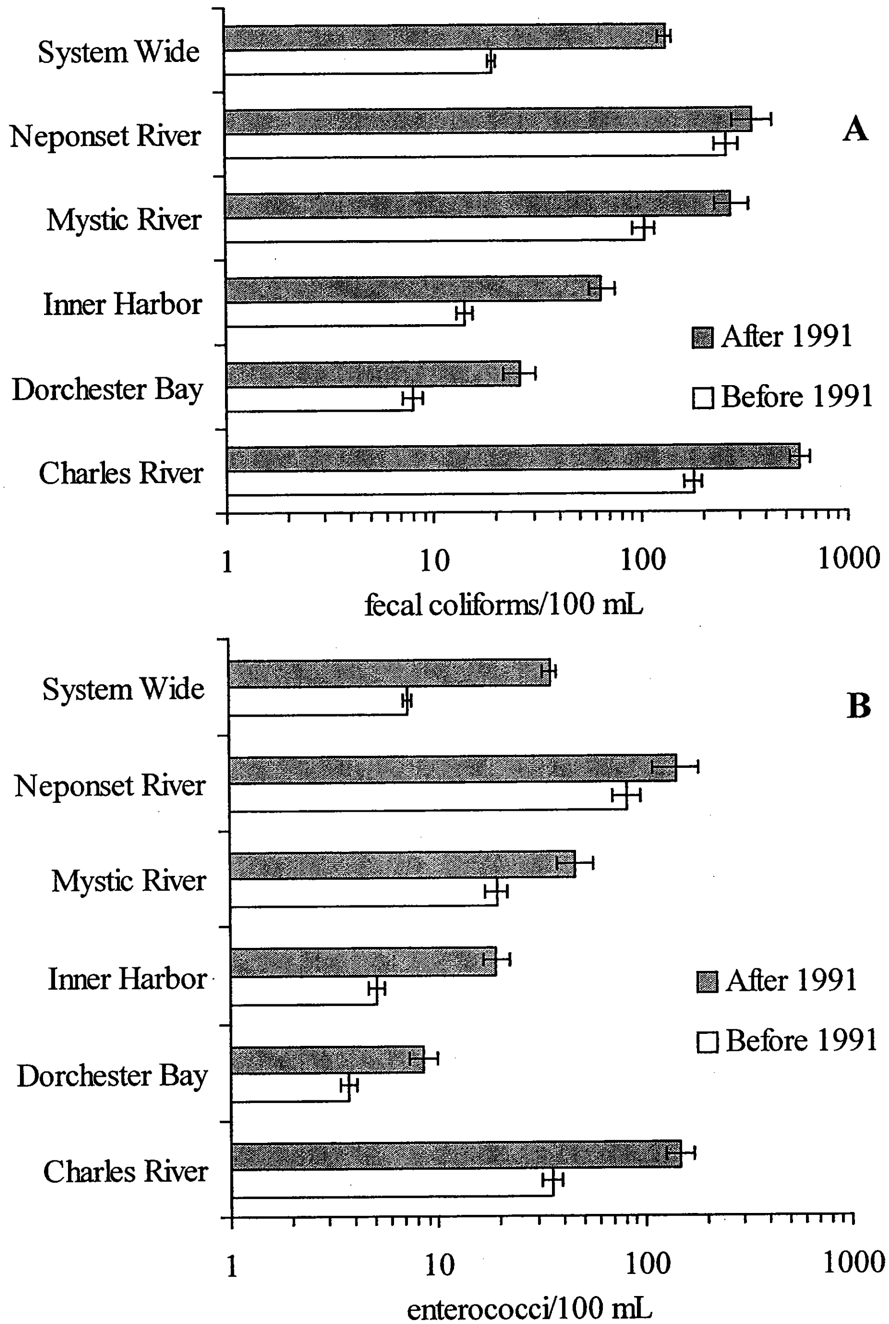

Figure 2.1. Grand means (geometric) for geographic regions and the system wide area of the Boston Region watersheds receiving stormwater and CSO discharge in the last 11 years. Panel A represents fecal coliforms and Panel B represents enterococci. Error bars are 95% confidence intervals based on back transformation of the log transformed data. See Table A.1 for p-values, n-values, and means.

2.2 Linear Regression Analyses

Subsequent to the comparative analyses in the previous section the yearly means were calculated and these data are shown in Figure 2.2. The individual raw data values were not plotted in this Figure (back transformed means and 95% confidence intervals are plotted), but the apparent trend of decreasing bacterial concentration with time on a system wide basis is shown. This decreasing trend is significant ($p < 0.0005$ for coliforms and $p = 0.005$ for enterococci) and linear regression fit lines and R^2-values are provided. The regression equations in Figure 2.2 refer to the $\log_{10}$ transformed data. Although the linear regressions are highly significant, one can note that the lines tend to flatten out with time. This linear relationship is almost totally lost as seen in Figure 2.3 when the regression analyses are categorized in the same manner as the regional before and after improvement comparisons and removing the data represented by the stations without before improvement information. The regression p-value for the fecal coliform data rises to 0.042 (adjusted $R^2 = 0.317$) while the enterococci p-value rises to an insignificant 0.436. This could be due to the presence of a background concentration of organisms present in the water due to circumstances beyond the control of MWRA and having nothing to do with the CSO system. Stormwater runoff and natural sources of coliforms would likely contribute to this background count. Natural sources from the presence of wildlife, such as birds and mammals (e.g., cats, dogs, and rats) would be extremely difficult to account for in this type of analysis. However, the potential for a natural baseline concentration of fecal coliforms or enterococc*i* in aquatic ecosystems implies that no matter how much improvement is carried out within the system of CSOs in the Boston Area, there may always be a detectable amount of indicator organisms present. Alternatively, this loss of significance for the regression analysis could be due to aggregating the data on a system wide basis.

