Total Maximum Daily Loads of Fecal Bacteria for the Non-tidal Cabin John Creek Basin in Montgomery County, Maryland

FINAL

DEPARTMENT OF THE ENVIRONMENT
1800 Washington Boulevard, Suite 540
Baltimore MD 21230-1718

Submitted to:

Watershed Protection Division
U.S. Environmental Protection Agency, Region III
1650 Arch Street
Philadelphia, PA 19103-2029

October 2006

EPA Submittal Date: October 30, 2006
EPA Approval Date: March 14, 2007

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ARA	Antibiotic Resistance Analysis
ARCC	Average rates of correct classification
BMP	Best Management Practice
BST	Bacteria Source Tracking
CAFO	Confined Animal Feeding Operations
cfs	Cubic Feet per Second
CFR	Code of Federal Regulations
CFU	Colony Forming Units
COMAR	Code of Maryland Regulations
CSO	Combined Sewer Overflow
CWA	Clean Water Act
DNR	Department of Natural Resources
EPA	Environmental Protection Agency
FA	Future Allocation
GIS	Geographic Information System
LA	Load Allocation
MACS	Maryland Agricultural Cost Share Program
MDE	Maryland Department of the Environment
MDP	Maryland Department of Planning
ml	Milliliter(s)
MOS	Margin of Safety
MPN	Most Probable Number
MPR	Maximum Practicable Reduction
MRLC	Multi-Resolution Land Cover
MS4	Municipal Separate Storm Sewer System
NPDES	National Pollutant Discharge Elimination System
RCC	Rates of Correct Classification
SSO	Sanitary Sewer Overflows
SSURGO	Soil Survey Geographic
TMDL	Total Maximum Daily Load
USGS	United States Geological Survey
WQIA	Water Quality Improvement Act
WLA	Wasteload Allocation
WQLS	Water Quality Limited Segment
WWTP	Wastewater Treatment Plant

EXECUTIVE SUMMARY

This document, upon approval by the U.S. Environmental Protection Agency (EPA), establishes a Total Maximum Daily Load (TMDL) for fecal bacteria in Cabin John Creek (basin number 02-14-02-07). Section 303(d) of the Federal Clean Water Act (CWA) and the EPA's implementing regulations direct each state to identify and list waters, known as water quality limited segments (WQLSs), in which current required controls of a specified substance are inadequate to achieve water quality standards. For each WQLS, states are required to either establish a TMDL of the specified substance that the waterbody can receive without violating water quality standards or demonstrate that water quality standards are being met.

The Maryland Department of the Environment (MDE) has identified the non-tidal waters of Cabin John Creek, a Use I-P waterbody [Code of Maryland Regulations (COMAR) 26.08.02.08O] in the State's 303(d) List as impaired by nutrients (1996), sediments (1996), fecal bacteria (2002) and impacts to biological communities (2002). This document proposes to establish a TMDL of fecal bacteria in the non-tidal portions of Cabin John Creek to allow for the attainment of beneficial use designation, primary contact recreation. The listings for nutrients, suspended sediments and impacts to biological communities will be addressed separately at a future date. A data solicitation for fecal bacteria was conducted by MDE in 2003, and all readily available data from the past five years was considered.

To establish baseline and allowable pollutant loads for this TMDL, USGS flow data and bacteria monitoring data collected by MDE were used to establish flow strata and flow duration curves, respectively. The pollutant loads set forth in this document are for the entire Cabin John Creek watershed. The sources of fecal bacteria are estimated at one representative station in the Cabin John Creek watershed above the confluence of Cabin John Creek and the Potomac River where samples were collected for one year. Multiple antibiotic resistance analysis (ARA) source tracking was used to determine the relative proportion of the following source categories: domestic (pets and human associated animals), human (human waste), livestock (agricultural related animals) and wildlife (mammals and waterfowl).

The allowable load is determined by estimating a baseline load from current monitoring data. For the baseline load there is a corresponding long-term geometric mean that is estimated using weighting factors from the flow duration curve. A reduction in concentration proportional to a reduction in load is assumed and thus the TMDL is equal to the current baseline load with the required reduction applied. The TMDL load for fecal bacteria entering Cabin John Creek is established after considering four different hydrological conditions: high flow and low flow annual conditions; and high flow and low flow seasonal conditions (the period between May 1st and September 30th where water contact recreation is more prevalent). This allowable load is reported in the units of Most Probable Number (MPN)/day and represents a long-term load estimated over a variety of hydrological conditions.

Two scenarios were developed; the first assessing whether attainment of current water quality standards could be achieved with the maximum practicable reductions (MPRs) applied and the second with the maximum practicable reduction constraints relaxed (*i.e.,* greater reductions than might be feasible). Scenario solutions were based on an optimization method where the objective was to minimize the overall risk to human health, assuming that the risk varies over the four source categories. In the Cabin John Creek watershed, it was estimated that water quality standards could be attained with the MPRs. Thus, for the Cabin John Creek watershed the first scenario was used to establish a TMDL.

The fecal bacteria (*E. coli*) TMDL developed for the Cabin John Creek watershed is 176 billion MPN/day. The TMDL is distributed between load allocations (LA) for nonpoint sources and waste load allocations (WLA) for point sources [waste water treatment plants (WWTPs) and municipal separate storm sewer systems (MS4)]. There are no WWTPs in Cabin John Creek. The WLA (MS4) is 108 billion MPN/day. The LA is 68 billion MPN/day. The margin of safety is explicit and is expressed as a 5% reduction of the *E. coli* water quality criterion concentration, from 126 MPN/100ml to 119.7 MPN/100ml.

Once the EPA has approved a TMDL, and it is known what measures must be taken to reduce pollution levels, implementation of best management practices (BMPs) is expected to take place. MDE intends for the required reduction to be implemented in an iterative process that first addresses those sources with the largest impact to water quality and risk to human health, with consideration given to ease and cost of implementation. In addition, follow up monitoring plans will be established to track progress and to assess the implementation efforts. As previously stated, when applying practicable reduction rates the water quality standards can be attained in the Cabin John Creek watershed.

1.0 INTRODUCTION

Section 303(d)(1)(C) of the federal Clean Water Act (CWA) and the U.S. Environmental Protection Agency (EPA) implementing regulations direct each State to develop a Total Maximum Daily Load (TMDL) for each impaired water quality limited segment (WQLS) on the Section 303(d) List, taking into account seasonal variations and a protective margin of safety (MOS) to account for uncertainty. A TMDL reflects the total pollutant loading of the impairing substance a waterbody can receive and still meet water quality standards.

TMDLs are established to determine the pollutant load reductions needed to achieve and maintain water quality standards. A water quality standard is the combination of a designated use for a particular body of water and the water quality criteria designed to protect that use. Designated uses include activities such as swimming, drinking water supply, and shellfish propagation and harvest. Water quality criteria consist of narrative statements and numeric values designed to protect the designated uses. Criteria may differ among waters with different designated uses.

Cabin John Creek (basin number 02-14-02-07) has been designated a Use I-P waterbody [Code of Maryland Regulations (COMAR) 26.08.02.08O]. The Maryland Department of the Environment (MDE) has identified the non-tidal waters of Cabin John Creek on the State's 303(d) List as impaired by the following: nutrients (1996); sediments (1996); bacteria (2002); and impacts to biological communities (2002). This document, upon approval by the EPA, establishes a TMDL of fecal bacteria in the non-tidal portions of the Cabin John Creek to allow for the attainment of the primary contact recreation designated use. The listings for nutrients, suspended sediments and impacts to biological communities will be addressed separately at a future date. A data solicitation for fecal bacteria was conducted by MDE in 2003, and all readily available data from the past five years was considered.

Fecal bacteria are microscopic single-celled organisms (primarily fecal coliforms and fecal streptococci) found in the wastes of warm-blooded animals. Their presence in water is used to assess the sanitary quality of water used for primary contact recreation, molluscan bivalve (shellfish) consumption and drinking water. Excessive amounts of fecal bacteria in surface water used for recreation are known to indicate an increased risk of pathogen-induced illness to humans. Infections due to pathogen-contaminated recreation waters include gastrointestinal, respiratory, eye, ear, nose, throat and skin diseases (EPA, 1986).

In 1986, EPA published "Ambient Water Quality Criteria for Bacteria" wherein three indicator organisms were assessed to determine their correlation with swimming-associated illnesses. Fecal coliform, *E. coli* and *Enterococci* were the indicators used in the analysis. Fecal coliform are a subgroup of total coliform bacteria and *E. coli* are a subgroup of fecal coliform. Most *E. coli* are harmless and are found in great quantities in the intestines of people and warm-blooded animals; however, certain pathogenic strains may cause illness. Enterococci are a subgroup of bacteria in the fecal streptococcus group. Fecal coliform, *E. coli* and *Enterococci*, can all be classified as fecal bacteria. The results of the EPA study (EPA, 1986) demonstrated that fecal coliform showed less correlation to swimming-associated gastroenteritis than either *E. coli* or *Enterococci*.

The non-tidal Cabin John Creek watershed was listed on the Maryland 303(d) List using fecal coliform as the indicator organism. The State of Maryland used the 1986 EPA guidance as the basis of a 2004 water quality standards change from an indicator organism of fecal coliform to *E.coli*/*Enterococci* to fulfill requirements of the Beaches Act of 2000. Because multiple monitoring datasets are available within this watershed for various pathogen indicators, the general term fecal bacteria will be used to refer to the impairing substance throughout this document. The TMDL will be based on the pathogen indicator organisms specified in Maryland's current bacteria water quality criteria, either *E. coli* or *Enterococci*. The indicator organism used in the Cabin John Creek TMDL analysis was *E. coli.*

2.0 SETTING AND WATER QUALITY DESCRIPTION

2.1 General Setting

Location

The Cabin John Creek watershed is located in southern Montgomery County, Maryland, just northwest of Washington, DC (see Figure 2.1.1). The headwaters of Cabin John Creek originate in the City of Rockville. The creek flows south about 10 miles, passing under Interstate 270, through Cabin John Regional Park under the Capital Beltway (I-495) and the historic Cabin John Bridge to its confluence with the Potomac River near the towns of Cabin John and Glen Echo. The major tributaries of the Creek are Bogley Branch, Booze Creek, Buck Branch, Congressional Branch, Ken Branch, Old Farm Branch, Snakeden Branch and Thomas Branch (also called Beltway Branch). Cabin John Creek and all its tributaries are non-tidal. (Friends of Cabin John Creek Watershed: www.cabinjohn.org)

Geology/Soils

The Cabin John Creek watershed encompasses 16,424 acres (25.7 sq. mi). The watershed lies entirely in the Piedmont physiographic province. This province is characterized by gentle to steep rolling topography, low hills and ridges. The surficial geology is characterized by crystalline igneous and metamorphic rocks of volcanic origin consisting primarily of schist and gneiss.

The Cabin John Creek watershed lies predominantly in the Baile soil series (see Figure 2.1.2). Soils in this series are fine-loamy, mixed, mesic Typic Ochraquults and are very deep and poorly drained soils (Montgomery County, Maryland Soil Conservation Service, 1995).

Land Use

The 2002 Maryland Department of Planning (MDP) land use/land cover data show that the watershed is primarily urban. The watershed has been significantly affected by high-density residential and commercial development. There are parks, trails and natural areas throughout the watershed. In addition to the Regional Park, there are wooded parklands and buffer areas along several miles of the creek mainstem and tributaries. The land use percentage distribution for Cabin John Creek watershed is shown in Table 2.1.1, and spatial distributions for each land use are shown in Figure 2.1.3.

Table 2.1.1: Land Use Percentage Distribution for Cabin John Creek Basin

Land Type	Acreage	Percentage
Commercial	2,178	13.2%
Forest	2,129	13.0%
Pasture	100	0.6%
Residential	11,987	73.0%
Water	30	0.2%
Totals	**16,424**	**100%**

Population

The total population in the Cabin John Creek watershed is estimated to be approximately 75,170. Figure 2.1.4 describes the population density in the watershed. The human population and the number of households were estimated based on a weighted average from the Geographic Information System (GIS) 2000 Census Block and the MDP Land Use 2002 Cover that includes the Cabin John Creek watershed. Since the Cabin John Creek watershed is a sub-area of the Census Block, percentages of each land use within the watershed were used to extract the areas from the 2000 Census Block. Table 2.1.2 shows the number of dwellings per acre in the Cabin John Creek watershed. The number of dwellings per acre was derived from information for residential density (low, medium, high) from MDP land use cover.

Table 2.1.2: Number of Dwellings Per Acre

Land Use Code	Dwelling Per Acres
11 Low Density Residential	1
12 Medium Density Residential	5
13 High Density Residential	8

