Receiver operating characteristic (ROC) curve analysis of antecedent rainfall and the Alewife/Mystic River receiving waters

Massachusetts Water Resources Authority

Environmental Quality Department
Report ENQUAD 2005-23

Citation

Morrison, Ann Michelle. 2005**. Receiver Operating Characteristic (ROC) Curve Analysis of Antecedent Rainfall and the Alewife/Mystic River Receiving Waters.** Boston: Massachusetts Water Resources Authority. Report ENQUAD 2005-23. 26 p.

Receiver Operating Characteristic (ROC) Curve Analysis of Antecedent Rainfall and the Alewife/Mystic River Receiving Waters

Submitted to

Massachusetts Water Resources Authority
Environmental Quality Department
100 First Avenue
Charleston Navy Yard
Boston, MA 02129
(617) 242-6000

prepared by

Ann Michelle Morrison
Morrison Environmental Data Services
376 Ridge Road
Northwood, NH 03261

submitted by

Battelle
397 Washington Street
Duxbury, MA 02332

May 2005

Report No. 2005-23

TABLE OF CONTENTS

LIST OF TABLES

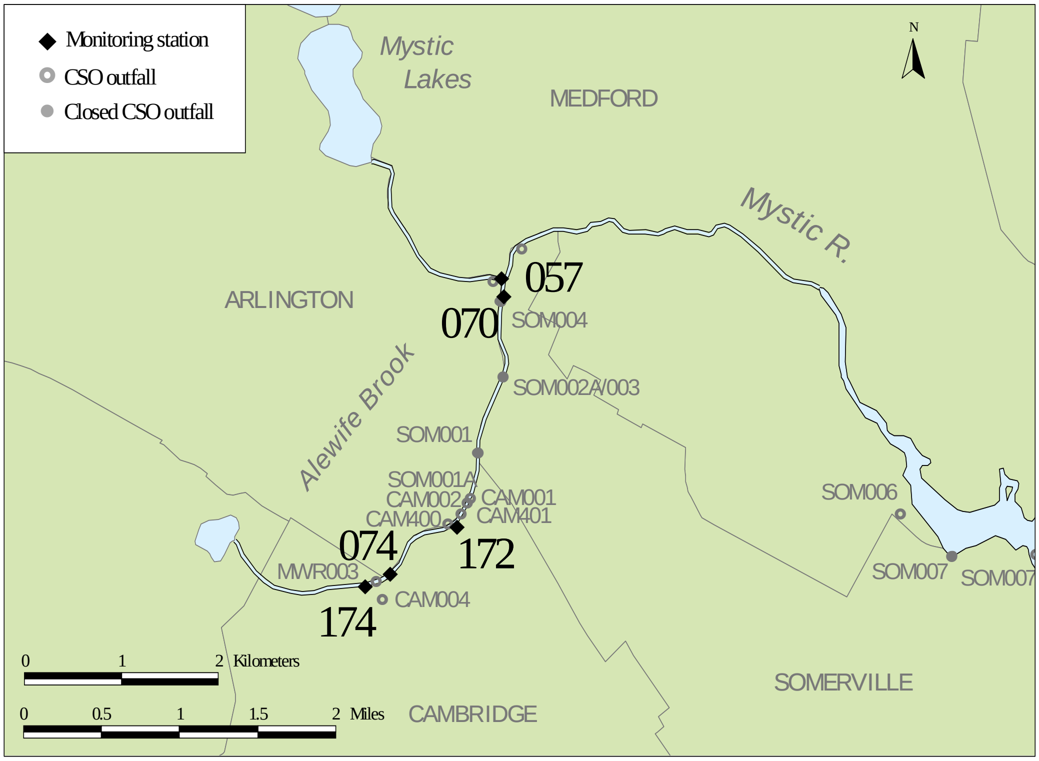
LIST OF FIGURES

1.0 INTRODUCTION

1.1 Background

Alewife Brook forms the boundary between Cambridge, Somerville, and Arlington and is a tributary of the Mystic River (Figure 1-1). This watershed is located in a heavily populated urban region and impacted by numerous pollution sources. Currently eight combined sewer overflows (CSOs) are located on Alewife Brook and discharge during moderate to heavy rainstorms to prevent sewer system backups into homes, businesses, and streets; the brook is also heavily impacted by stormwater (MWRA 2003).

Figure 1 -1. Alewife Brook/Mystic River Watershed. Sampling locations and CSO outfalls are indicated on the map.



Since 1989, the Massachusetts Water Resources Authority (MWRA) has performed an intensive water quality monitoring program of CSO receiving waters, which has included Alewife Brook and the Mystic River. Water samples from the CSO receiving waters are collected regularly and analyzed for fecal coliform and *Enterococcus*. In April, 2002 *Escherichia coli* was added to the list of monitored bacteria species.

1.2 Previous Work

Several previous reports characterized the water quality of Alewife Brook (Ellis and Rosen 2000; Gong et al. 1997; Gong et al. 1998; Leo et al. 1994; MWRA 2003; MWRA 2004; Rex 1991; Rex 1993). They report a strong relationship between antecedent rainfall and increased indicator bacteria densities.

1.3 Data Characteristics

This report uses data collected from five sampling locations between June 1989 and May 2004. The five sampling locations are described in Table 1-1. 971 water samples were collected from the five sites during this period. Samples were collected throughout the year, but more samples were collected in the warmer months than the colder months due to weather conditions and accessibility.