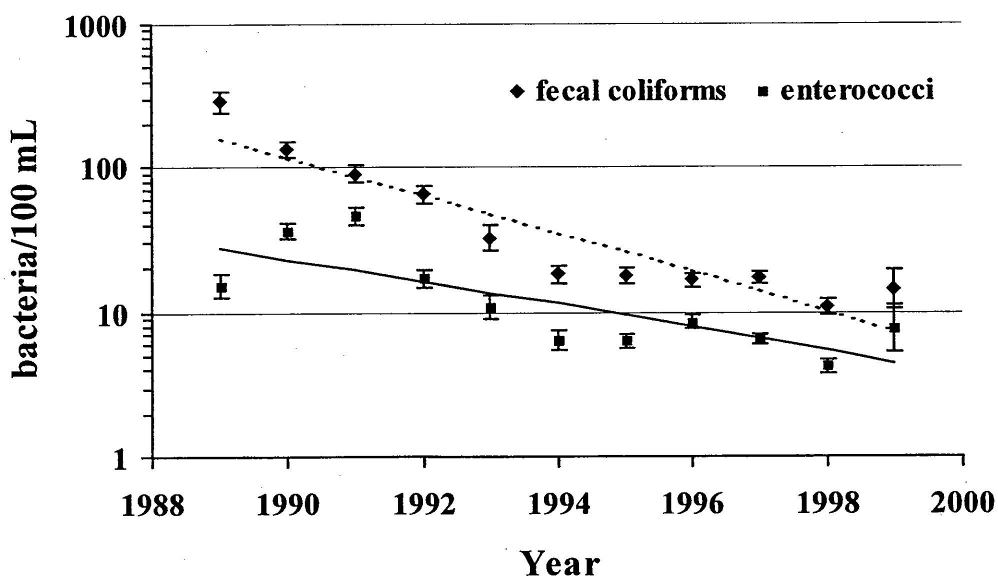

Figure 2.2. Linear regression results on the grand means for all data over the course of the last eleven years. Diamonds represent fecal coliforms and the fit line is represented by the dashed line (adjusted $R^2 = 0.849$, $\log_{10}$(cells/100 mL) = 256 – 0.128 $\log_{10}$ (cells/100 mL/year) x Year). Squares refer to the *enterococci* counts and the fit line is represented by the solid line (adjusted $R^2 = 0.566$, $\log_{10}$(cells/100 mL) = 144 – 0.072 $\log_{10}$ (cells/100 mL/year) x Year). Error bars are 95% confidence intervals for the yearly means.

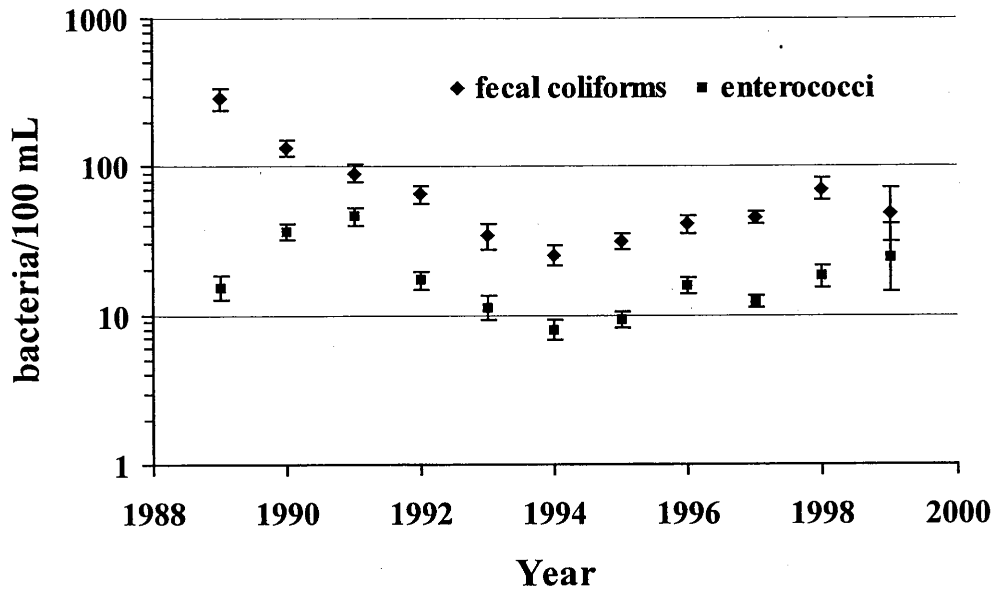

Figure 2.3. Linear regression results on the grand means for all fecal coliforms and *Enterococci* data over the course of the last eleven years. Only stations with data from the before and after periods was used. Note the loss of linearity as compared to Figure 2. Symbols are the same as for Figure 2.

Given the multiple potential reasons for this trend in the bacterial concentrations over time, we reanalyzed the data at the level of individual geographic regions and these results are reported in Table 2.1. This Table shows that the yearly means for Dorchester Bay and the Neponset River Estuary have not decreased linearly over the past eleven years as compared to the other waters within the MWRA sampling region. The counts for all other areas have decreased on a yearly basis. One concern, in regards to decreasing bacterial counts, is the relationship between rainfall and concentration of indicator organisms. We have shown that bacterial counts have decreased in certain regions yet this could be due to variations in the pattern of rainfall over the last eleven years. Figure 2.4 shows the yearly average for three of the rainfall variables measured by MWRA on those days in which bacteriological samples were taken. This graph does not represent the averages for the whole year. There was a pattern of decreasing rainfall from 1990 to 1993, however there did not seem to be any discernible trend in the rainfall events between 1993 and 1999.