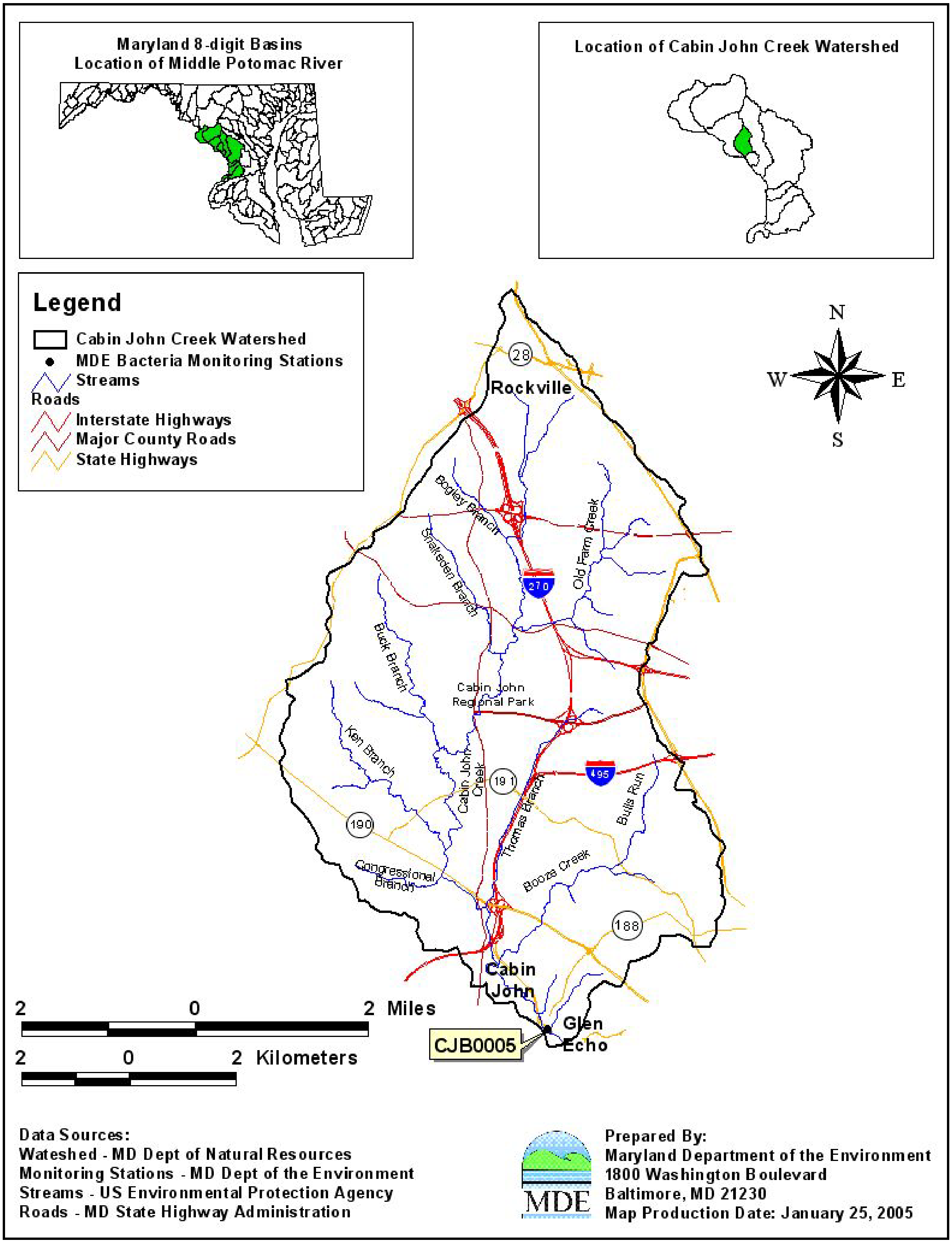

Figure 2.1.1: Location Map of the Cabin John Creek Watershed

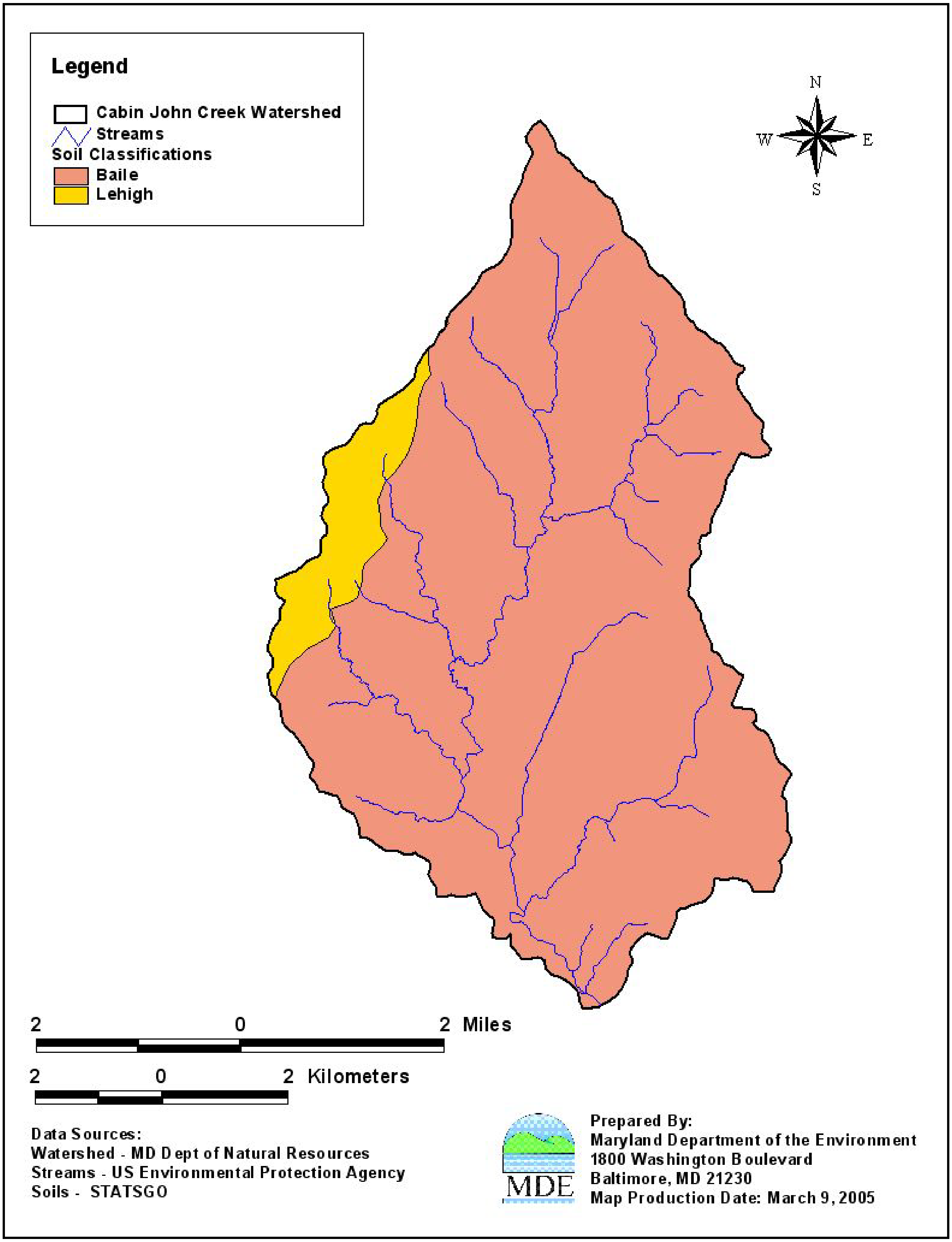

Figure 2.1.2: General Soil Series in the Cabin John Creek Watershed

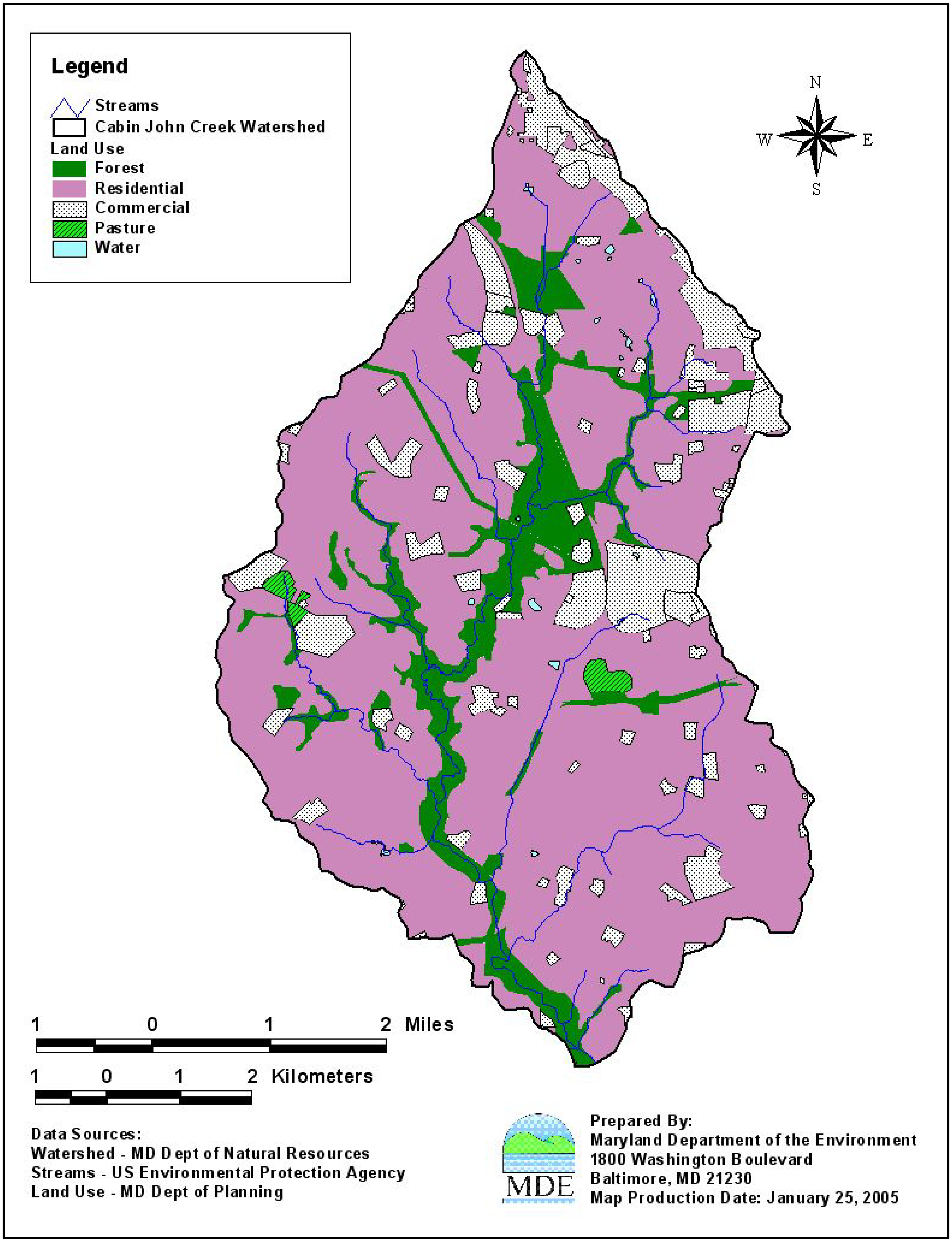

Figure 2.1.3: Land Use of the Cabin John Creek Watershed

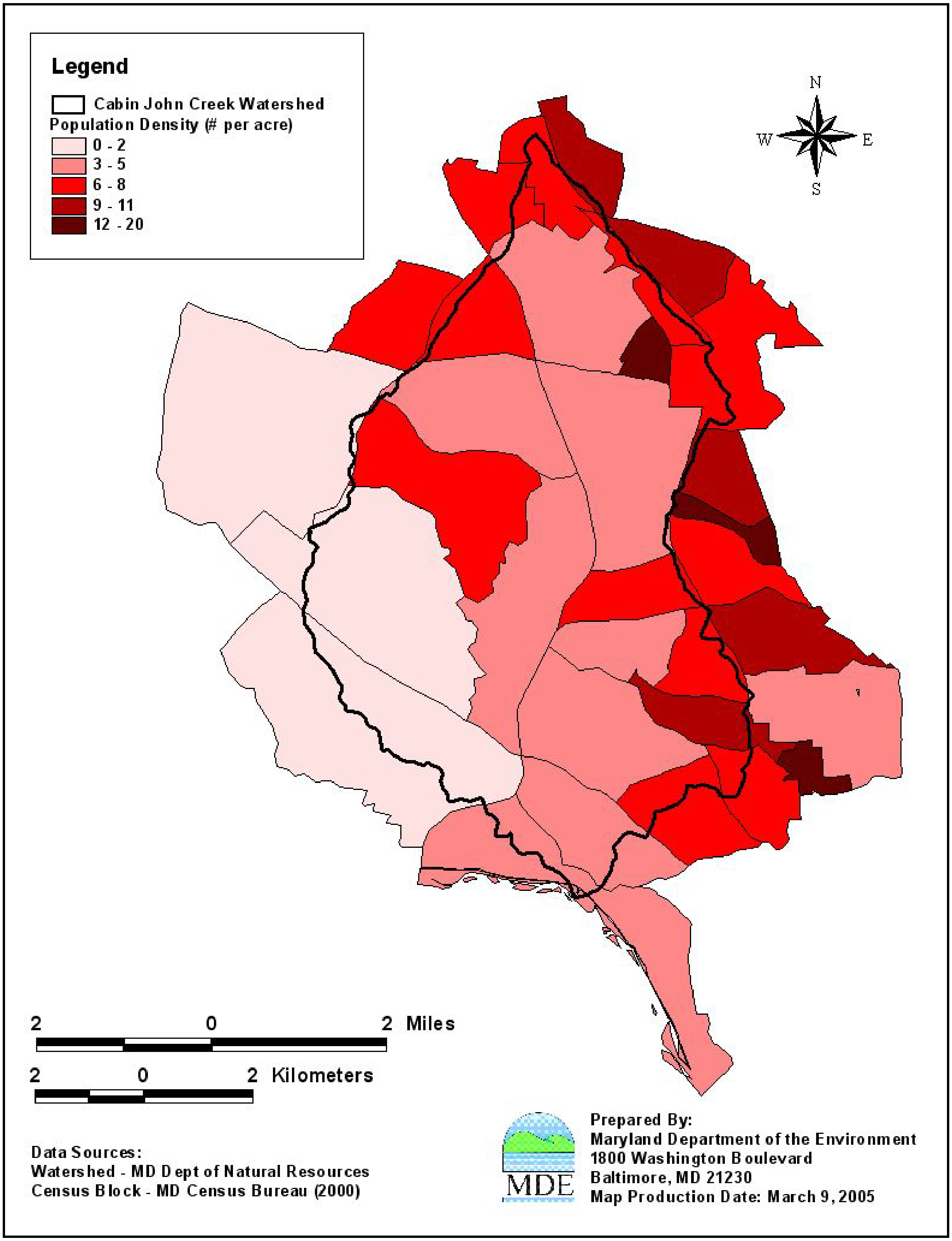

Figure 2.1.4: Population Density in the Cabin John Creek Watershed

2.2 Water Quality Characterization

From EPA's guidance document, "Ambient Water Quality Criteria for Bacteria" (1986), fecal bacteria, *E. coli* and *Enterococci* were assessed as indicator organisms for predicting human health impacts. A statistical analysis found that the highest correlation to gastrointestinal illness was linked to elevated levels of *E. coli* and *Enterococci* in fresh water (*Enterococci* in salt water), leading EPA to propose that States use *E. coli* or *Enterococci* as pathogen indicators. Maryland has adopted the EPA recommended bacterial indicators, *E. coli* and *Enterococcus*. Although the criteria numbers are different, the risk to the recreational bathers at the criteria levels are the same, thus the new indicators can better address this impairment although the impairment was identified using fecal coliform.

Bacteria Monitoring

Table 2.2.1 lists the historical monitoring data for the Cabin John Creek watershed. Monitoring Station CJB0005 (CORE) was used by the Maryland Department of Natural Resources (DNR) to identify the bacterial impairment. MDE conducted intensive monitoring from October 2002 through October 2003. There is one MDE monitoring station in the Cabin John Creek watershed, which is in the same location as the DNR CORE station CJB0005. There are no United States Geological Survey (USGS) gage stations in the Cabin John Creek watershed. A USGS gage station (01650500) located in the Anacostia River watershed with a drainage basin of similar size, land use and soils was chosen to represent the surface flow in the Cabin John Creek watershed. A comparison of the land use and soil types is found in Appendix B.

The locations of these stations are shown in Table 2.2.2 and Table 2.2.4 and illustrated in Figure 2.2.1. Observations recorded during the period 2002-2003 from MDE's monitoring station are displayed in Table A-1 and illustrated in Figure A-1 in Appendix A.

Table 2.2.1: Historical Monitoring Data in the Cabin John Creek Watershed

Sponsor	Location	Date	Design	Summary
Maryland Department of Natural Resources (DNR) Core Monitoring*	MD	1/8/97 – 4/1/98	Fecal Coliform	CJB0005 at Bridge at MacArthur Boulevard
MDE	MD	11/02 to 10/03	*E. coli*	1 station Enumeration 2x per month
MDE	MD	11/02 to 10/03	BST (*E. coli*)	1 station ARA Microbial Source Tracking (BST) 1x per month

* Data used for original listing; only *E. coli* data used for this analysis.

Table 2.2.2: Location of DNR (CORE) Monitoring Station in the Cabin John Creek Watershed

Monitoring Station	Observation Period	Total Observations	LATITUDE Deg-min	LONGITUDE Deg-min
CJB0005	1997 - 1998	16	38 58.4069	77 8.9301

Table 2.2.3: Location of MDE Monitoring Station in the Cabin John Creek Watershed

Monitoring Station	Obs. Period	Total Observations	LATITUDE Deg-min	LONGITUDE Deg-min
CJB0005	2002-2003	26	38 58.4069	77 8.9301

Table 2.2.4: Location of USGS Gage Station (Anacostia River)

Monitoring Station	Obs. Period	Total Observations	LATITUDE Deg-min	LONGITUDE Deg-min
1650500	1998-2003	5453	39 03.60	77 01.80

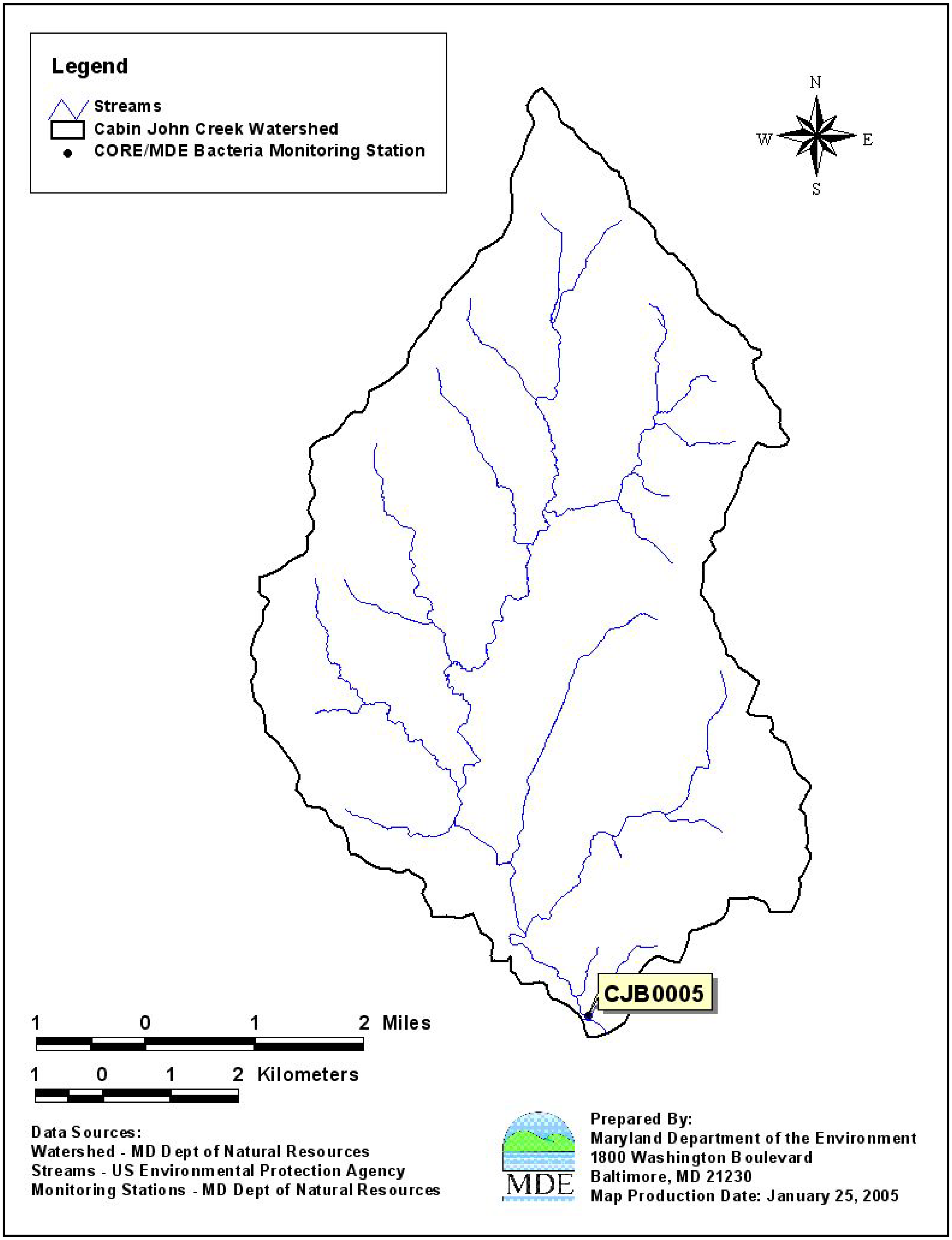

Figure 2.2.1: Monitoring Station in the Cabin John Creek Watershed

2.3 Water Quality Impairment

Designated Uses and Water Quality Standard

The Maryland water quality standards Surface Water Use Designation for the Cabin John Creek watershed area is Use I-P – Water Contact Recreation, and Protection of Aquatic Life and Public Water Supply (COMAR 26.08.02.08O). Cabin John Creek has been included on the final 2004 Integrated 303(d) List as impaired by bacteria.