Table 1-1. Sampling locations.

Sampling Location (upstream to downstream)	MWRA Location Code
Alewife Brook, Alewife T Station	174
Alewife Brook, Massachusetts Avenue Bridge	074
Little River	172
Alewife Brook, Mystic Valley Parkway	070
Mystic River, Mystic/Alewife Confluence	057

Data for three types of sewage indicator bacteria were used in this analysis. Fecal coliform was measured from June 1989 to July 2003. *Enterococcus* was measured throughout the study period, June 1989 to May 2004. *Escherichia coli* was measured from April 2002 to May 2004. Methods used to quantify the bacteria density in a water sample are listed in Table 1-2.

Table 1-2. Indicator bacteria variables measured in this study.

Variable	Method
Enterococcus	Standard Methods 9230C 2c, membrane filtration (for samples collected 1989 – 1998) EPA Method 1600 (for samples collected 1999–2004)
Fecal coliform	Standard Methods 9222D, membrane filtration
E. coli	EPA method 1603, membrane filtration on modified M-TEC

Antecedent rainfall variables were computed using rain data from the National Weather Service rain gauge at Logan International Airport. The objective in selecting appropriate antecedent rainfall measures was to utilize data that will be readily available for rapid reporting of water quality of Alewife Brook. Consequently, rain variables do not include rainfall totals from the day the sample was collected, which would be only available for a retrospective analysis and not daily management. The rainfall variables used in this study are described in Table 1-3.

Table 1-3. Rainfall variables used in this study.

Variable	Description
Day – 1 Rain	Total rainfall collected at the Logan Airport National Weather Service rain gauge during the previous calendar day, excluding the day the sample was taken
Day – 2 Rain	Total rainfall collected at the Logan Airport National Weather Service rain gauge during the previous 2 calendar days, excluding the day the sample was taken
Day – 3 Rain	Total rainfall collected at the Logan Airport National Weather Service rain gauge during the previous 3 calendar days, excluding the day the sample was taken

1.4 Study Objectives

Previous studies have described a strong association between antecedent rainfall and indicator bacteria levels. Bacteria enumeration requires at least 24 hours; however, a real-time tool is needed to predict water quality. Receiver operating characteristic (ROC) analysis quantifies the relationship between antecedent rainfall and exceedence of bacteria indicator variables. It provides a common metric to compare the different analyses and facilitates the selection of a rainfall threshold value. The objectives of this study are to examine the relationship between antecedent rainfall and indictor bacteria using ROC analysis and determine a threshold level of rainfall that is likely to be associated with exceedence of safe swimming and/or boating standards.

2.0 ANALYTICAL METHODS

2.1 Receiver Operating Characteristic (ROC) Curves

Receiver Operating Characteristic (ROC) curves were developed in the field of statistical decision theory, and later used in the field of signal detection for analyzing radar images during World War II (Collison 1998). ROC curves enabled radar operators to distinguish between an enemy target, a friendly ship, or noise. ROC curves assess the value of diagnostic tests by providing a standard measure of the ability of a test to correctly classify subjects. The biomedical field uses ROC curves extensively to assess the efficacy of diagnostic tests in discriminating between healthy and diseased individuals (Metz 1978). ROC curves can (1) assess the overall discriminatory ability of different potential indicator variables by generating a common metric for comparison and (2) aid in the selection of a specific value of an indicator variable to use as a threshold, or limit, that provides a desired trade-off in the true positive rate and false positive rate. With respect to water quality, ROC curves can quantify the overall effectiveness of different indicator variables, such as antecedent rainfall, to correctly classify recreational waters as suitable for swimming or boating and generate a single metric by which the different indicator variables can be compared.

ROC curves use the true positive rate and false positive rate associated with each level of an indicator variable to generate the curve. In this study, the true positive rate (TPR) is the proportion of samples that are correctly classified by the antecedent rainfall indicator variable as unsuitable for swimming or boating. The false positive rate (FPR) is the proportion of samples that are incorrectly classified by the antecedent rainfall indicator variable as unsuitable for swimming or boating. The TPR and FPR are expressed as percentages by multiplying the rate by 100. A detailed description of how to construct an ROC curve using Microsoft Excel software is provided in the appendix. Once the ROC curve is constructed, the area under the ROC curve (AUC) can be calculated and used as a common metric by which to determine the effectiveness of antecedent rainfall to predict whether or not the levels of indicator bacteria are likely to meet water quality standards. Figure 2-1 shows two sample ROC curves. A line of “no information” follows the diagonal of the graph with an area under the curve of 0.5 and denotes an indicator that is not effective. An ideal indicator variable has an AUC of 1.0. An AUC between 0.5 and 0.6 shows some relationship between the indicator variable and the outcome of interest (exceedance of the regulatory standard), but this indicator variable is not very useful in practice. AUCs above 0.7 show a strong relationship between the indicator variable and the outcome of interest and are potentially very useful for recreational water management if desired TPR and FPR values are associated with reasonable threshold bacteria counts.

Several CSOs were closed between 1997 and 1998, and these closings may have had an effect on the relationship between antecedent rainfall and exceedance of bacteria indicator organism swimming and boating standards. A preliminary ROC analysis was conducted using only data from 1997-2004, and the AUCs of the ROC curves generated for the restricted dataset were not significantly different from those of the complete dataset (t-Test, $p>0.05$). Consequently, all data available were used in the construction of ROC curves in this study (June 1989 – May 2004).