Geographic Region	Indicator Organism	Intercept	Slope	R^2 (adjusted)	p-value (regression)
Charles River	coliforms	2.85	-0.070	0.52	0.007
	enterococci	2.32	-0.069	0.42	0.018
Dorchester Bay	coliforms	1.36	-0.002	0	0.96*
	enterococci	0.76	0.031	0.033	0.28*
Inner Harbor	coliforms	1.96	-0.11	0.64	0.002
	enterococci	1.39	-0.078	0.55	0.005
Mystic River	coliforms	2.76	-0.143	0.66	0.001
	enterococci	1.85	-0.095	0.46	0.001
Neponset River	coliforms	2.73	-0.016	0	0.65*
	enterococci	2.21	-0.002	0	0.95*

Table 2.1. Linear regression analysis for yearly mean bacterial counts in the indicated geographic region. In this analysis 1989 was designated as Year One and 1999 was designated Year 11.*- indicates that the regression was not significant.

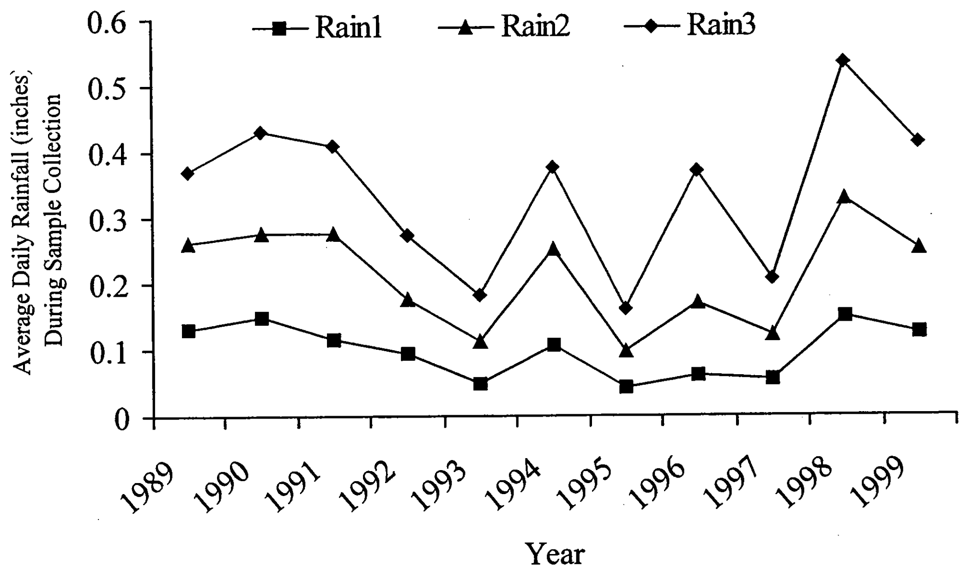

Figure 2.4. Average yearly 1-day, 2-day, and 3-day summed rainfall for bacteriological sampling over the last eleven years in the Boston Area CSO receiving water system. The 1-day and 2-day rainfall has not changed significantly (Wilcoxon Ranked Sums, $p = 0.795$ and 0.058 respectively) whereas the 3 day rainfall is significantly different over the years ($p = 0.005$).

2.3 Correlations between Rainfall and Indicator Bacteria

Correlation analyses were performed on the $\log_{10}$ transformed data and the results are reported in Table 2.2. In this case we used the correlation analyses to investigate the strength of association between rainfall and the magnitude of the bacterial counts. The rainfall data were collected over the course of 2 days prior to sampling (plus the day of sampling) to take into account the possibility of a delayed effect due to the amount of rainfall required to produce a CSO discharge. As such, the correlation analyses used all rainfall and $\log_{10}$ transformed fecal coliform and enterococci data. Given that larger rainfall events are likely to be more highly correlated with indicator bacteria concentration, the correlation analysis can also be subset according to how much rain had fallen and when the event occurred with relation to the sampling time. For example, Gong et al., in their randomized-block ANOVA split the root mean square of the 3-day rainfall into 3 groups: (1) RMS = 0, dry conditions, (2) 0< RMS < 0.25 inches, light rain, and (3) RMS > 0.25 inches, heavy rain. To define a rainfall event in a more systematic manner, any of a number of statistics such as the mean, median, or percentiles such as the 75th, 90th, 95th, etc. could be chosen. It is interesting to note the cut value of 0.25 inches of rainfall chosen by Gong et al. (1996, 1998) represents the 90th percentile for the one day rainfall (as an average) over all sampling dates throughout the eleven year course of MWRAs survey study. Thus, this value would represent an extreme 1-day event but not such an extreme event if one considers the three other rainfall measurements. For a 2-day rain this value would represent the 80th percentile, for a 3-day rain it would be the 69th percentile and for
the RMS (3-day rain) it would be the 70th percentile. To maintain continuity with previous reports and analyses the value of 0.25 inches of rainfall was maintained as a cutoff point for delineating rainfall events within the correlation analysis.

Starting at a system-wide level, divided into before and after improvement periods, a correlation analysis using all of the data indicates a definite change in the correlation between rainfall and bacterial counts as evinced in Table 2.2. The correlation coefficient refers to the correlation between the 3-day summed rainfall and the log transformed bacterial concentrations and does not use the RMS as was used by Gong, et al. (1996).