Water Quality Criteria

The State water quality standard for bacteria (*E. coli*) used in this study is as follows (COMAR Section 26.08.02.03-3):

Table 2.3.1: Bacteria Criteria Values from Table 1 COMAR 26.08.02.03-3 Water Quality Criteria Specific to Designated Uses.

Indicator	Steady-State Geometric Mean Indicator Density
Freshwater	
E. coli	126 MPN/100ml

Interpretation of Bacteria Data for General Recreational Use

The relevant portion (for freshwater) of the listing methodology pursuant to the 2006 integrated 303(d) List for all Use Waters - Water Contact Recreation and Protection of Aquatic Life is as follows (MDE, January 2006):

Recreational Waters

A steady-state geometric mean will be calculated with available data where there are at least five representative sampling events. The data shall be from samples collected during steady-state conditions and during the beach season (Memorial Day through Labor Day) to be representative of the critical condition. If the resulting steady-state geometric mean is greater than 126 cfu/100 ml *E. coli* in freshwater, the waterbody will be listed as impaired. If fewer than five representative sampling events for an area being assessed are available, data from the previous two years will be evaluated in the same way. The single sample maximum criterion applies only to beaches and is to be used for closure and advisory decisions based on short term exceedances of the geometric mean portion of the standard.

Water Quality Assessment

A water quality impairment was assessed by comparing both the annual and the seasonal (May 1st –September 30th) steady-state geometric means of *E. coli* concentrations with the water quality criterion. The steady-state condition is defined as unbiased sampling targeting average flow conditions and/or equally sampling or providing for unbiased sampling of high and low flows. The 1986 EPA criteria document assumed steady-state flow in determining the risk at various bacterial concentrations, and therefore the chosen criterion value also reflects steady-state conditions for bacteria (EPA, 1986). The steady-state geometric mean condition can be estimated either by monitoring design or more practically by statistical analysis as follows:

1. A stratified monitoring design is used where the number of samples collected is proportional to the duration of high flows, mid flows and low flows within the watershed. This sample design allows a geometric mean to be calculated directly from the monitoring data.

2. Routine monitoring typically results in samples from varying hydrologic conditions (*i.e*., high flows, mid flows and low flows) where the numbers of samples are not proportional to the duration of those conditions. Averaging these results without consideration of the sampling conditions results in a biased estimate of the steady-state geometric mean. The potential bias of the steady-state geometric means can be reduced by weighting the samples results collected during high flow, mid flow and low flow regimes by the proportion of time each flow regime is expected to occur. This ensures that the high flow and low flow conditions are proportionally balanced.

3. If (1) the monitoring design was not stratified based on flow regime or (2) flow information is not available to weight the samples accordingly, then a geometric mean of sequential monitoring data can be used as an estimate of the steady-state geometric mean condition for the specified period.

A routine monitoring design was used to collect bacteria data in the Cabin John Creek watershed. To estimate the steady-state geometric means, the monitoring data was first reviewed by plotting the sample results versus their corresponding daily flow duration percentile. Graphs illustrating these results can be found in Appendix B.

To calculate the steady-state geometric means with routine monitoring data, a conceptual model was developed by dividing the daily flow frequency for the stream segment into strata that are representative of hydrologic conditions. A conceptual continuum of flows is illustrated in Figure 2.3.1.

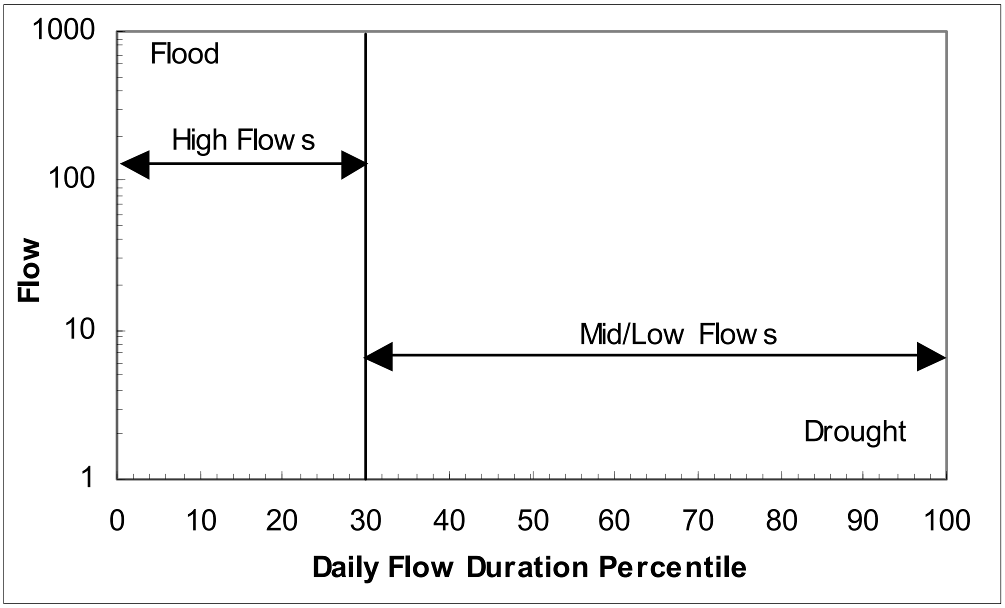

Figure 2.3.1: Conceptual Diagram of Flow Duration Zones

During high flows a significant portion of the total stream flow is from surface flow contributions. Low flow conditions represent periods with minimal rainfall and surface runoff. There is typically a transitional period (mid flows) between the high and low flow durations that is representative of varying contributions of surface flow inputs that result from differing rainfall volumes and antecedent soil moisture conditions. The division of the entire flow regime into strata enables the estimation of a less biased geometric mean from routine monitoring data that more closely approaches steady state. The daily flow duration intervals that define these regions and supporting details of how these zones were developed are presented in Appendix B.

Factors for estimating a steady-state geometric mean are based on the frequency of each flow stratum. The weighting factor accounts for the proportion of time that each stratum represents. The weighting factors for an average hydrological year are presented in the following table (Table 2.3.2).

Table 2.3.2: Weighting Factors Used for Estimation of Geometric Mean in the Cabin John Creek Watershed (Average Hydrological Year)

Flow Duration Zone	Duration Interval	Weighting Factor
High Flows	0 – 30%	0.30
Low Flows	30 – 100%	0.70

Bacteria enumeration results for samples within a specified stratum will receive their corresponding weighting factor. The steady-state geometric mean is calculated as follows:

$$M_i = \frac{\sum_{j=1}^{n_i} \log_{10}(C_{i,j})}{n_i} \qquad (1)$$

where

$$M = \sum_{i=1}^{2} M_i * W_i \qquad (2)$$

M_i = log mean concentration for stratum i
$C_{i,j}$ = Concentration for sample j in stratum i
n_i = number of samples in stratum I
M = weighted mean
W_i= Proportion of stratum i

Finally the weighted log mean is back transformed from log space using the following equation.

$$C_{gm} = 10^M \qquad (3)$$

C_{gm} = steady state geometric concentration

The results of the geometric mean by stratum and overall steady-state geometric mean in Cabin John Creek for annual and seasonal periods are presented in Table 2.3.3.

Table 2.3.3: Cabin John Creek Steady-State Geometric Mean by Stratum (Annual and Seasonal Period)

Tributary	Station	Period	Flow Stratum	Steady-State Geometric Mean MPN/100ml	Overall Steady-State Geometric Mean MPN/100ml
Cabin John Creek	CJB0005	Annual	High	306	210
			Low	179	
		Seasonal	High	283	250
			Low	237	

Summary of Water Quality Data

The water quality impairment was assessed by comparing the steady-state geometric mean concentrations of *E. coli* for the annual and critical condition with the water quality criteria. Graph illustrating these results can be found in Appendix B. Steady-state geometric means of the monitoring data and the water quality criterion for the annual and seasonal periods are shown in Table 2.3.4.

Table 2.3.4: Cabin John Creek Monitoring Data and Steady-State Geometric Mean (Annual and Seasonal Period)

Tributary	Station	Period	# Samples	*E. coli* Minimum Concentration MPN/100ml	*E. coli* Maximum Concentration MPN/100ml	*E. coli* Steady-State Geometric Mean Concentration MPN/100ml	*E. coli* Criterion MPN/100ml
Cabin John Creek	CJB0005	Annual	26	30	2850	210	126
		Seasonal	12	30	2850	250	

2.4 Source Assessment

Nonpoint Source Assessment

Nonpoint sources of fecal bacteria do not have one discharge point but occur over the entire length of a stream or waterbody. Many types of nonpoint sources introduce fecal bacteria to the land surface including the manure spreading process, direct deposition from livestock during the grazing season, and excretions from pets and wildlife. As the runoff occurs during rain events, surface runoff transports water and fecal bacteria over the land surface and discharges to the stream system. The deposition of non-human fecal bacteria directly to the stream occurs when livestock or wildlife have direct access to the waterbody. Nonpoint source contributions from human activities generally arise from failing septic systems and their associated drain fields or leaking infrastructure (*i.e.,* sewer systems). The transport of fecal bacteria from the land surface to the stream system is dictated by the rainfall, soil type, land use, and topography of the watershed.

Sewer and Septic Systems

The Cabin John Creek watershed is primarily serviced by sewers covering over 98% (16,246 acres) of the watershed area. Figure 2.4.1 depicts the areas that are serviced by sewers and septics. There are only three on-site disposal (septic) systems located in the southwest portion watershed.

Sanitary Sewer Oveflows (SSOs) occur when the capacity of a separate sanitary sewer is exceeded, normally during storm events. There are several factors that may contribute to SSOs from a sewerage system, including pipe capacity, operations and maintenance effectiveness, sewer design, age of system, pipe materials, geology and building codes. SSOs are prohibited by the facilities' permit and therefore must be reported to MDE's Water Management Administration in accordance to COMAR 26.08.10 to be addressed under the State's enforcement program.

There were a total of 29 SSOs in the Cabin John Creek watershed reported to MDE between January 26, 2001 and September 21, 2004. Approximately 119,924 gallons of SSO discharge was released into Cabin John Creek through various waterways (surface water, groundwater, sanitary sewers, etc.) (MDE, Water Management Agency). Figure 2.4.2 depicts the location of sanitary sewer overflows in the Cabin John Creek watershed.

Point Source Assessment

Stormwater

The Cabin John Creek watershed is located in Montgomery County, a Phase I National Pollutant Discharge Elimination System (NPDES) Municipal Separate Storm Sewer System (MS4) permit jurisdiction. The MS4 permit covers stormwater discharges from the municipal separate stormwater sewer system in the County.

Municipal and Industrial Wastewater Treatment Plants (WWTPs)

There are no municipal or industrial NPDES WWTPs with permits regulating the discharge of fecal bacteria directly into Cabin John Creek or its tributaries.

Bacteria Source Tracking

Bacteria source tracking (BST) was used to identify the relative contribution of bacteria in in-stream water samples. BST monitoring was conducted at one station in the Cabin John Creek watershed with 13 samples collected for a one-year duration. Sources are defined as domestic (pets and human associated animals), human (human waste), livestock (agricultural animals), wildlife (mammals and waterfowl) and unknown. To identify sources, samples are collected within the watershed from known fecal sources and the patterns of antibiotic resistance of these known sources are compared to isolates of unknown bacteria from ambient samples. Details of the BST methodology and data can be found in Appendix C.

An accurate representation of the expected average source observed at each monitoring station is estimated by using a stratified weighted mean of the identified sample results over the specified averaging period. The weighting factors are based on the $\log_{10}$ of the bacteria concentration and the percent of time that represents the high stream flow or low stream flow (see Appendix B). The procedure for calculating the stratified weighted mean of the sources per monitoring station is as follows:

1. Calculate the percentage of isolates per source per each sample date (S).
2. Calculate the weighted percentage (MS) of each source per flow strata (high/low). The weighting is based on the $\log_{10}$ bacteria concentration for the water sample.
3. The final weighted mean source percentage, for each source category, is based on the proportion of time in each flow duration zone (*i.e.*, high flow=0.3, low flow=0.7).

The weighted mean of each source category for the annual and critical condition is calculated using the following equations:

$$MS_{i,k} = \frac{\sum_{j=1}^{n_i} \log_{10}(C_{i,j}) * S_{i,j,k}}{n_i} \qquad (4$$

where

$MS_{i,k}$ = Weighted mean proportion of isolates for source k in stratum i
i = stratum
j = samples
k = Source category (1 = human, 2 = domestic, 3 = livestock, 4 = wildlife, 5 = unknown)
$C_{i,j}$ = Concentration for sample j in stratum i
$S_{i,j,k}$ = Proportion of isolates for sample j, of source k in stratum i
n_i = number of samples in stratum I

$$MS_k = \sum_{i=1}^{2} M_{i,k} * W_i \qquad (5$$

MS = weighted mean proportion of isolates of source k
W_i= Proportion covered by stratum i

The complete distributions of the annual and seasonal period source loads are listed in Table 2.4.1. Details of the BST data can be found in Appendix C.

Table 2.4.1: Distribution of Fecal Bacteria Source Loads in the Cabin John Creek Watershed (Annual and Seasonal Period)

Station	Period	Flow Stratum	% Domestic Animals	% Human	% Livestock	% Wildlife	% Unknown
CJB0005	Seasonal	High	35	28	21	1	15
		Low	25	41	11	9	14
		Weighted	28	37	14	7	14
	Annual	High	33	24	23	9	11
		Low	25	35	12	11	17
		Weighted	28	32	15	10	15

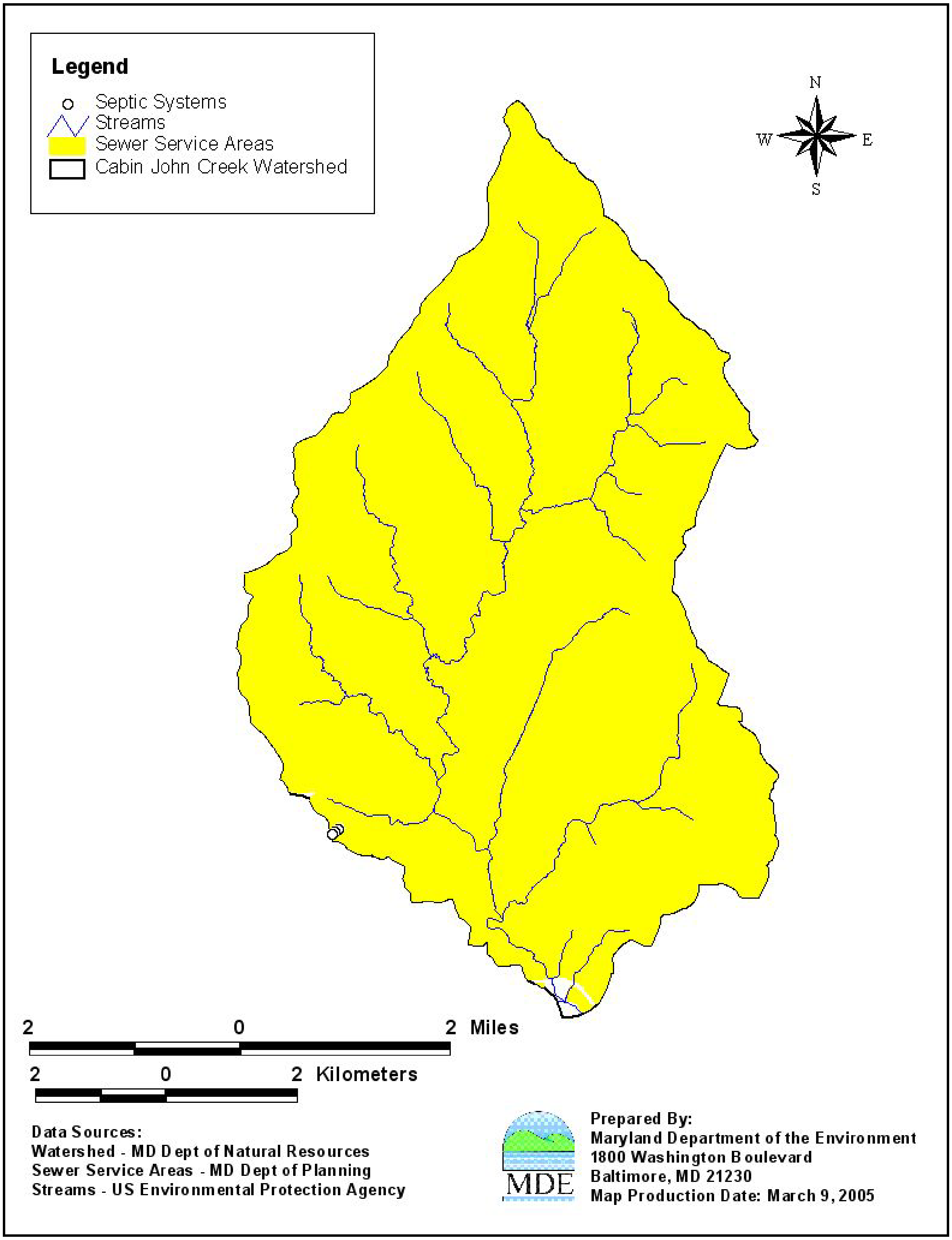

Figure 2.4.1: Sanitary Sewer Service and Septics Areas in the Cabin John Creek Watershed