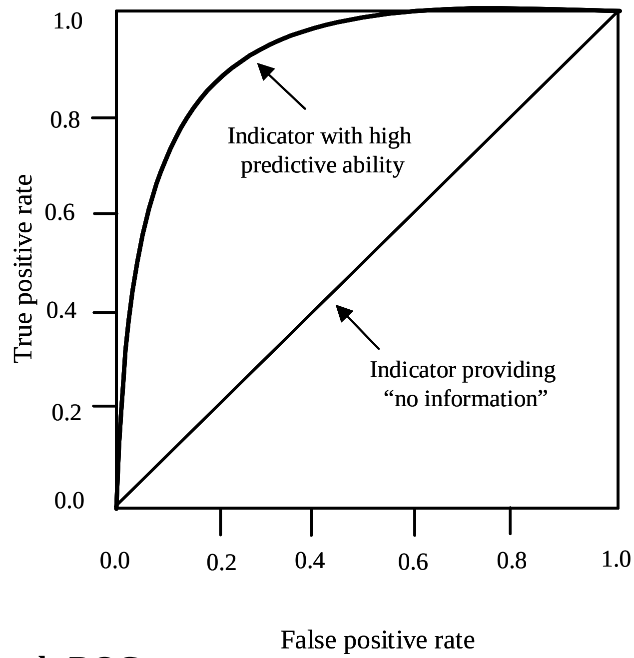


Figure 2-1. Sample ROC curves.
A line of "no information" follows the diagonal of the curve and has an AUC close to 0.5. A good indicator variable has high predictive ability and an AUC close to 1.0.

2.2 Safe Swimming and Boating Criteria

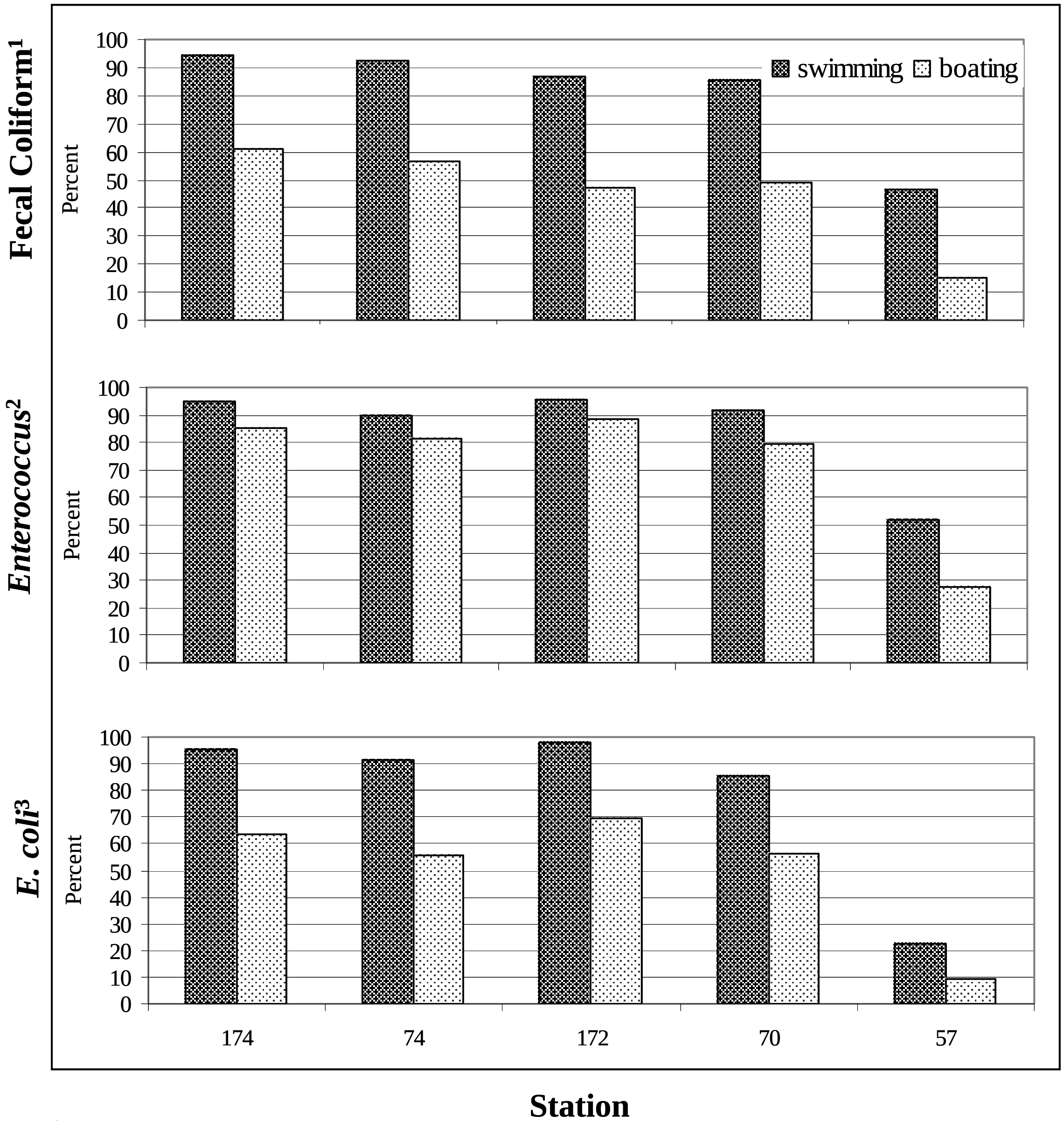
Criteria for bacterial indicators have been developed by EPA and state agencies for recreational waters, including guidelines for swimming and boating (Table 2-1). The swimming criteria limit the level of bacteria indicator organism densities to a lower level, and are, therefore, more restrictive. The boating criteria have higher indicator organism densities and are less restrictive. Initial analysis of the Alewife Brook data revealed that water samples frequently fail to meet the swimming criteria at all sampling locations in both wet and dry weather (Figure 2-2). Samples more frequently met the boating criteria. ROC analyses for each sampling location presented in this report used the boating criteria for each of the indicator organisms to determine if a sample exceeded the regulatory limit.