Geographic Region	Indicator Organism	Before		After		p-value
		n	r	n	r	
Charles River	coliform	645	0.40	1,661	0.25	<0.0005
	enterococci	751	0.44	1,672	0.26	<0.0005
Dorchester Bay	coliform	760	0.31	1,868	0.30	0.83
	enterococci	759	0.33	1,859	0.27	0.14
Inner Harbor	coliform	1,028	0.43	2,496	0.25	<0.0005
	enterococci	1,030	0.39	2,465	0.28	0.0017
Mystic River	coliform	354	0.25	1,469	0.17	0.14
	enterococci	353	0.28	1,462	0.22	0.31
Neponset River	coliform	317	0.22	1,007	0.22	0.94
	enterococci	317	0.28	942	0.27	0.85
System Wide	coliform	3,404	0.30	12,629	0.16	<0.0005
	enterococci	3,210	0.33	12,459	0.19	<0.0005
System Wide[a]	coliform	3,404	0.30	8,501	0.20	<0.0005
	enterococci	3,210	0.33	8,400	0.22	<0.0005

Table 2.2. Correlation coefficients between 3-day summed rainfall and indicator bacteria in the Boston Area CSO receiving water systems. a – data subset does not include the Outer Harbor area (see text for further details). P-values indicate significant ($p < 0.05$) or insignificant ($p > 0.05$) changes in the correlation coefficient between the before and after improvement sampling periods.

While the correlation coefficients are generally low, they tend to be highly significant (i.e., $p < 0.0005$) and positive, supporting the assumption of higher bacterial counts in the presence of higher rainfall. Without any further investigation of the various regions, it is apparent that some of the correlations have decreased significantly from the designated baseline period prior to improvements in the CSO receiving water system. A few of the regions show no change in the correlation between rainfall and bacterial counts, either positive or negative. Since these correlations are based on all of the data, dry, moist and wet weather included, we opted for analyzing the data in a more selective manner. Since it is mostly assumed that higher rainfall coincides with higher bacterial counts due to increased CSO discharge, segregating those days when rainfall was beyond some defined threshold value should enable us to ascertain if there has been a decrease in the correlation between wet weather and bacterial counts. A correlation analysis on 3- day rainfall < 0.25 inches showed that there was very little correlation between low/no rainfall and indicator bacteria throughout the greater Boston Area in either the before or after period supporting our assumption for partitioning the data in this manner. Performing the same correlation analysis with the 3-day rainfall restricted to greater than 0.25 inches, as provided in Table 2.3 shows that the correlations were significantly higher prior to 1992.

Tables 2.2 and 2.3 reveal that there are areas where the correlation between rainfall and indicator organisms has not changed significantly. These areas are the Neponset and Mystic River watersheds and Dorchester Bay. The Neponset River feeds into Dorchester Bay while the Mystic River watershed could be significantly impacted by Alewife Brook (Station 70) where there are notably high yearly mean concentrations of both indicator organisms.

Dorchester Bay and the Neponset River Estuary bacterial concentrations have not changed to any large degree between the before and after period. Thus, both the correlation between rainfall and bacterial counts and the temporal reduction in bacterial counts seen in other parts of the Boston Area receiving waters are not duplicated in these regions.

Geographic Region	**Indicator Organism**	**Before**		**After**		**p-value**
		n	**r**	**n**	**r**	
Charles River	Coliform	350	0.50	397	0.22	<0.0005
	enterococci	304	0.37	397	0.22	0.033
Dorchester Bay	Coliform	256	0.24	612	0.24	0.943
	enterococci	255	0.35	609	0.24	0.128
Inner Harbor	Coliform	354	0.47	667	0.25	<0.0005
	enterococci	356	0.47	659	0.25	<0.0005
Mystic River	Coliform	122	b	419	b	c
	enterococci	122	0.31	416	0.11	0.047
Neponset River	Coliform	112	b	328	b	c
	enterococci	112	b	321	0.15	0.171
System Wide	Coliform	1,194	0.30	3,760	0.15	<0.0005
	enterococci	1,149	0.31	3,719	0.17	<0.0005
System Wide[a]	Coliform	1,194	0.3	2,423	0.17	<0.0005
	enterococci	1,149	0.31	2,402	0.18	<0.0005

Table 2.3. Correlation coefficients in the various Boston Area CSO receiving waters when the 3 day summed rainfall exceeds 0.25 inches. a- subset does not include outer harbor and associated areas (see text for further details). b correlation is not significantly different from zero. c- test is not applicable.

2.4 Identification of Problems Areas

2.4.1 Neponset River and Dorchester Bay

One rather simple graphical method for identifying stations or regions where high counts exist, would be to plot the fecal coliform vs. enterococci data on a scatter plot and identify stations or geographical areas that exhibit high counts for both indicator organisms. This has been done in Figures 2.5-2.9 by geographical region partitioned into before and after improvement. A number of the sampling stations within each region have been noted in the figures in order to highlight potential problem areas. Almost all of these stations are associated with active CSOs. For example, Figures 2.5 and 2.6 show the Neponset River Estuary and Dorchester Bay plotted in this manner. There are six sampling stations (41, 42, 54, 55, 89, and 100) positioned within the Neponset River Watershed and they are interspersed from above Baker Dam (Station 55) to Commercial Point (Stations 41 and 89), oceanward of Tenean Beach. Three of these stations are located downstream of CSOs (Stations 54, 42, and 89). There has been improvement in the yearly means at the stations, which tended toward higher bacterial concentrations prior to 1991, but there still tends to be higher yearly counts associated with these particular stations.

Dorchester Bay, which is contiguous with the Neponset River, was included in the analysis of this region. Station 39, which is associated with the Fox Point (BOS089) CSO discharge and station 40 which is associated with BOS088 in the Malibu Beach area tend to higher counts than other stations in this region. Dorchester Bay indicator bacteria counts are the second lowest in the Boston Area (lowest counts are found in the Outer Harbor region). However this does not allow us to state that there has been a specific improvement in this area over the before and after period (however, see comparison of means section 2.1).