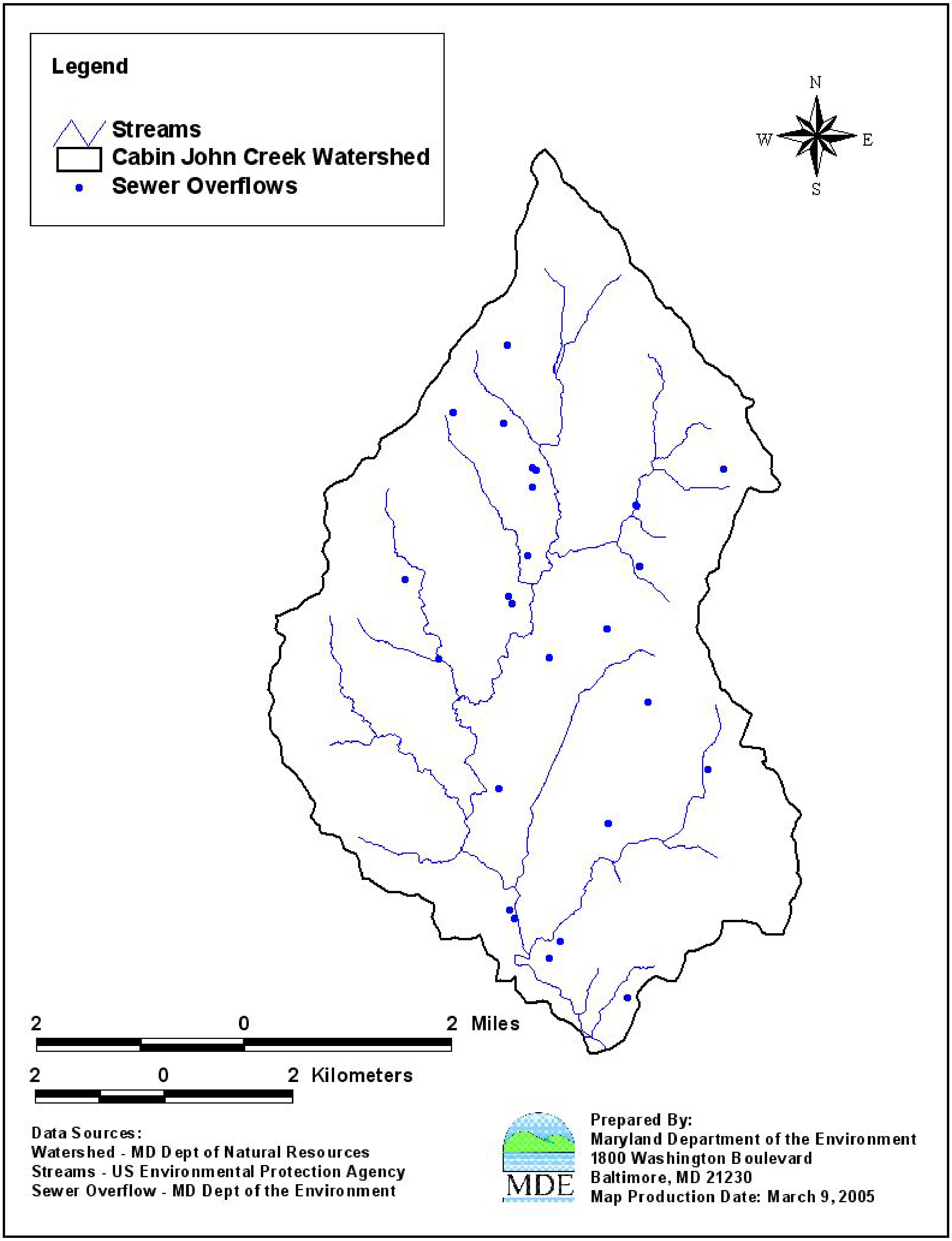

Figure 2.4.2: Sanitary Sewer Overflows in the Cabin John Creek Watershed

3.0 TARGETED WATER QUALITY GOAL

The overall objective of the fecal bacteria TMDL set forth in this document is to establish the loading caps needed to ensure attainment of water quality standards in the Cabin John Creek watershed area. These standards are described fully in Section 2.3, "Water Quality Impairment."

4.0 TOTAL MAXIMUM DAILY LOADS AND SOURCE ALLOCATION

4.1 Overview

This section provides an overview of the non-tidal fecal bacteria TMDL development, with a discussion on the many complexities involved with the estimation of bacteria concentrations, loads and sources. The second section presents the analysis framework and how the hydrological, water quality and BST data are linked together in the TMDL process. The third section describes the analysis for estimating a representative geometric mean fecal bacteria concentration and baseline loads. The analysis methodology is based on available monitoring data and specific to a free flowing stream system. The fourth section addresses the critical condition and seasonality. The fifth section presents the margin of safety. The sixth section discusses TMDL loading caps. The seventh section presents TMDL scenario descriptions. The eighth section presents the load allocations. Finally, in section nine, the TMDL equation is summarized.

To be most effective, the TMDL provides a basis for allocating loads among the known pollutant sources in the watershed so that appropriate control measures can be implemented and water quality standards achieved. By definition, the TMDL is the sum of the individual waste load allocations (WLA) for point sources, load allocations (LA) for nonpoint sources and natural background sources. A margin of safety (MOS) is also included and accounts for the uncertainty in the analytical procedures used for water quality modeling, and the limits in scientific and technical understanding of water quality in natural systems. Although this formulation suggests that the TMDL be expressed as a load, the Code of Federal Regulations (40 CFR 130.2(i)) states that the TMDL can be expressed in terms of "mass per time, toxicity or other appropriate measure".

For many reasons, bacteria are difficult to simulate in water quality models. They reproduce and die off in a non-linear fashion as a function of many environmental factors, including temperature, pH, turbidity (UV light penetration), and settling. They occur in concentrations that vary widely (*i.e.,* over orders of magnitude) and accurate estimation of source inputs are difficult to develop. Finally, limited data are available to characterize the effectiveness of any program or practice at reducing bacteria loads (Schueler, 1999).

Bacteria concentrations, determined through laboratory analysis of in-stream water samples for bacteria indicators (*e.g*., *Enterococci*), are expressed in either colony forming units (CFU) or most probable number (MPN) of colonies. The first method (Method 1600) is a direct estimate of the bacteria colonies (EPA, 1985), and the second (Method 9223B) is a statistical estimate of the number of colonies (APHA, 1998). Enumeration results indicate the extreme variability in the total bacteria counts. The distribution of the enumeration results from water samples tends to be lognormal, with a strong positive skew of the data. Estimating loads of constituents that vary by orders of magnitude can introduce much uncertainty and result in large confidence intervals around the final results.

Estimating bacteria sources can be problematic due to the many assumptions required and the limited available data. For example, when considering septic systems, information is required on spatial location of failing septic systems, consideration of transport to in-stream assessment location and estimation of the load from the septic system (degree of failure). Secondary sources, such as illicit discharges, also add to the uncertainty in a bacteria water quality model.

Estimating domestic animal sources requires information regarding the pet population in a watershed, how often the owners clean up after them, and the spatial location of the pet waste relative to the stream (for near-field upland transport). Livestock sources are limited by spatial resolution of Agricultural Census information (available at the county level), site-specific issues relating to animals' confinement, and confidentiality of data related to the development of Nutrient Management Plans. The most uncertain source category is wildlife. In an urban environment, this can result from the increased deer populations near streams to rat populations in storm sewers. In rural areas, estimation of wildlife populations and habitat locations in a watershed is required.

MDE recognizes the inherent uncertainty in developing traditional water quality models for the calculation of bacteria TMDLs. In this TMDL, MDE applies an analytical method which, when combined with BST, provides reasonable results (Cleland, 2003) and allows impaired streams to be addressed expeditiously.

4.2 Analysis Framework

This TMDL analysis uses flow duration curves to identify flow intervals that are used as indicator hydrological conditions (*i.e.*, annual average, critical conditions). As explained previously, this analytical method combined with water quality monitoring data and BST provides a better description of water quality and meets TMDL requirements.

Figure 4.2.1 illustrates how the hydrological (flow duration curve), water quality and BST data are linked together for the TMDL development.

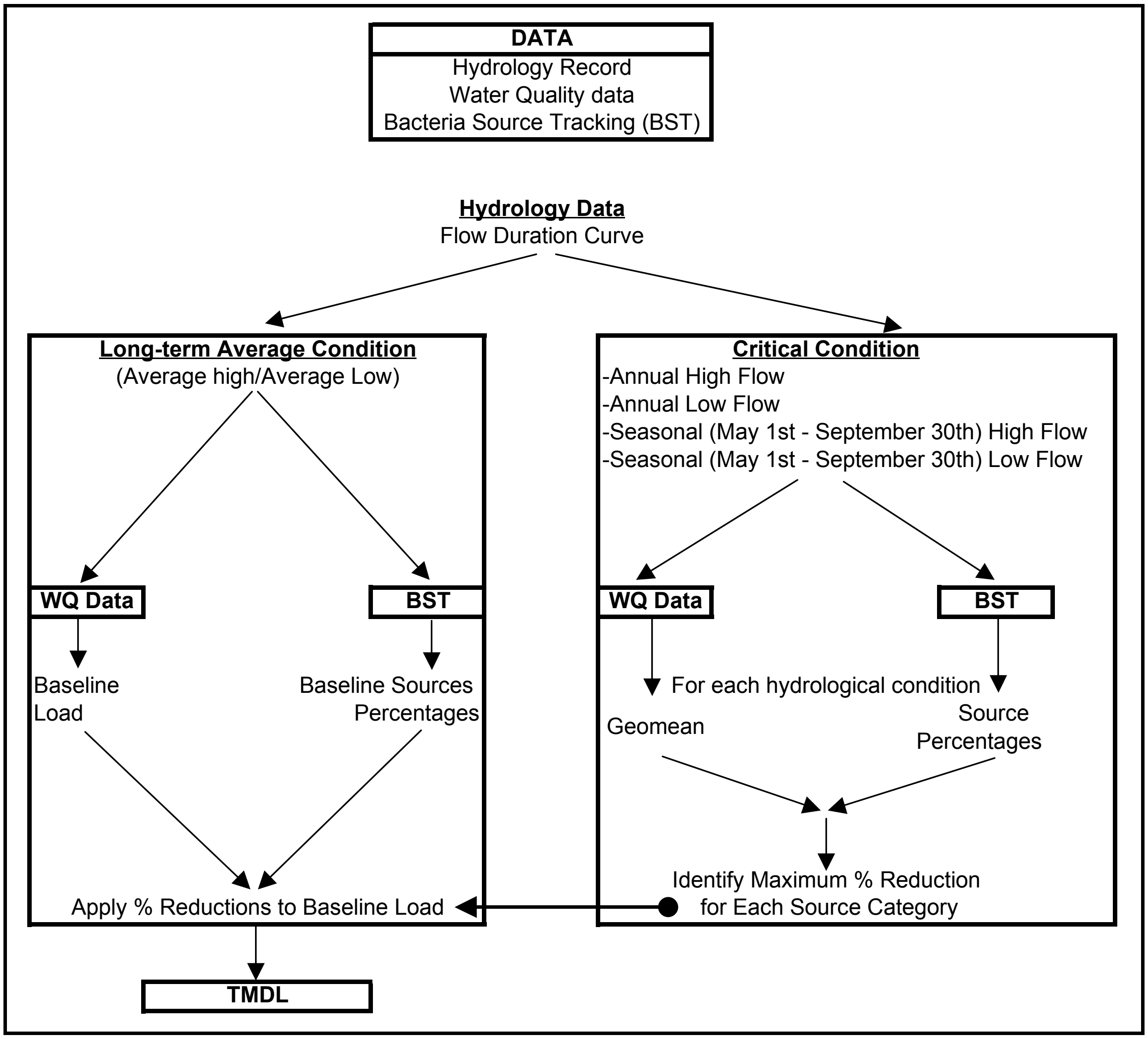

Figure 4.2.1: Diagram of Non-tidal Bacteria TMDL Analysis Framework

4.3 Estimating Baseline Loads

Loads estimated in this TMDL are based on the geometric mean concentration and a daily average flow.

The geometric mean concentration is calculated from the log transformation of the raw data. When back transformed values are used to calculate average daily loads or total annual loads, the loads will be biased low (Richards, 1998). To avoid this bias, a factor should be added to the log-concentration before it is back transformed. There are several methods of determining this bias correction factor ranging from parametric estimates resulting from the theory of the log-normal distribution to non-parametric estimates using a smearing factor. [Ferguson, 1986; Cohn *et al*., 1989; Duan, 1983]. There is much literature on the applicability and results from these various methods with a summary provided in Richards (1998). Each has advantages and conditions of applicability. A non-parametric estimate of the bias correction factor (Duan, 1983) was used in this TMDL analysis.

Daily average flows are estimated for each flow stratum using the watershed area ratio approach, since nearby long-term flow monitoring data are available.

The loads for each stratum are estimated as follows:

$$L_i = Q_i * C_i * F_1 * F_2 \tag{6}$$

where

L_i = Daily average load (MPN/day) at each station for stratum i
Q_i = Daily average flow (cfs) for stratum i
C_i = long term annual geometric mean for stratum i
F_1= Unit conversion factor from cfs*MPN/100ml to MPN/day (2.4466×10^7)
F_2= Bias correction factor.

To total baseline load is estimated as follows:

$$L_t = \sum_{i=1}^{2} L_i * W_i \tag{7}$$

L_t = Daily average load at station (MPN/day)
W_i= Proportion or weighting factor of stratum i

A weighting factor of 0.3 for high flow and .0.7 for low flow was used to estimate the average annual baseline load.

Results for the Cabin John Creek watershed are found in Table 4.3.1:

Table 4.3.1: Baseline Load Calculations

Station	Area (sq. miles)	USGS Reference Gage	High Flow			Low Flow			Baseline Load (billion MPN/day)
			Unit flow (cfs/sq, mile)	Q (cfs)	*E. coli* Concentration (MPN/100ml)	Unit flow (cfs/sq. mile)	Q (cfs)	*E. coli* Concentration (MPN/100ml)	
CJB0005	25.7	1650500 (est)	3.08	79.1	305.6	0.42	10.8	179.0	369.9

4.4 Critical Condition and Seasonality

Federal regulations (40 CFR 130.7(c)(1)) require TMDLs to take into account critical conditions for stream flow, loading, and water quality parameters. The intent of this requirement is to ensure that the water quality of the waterbody is protected during times when it is most vulnerable.

For this TMDL the critical condition is determined by assessing various hydrological conditions (high flow/low flow) and seasonality is captured by assessing the time period when water contact recreation is expected (May 1st – September 30th). The average hydrological condition over a 15-year period is approximately 30% high flow and 70% low flow as defined in Appendix B. Using the definition of a high flow condition occurring when the daily flow duration interval is less than 30% and a low flow condition occurring when the daily flow duration interval is greater than 70%, critical hydrological condition can be estimated by the percent of high or low flows during a specific period.

The following five conditions were used to account for the critical condition and include the effects of seasonality.

Table 4.4.1: Hydrological Conditions Used to Account for Critical Condition and Seasonality

Time Period	Hydrological Condition	Water Quality Data Used	Fraction High Flow	Fraction Low Flow	Condition Period
Annual (365 days)	Average	All	0.3	0.7	May 1995 - Apr 1996
	High flow	All	0.55	0.45	May 1996 - Apr 1997
	Low flow	All	0.07	0.93	Nov 2001 - Oct 2002
Seasonal (May 1st- Sept 30th)	High flow	May 1st – Sept 30th	0.51	0.49	May 2003 - Sep 2003
	Low flow	May 1st – Sept 30th	0.09	0.91	May 2002 - Sep 2002

The critical condition is determined by the maximum reduction per source that satisfy all five conditions and is required to meet the water quality standard while minimizing the risk to water contact recreation. It is assumed that the reduction that can be implemented to a bacteria source category will be constant through all conditions (*e.g.*, pets waste can be reduced to 75%).

The monitoring data for all stations located in the Cabin John Creek watershed cover a sufficient temporal span (at least one year), to estimate annual and seasonal conditions loads; therefore, seasonality is also included. The required reductions for each condition and maximum reductions for each source category overall is displayed in Table 4.4.2.

Table 4.4.2: Required Reductions to Meet Water Quality Standards

Time Period	Hydrological Condition	Domestic %	Human %	Livestock %	Wildlife %
Annual	Average	17%	95%	9%	0%
	High flow	40%	95%	20%	0%
	Low flow	0%	86%	0%	0%
Seasonal	High flow	35%	95%	17%	0%
	Low flow	16%	95%	8%	0%
Maximum Source Reduction		**40%**	**95%**	**20%**	**0%**

4.5 Margin of Safety

A Margin of Safety (MOS) is required as part of this TMDL in recognition of the many uncertainties in the understanding and simulation of bacteriological water quality in natural systems and in statistical estimates of indicators. As mentioned in Section 4.2, it is difficult to estimate stream loadings for fecal bacteria due to the variation in loadings across sample locations and time. Load estimation methods should be both precise and accurate to obtain the true estimate of the mean load. Refined precision in the load estimation is due to using a stratified approach along the flow duration intervals thus reducing the variation in the estimates. Moreover, Richards (1998) reports that averaging methods are generally biased, and the bias increases as the size of the averaging window increases. Finally, accuracy in the load estimation is based on minimal bias in the final result when compared to the true value.

Based on EPA guidance, the MOS can be achieved through two approaches (EPA, April 1991). One approach is to reserve a portion of the loading capacity as a separate term in the TMDL (i.e., TMDL = LA + WLA + MOS). The second approach is to incorporate the MOS as conservative assumptions used in the TMDL analysis. For this TMDL, the second approach was used by estimating the loading capacity of the stream based on a reduced-more stringent water quality criterion concentration. The *E. coli* water quality criterion concentration was reduced by 5%, from 126 *E. coli* MPN/100ml to 119.7 *E. coli* MPN/100ml.