Table 2-1. Swimming and boating regulatory criteria*. Values are given in colony forming units (CFU)/100 mL.

Bacteria Indicator	Swimming Criteria	Boating Criteria
Enterococcus	61	151
Fecal coliform	200	1000
E. coli	235	576

* Maximum value. EPA's "infrequently used full-body contact" is assumed to be analogous with "boating".

Figure 2-2. Percent exceedance of swimming and boating criteria for all sampling locations on Alewife Brook, 1989-2004.



[1]Fecal coliform results are from June, 1989-July, 2003.
[2] *Enterococcus* results are from June, 1989 – May, 2004.
[3]*E. coli* results are from April, 2002- May, 2004.

2.3 Data preparation

Data were obtained from the Massachusetts Water Resources Authority and manipulated in a Microsoft Excel spreadsheet (Microsoft Corporation). Columns of data were created to describe whether a water sample was above the boating standard for each indicator organism. These data were then used to compute ROC curves using AccuROC for Windows (Accumetric Corporation). Standard error values were calculated within the AccuROC software based on the methods of DeLong, DeLong, and Clarke-Pearson (DeLong et al. 1988).

3.0 RESULTS AND DISCUSSION

3.1 ROC analysis of Mystic River, Mystic/Alewife Confluence

The Mystic River sampling location, station 057, had the fewest indicator bacteria exceedances of the five sites examined in this study. Station 057 is the most "downstream" sampling location on the Alewife Brook and is influenced by the upper Mystic River. 274 samples were measured for *Enterococcus;* 220 samples were measured for fecal coliform; and 44 samples were measured for *E. coli*. Based on ROC analysis, this location showed a strong relationship between antecedent rainfall and both fecal coliform and *Enterococcus* densities. Three ROC curves were generated for each bacteria indicator, "Day -1 Rain", "Day -2 Rain", and "Day -3 Rain", based on antecedent rainfall collected in the midnight to midnight period one, two and three days prior to sample collection. The AUCs for each of the antecedent rainfall variables were significantly different ($p<0.05$) from a line of no information, AUC = 0.5 (Table 3-1). "Day -2 Rain" produced an ROC curve with the largest AUC for *Enterococcus,* 0.72. "Day -1" produced an ROC curve with the largest AUC for fecal coliform, 0.73. The ROC curves generated for *E. coli* were not significantly different from a line of no information, suggesting that rainfall may be poorly associated with *E. coli* exceedances at this sampling location. *E. coli* had the smallest sample size at all locations in the study. More measurements collected in the future may help refine the association between antecedent rainfall and *E. coli* at the Alewife Brook. The ROC curves that produced the largest AUC for each bacteria indicator at the Mystic/Alewife confluence are shown in Figure 3-1.

Table 3-1. Areas under the ROC curves (AUC) associated with the boating standard for the three indicator bacteria and different rainfall variables at station 057, Mystic River, Mystic/Alewife Confluence. Standard error (SE) values for the areas under the curves are shown. * Indicates an AUC significantly different ($p<0.05$) from a line of no information (AUC = 0.5)

	***Enterococcus* n = 274**		**Fecal Coliform n = 220**		***E. coli* n = 44**	
	AUC	**SE**	**AUC**	**SE**	**AUC**	**SE**
Day -1 Rain	0.69 *	0.03	0.73 *	0.05	0.59	0.17
Day -2 Rain	0.72 *	0.03	0.70 *	0.05	0.60	0.20
Day -3 Rain	0.69 *	0.04	0.66 *	0.05	0.57	0.11

ROC curves provide information on the true positive rate and false positive rate associated with each unique value of the indicator variable antecedent rainfall. An ideal indicator variable will have a high true positive rate and a low false positive rate. "Day-2 Rain" produced the largest AUC for *Enterococcus*. From the information provided in the ROC analysis, 0.04 inches of rainfall two days prior to sample collection has a true positive rate of 0.68 and a false positive rate of 0.31 for *Enterococcus*. Therefore, if the location is posted as unsuitable for boating after 0.04 inches of 2 day antecedent rainfall, 68% of the time the water will be correctly classified as unsuitable for boating and 31% of the time the water will be incorrectly classified as unsuitable for boating because the *Enterococcus* density is, in fact, below the regulatory criteria for safe boating.

A potential threshold level for fecal coliform exceedence based on "Day-1 Rain" is 0.01 inches of 1 day antecedent rainfall. The true positive rate and false positive rate for fecal coliform exceedance associated with 0.01 inches of 1 day antecedent rainfall is 0.64 and 0.25, respectively. For practical purposes, any 1-day antecedent rainfall would indicate a likely exceedance of the fecal coliform criteria for boating.