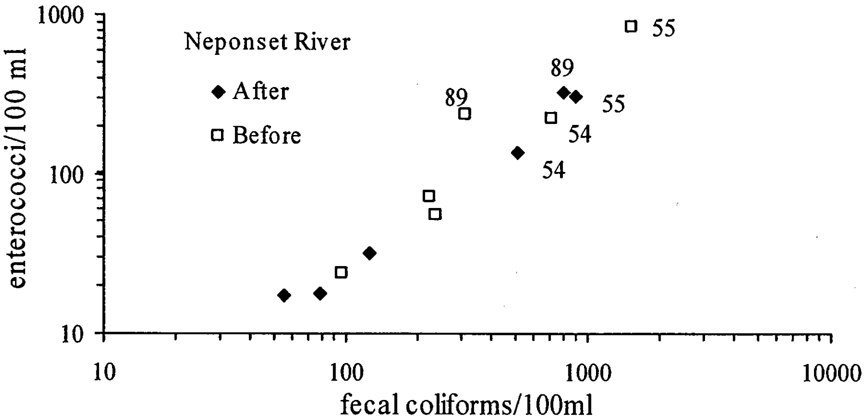

Figure 2.5. Before and after fecal coliform geometric mean concentrations vs. enterococci geometric mean concentrations in the Neponset River estuary plotted by station.

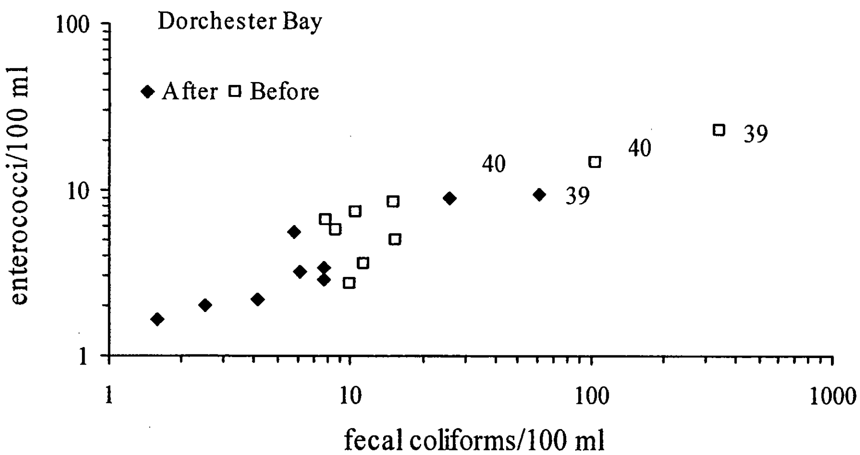

Figure 2.6. Before and after fecal coliform geometric mean concentrations vs. enterococci geometric mean concentrations in Dorchester Bay plotted by station.

2.4.2 *Charles River*

Figure 2.7 shows that a few of the stations in the Lower Charles River seem to pose particular problems with respect to bacterial counts. Station 2 and Station 6 (located near a CSO treatment facility) and to a lesser extent Stations 1 and 12 both of which are upstream of CSOs) are the only stations that have not shown a highly visible improvement in yearly means. Station 2 is downstream of BOS033 and BOS032 (closed 11/97) and changes may not be detectable from the number of samples taken since 1997 (only 5 samples since 9/2/97). However, it is apparent that there has been a major shift in the distribution of the counts since prior to 1992. The values represented by stations 1, 6, and 12 represent significant changes (46 to 84% reduction) from prior to 1991, yet they are still fairly high. Station 6 is just downstream of the Cottage Farm (MWR201) CSO while stations 1 and 12 are just below Watertown Dam at the upper reaches of the Lower Charles River.

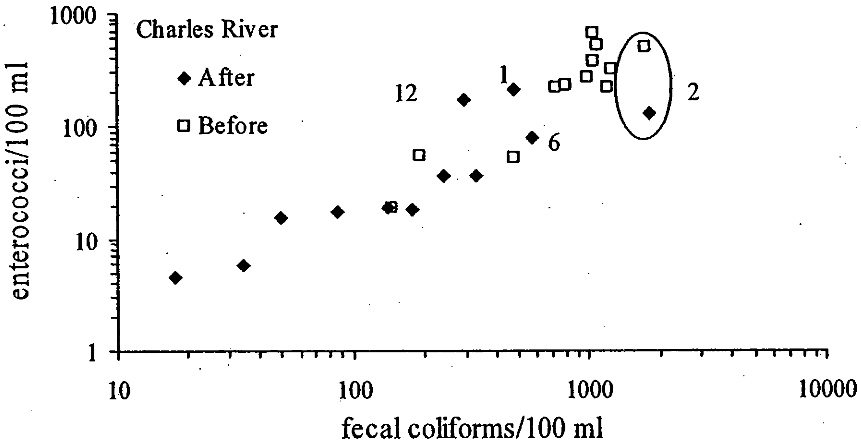

Figure 2.7. Before and after fecal coliform geometric mean concentrations vs. enterococci geometric mean concentrations in the Charles River plotted by station.

2.4.3 *Inner Harbor*

The Inner Harbor shows low indicator bacteria (as annual averages) concentrations for both of the primary periods in question. This may be indicative of the effect of other factors on bacterial survival such as pollution, salinity, and temperature, or simply reflect the greater dilution and flushing of harbor water. The decrease in bacterial counts, reflected in the before and after improvement periods is not seen at Fort Point Channel Stations 18 and 75 as shown in Figure 2.8. This result may be due, in part to the relative stagnancy of the water in Fort Point Channel, as compared to the rest of the Inner Harbor region. Fort Point Channel receives the largest amount of untreated combined sewer overflow in the Boston Harbor Watershed. As of 1998, approximately 8 CSOs discharge into this rather small Channel (Rex, et al. 2000) with one receiving treatment. Excepting stations 18 and 75 there is a clear delineation in the before and after improvement annual means for stations within the Inner Harbor.