4.6 TMDL Loading Caps

The TMDL loading cap is an estimate of the assimilative capacity of the monitored watershed and is provided in MPN/day. This loading is for the watershed upstream of monitoring station CJB0005.

The TMDL is based on a long-term average hydrological condition. Estimation of the TMDL requires knowledge of how the bacteria concentrations vary with flow rate or the flow duration interval. This concentration versus flow relationship is accounted for by using the strata defined on the flow duration curve.

The TMDL loading cap is estimated by first determining the baseline or current condition load and the associated geometric mean from the available monitoring data. The baseline load is estimated using the geometric mean concentration and average daily flow for each flow stratum. The loads from these two strata are then weighted to represent average conditions (see Table 4.3.1), based on the proportion of each stratum, to estimate the total long-term loading rate.

Next, the percent reduction, based on the critical condition, required to meet the water quality criterion is estimated from the observed bacteria concentrations. It is assumed that a reduction in bacteria concentrations is proportional to a reduction in load and thus the TMDL is equal to the current baseline load multiplied by one minus the required reduction.

$$TMDL = L_b * (1 - R) \qquad (1)$$

where

L_b = Current or baseline load estimated from monitoring data
R = Reduction required from baseline to meet water quality criterion

The bacteria TMDL for the watershed upstream of monitoring station CJB0005 is:

Table 4.6.1: Cabin John Creek Watershed TMDL Summary

Station	Baseline Load *E. coli* (billion MPN/day)	TMDL Load *E. coli* (billion MPN/day)	% Target Reduction
CJB0005	369.9	176.4	52.3%

4.7 Scenario Descriptions

Source Distribution

The final annual source distribution is derived from the source proportions listed in Table 2.4.1. For the purposes of the TMDL analysis and allocations, the percentage of sources identified as "unknown" was removed and the known sources were then scaled up proportionally so that they totaled 100%. The annual source distribution is presented in Table 4.7.1.

Table 4.7.1: Source Distributions for Average Conditions

Station	Domestic %	Human %	Livestock %	Wildlife %
CJB0005	32.5%	37.7%	17.8%	11.9%

Practicable Reduction Targets

The maximum practicable reduction (MPR) per each of the four source categories is listed in Table 4.7.2. These values are based on best professional judgment and a review of the available literature. It is assumed that human sources would potentially have the highest risk of causing gastrointestinal illness and therefore should have the highest reduction. The domestic animal category includes sources from pets (*e.g.*, dogs), and the MPR is based on an estimated success of education and outreach programs.

Table 4.7.2: Maximum Practicable Reduction Targets

	Human	Domestic	Livestock	Wildlife
Max Practical Reduction per Source	95%	75%	75%	0%
Rationale	(a) Direct source inputs (b) Human pathogens more prevalent in humans than animals. (c) Enteric viral diseases spread from human to human[1]	Target goal reflects uncertainty in effectiveness of urban BMP's[2] and is also based on best professional judgment	Target goal based on sediment reductions from BMP's[3] and best professional judgment	No programmatic approaches for wildlife reduction to meet water quality standards Waters contaminated by wild animal waste offers a public health risk that is orders of magnitude less than that associated with human waste.[4]

[1]USEPA. 1984. Health Effects Criteria for Fresh Recreational Waters. EPA-600/1-84-004. U.S. Environmental Protection Agency, Washington, D.C.
[2]USEPA. 1999. Preliminary Data Summary of Urban Storm Water Best Management Practices. EPA-821-R-99-012. U.S. Environmental Protection Agency, Washington, DC.
[3]USEPA. 2004. Agricultural BMP Descriptions as Defined for The Chesapeake Bay Program Watershed Model. Nutrient Subcommittee Agricultural Nutrient Reduction Workshop.
[4]Environmental Indicators and Shellfish Safety. 1994. Edited by Cameron, R., Mackney and Merle D. Pierson, Chapman & Hall.

As previously stated, these practicable reduction targets are based on the available literature and best professional judgment. There is much uncertainty with estimated reductions from best management practices (BMPs). The BMP efficiency for bacteria reduction ranged from –6% to +99% based on a total of 10 observations (EPA, 1999). The MPR to agricultural lands was based on sediment reductions identified by the EPA (EPA, 2004).

The practicable reduction scenario was developed based on an optimization analysis whereby a subjective estimate of risk was minimized and constraints were set on maximum reduction and allowable background conditions. Risk was defined on a scale of one to five, where it was assumed that human sources had the highest risk (5), domestic animals and livestock next (3), and wildlife the lowest (1) (See Table 4.7.2). The objective is to minimize the sum of the risk for all four conditions (high and low flow annual and seasonal conditions) while meeting the maximum practicable reduction constraints. The model was defined as follows:

$$\text{Min} \sum_{i=1}^{5} (P_h*5 + P_d*3 + P_l*3 + P_w*1)_i \qquad i = \text{hydrological condition}$$

Subject to

$C = C_{cr}$
$0 <= R_h <= 95\%$
$0 <= R_l <= 75\%$
$0 <= R_d <= 75\%$
$R_w = 0$
$P_h, P_l, P_d, P_w >= 1\%$

Where

P_h = % human source in final allocation
P_d = % domestic animal source in final allocation
P_l = % livestock source in final allocation
P_w = % wildlife source in final allocation
C = Instream concentration
C_{cr} = Water quality criterion
R_h = Reduction applied to human sources
R_l = Reduction applied to livestock sources
R_d = Reduction applied to domestic animal sources

The constraints of this scenario are satisfied for Cabin John Creek, indicating there is a feasible solution. A summary of the practicable reductions is presented in the Table 4.7.3.

Table 4.7.3: Practicable Reduction Results

Station	Applied Reductions				Achievable
	Domestic %	Human %	Livestock %	Wildlife %	
CJB0005	39.8%	95.0%	19.9%	0.0%	Yes

4.8 TMDL Allocation

The TMDL allocation includes waste load allocations (WLA) for point sources, for stormwater (where MS4 permits are required), and the LA for nonpoint sources. The margin of safety is explicit and is expressed as a 5% reduction of the *E. coli* water quality criterion concentration, from 126 MPN/100ml to 119.7 MPN/100ml. TMDL allocations are based on Maryland's bacteria water quality criteria and represent loads based on annual conditions. The load reduction scenario results in a load allocation by which the TMDL can be implemented to achieve water quality standards. The State reserves the right to revise these allocations provided such allocations are consistent with the achievement of water quality standards.

The bacteria sources are grouped into four categories that are also consistent with divisions for various management strategies. The categories are human, domestic animal, livestock and wildlife. TMDL allocation rules are presented in Table 4.7.1. This table identifies how the TMDL will be allocated among MS4 permits and the LA.

Table 4.8.1: Potential Source Contributions for TMDL Allocations

Allocation Category	Human	Domestic	Livestock	Wildlife
WLA				
MS4		X		X
LA	X	X	X	X

With no wastewater treatment plants (WWTPs) present in the Cabin John Creek basin, the human allocation is assigned entirely to LA. Where the entire watershed is covered by a MS4 permit(s), the domestic pet allocation is assigned to the MS4 WLA. Livestock is not covered by MS4 permits and will therefore, be part of the LA when it is not included as part of a CAFO. Wildlife is split between MS4 and LA. This wildlife ratio is estimated based on the amount of urban pervious land (*e.g.*, residential) compared to other pervious land (*e.g.*, pasture, forest). Note that only the final LA or WLA is reported in this TMDL.

Municipal Separate Storm Sewer System (MS4) Allocations

Both individual and general NPDES MS4 Phase I and Phase II permits are point sources subject to WLA assignment in the TMDL. Quantification of rainfall-driven nonpoint source loads is uncertain. EPA recognized this in its guidance document entitled "Establishing Total Maximum Daily Load (TMDL) Wasteload Allocations (WLAs) for Storm Water Sources and NPDES Permit Requirements Based on Those WLAs" (November 2002), which states that available data and information usually are not detailed enough to determine WLAs for NPDES-regulated stormwater discharges on an outfall-specific basis. Therefore, in watersheds with an existing MS4 permit, domestic animal bacteria loads will be lumped into a single WLA-MS4 load. In watersheds with no existing individual MS4 permits, these loads will be included in the LA.

The county within the Cabin John Creek watershed, Montgomery County, is covered by individual Phase I MS4 program regulations. Based on EPA's guidance, the MS4 WLA is presented as one combined load for the entire land area of the county. In the future, when more detailed data and information become available, it is anticipated that MDE will revise the WLA into appropriate WLAs and LAs, and may also revise the LAs accordingly. Note that the overall reductions in the TMDL will not change.

4.9 Summary

The TMDL for the Cabin John Creek watershed is presented below.

Table 4.9.1: Cabin John Creek Watershed TMDL

Station	TMDL Load (billion MPN/day)	LA Load (billion MPN/day)	WLA-PS Load (billion MPN/day)	WLA – MS4 Load (billion MPN/day)
CJB0005	176.36	68.17	0.00	108.19

In Cabin John Creek, based on the practicable reduction rates specified, water quality standards could be achieved.

5.0 ASSURANCE OF IMPLEMENTATION

This section provides the basis for reasonable assurances that the fecal bacteria TMDL will be achieved and maintained. The appropriate measures to reduce pollution levels in the impaired segments include, where appropriate, the use of better treatment technology or installation of BMPs. Details of these methods are to be described in the implementation plan.

The final scenario is based on reductions that meet the MPR targets. These MPR targets were defined based on a literature review of BMPs effectiveness and assuming a zero reduction for wildlife sources. The uncertainty of BMPs effectiveness for bacteria, reported within this literature, is quite large. As an example, pet waste education programs have varying results based on stakeholders' involvement. Additionally, the extent of wildlife reduction associated with various BMPs methods (*e.g.,* structural, non-structural, etc) is uncertain. Therefore, MDE intends for the required reductions to be implemented in an iterative process that first addresses those sources with the largest impact on water quality, with consideration given to ease of implementation and cost. The iterative implementation of BMPs in the watershed has several benefits: tracking of water quality improvements following BMP implementation through follow-up stream monitoring; providing a mechanism for developing public support through periodic updates on BMP implementation; and helping to ensure that the most cost-effective practices are implemented first.

Potential funding sources for implementation include the Maryland's Agricultural Cost Share Program (MACS) which provides grants to farmers to help protect natural resources and the Environmental Quality and Incentives Program which focuses on implementing conservation practices and BMPs on land involved with livestock and production. Funding sources available for local governments include the State Water Quality Revolving Loan Fund and the Stormwater Pollution Cost Share Program. Details of these programs and additional funding sources can be found at http://www.dnr.state.md.us/bay/services/summaries.html.

In 1983, the EPA Nationwide Urban Runoff Program, found that stormwater runoff from urban areas contains the same general types of pollutants found in wastewater, and that 30% of identified cases of water quality impairment were attributable to stormwater discharges. In November 1990, EPA required jurisdictions with a population greater than 100,000 to apply for NPDES Permits for stormwater discharges. The jurisdiction where the Cabin John Creek watershed is located (Montgomery County) is required to participate in the stormwater NPDES program, and has to comply with the NPDES Permit regulations for stormwater discharges. The permit-required management programs are being implemented in the County to meet locally established watershed protection and restoration goals and to control stormwater discharges to the maximum extent practicable. Funding sources for implementation include the State Water Quality Revolving Loan Fund and the Stormwater Pollution Cost Share Program. Details of this program and additional funding sources can be found at http://www.dnr.state.md.us/bay/services/summaries.html.

Additionally, MDE's "Managing Maryland for Results" document (MDE, 2005) states the following related to sewage overflows:

> Objective 4.5: Reduce the quantity in gallons of sewage overflows [total for Combined Sewer System Overflows (CSO) and Separate Sewer System Overflows (SSO)] equivalent to a 50% reduction of 2001 amounts (50, 821,102 gallons) by the year 2010 through implementation of EPA's minimum control strategies, long term control plans (LTCP), and collection system improvements in capacity, inflow and infiltration reduction, operation and maintenance.
>
> Strategy 4.5.1: MDE will implement regulations adopted in FY 2004 to ensure that all jurisdictions are reporting all sewage overflows to the Department, notifying the public about significant overflows, and are taking appropriate steps to address the cause(s) of the overflows.
>
> Strategy 4.5.2: MDE will inspect and take enforcement actions against those CSO jurisdictions that have not developed long-term control plans with schedules for completion and require that enforceable schedules are incorporated in consent decrees or judicial orders.
>
> Strategy 4.5.3: MDE will take enforcement actions to require that jurisdictions experiencing significant or repeated SSOs take appropriate steps to eliminate overflows, and will fulfill the commitment in the EPA 106 grant for NPDES enforcement regarding the initiation of formal enforcement actions against 20% of jurisdictions in Maryland with CSOs and significant SSO problems annually.

In 2004, the United States and the State of Maryland brought suit against Washington Suburban Sanitary Commission (WSSC) in the U.S. District Court for the District of Maryland to remedy recurrent SSOs from the WSSC system, *United States et al. v. Washington Suburban Sanitary Commission,* C.A. No. PJM 04-3679 (Greenbelt Division). A consent decree was negotiated among the United States, Maryland, several intervenor citizen groups and WSSC and lodged on July 26, 2005. It is now before the court for approval. WSSC already reports overflows to MDE as required by Environment Article, Section 9-331.1, Annotated Code of Maryland and COMAR 26.08.10.

REFERENCES

American Society of Agricultural Engineers (ASAE) (1998). *ASAE Standards, 45th edition:Standards, Engineering Practices, Data*. St. Joseph, MI.

Cleland, Bruce. 2003. TMDL Development from the "Bottom Up" – Part III: Duration Curves and Wet-weather Assessments. America's Clean Water Foundation. Washington D.C.

Code of Federal Regulations (CFR) 40 CFR 130.2(i). TMDL and individual water quality-based effluent limitations – Website http://www.access.gpo.gov/nara/cfr/waisidx_98/40cfr130_98.html

Code of Federal Regulations (CFR) 40 CFR 130.7(c)(1). Definitions - Website http://www.access.gpo.gov/nara/cfr/waisidx_98/40cfr130_98.html 40 CFR 130.7(c)(1)

Code of Maryland Regulations (COMAR), 26.08.02.03-3A(1). Bacteriological Criteria for Use I, III, IV-P Waters - Website http://www.dsd.state.md.us/comar/26/26.08.02.03-3.htm.

Code of Maryland Regulations (COMAR), 26.08.02.08N. Stream Segment Designations – Website http://www.dsd.state.md.us/comar/26/26.08.02.08.htm

Code of Maryland Regulations (COMAR), 26.08.10. Overflows or Bypasses - Website http://www.dsd.state.md.us/comar/26/26.08.02.08.htm

Dillow, J.J.A., (1996). Technique for Estimating Magnitude and Frequency of Peak Flow in Maryland: U.S. Geological Survey Water –Resources Investigations Report 97-4279.

Friends of Cabin John Creek Watershed http://www.cabinjohn.org, sponsored by the Cabin Johns Citizen Association, last visited April 25, 2005.

Natural Resources Conservation Service. Montgomery County SSURGO Coverage.

Geldrenich, E. and E. A. Kenner. (1969). Concepts of Fecal Streptococci in Stream Pollution. Journal of Water Pollution Federation. 41:R336-R352.

Hirsch, R.M. 1982. "A Comparison of Four Record Extension Techniques". Water Resources Research, V. 18, No. 4, p. 1082-1088.

Maryland Department of the Environment, *2002 List of Impaired Surface Waters [303(d) List] and Integrated Assessment of Water Quality in Maryland.*

Maryland Department of the Environment, *2004 FINAL List of Impaired Surface Waters [303(d) List] and Integrated Assessment of Water Quality in Maryland.*

Maryland Department of Natural Resources, GIS shapefile of Stream Corridor Assessment Survey For Sligo Creek, January 2005.

Maryland Department of Planning, 2000, 2000 Land Use, Land Cover Map Series.

Maryland Department of Planning. Estimates of Septic Systems (2003). Baltimore: Maryland Department of Planning, Comprehensive Planning Unit.

Moglen, G.E., Thomas, W.O. and Miller, A.C. (2002). Evaluation of Alternative Statistical Method for Estimating Frequency of Peak Flows in Maryland: Maryland Department of Transportation State Highway Administration Final Report SP907CRB.

Pettyjohn, W.A., and Henning R. 1979. Preliminary Estimate of Ground-Water Recharge Rates, Related Streamflow and Water Quality in Ohio: Ohio State University Resources Center Project Completion Report Number 552. 323p.

Reiss, K.G. and Friesz, P.J., (2000). Methods for Estimating Low-Flow Statistics for Massachusetts Streams: U.S. Geological Survey Water –Resources Investigations Report 00-4135.

Richards, R.P. 1998. Estimation of pollutant loads in rivers and streams: A guidance document for NPS programs. Project report prepared under Grant X998397-01-0, U.S. Environmental Protection Agency, Region VIII, Denver. 108 p.

Schueler, T. 1999. "Microbes and Urban Watersheds". Watershed Protection Techniques. 3(1): 551-596.

Soil Conservation Service (SCS). *Soil Survey of Montgomery County, MD*, 1995.

Swann, C, 1999. A survey of Residential Nutrient Behaviors in the Chesapeake Bay. Widener Burrows, Inc. Chesapeake Research Consortium. Center For Watershed Protection. Ellicott City, MD. 112 pp.

US Department of Commerce. United States Census Bureau's GIS Coverage (2000). Washington DC: US Bureau of the Census.

US Department of Agriculture. (1995). State Soil Geographic (STATSGO) DataBase.