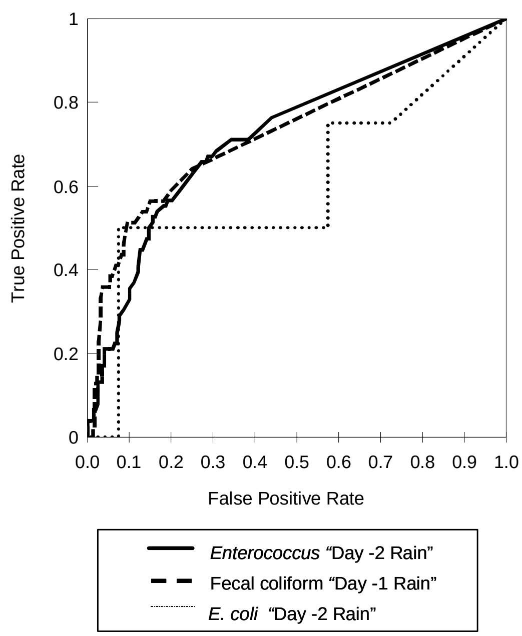


Figure 3-1. ROC curves for the Mystic River, Mystic/Alewife Confluence, station 057. The ROC curve that produced the largest AUC for each bacteria indicator is shown**.**

3.2 ROC analysis of Alewife Brook, Mystic Valley Parkway

The Alewife Brook, Mystic Valley Parkway sampling location, station 070, is downstream of all Alewife CSOs and storm drains but is less influenced by the Mystic River flow than station 057. Thus, this location may reasonably be considered the "most representative" station on the Alewife Brook. Station 057 had the largest number of samples collected during the 16 years of monitoring. 298 samples were measured for *Enterococcus*; 278 samples were measured for fecal coliform; and 48 samples were measured for *E. coli*. This location showed a consistent association of bacteria indicator organism exceedence and antecedent rainfall based on ROC analysis. Within each bacteria indicator organism group AUCs vary little between the antecedent rainfall variables and all AUCs are significantly different from a line of no information (Table 3-2). The largest AUCs were calculated with the antecedent rainfall variable

"Day-2 Rain." Interestingly, *E. coli* shows the strongest association with antecedent rainfall, which was not seen at the Mystic River, Alewife/Mystic River Confluence sampling location. The "Day -2 Rain" ROC curve for *E. coli* generated the largest AUC for the entire study at 0.78. The "Day-2 Rain" ROC curves for each of the bacteria indicator organisms are shown in Figure 3-2.

Table 3-2. Areas under the ROC curves (AUC) associated with the boating standard for the three indicator bacteria and different rainfall variables at station 070, Alewife Brook, Mystic Valley Parkway. Standard error (SE) values for the areas under the curves are shown. * Indicates an AUC significantly different ($p<0.05$) from a line of no information (AUC = 0.5).

	***Enterococcus* n =298**		**Fecal Coliform n=278**		***E. coli* n = 48**	
	AUC	**SE**	**AUC**	**SE**	**AUC**	**SE**
Day -1 Rain	0.65 *	0.02	0.65 *	0.03	0.72 *	0.06
Day -2 Rain	0.66 *	0.03	0.66 *	0.03	0.78 *	0.07
Day -3 Rain	0.65 *	0.04	0.65 *	0.03	0.76 *	0.07

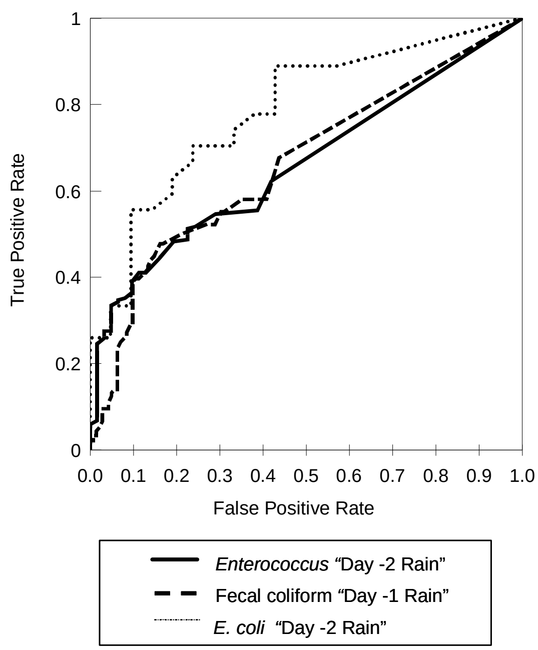


Figure 3-2. ROC curves for the Alewife Brook, Mystic Valley Parkway sampling location, station 070. The ROC curve that produced the largest AUC for each bacteria indicator is shown.

ROC curves from each of the three bacteria indicator organisms suggest that antecedent rainfall is potentially a good indictor variable; however, it is important to consider the threshold antecedent rainfall values and the true positive and false positive rates associated with each. Though "Day-2 Rain" produced the largest AUC for both *Enterococcus* and fecal coliform, there was very little difference between the AUCs for each of the antecedent rainfall variables. Antecedent rainfall thresholds and associated true positive rate and false positive rate values for each bacteria indicator organism and antecedent rainfall indicator variable are shown in Table 3-3.

Table 3-3. Antecedent rainfall thresholds and associated true positive rate (TPR) and false positive rate (FPR) values for each bacteria indicator organism and antecedent rainfall indicator variable at Alewife Brook, Mystic Valley Parkway.