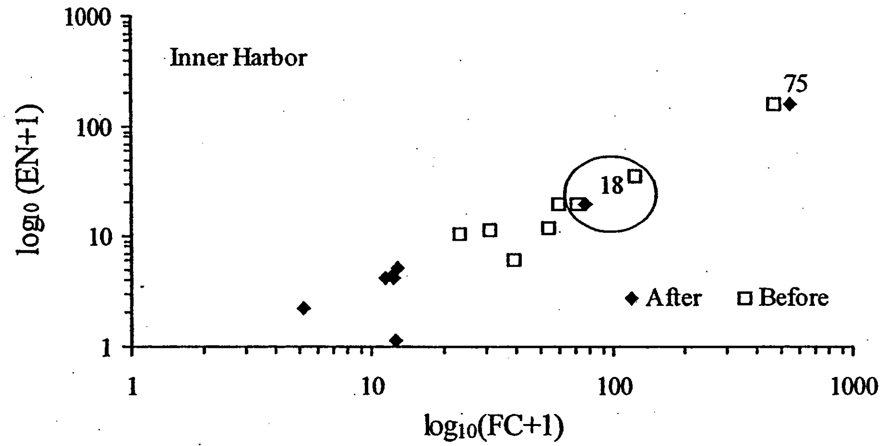

Figure 2.8. Before and after fecal coliform geometric mean concentrations vs. enterococci geometric mean concentrations in the Inner Harbor plotted by station.

2.4.4 Mystic River

The Mystic River Estuary has been of particular concern, mostly with respect to Alewife Brook. This rather small brook has a significant number of CSO discharges. The absence of appreciable herring runs in the last decade compared to previous decades has brought considerable focus onto Alewife Brook and the potential impact that CSO discharge has on the spawning of the herring. Figure 2.9 shows that there has little, if any, improvement in the bacteriological water quality flowing from Alewife Brook (i.e. Station 70). Minimal improvement has been noted in the individual stations situated along the Mystic River. Figure 2.9 shows the distribution of the bacterial counts based on before and after improvement averages.

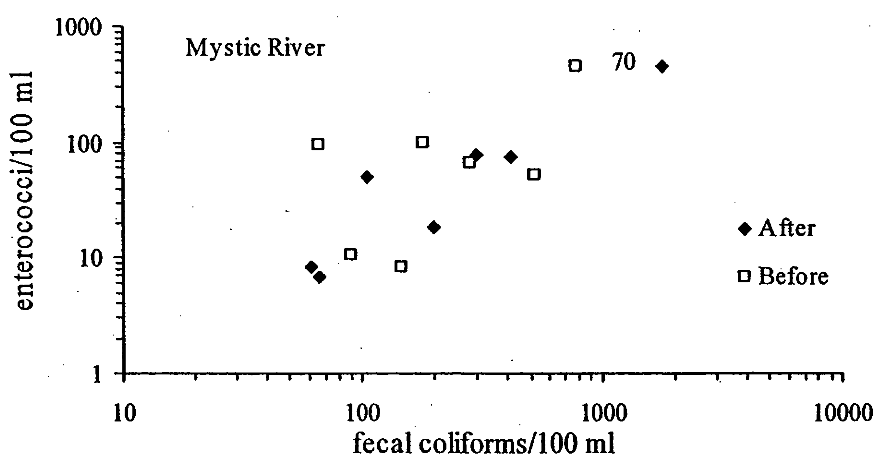

Figure 2.9. Before and after fecal coliform geometric mean concentrations vs. enterococci geometric mean concentrations in the Mystic River Estuary plotted by station.

3.0 CONCLUSIONS

The objective of this report was to analyze the database compiled by MWRA and containing information on the concentration of indicator bacteria in the combined sewer overflow receiving waters in the Boston Harbor area and its tributary rivers. An attempt to answer two main questions was the aim of this study. These two questions were the following: (1) Has there been a decline in bacterial occurrence concentrations within the Boston area CSO receiving waters and (2) Has the association between rainfall and bacterial pollution in the Harbor and rivers changed due to the pollution abatement projects completed over the past 10 years?

A significant and detectable decline in bacterial concentrations in Boston Area receiving waters since 1991 was shown in sections 2.1 and 2.4. This improvement can be seen, both system wide and within individual watersheds or regions with respect to annual averages for both indicator organisms. All of the regions with data for the before and after improvement periods, including Dorchester Bay, the Neponset River, the Inner Harbor, the Mystic River, and the Charles River have seen a significant reduction in the number of indicator organisms since 1991. The Inner Harbor, Mystic River and Charles River have seen a steady and significant decline in both fecal coliforms and enterococci concentrations on a year to year basis. This was shown by the significant regressions in section 2.2 whereas the regressions for Dorchester Bay and Neponset River were not significant, indicating that these regions did not experience a linear decline in indicator bacteria counts over the same time period.

It has been noted that there tended to be a correlation between rainfall events and bacterial concentrations in the Boston Harbor area CSO receiving waters over the years of sampling. The correlation between rainfall and the magnitude of the bacterial concentrations has followed a similar trend to that noted above for the steady, yearly decline in bacterial counts. Dorchester Bay, the Neponset River, and the Mystic River have seen little or no decline in the correlation between indicator bacteria and rainfall with respect to all of the data sampled by the MWRA over the last 11 years (i.e. no partitioning by the intensity of rainfall events). The Inner Harbor, Charles River, and System Wide have seen a significant reduction in this correlation when comparing before and after periods. Analysis of the data, only in the presence of significant (>0.25 inches) 3-day rainfall events provides the same results, but with a higher degree of certainty. The relationship between rainfall and the presence of indicator bacteria such as Fecal coliforms and enterococci has decreased significantly in most of the Boston area CSO receiving waters with the exceptions noted above. These exceptions, while not known to be due to specific sampling stations within each individual region, can be hypothesized to be due to stations, which seem to be the exceptions to the general trend. These include specific stations in the Mystic River (70), the Neponset River (55, 89, 44), and Dorchester Bay (39, 40). Station 2, and to a lesser degree stations 1, 6, and 12 in the Charles River, continue to be a significant source of high bacterial counts in this watershed. The Inner Harbor stations 18 and 75 in the Fort Point Channel are significant sources of high indicator bacteria counts.