US Department of Agriculture. (1997). Census of Agriculture: Maryland State and County Data. Washington, DC: National Agricultural Statistic Service.

U.S. Environmental Protection Agency. *Ambient Water Quality Criteria for Bacteria-1986*. EPA-440/5-84-002, 1986.

U.S. Environmental Protection Agency. *National Recommended Water Quality Criteria: 2002*. EPA-822-R-02-047. November 2002.

U.S. Environmental Protection Agency. *Implementation Guidance for Ambient Water Quality Criteria for Bacteria: Draft*, U.S. Environmental Protection Agency. Office of Water, Washington, D.C. EPA-823-B-02-003. 2003.

USEPA. 1985. Test Methods for Eschericichia coli and Enterococci in Water by the Membrane Filter Procedure. EPA600/4-85-076. U.S. Environmental Protection Agency, Washington, DC. NTIS PB86-158052.

USGS. 1996. HYSEP: A Computer Program for Streamflow Hydrograph Separation and Analysis. USGS Water-Resource Investigations Report 96-4040.

US EPA, Office of Water (2000). Bacteria Indicator Tool User's Guide. EPA-823-B-01-003.

US EPA (2001). Protocol for developing Pathogen TMDLs, EPA 841-R-00-002, Office of Water (4503F), United States Environmental Protection Agency, Washington, DC. 134pp.

US EPA Stormwater Manager Resource Center. Website: http://www.stormwatercenter.net/.

US EPA Chesapeake Bay Program (1996). Chesapeake Bay Program:Watershed Model Application to Calculate Bay Nutrient Loadings: Final Findings and Recommendations, and Appendices, Annapolis, MD.

VA DEQ (2002) Fecal Coliform TMDL for Dodd Creek Watershed, Virginia, June 2002.

Appendix A – Bacteria Monitoring Data

Table A-1: *E. coli* data for Cabin John Creek Monitoring Station CJB0005

SAMPLING STATION IDENTIFIER	Date	Daily flow frequency	*E. coli* MPN/100 ml
CJB0005	10/07/2002	97.8372	140
CJB0005	10/21/2002	87.2410	190
CJB0005	11/06/2002	7.9789	1780
CJB0005	11/18/2002	6.0705	1720
CJB0005	12/02/2002	70.0109	120
CJB0005	12/16/2002	28.6441	240
CJB0005	01/06/2003	20.7743	130
CJB0005	01/21/2003	50.5998	130
CJB0005	02/03/2003	47.0738	200
CJB0005	03/03/2003	5.1981	290
CJB0005	03/17/2003	24.1912	160
CJB0005	04/21/2003	26.9357	90
CJB0005	05/05/2003	32.9335	70
CJB0005	05/19/2003	17.9026	350
CJB0005	06/02/2003	25.4998	199
CJB0005	06/16/2003	25.4998	260
CJB0005	06/23/2003	18.7023	230
CJB0005	07/07/2003	11.2323	1020
CJB0005	07/21/2003	50.5998	170
CJB0005	08/04/2003	39.7492	2850
CJB0005	08/18/2003	43.8023	1020
CJB0005	08/25/2003	70.1927	30
CJB0005	09/08/2003	66.1032	170
CJB0005	09/22/2003	22.7190	120
CJB0005	10/06/2003	55.1618	120
CJB0005	10/20/2003	51.9447	110

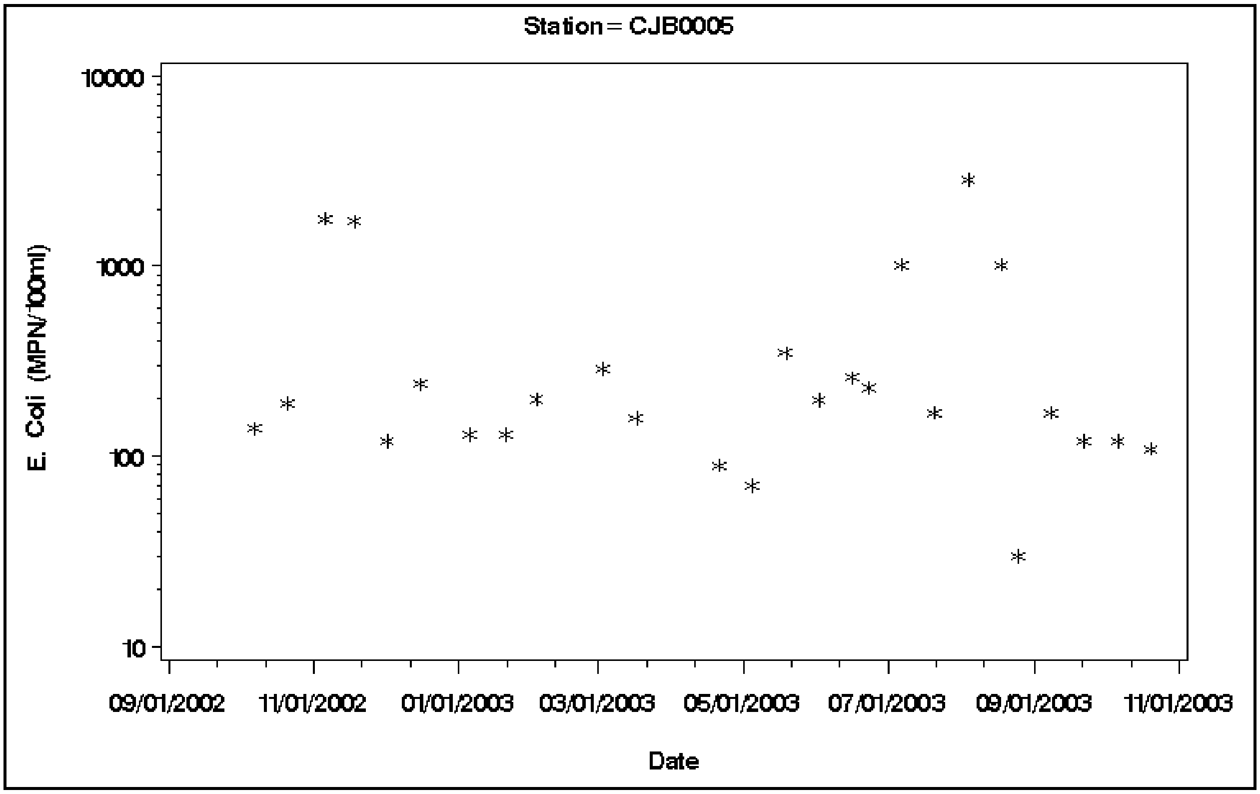

Figure A-1: *E. coli* Concentration vs. Time for Cabin John Creek Monitoring Station CJB0005

Appendix B - Flow Duration Curve Analysis to Define Strata

The Cabin John Creek watershed was assessed to determine hydrologically significant strata. The purpose of these strata is to apply weights to monitoring data and thus (1) reduce bias associated with the monitoring design and (2) approximate a critical condition for TMDL development. The strata group hydrologically similar water quality samples and provide a better estimate of the mean concentration at the monitoring station.

The flow duration curve for a watershed is a plot of all possible daily flows, ranked from highest to lowest, versus their probability of exceedance. In general, the higher flows will tend to be dominated by excess runoff from rain events and the lower flows will result from drought type conditions. The mid range flows are a combination of high base flow with limited runoff and lower base flow with excess runoff. The range of these mid level flows will vary with soil antecedent conditions. The purpose of the following analysis is to identify hydrologically significant groups, based on the previously described flow regimes, within the flow duration curve.

Flow Analysis

There are no United States Geological Survey (USGS) gage stations in the Cabin John Creek watershed. A USGS gage station (01650500) located in the Anacostia River watershed with a drainage basin of similar size, land use and soils was chosen to represent the surface flow in the Cabin John Creek watershed. Soil types were calculated using the Soil Survey Geographic (SSURGO) coverage for Montgomery County. A comparison of the land use and soil types is displayed in Table B-1 and Table B-2.

Table B-1: USGS/CORE Station Drainage Basin Land Use Comparison

Station	Commercial	Crops	Forest	Pasture	Residential	Water	Area (Acres)
1650500	4.9%	4.3%	23.5%	3.4%	63.7%	0.2%	13,575
CJB0005	13.3%	0.0%	13.0%	0.6%	73.0%	0.2%	16,424

Table B-2: USGS/CORE Station Drainage Basin Soil Type Comparison

Station	Soil Type				
	A	B	C	C/D	D
1650500	0.19%	78.37%	6.09%	1.02%	14.34%
CJB0005	0.22%	67.58%	4.16%	4.98%	23.06%

The gage flow data is incomplete for this station therefore the flow for unobserved periods (10/01/1988 to 11/26/1997) was estimated using USGS gage station (0165100). The gage and dates of information used are as follows:

Table B-3: USGS Gages used in the Cabin John Creek Watershed

USGS Gage #	Dates used	Description
01650500 (estimate)	Oct 1, 1988 to Sep 30, 2003	Estimated flow based on USGS Gage 0165100 using MOVE.1 (Hirsch, 1982)

The flow duration curve for the estimated gage is presented in figure B-1.

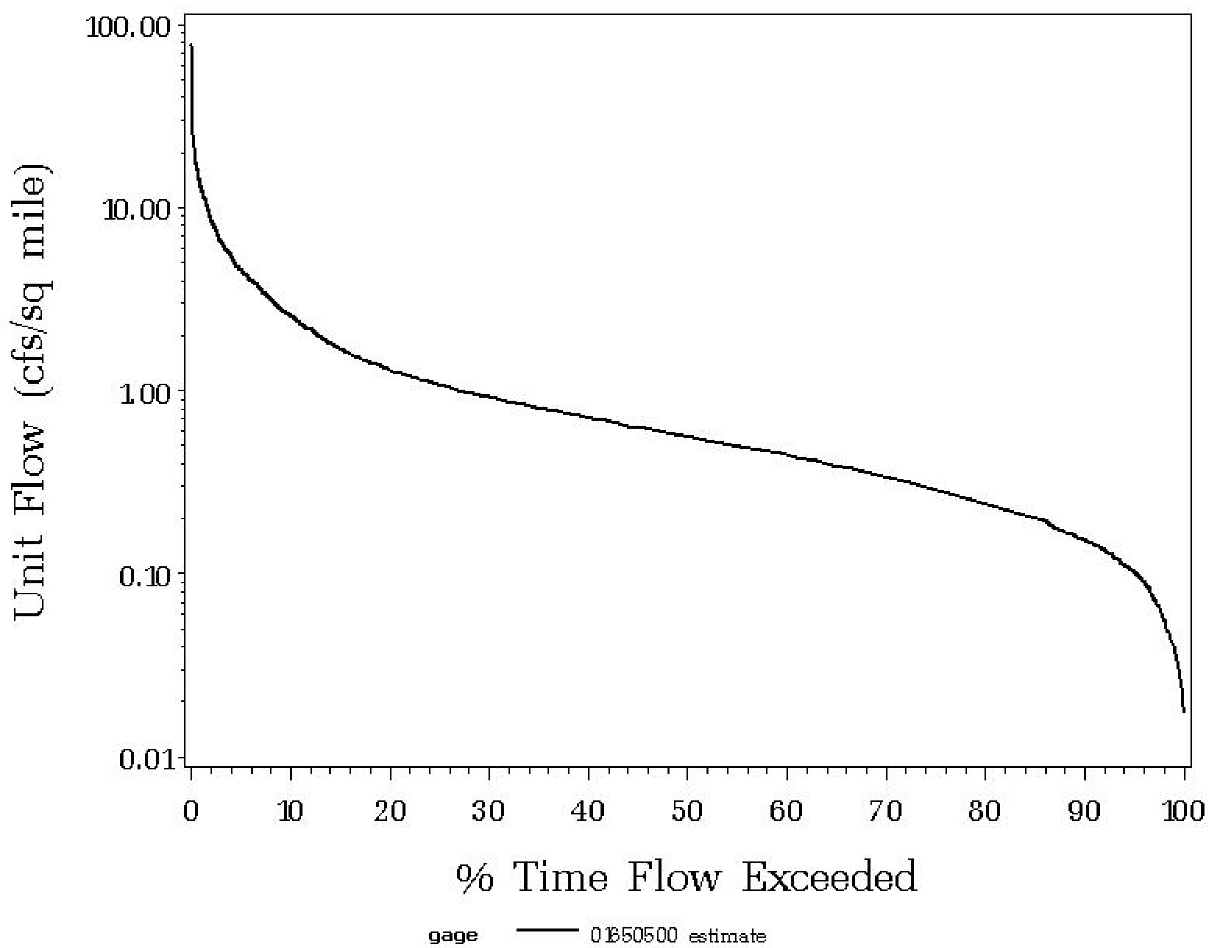

Figure B-1: Cabin John Creek Flow Duration Curves

The separation of high flow and low flow was based on the analysis of flow data for USGS gage 01650500 (estimated flow). The hydrograph separation technique is equivalent to the sliding interval technique use in the USGS HYSEP program (USGS, 1996) and the interval is based on the duration of surface runoff estimated from Linsley, *et. al.* (1982) and Pettyjohn and Henning

(1979). Following hydrograph separation, the percent of surface runoff vs. the daily flow duration interval is plotted and a non-parametric smoothing method (LOESS) was used to identify general patterns.

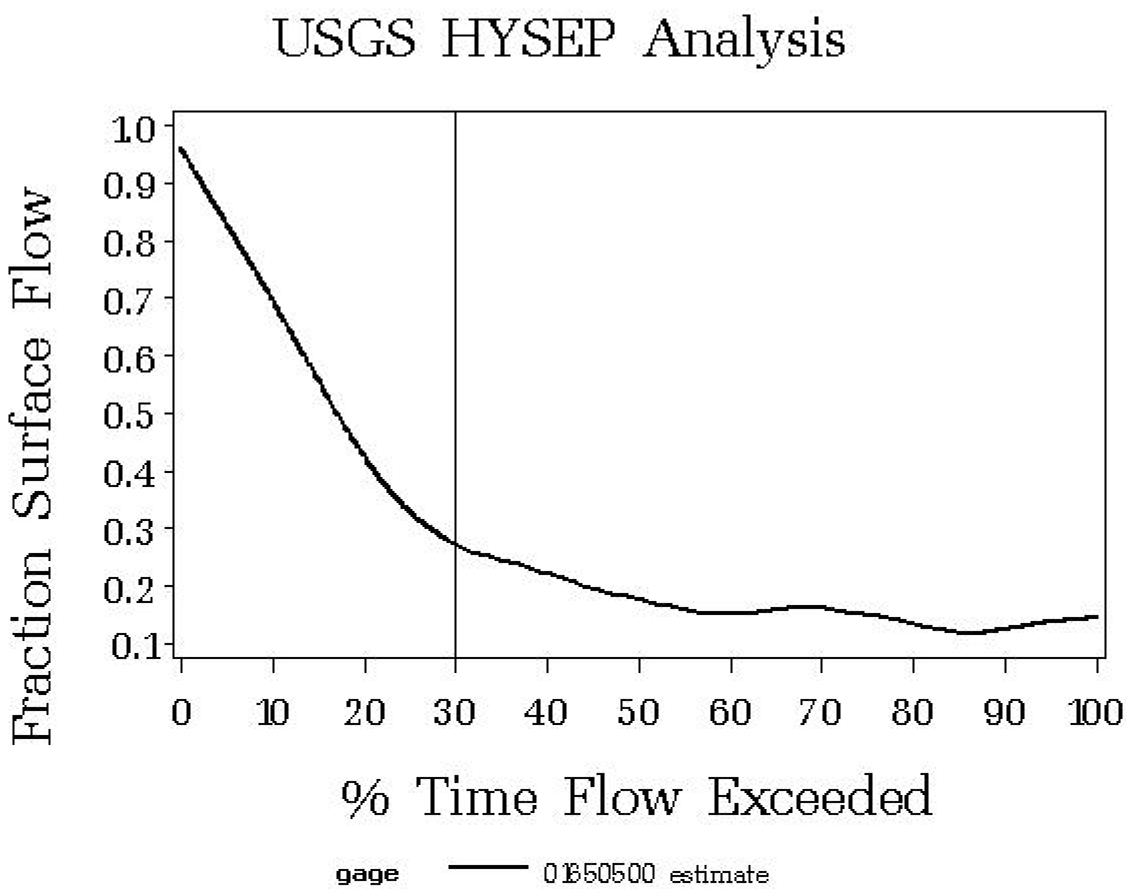

Figure B-2: Cabin John Creek - LOESS smoothing of hydrograph separation

These patterns are illustrated in Figure B-2. From this figure it can be seen that a significant change in slope occurs at approximately the 30 percent daily flow interval of Cabin John Creek.

It was observed that no significant change in slope or mean fraction of surface runoff occurs in Cabin John Creek *below* the 30 percent daily flow interval and that this area is representative of a region of significant and increasing surface flow contribution to the stream. Therefore, the 30 percentile threshold was used to define the limits between high flow and low flows as appropriate. Using these thresholds, definitions of high and low range flows are presented in Table B-4.

Table B-4: Definition of flow regimes

High flow	Represents conditions where stream flow tends to be dominated by surface runoff.
Low flow	Represents conditions where stream flow tends to be more dominated by groundwater flow.

Flow-Data Analysis

The final analysis to define the daily flow duration intervals (flow regions, strata) includes the bacteria monitoring data. Bacteria (*Enterococci* or *E. coli*) monitoring data are "placed within the regions (strata based on the daily flow duration percentile of the date of sampling)." Figure B-3 shows the Cabin John Creek *E. coli* monitoring data with corresponding flow frequency.