		Rain Threshold (in.)	TPR	FPR
Enterococcus	Day -1 Rain	0.01	0.41	0.15
	Day -2 Rain	0.01	0.62	0.42
	Day -3 Rain	0.06	0.60	0.40
Fecal Coliform	Day -1 Rain	0.01	0.46	0.22
	Day -2 Rain	0.01	0.68	0.44
	Day -3 Rain	0.09	0.63	0.36
E. coli	Day -1 Rain	0.01	0.59	0.24
	Day -2 Rain	0.23	0.70	0.24
	Day -3 Rain	0.31	0.70	0.33

The Alewife Brook, Mystic Valley Parkway area shows the most promise to successfully implement an antecedent rainfall indicator program for *E. coli* exceedance. Because all ROC curves produced for this sampling location had AUCs significantly different from a line of no information, the ROC curves were constructed again using the swimming standard for each bacteria indicator organism. Surprisingly, "Day-3 Rain" produced an ROC curve for *E. coli* with the highest AUC in the study. AUCs for other indicator organisms and antecedent rainfall variables changed little and usually decreased (Table 3-4). This suggests that "Day -3 Rain" may successfully be used to notify the community of likely poor water quality in Alewife Brook, Mystic Valley Parkway. A TPR of approximately 0.75 is associated with 0.09 inches of antecedent 3 day rain for the *E. coli* swimming standard and 0.27 inches of antecedent 3 day rain for the *E. coli* boating standard. The FPRs for the *E. coli* swimming and boating standards are 0.29 and 0.38, respectively.

Table 3-4. Areas under the ROC curves (AUC) associated with the boating and swimming standards for the three indicator bacteria and different rainfall variables at station 070, Alewife Brook, Mystic Valley Parkway. Standard error (SE) values for the areas under the curves are shown. * Indicates an AUC significantly different ($p<0.05$) from a line of no information (AUC = 0.5)

Boating	***Enterococcus*** **n =298**		**Fecal Coliform n=278**		***E. coli*** **n = 48**	
	AUC	**SE**	**AUC**	**SE**	**AUC**	**SE**
Day -1 Rain	0.65 *	0.02	0.65 *	0.03	0.72 *	0.06
Day -2 Rain	0.66 *	0.03	0.66 *	0.03	0.78 *	0.07
Day -3 Rain	0.65 *	0.04	0.65 *	0.03	0.76 *	0.07
Swimming						
Day -1 Rain	0.60 *	0.04	0.62 *	0.03	0.64 *	0.08
Day -2 Rain	0.65 *	0.04	0.68 *	0.03	0.79 *	0.07

Day -3 Rain	0.62 *	0.05	0.67 *	0.04	0.82 *	0.06

3.3 ROC analysis of Alewife Brook, Massachusetts Avenue

The third sampling location studied was the Alewife Brook, Massachusetts Avenue location, station 074. This station showed little relationship between antecedent rainfall and indicator bacteria exceedance based on ROC analysis. Table 3-5 shows the AUC values computed from the ROC curves generated. Only three curves were significantly different from a line of no information, but the AUCs for these curves are still very low. *Enterococcus* showed the best association between antecedent rainfall and exceedance of the *Enterococcus* boating standard. The AUCs for both "Day – 2 Rain" and "Day – 3 Rain" are 0.64. Only "Day – 1 Rain" produced an ROC curve significantly different from a line of no information for fecal coliform. The AUC associated with this curve is 0.60. No ROC curve generated for *E. coli* was significantly different from a line of no information. The ROC curves that produced the largest AUC for each bacteria indicator are shown in Figure 3-3. The "Day – 2 Rain" ROC curve is shown for *Enterococcus* in Figure 3-3.

Table 3-5. Areas under the ROC curves (AUC) associated with the boating standard for the three indicator bacteria and different rainfall variables at station 074, Alewife Brook, Massachusetts Avenue. Standard error (SE) values for the areas under the curves are shown. * Indicates an AUC significantly different ($p<0.05$) from a line of no information (AUC = 0.5)

	***Enterococcus* n = 144**		**Fecal Coliform n = 128**		***E. coli* n = 45**	
	AUC	**SE**	**AUC**	**SE**	**AUC**	**SE**
Day -1 Rain	0.59	0.05	0.60 *	0.04	0.63	0.08
Day -2 Rain	0.64 *	0.05	0.53	0.05	0.60	0.09
Day -3 Rain	0.64 *	0.05	0.51	0.05	0.56	0.09

From the information provided in the ROC analysis, antecedent rainfall and indicator bacteria exceedance are poorly related at the Alewife Brook, Massachusetts Avenue sampling location. 0.11 inches of rainfall three days prior to sample collection has a TPR of 0.60 and a FPR of 0.33 for *Enterococcus*. 0.03 inches of rainfall two days prior to sample collection has a TPR of 0.60 and a FPR of 0.41 for *Enterococcus*. For fecal coliform a "Day-1 Rain" threshold of 0.01 has a TPR of 0.44 and a FPR of 0.27. All of these TPR values are poor because the unsuitable water quality is not identified by the antecedent rainfall volume 40 – 56 % of the time. Poor TPR values and low AUCs indicate that antecedent rainfall is not a particularly useful indicator variable for recreational water management at the Alewife Brook, Massachusetts Avenue sampling location.