3.1 Recommendations

1) Institution of a sampling strategy aimed at delineating and tracking improvements within specific regions would greatly enhance the ability for MWRA to detect significant reductions in indicator bacteria concentrations (i.e. less random, more often, and better geographical coverage).
2) A closer inspection or survey of the stations (or areas around the stations) that are noted in Section 2.4 above might enable MWRA to pinpoint undesirable sources of indicator organisms in receiving waters (if they are not already known).

4.0 REFERENCES

Gong, G., Lieberman, J., and McLaughlin, D. 1997. Statistical Analysis of Combined Sewer Overflow Receiving Water Data, 1989-1995. Massachusetts Water Resources Authority, Environmental Quality Department.

Gong, G., Lieberman, J., and McLaughlin, D. 1998. Statistical Analysis of Combined Sewer Overflow Receiving Water Data, 1989-1996. Massachusetts Water Resources Authority, Environmental Quality Department. July 1998.

Leo, W.S., Rex, A.C., Carroll, S.R., and Connor, M.S. 1994. The State of Boston Harbor, Connecting the Harbor to its Watersheds. Massachusetts Water Resources Authority, Environmental Quality Department. 1994.

Rex, A.C. 2000. 2000. The State of Boston Harbor 1997-1998, Beyond the Boston Harbor Project. Massachusetts Water Resources Authority, Environmental Quality Department. Technical Report No. 00-05.

Rex, A.C. and Taylor, D.I. 1998. CW/QAPP for Water Quality Monitoring and Combined Sewer Overflow Receiving Water Monitoring in Boston Harbor and Its Tributary Rivers. Massachusetts Water Resources Authority, Environmental Quality Department. Draft, April 30, 1998.

Sokal R.R. and Rohlf, F.J. 1995. Biometry, The Principles and Practice of Statistics in Biological Research. 3rd edition. W.H. Freeman and Company. New York.

Solow, A.R. 1993. Letter Report on Statistical Analysis of CSO Data. Submitted to MWRA October 1993.

Standard Methods for the Examination of Water and Wastewater. 1995. (eds) A.D. Eaton, L.S. Clesceri, and Greenberg, A.E. 19th Edition. APHA, AWWA, and WEF.

APPENDIX A

Geographic Region	Indicator Organism	Before (1989-1991)			After (1992-2000)			p
		n value	mean	95% C.I.	n value	mean	95% C.I.	
Dorchester Bay	coliform	760	26	22, 31	1868	8.0	7.2, 8.8	0.0001
	enterococci	759	8.5	7, 10	1859	4.0	3.4, 4.1	0.0001
Charles River	coliform	945	588	529, 654	1661	179	162, 197	0.0001
	enterococci	751	147	126,172	1672	35	32, 39	0.0001
Mystic River	coliform	354	276	229, 334	1469	104	93, 117	0.0001
	enterococci	353	46	38, 57	1462	19	17, 22	0.0001
Inner Harbor	coliform	1028	65	57, 75	2496	14	13, 15	0.0001
	enterococci	1030	19	17, 22	2465	5.1	4.7, 5.6	0.0001
Neponset River	coliform	317	351	278, 441	1007	263	229, 303	0.0074
	enterococci	317	143	110, 184	942	81.8	70, 96	0.0004
System Wide	coliform	3404	134	123, 145	12629	19	18, 20	0.0001
	enterococci	3210	35	32, 38	12459	7.3	7.3, 7.6	0.0001
System Wide[1]	coliform	3404	134	123, 145	8501	42	40, 44	0.0001
	enterococci	3210	35	32, 38	8400	12.6	11.8, 13.4	0.0001
Deer Island TP	coliform	nd	nd	nd	1328	2.4	2.1, 2.6	na
	enterococci	nd	nd	nd	1325	1.1	1.0, 1.13	na
Nut Island TP	coliform	nd	nd	nd	922	6.1	5.2, 7.0	na
	enterococci	nd	nd	nd	932	5.4	4.6, 6.3	na
Winthrop Bay	coliform	nd	nd	nd	93	9.8	6.0, 15.8	na
	enterococci	nd	nd	nd	94	3.3	2.0, 5.2	na
Outer Harbor	coliform	nd	nd	nd	1194	2.9	2.5, 3.3	na
	enterococci	nd	nd	nd	1144	1.7	1.4, 1.9	na
Station 77-Quincy Bay	coliform	nd	nd	nd	268	2.6	2.0, 3.4	na
	enterococci	nd	nd	nd	255	0.9	0.7, 1.2	na

Table A1. Geometric means and 95% confidence intervals along with the associated n-value according to geographic regions, indicator organism and year of testing. P - p-value results from a nonparametric Wilcoxon rank sums test comparing all individual observations from the before and after periods. 1- this system wide subset is due to removing data points arising from the Outer Harbor, Deer Island, and Nut Island (also stations 130, 133, 134, 124, and 77) due to the lower counts represented by these marine samples and lack of data available prior to 1991. nd – no data. na – not applicable. Cell concentrations are in units of cells/100 mL.