Maryland's water quality standards for bacteria state that a steady-state geometric mean will be calculated with available data where there are at least five representative sampling events. The data shall be from samples collected during steady-state conditions and during the beach season (Memorial Day through Labor Day) to be representative of the critical condition. If fewer than five representative sampling events are available, the previous two years will be evaluated. In Cabin John Creek, there are sufficient samples in the high flow strata to estimate the geometric means. For the low flow strata, less than five examples exist; therefore, the mid and low flow strata will be combined to calculate the geometric mean.

Weighting factors for estimating a weighted geometric mean are based on the frequency of each flow stratum during the averaging period. The weighting factors for the averaging period and hydrologic conditions are presented in Table B-5. Weighted geometric means are plotted with the monitoring data on Figure B-3 and Figure B-4 for the annual and seasonal periods.

Table B-5: Weighting factors for estimation of geometric mean

Time Period	Hydrologic Condition	Weighting Factor (High Flow)	Weighting Factor (Low Flow)
Annual (365 days)	Average	0.30	0.70
	High	0.55	0.45
	Low	0.07	0.93
Seasonal (May 1st – Sept 30th)	High	0.51	0.49
	Low	0.09	0.91

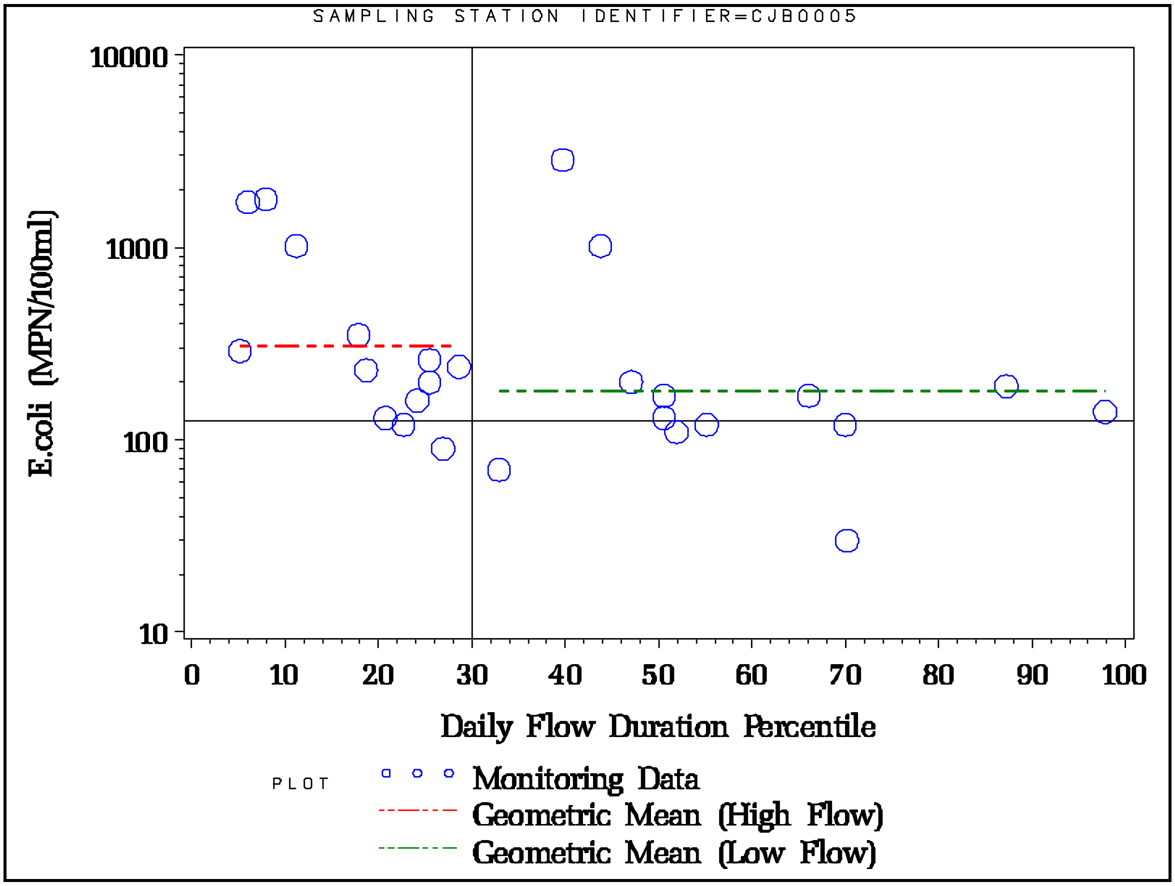

Figure B-3: *E. coli* Concentration vs. Flow Duration for Station CJB0005 (Annual Period)

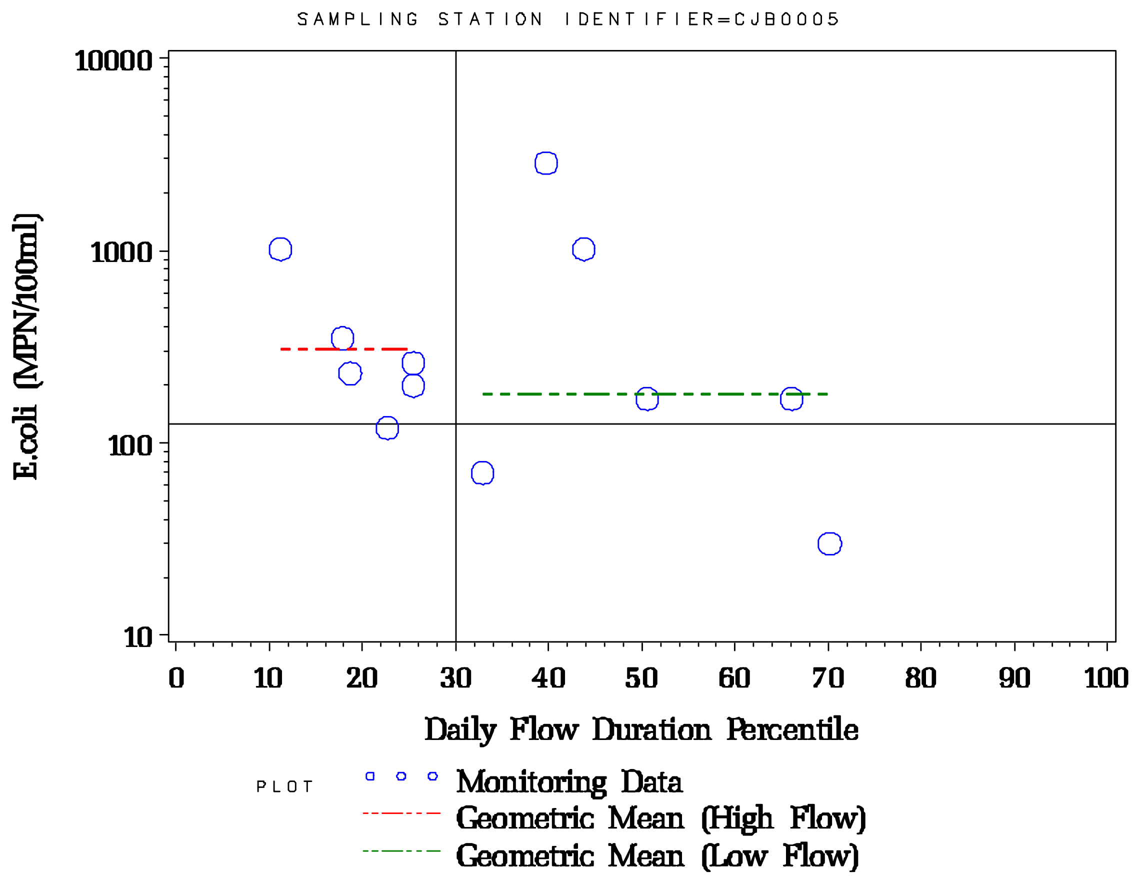

Figure B-4: *E. coli* Concentration vs. Flow Duration for Station CJB0005 (Seasonal Period)

Appendix C - Bacterial Source Tracking Report:

Identifying Sources of Fecal Pollution in the Cabin John Creek Watershed, Maryland

Mark F. Frana, Ph.D. and Elichia A. Venso, Ph.D.
Department of Biological Sciences and Environmental Health Science
Salisbury University, Salisbury, MD 21801

May 9, 2005

INTRODUCTION

Microbial Source Tracking. Microbial Source Tracking (MST) is a relatively recent scientific and technological innovation designed to distinguish the origins of enteric microorganisms found in environmental waters. Several different methods and a variety of different indicator organisms (both bacteria and viruses) have successfully been used for MST, as described in recent reviews (Scott *et al.*, 2002; Simpson *et al.*, 2002). When the indicator organism is bacteria, the term Bacterial Source Tracking (BST) is often used. Some common bacterial indicators for BST analysis include: *E. coli , Enterococcus* spp., *Bacteroides-Prevotella,* and *Bifidobacterium* spp.

Techniques for MST can be grouped into one of the following three categories: molecular (genotypic) methods, biochemical (phenotypic) methods, or chemical methods. Ribotyping, Pulsed-Field Gel Electrophoresis (PFGE), and Randomly-Amplified Polymorphic DNA (RAPD) are examples of molecular techniques. Biochemical methods include Antibiotic Resistance Analysis (ARA), F-specific coliphage typing, and Carbon Source Utilization (CSU) analysis. Chemical techniques detect chemical compounds associated with human activities, but do not provide any information regarding nonhuman sources. Examples of this type of technology include detection of optical brighteners from laundry detergents or caffeine (Simpson *et al*., 2002).

Many of the molecular and biochemical methods of MST are “library-based,” requiring the collection of a database of fingerprints or patterns obtained from indicator organisms isolated from known sources. Statistical analysis determines fingerprints/patterns of known-source species or categories of species (i.e., human, livestock, pets, wildlife). Indicator isolates collected from water samples are analyzed using the same MST method to obtain their fingerprints or patterns, which are then statistically compared to those in the library. Based upon this comparison, the final results are expressed in terms of the “statistical probability” that the water isolates came from a given source (Simpson *et al*. 2002).

In this BST study of the Cabin John Creek Watershed, we used the ARA method with *Enterococcus* spp. as the indicator organism. Previous BST publications have demonstrated the predictive value of using this particular technique and indicator organism (Hagedorn, 1999; Wiggins, 1999).

Antibiotic Resistance Analysis. A variety of different host species can potentially contribute to the fecal contamination found in natural waters. Many years ago, scientists speculated on the possibility of using resistance to antibiotics as a way of determining the sources of this fecal contamination (Bell *et al.*, 1983; Krumperman, 1983). In ARA, the premise is that bacteria isolated from different hosts can be discriminated based upon differences in the selective pressure of microbial populations found in the gastrointestinal tract of those hosts (humans, livestock, pets, wildlife) (Wiggins, 1996). Microorganisms isolated from the fecal material of wildlife would be expected to have a much lower level of resistance to antibiotics than isolates collected from the fecal material of humans, livestock and pets. In addition, depending upon the specific antibiotics used in the analysis, isolates from humans, livestock and pets could be differentiated from each other.

In ARA, isolates from known sources are tested for resistance or sensitivity against a panel of antibiotics and antibiotic concentrations. This information is then used to construct a library of antibiotic resistance patterns from known-source bacterial isolates. Microbial isolates collected from water samples are then tested and their resistance results are recorded. Based upon a comparison of resistance patterns of water and library isolates, a statistical analysis can predict the likely host source of the water isolates. (Hagedorn 1999; Wiggins 1999).

LABORATORY METHODS

Isolation of Enterococci from Known-Source Samples. Fecal samples, identified to source, were delivered to the Salisbury University (SU) BST lab by Maryland Department of the Environment (MDE) personnel. Fecal material suspended in phosphate buffered saline was plated onto selective m-Enterococcus agar. After incubation at 37^{o} C, up to 10 *Enterococci* isolates were randomly selected from each fecal sample for ARA testing.

Isolation of Enterococci from Water Samples. Water samples were collected by MDE staff and shipped overnight to MapTech Inc, Blacksburg, Va. Bacterial isolates were collected by membrane filtration. Up to 24 randomly selected *Enterococci* isolates were collected from each water sample and all isolates were then shipped to the SU BST lab.

Antibiotic Resistance Analysis. Each bacterial isolate from both water and scat were grown in Enterococcosel® broth (Becton Dickinson, Sparks, MD) prior to ARA testing. Enterococci are capable of hydrolyzing esculin, turning this broth black. Only esculin-positive isolates were tested for antibiotic resistance.

Bacterial isolates were plated onto tryptic soy agar plates, each containing a different concentration of a given antibiotic. Plates were incubated overnight at 37^{o} C and isolates then scored for growth (resistance) or no growth (sensitivity). Data consisting of a “1” for resistance or “0” for sensitivity for each isolate at each concentration of each antibiotic was then entered into a spread-sheet for statistical analysis.

The following includes the antibiotics and concentrations used for isolates in the Cabin John Creek Watershed analysis.

Table C-1: Antibiotics and concentrations used for ARA

Antibiotic	Concentration (µg ml^{-1})
Amoxicillin	0.625
Cephalothin	10, 15, 30, 50
Chloramphenicol	1, 2.5, 5, 10
Chlortetracycline	60, 80, 100
Erythromycin	10, 15, 30, 50
Gentamycin	5, 10, 15, 20
Neomycin	40, 60, 80
Oxytetracycline	20, 40, 60, 80, 100
Salinomycin	1, 2.5, 5, 10
Streptomycin	40, 60, 80, 100
Tetracycline	10, 15, 30, 50, 100
Vancomycin	2.5

KNOWN-SOURCE LIBRARY

Construction and Use. Fecal samples (scat) from known sources in the watershed were collected during the study period by MDE personnel and delivered to the BST Laboratory at SU. Enterococci isolates were obtained from known sources, which included human, dog, cow, goat, horse, pig, sheep, chicken, deer, rabbit, fox, and goose. A library of patterns of *Enterococcus* isolate responses to the panel of antibiotics was analyzed using the statistical software CART® (Salford Systems, San Diego, CA). The library consisted of response patterns of 2643 *Enterococcus* isolates from Cabin John Creek and the adjacent watersheds of Anacostia River and Piscataway River. The combination of watersheds from which isolates were obtained was chosen after examination of possible library combinations (Figure 1). The classification models in Figure 1 show the sharp increase in percent unknown isolates for Cabin John Creek and Piscataway River at a cutoff probability > 50%.

Enterococci isolate response patterns were also obtained from bacteria in water samples collected at the one (1) monitoring station, CBJ0005, in the Cabin John Creek basin. Using statistical techniques, these patterns were then compared to those in the combined library to identify the probable source of each water isolate.

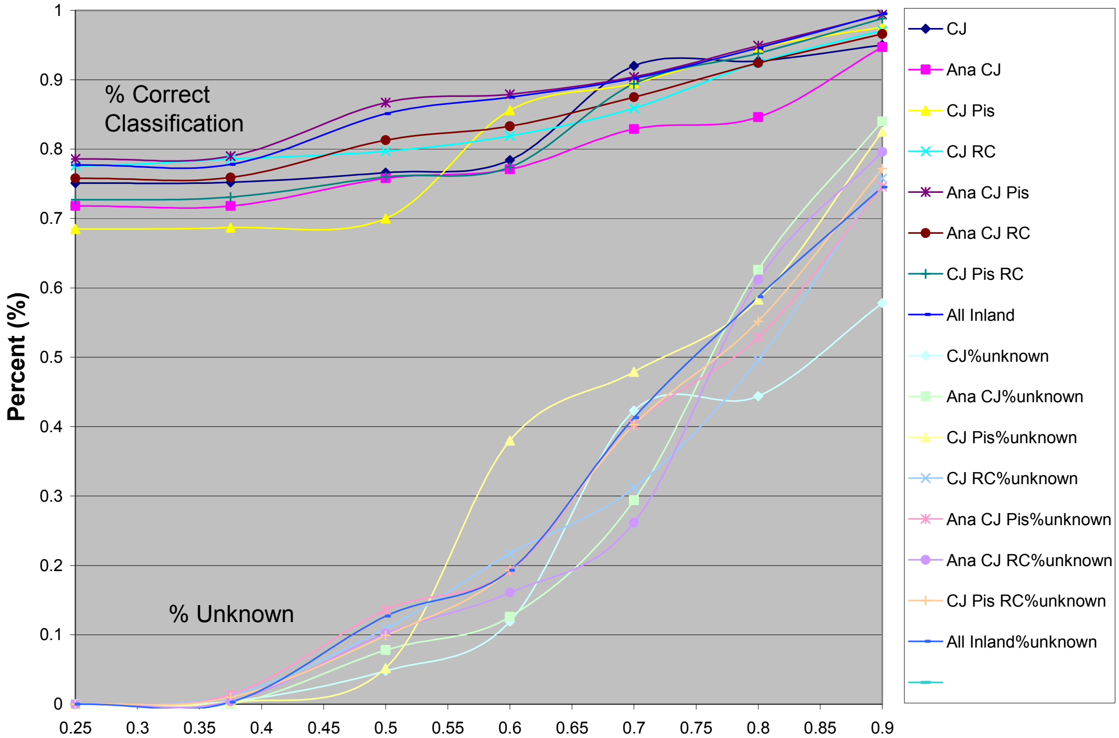

Figure C-1: Classification models for determination of composition of known-source library for identification of Cabin John Creek water isolates

STATISTICAL ANALYSIS

We applied a tree classification method, [1]CART®, to build a model that classifies isolates into source categories based on ARA data. CART® builds a classification tree by recursively splitting the library of isolates into two nodes. Each split is determined by the antibiotic variables (antibiotic resistance measured for a collection of antibiotics at varying concentrations). The first step in the tree-building process splits the library into two nodes by considering every binary split associated with every variable. The split is chosen that maximizes a specified index of homogeneity for isolate sources within each of the nodes. In subsequent steps, the same process is applied to each resulting node until a *stopping* criterion is satisfied. Nodes where an additional split would lead to only an insignificant increase in the *homogeneity index* relative to the *stopping* criterion are referred to as *terminal* nodes.[2] The collection of *terminal* nodes

---

[1] The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Hastie T, Tibshirani R, and Friedman J. Springer 2001.