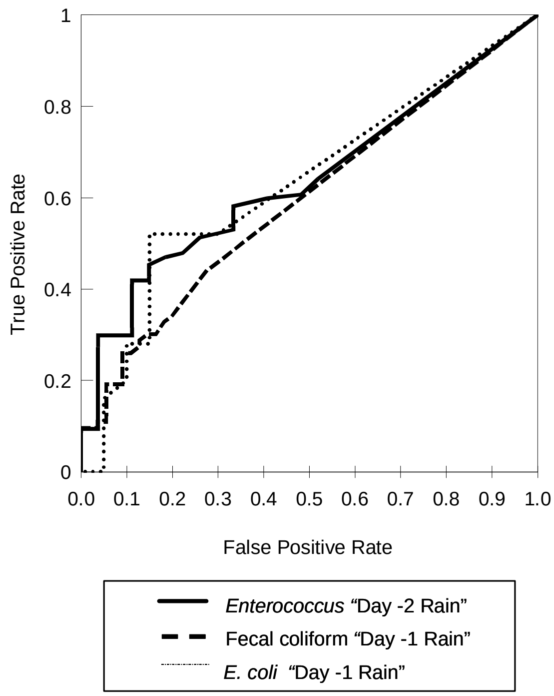


Figure 3-3. ROC curves for the Alewife Brook, Massachusetts Avenue sampling location, station 074. The ROC curve that produced the largest AUC for each bacteria indicator is shown.

3.4 ROC analysis of Little River

Based on ROC analysis, the Little River sampling location, station 172, also showed a poor relationship between antecedent rainfall and exceedance of the boating standard for each of the three bacteria indicator variables. "Day -1 Rain" produced ROC curves for fecal coliform and *E. coli* that were significantly different from a line of no information (Table 3-6). *Enterococcus* showed no relationship between the antecedent rainfall indicator variables and exceedance of the *Enterococcus* boating standard. The AUC for fecal coliform and "Day -1 Rain" was 0.63, and the AUC for *E. coli* and "Day -1 Rain" was 0.67. The ROC curves that produced the largest AUC for each bacteria indicator are shown in Figure 3-4.

Table 3-6. Areas under the ROC curves (AUC) associated with the boating standard for the three indicator bacteria and different rainfall variables at station 172, Little River. Standard error (SE) values for the areas under the curves are shown. * Indicates an AUC significantly different ($p<0.05$) from a line of no information (AUC = 0.5).

	***Enterococcus* n = 138**		**Fecal Coliform n = 122**		***E. coli* n = 46**	
	AUC	**SE**	**AUC**	**SE**	**AUC**	**SE**
Day -1 Rain	0.49	0.06	0.63 *	0.04	0.67 *	0.07
Day -2 Rain	0.55	0.06	0.59	0.05	0.65	0.09
Day -3 Rain	0.60	0.05	0.59	0.05	0.65	0.09

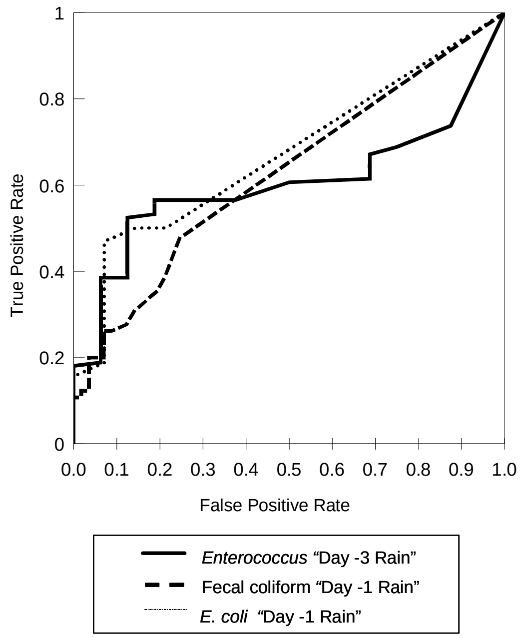


Figure 3-4. ROC curves for the Little River sampling location, station 172. The ROC curve that produced the largest AUC for each bacteria indicator is shown.

The TPR values associated with "Day -1 Rain" threshold levels are very poor. 0.01 inches of 1 day antecedent rain is associated with a TPR of 0.48 and an FPR 0.25 for fecal coliform. For *E. coli,* 0.01 inches of rain is associated with a TPR of 0.50 and an FPR of 0.21. The TPR values indicate that the association between antecedent rainfall and exceedance of these bacteria indicator organisms is the same as chance, and antecedent rainfall is not a useful indictor variable for recreation water management at the Little River sampling location.

3.5 ROC analysis of Alewife Brook, Alewife T Station

None of the ROC analyses performed for the Alewife Brook, Alewife T Station sampling location, station 174, produced an ROC curve that was significantly different from a line of no information ($p<0.05$) (Table 3-7). The *E. coli* analyses actually produced AUC values below 0.5, which suggests an inverse relationship between *E. coli* exceedance of the boating standard and 2 and 3 day antecedent rainfall such that *E. coli* exceedance of the boating standard actually decreases with increased rainfall. These results argue that the elevated *E. coli* densities seen at the Alewife Brook, Alewife T Station sampling location are the result of dry-weather contamination and not related to the discharge of the CSO near this sampling location. *Enterococcus* and fecal coliform ROC analyses produced AUC values around 0.5, suggesting no clear relationship between antecedent rainfall and exceedance of the boating standards for these bacteria indicator organisms. The ROC curves that produced the largest AUC for each bacteria indicator are shown in Figure 3-5. Interestingly, "Day-1 Rain" produced the largest AUCs for each bacteria indicator organism and the only ROC curve for *E. coli* above 0.5.