Year	fecal coliforms		enterococci	
	Marine[1]	Fresh/Estuarine[2]	Marine[1]	Fresh/Estuarine[2]
1992	nd	64.77	nd	17.16
1993	1.78	33.73	1.04	11.17
1994	2.75	25.02	1.71	7.98
1995	3.00	31.04	1.75	9.39
1996	5.18	40.57	3.48	15.81
1997	2.52	45.13	1.77	12.26
1998	3.12	69.96	1.40	18.30
1999	5.11	47.43	2.72	24.57

Table A2. Yearly mean bacterial counts (cells/100mL) in Marine vs. Freshwater/Estuarine regions. Note: The means only include data from the after improvement period. 1- Marine regions are defined as those that do not contain any sampling data prior to 1992 and include Deer Island, Nut Island, Winthrop Bay, Outer Harbor, and Quincy Bay. These regions correspond to the regions in Table A.1 where no data is available prior to 1992. 2 – Freshwater/Estuarine regions include regions where data is available for both the before and after improvement stratification strategy. See table A.3 below for a T-test performed on these means with marine compared to freshwater/estuarine regions. nd- no data available.

Organism	Region	Mean[1]	95%C.I.	P-value
fecal coliforms	Marine	3.4	2.2, 4.6	0.0002
	Fresh/Estuarine	44.7	31.4, 58.0	
enterococci	Marine	2.0	1.2, 2.8	0.0004
	Fresh/Estuarine	14.6	10.0, 19.2	

Table A3. Results for a T-test performed on the aggregated means from Table A.2. (all data was normally distributed and the variances were not homogeneous). 1- cell concentrations are in units of cells/100 mL. Regions conform to Table A.2. See text for further details.

General Linear Model for Yearly Averages

Results from a nested ANOVA on the yearly means for each station

The yearly mean for each station within each geographic region was nested within the before- and after-improvement periods. While the means are slightly different from those provided in Table A.I due to the effect of averaging and changes in the n-value associated with each station/region/period combination, the results provide the same answer. For both enterococci and Fecal coliforms there has been a significant change in the before and after levels of indicator bacteria measured as system wide grand means (before/after p-values given below are 0.002 and 0.004) and for regions within the major division of before and after improvement years (p values are <0.0005).

Table A4. Factors

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Before/After	1	3.4265	2.5387	2.5387	10.76	0.002
Region(before /after)	8	20.0214	20.0214	2.5027	10.60	0.000
Error	74	17.4650	17.4650	0.2360		
Total	83	40.9129				

Analysis of Variance for fecal coliforms, using Adjusted SS for Tests

Source	**DF**	**Seq SS**	**Adj SS**	**Adj MS**	**F**	**P**
before/after	1	2.9049	2.1085	2.1085	9.01	0.004
Region(before/after)	8	28.2304	28.2304	3.5288	15.09	0
Error	74	17.3088	17.3088	0.2339		
Total	83	48.4441				

Unusual Observations for fecal coliforms

Obs	**log(cells)**	**Fit**	**StdFit[1]**	**Residual**	**St Resid[2]**	**Region**
24	2.54	1.35	0.16	1.19	2.61R	Dorchester
44	3.26	2.25	0.14	1.01	2.19R	Charles
52	1.27	2.25	0.14	-0.97	-2.10R	Charles
79	3.25	2.33	0.18	0.92	2.07R	Mystic
80	2.74	1.38	0.17	1.36	3.01R	Inner Har.

1- standard deviation of the fit, 2- standardized residual, R – denotes an observation with a large standardized residual.

Factor	Type	Levels	Values
before/after	fixed	2	1 2
Region(before/after)	fixed	10	Charles River Dorchester Bay Inner Harbor Mystic River Neponset River Charles River Dorchester Bay Inner Harbor Mystic River Neponset River

Analysis of Variance for *enterococci*, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
before/after	1	3.4265	2.5387	2.5387	10.76	0.002
Region(before/after)	8	20.0214	20.0214	2.5027	10.60	0.000
Error	74	17.4650	17.4650	0.2360		
Total	83	40.9129				

Unusual Observations for *enterococci*

Obs	log(cells)	Fit	FitStd	Residual	St Resid	Region	Station
10	1.29	2.29	0.14	-1.00	-2.16R	Charles	10 (before)
37	2.65	1.74	0.18	0.90	2.01R	Mystic	70 (before)
79	2.65	1.61	0.18	1.04	2.32R	Mystic	70 (after)
80	2.21	0.91	0.17	1.30	2.85R	Inner Har.	75 (after)

R denotes an observation with a large standardized residual.

Large standardized residuals are indicative of outliers that unduly influence the results of the ANOVA.

Analysis of Variance for fecal coliforms, using Adjusted SS for Tests

Source	DF	Seq SS	Adj SS	Adj MS	F	P
before/after	1	2.9049	2.1085	2.1085	9.01	0.004
Region(before/after)	8	28.2304	28.2304	3.5288	15.09	0.000
Error	74	17.3088	17.3088	0.2339		
Total	83	48.4441				

Unusual Observations for fecal coliforms

Obs	log(cells)	Fit	StdFit[1]	Residual	St Resid[2]	Region	Station
24	2.54	1.35	0.16	1.19	2.61R	Dorchester	39 (before)
44	3.26	2.25	0.14	1.01	2.19R	Charles	2(after)
52	1.27	2.25	0.14	-0.97	-2.10R	Charles	9(after)
79	3.25	2.33	0.18	0.92	2.07R	Mystic	70(after)
80	2.74	1.38	0.17	1.36	3.01R	Inner Har.	75(after)

1- standard deviation of the fit, 2- standardized residual, R – denotes an observation with a large standardized residual.

Massachusetts Water Resources Authority
Charlestown Navy Yard
100 First Avenue
Boston, MA 02129
(617) 788-4719