[2] An ideal split, i.e., a split that achieves the theoretical maximum for homogeneity, would produce two nodes each containing library isolates from only one source.

defines the classification model. Each *terminal* node is associated with one source, the source that is most populous among the library isolates in the node. Each water sample isolate (*i.e*., an isolate with an unknown source), based on its antibiotic resistance pattern, is identified with one specific *terminal* node and is assigned the source of the majority of library isolates in that *terminal* node.[3]

We imposed an additional requirement in our classification method for determining the sources of water sample isolates. We interpreted the proportion of the majority source among the library isolates in a *terminal* node as a probability. This proportion is an estimate of the probability that an isolate with unknown source, but with the same antibiotic resistance pattern as the library isolates in the *terminal* node, came from the source of the majority of the library isolates in the *terminal* node. If that probability was less than a specified *acceptable source identification probability*, we did not assign a source to the water sample isolates identified with that *terminal* node. Instead we assigned "Unknown" as the source for that node and "Unknown" for the source of all water sample isolates identified with that node. For the Cabin John Creek tree-classification model, the *acceptable source identification probability* was set at 0.50 (50%).

RESULTS: LIBRARY

Known-Source Library. The known-source isolates in the combined Cabin John Creek, Anacostia River, and Piscataway Creek known-source library were grouped into four categories: pet (specifically dogs), human, livestock, and wildlife (Table C-2). The library was analyzed for its ability to take a subset of the library isolates and correctly predict the identity of their host sources when they were treated as unknowns. Average rates of correct classification (ARCC) for the combined library were found by repeating this analysis using several probability cutoff points, as described above. From these results, the percent unknown and percent correct classification (ARCC) was calculated (Table C-3).

---

3 The CART® tree-classification method we employed includes various features to ensure the development of an optimal classification model. For brevity in exposition, we have chosen not to present details of those features, but suggest the following sources: Breiman L, et al. *Classification and Regression Trees*. Pacific Grove: Wadsworth, 1984; and Steinberg D and Colla P. *CART—Classification and Regression Trees.* San Diego, CA: Salford Systems, 1997.

Table C-2: Category, number of isolates, and number of unique patterns by watershed in the combined known-source library

Category	Anacostia River Isolates	Cabin John Creek Isolates	Piscataway Creek Isolates	Total No. Isolates	Total No. Unique Patterns
Pet	257	126	173	556	212
Human	399	155	154	708	364
Livestock	287	63	180	530	252
Wildlife	212	370	267	849	271
Total	1155	714	774	2643	1099

Table C-3: Percent unknown and percent correct for seven (7) cutoff probabilities for the combined Cabin John Creek, Anacostia River, and Piscataway Creek libraries used to identify potential sources of Cabin John Creek water isolates

Cutoff Probability	Percent Unknown	Percent Correct
0.25	0.0%	78.6%
0.375	1.4%	79.0%
0.50	13.6%	86.7%
0.60	19.2%	87.9%
0.70	40.3%	90.4%
0.80	52.9%	94.9%
0.90	74.6%	99.4%

A cutoff probability of 0.50 (50%) was shown to yield a high ARCC of 87%. The percent correct using no cutoff was 79%. Using a cutoff probability of 0.50 (50%), the combined library isolates that were not classified and thus were unknown were removed. The library containing the remaining isolates was then used to test the ability of the combined library to correctly predict the known-source isolates obtained from the Cabin John Creek Watershed. The rates of correction classification for the four categories of sources in the Cabin John Creek known-source isolate library are shown in Table C-4 below. The combined library was then used in the statistical prediction of probable sources of bacteria in water samples collected from the Cabin John Creek.

Table C-4: Actual source categories versus predicted categories of Cabin John Creek known-source isolate library using the three-watershed combined library, with total number of unknown isolates, total isolates, total classified, and rates of correct classification (RCC) for each category

	Predicted →							
Actual ↓	Pet	Human	Livestock	Wildlife	Unknown	Total	Total Classified	RCC[1]
Pet	**107**	5	0	3	11	126	115	93%
Human	9	**122**	3	3	18	155	137	89%
Livestock	0	2	**32**	16	13	63	50	64%
Wildlife	12	19	10	**274**	55	370	315	87%
Total	128	148	45	296	97	714	617	

[1]RCC = Number of correctly predicted species category / Total number classified (predicted). Example: One hundred seven (107) Pet correctly predicted / 115 total number classified for Pet = 107/115 = 93% RCC.

RESULTS: WATER

Cabin John Creek Water Samples. Monthly monitoring from the Cabin John Creek monitoring station was the source of water samples. If weather conditions prevented sampling at a station, a second collection in a later month was performed. The maximum number of *Enterococci* isolates per water sample was 24, although the number of isolates that actually grew was sometimes fewer than 24. A total of 296 *Enterococci* isolates were analyzed by statistical analysis. The BST results by category, Table 5 below shows the number of isolates and percent isolates classified at the 0.50 (50%) cutoff probability, as well as the percent classified overall.

Table C-5: Potential host sources of water isolates by category, number of isolates, percent isolates classified at cutoff probabilities of 50%

Category	No.	% Isolates Classified 50% Prob.
Pet	46	15.5%
Human	135	45.6%
Livestock	35	11.8%
Wildlife	59	19.9%
Unknown	21	7.1%
Missing Data	0	
Total w/ Complete Data	296	
Total	296	
% Classified		92.9%

The relative contributions of potential sources of *Enterococci* contamination in the watershed are shown below in Figure C-2.

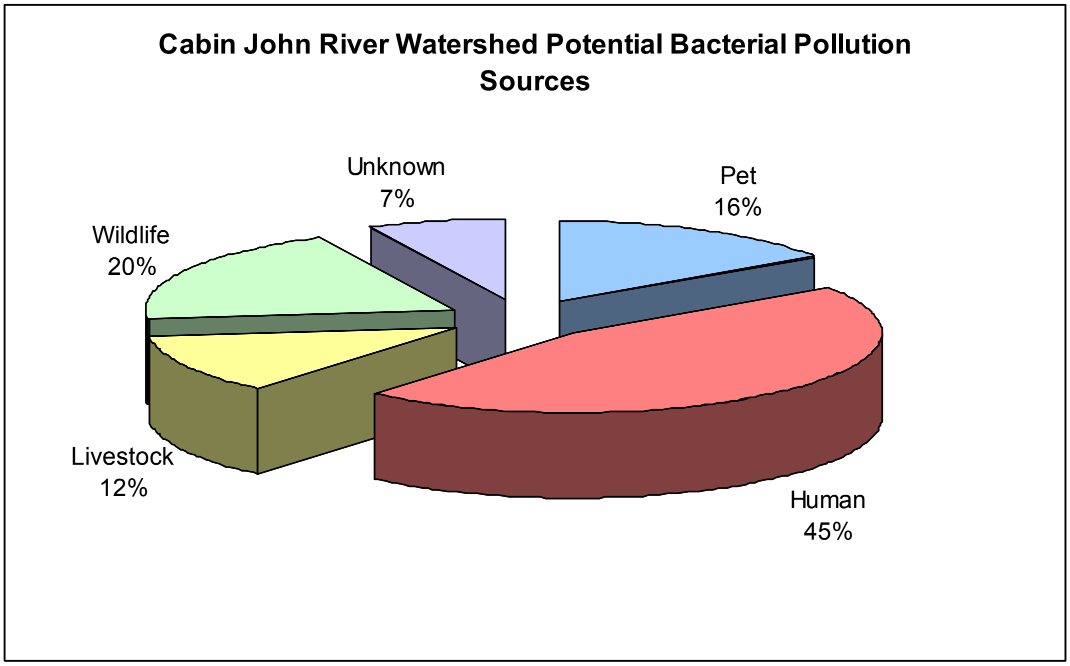

Figure C-2: Cabin John Creek Watershed relative contributions by potential sources of *Enterococci* contamination

The seasonal distribution of water isolates from samples collected at each sampling station is shown below on Table C-6.

Table C-6: Enterococci isolates from water collected and analyzed during the fall, winter, spring, and summer seasons for Cabin John Creek's one (1) monitoring station

Station	Fall	Winter	Spring	Summer	Total
CBJ0005	86	71	72	67	296

Tables C-7 through C-11 on the following pages show the results of BST analysis from the estimation of number of isolates per station per date to the final estimation of the overall percentage of bacteria sources by subwatershed.

Table C-7: BST Analysis – Number of Isolates per Station per Data

Station	Date	# Domestic animals	# Human	# Livestock	# Wildlife	# Unknown
CJB0005	**11/18/2002**	6	5	6	4	2
CJB0005	**12/02/2002**	4	5	5	4	4
CJB0005	**01/06/2003**	3	3	10	5	3
CJB0005	**02/03/2003**	8	10	4	1	1
CJB0005	**03/03/2003**	11	4	4	3	1
CJB0005	**04/21/2003**	10	9	2	1	2
CJB0005	**05/05/2003**	4	7	4	3	6
CJB0005	**06/02/2003**	4	6	4	0	10
CJB0005	**07/07/2003**	12	8	4	0	0
CJB0005	**08/04/2003**	8	9	2	3	1
CJB0005	**09/08/2003**	3	11	2	0	4
CJB0005	**09/22/2003**	7	5	7	1	2
CJB0005	**10/06/2003**	5	3	0	3	8

Table C-8: Percentage of Sources per Station per Date

Station	Date	% Domestic animals	% Human	% Livestock	% Wildlife	% Unknown
CJB0005	**11/18/2002**	26.0870	21.7391	26.0870	17.3913	8.6957
CJB0005	**12/02/2002**	18.1818	22.7273	22.7273	18.1818	18.1818
CJB0005	**01/06/2003**	12.5000	12.5000	41.6667	20.8333	12.5000
CJB0005	**02/03/2003**	33.3333	41.6667	16.6667	4.1667	4.1667
CJB0005	**03/03/2003**	47.8261	17.3913	17.3913	13.0435	4.3478
CJB0005	**04/21/2003**	41.6667	37.5000	8.3333	4.1667	8.3333
CJB0005	**05/05/2003**	16.6667	29.1667	16.6667	12.5000	25.0000
CJB0005	**06/02/2003**	16.6667	25.0000	16.6667	0.0000	41.6667

Station	Date	% Domestic animals	% Human	% Livestock	% Wildlife	% Unknown
CJB0005	**07/07/2003**	50.0000	33.3333	16.6667	0.0000	0.0000
CJB0005	**08/04/2003**	34.7826	39.1304	8.6957	13.0435	4.3478
CJB0005	**09/08/2003**	15.0000	55.0000	10.0000	0.0000	20.0000
CJB0005	**09/22/2003**	31.8182	22.7273	31.8182	4.5455	9.0909
CJB0005	**10/06/2003**	26.3158	15.7895	0.0000	15.7895	42.1053

Table C-9: *E. coli* Concentration and Percentage of Sources by Stratum

SAMPLING STATION IDENTIFIER	DATE START SAMPLING	Flow Stratum (1=high/2=low)	*E. coli* conc. MPN/100ml	log mean conc	% domestic animals	% human	% livestock	% wildlife	% unknown
CJB0005	**10/07/2002**	2	140	2.14613	.	.	.	.	.
CJB0005	**10/21/2002**	2	190	2.27875	.	.	.	.	.
CJB0005	**11/06/2002**	1	1780	3.25042	.	.	.	.	.
CJB0005	**11/18/2002**	1	1720	3.23553	26.0870	21.7391	26.0870	17.3913	8.6957
CJB0005	**12/02/2002**	2	120	2.07918	18.1818	22.7273	22.7273	18.1818	18.1818
CJB0005	**12/16/2002**	1	240	2.38021	.	.	.	.	.
CJB0005	**01/06/2003**	1	130	2.11394	12.5000	12.5000	41.6667	20.8333	12.5000
CJB0005	**01/21/2003**	2	130	2.11394	.	.	.	.	.
CJB0005	**02/03/2003**	2	200	2.30103	33.3333	41.6667	16.6667	4.1667	4.1667
CJB0005	**03/03/2003**	1	290	2.46240	47.8261	17.3913	17.3913	13.0435	4.3478
CJB0005	**03/17/2003**	1	160	2.20412	.	.	.	.	.
CJB0005	**04/21/2003**	1	90	1.95424	41.6667	37.5000	8.3333	4.1667	8.3333
CJB0005	**05/05/2003**	2	70	1.84510	16.6667	29.1667	16.6667	12.5000	25.0000
CJB0005	**05/19/2003**	1	350	2.54407	.	.	.	.	.
CJB0005	**06/02/2003**	1	199	2.29885	16.6667	25.0000	16.6667	0.0000	41.6667

SAMPLING STATION IDENTIFIER	DATE START SAMPLING	Flow Stratum (1=high/2=low)	*E. coli* conc. MPN/100ml	log mean conc	% domestic animals	% human	% livestock	% wildlife	% unknown
CJB0005	06/16/2003	1	260	2.41497	.	.	.	.	.
CJB0005	06/23/2003	1	230	2.36173	.	.	.	.	.
CJB0005	07/07/2003	1	1020	3.00860	50.0000	33.3333	16.6667	0.0000	0.0000
CJB0005	07/21/2003	2	170	2.23045	.	.	.	.	.
CJB0005	08/04/2003	2	2850	3.45484	34.7826	39.1304	8.6957	13.0435	4.3478
CJB0005	08/18/2003	2	1020	3.00860	.	.	.	.	.
CJB0005	08/25/2003	2	30	1.47712	.	.	.	.	.
CJB0005	09/08/2003	2	170	2.23045	15.0000	55.0000	10.0000	0.0000	20.0000
CJB0005	09/22/2003	1	120	2.07918	31.8182	22.7273	31.8182	4.5455	9.0909
CJB0005	10/06/2003	2	120	2.07918	26.3158	15.7895	0.0000	15.7895	42.1053
CJB0005	10/20/2003	2	110	2.04139	.	.	.	.	.

Table C-10: Percentage of Sources per Station by Stratum (Annual Condition)

Station	Flow Stratum (1=high/2=low)	% Domestic Animals	% Human	% Livestock	% Wildlife	% Unknown
CJB0005	1	32.9348	24.3624	22.5159	8.7463	11.4406
	2	25.2753	34.8568	12.0590	10.6040	17.2049

Table C-11: Overall Percentage of Sources per Station (Annual Condition)

Station	% Domestic Animals	% Human	% Livestock	% Wildlife	% Unknown	Total
CJB0005	27.5732	31.7085	15.1961	10.0466	15.4756	100

SUMMARY

The use of ARA was successful for identification of probable bacterial sources in the Cabin John Creek Watershed as evidenced by the RCCs (a range of from a usable 64% to a high of 93%) in the library. When water isolates were compared to the library and potential sources predicted, 93 % of the water isolates were classified by statistical analysis. The largest category of potential sources in the watershed was human (45%). The remaining potential sources included wildlife (20%), pet (16%) and livestock (12%).

REFERENCES

Bell, J.B., Elliott, G.E. & Smith, D.W. (1983) Influence of Sewage Treatment an Urbanization on Selection of Multiple Resistance in Fecal Coliform Populations. *Appl. Environ. Microbiol.* 46, 227-32

Hagedorn, C., Robinson, S.L., Filtz, J.R., Grubbs, S.M., Angier, T.A. & Beneau, R.B. (1999) Determining Sources of Fecal Pollution in a Rural Virginia Watershed with Antibiotic Resistance Patterns in Fecal Streptococci. *Appl. Environ. Microbiol.* 65, 5522-5531.

Krumperman, P.H. (1983) Multiple Antibiotic Resistance Indexing of *Escherichia coli* to Identify High-Risk Sources of Fecal Contamination of Foods. *Appl. Environ. Microbiol.* 46, 165-70

Scott, T.M., Rose, J.B., Jenkins, T.M., Farrah, S.R. & Lukasik, J. 2002 Microbial Source Tracking: Current Methodology and Future Directions. *Appl. Environ. Microbiol.* 68(12), 3373-3385.

Simpson, J.M., Santo Domingo, J.W. & Reasoner, D.J. 2002 Microbial Source Tracking: State of the Science. *Environ. Sci. Technol.* 36(24), 5279-5288.

Wiggins, B.A. (1996) Discriminant Analysis of Antibiotic Resistance Patterns in Fecal Streptococci, a Method to Differentiate Human and Animal Sources of Fecal Pollution in Natural Waters. *Appl. Environ. Microbiol.* 62,3997-4002.

Wiggins, B.A., Andrews, R.W., Conway, R.A., Corr, C.L., Dobratz, E. J., Dougherty, D.P., Eppard, J.R., Knupp, S.R., Limjoco, M.C., Mettenburg, J.M., Rinehardt, J.M., Sonsino, J., Torrijos, R.L. & Zimmerman, M.E. (1999) Use of Antibiotic Resistance Analysis to Identify Nonpoint Sources of Fecal Pollution. *Appl. Environ. Microbiol.* 65, 3483-3486.

ACKNOWLEDGEMENTS

We wish to thank the Richard A. Henson School of Science and Technology of Salisbury University, Salisbury, MD. We also want to acknowledge Dr. Bertram Price of Price Associates, Inc., for his contribution to the statistical analysis in this project.