Table 3-7. Areas under the ROC curves (AUC) associated with the boating standard for the three indicator bacteria and different rainfall variables at station 174, Alewife Brook, Alewife T Station. Standard error (SE) values for the areas under the curves are shown. * Indicates an AUC significantly different ($p<0.05$) from a line of no information (AUC = 0.5)

	Enterococcus **n = 117**		**Fecal Coliform n = 102**		***E. coli*** **n = 44**	
	AUC	**SE**	**AUC**	**SE**	**AUC**	**SE**
Day -1 Rain	0.60	0.05	0.54	0.05	0.51	0.08
Day -2 Rain	0.57	0.07	0.51	0.06	0.39	0.10
Day -3 Rain	0.53	0.07	0.48	0.06	0.38	0.10

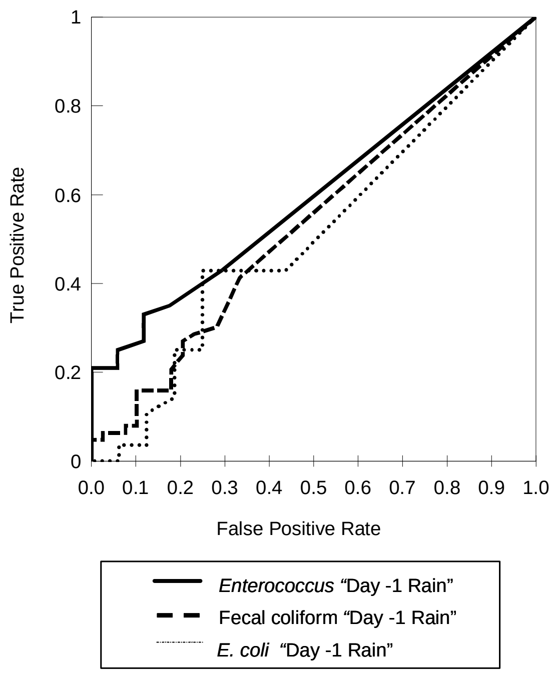


Figure 3-5. ROC curves for the Alewife Brook, Alewife T Station sampling location, station 174. The ROC curve that produced the largest AUC for each bacteria indicator is shown.

Because none of the ROC analyses produced an ROC curve significantly different from a line of no information ($p< 0.05$), there is no need to examine potential antecedent rainfall threshold values. Increased antecedent rainfall does not appear to be directly related to increased exceedance of the boating standard for any of the bacteria indicator organisms at the Alewife Brook, Alewife T Station sampling location.

4.0 CONCLUSIONS

Alewife Brook frequently has bacteria densities that exceed swimming and boating standards. Interestingly, the most downstream station on Alewife Brook had the most predictable relationship between antecedent rainfall and bacteria levels. At this station the calculated rain threshold of 0.27 inches occurring in the three previous days predicted whether *E. coli* would exceed the boating standard with a true positive rate of 75% and a false positive rate of 38%. Thus, an effective public notification process for Alewife Brook could entail a daily calculation of antecedent 3-day rainfall. If the rain total is greater than or equal to 0.27 inches, then posting public advisories to avoid contact with the water would, on average, be protective 75% of the time. This ROC study of five sampling locations along Alewife Brook has revealed that each sampling location is affected by antecedent rainfall differently and the utility of antecedent rainfall as a real-time indicator variable of exceedance of a bacteria indicator organism is dependent on the sampling location.

5.0 REFERENCES

Collison P. 1998. Of bombers, radiologists, and cardiologists: time to ROC. Heart 80(3):215-217.

DeLong ER, DeLong DM, Clarke-Pearson DL. 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. Biometrics 44:837-845.

Ellis B, Rosen J. 2000. Statistical Analysis of Combined Sewer Overflow Receiving Water Data, 1989-1999. Massachusetts Water Resources Authority, Environmental Quality Department. Report nr Technical Report No. 2001-06.

Gong G, Lieberman J, 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, McLaughlin D. 1998. Statistical Analysis of Combined Sewer Overflow Receiving Water Data, 1989-1996. Massachusetts Water Resources Authority, Environmental Quality Department. Report nr Technical Report No. 98-09.

Leo W, Collins M, Domenica M, Kirschen P, Marx L, Rex A. 1994. Master planning and CSO facility planning baseline water quality assessment Master planning and CSO facility planning baseline water quality assessment. Boston, MA: Massachusetts Water Resources Authority. Report nr ms-024. 465 p.

Metz C. 1978. Basic principles of ROC analysis. Seminars in Nuclear Medicine 8:283-298.

MWRA. 2003. Final variance report for Alewife Brook and the Upper Mystic River. Boston, MA: Massachusetts Water Resources Authority. Report nr EOEA No. 10335.

MWRA. 2004. Combined sewer overflow control plan, annual progress report - 2003. Boston, MA: Massachusetts Water Resources Authority.

Rex A. 1991. Combined Sewer Overflow Receiving Water Monitoring: Boston Harbor and Tributary Rivers. June 1989-October 1990. Massachusetts Water Resources Authority, Environmental Quality Department. Report nr Technical Report No. 91-2.

Rex A. 1993. Combined Sewer Overflow Receiving Water Monitoring: Boston Harbor and its Tributary Rivers. October 1990-September 1991. Massachusetts Water Resources Authority. Technical Report No. 93-4.