Awwa Research Foundation
Advancing the Science of Water®

EPA

# Long-Term Variability of BDOM and NOM as Precursors in Watershed Sources

Subject Area:
High-Quality Water

# Long-Term Variability of BDOM and NOM as Precursors in Watershed Sources

©2007 AwwaRF. All Rights Reserved.

## About the Awwa Research Foundation

The Awwa Research Foundation (AwwaRF) is a member-supported, international, nonprofit organization that sponsors research to enable water utilities, public health agencies, and other professionals to provide safe and affordable drinking water to consumers.

The Foundation's mission is to advance the science of water to improve the quality of life. To achieve this mission, the Foundation sponsors studies on all aspects of drinking water, including supply and resources, treatment, monitoring and analysis, distribution, management, and health effects. Funding for research is provided primarily by subscription payments from approximately 1,000 utilities, consulting firms, and manufacturers in North America and abroad. Additional funding comes from collaborative partnerships with other national and international organizations, allowing for resources to be leveraged, expertise to be shared, and broad-based knowledge to be developed and disseminated. Government funding serves as a third source of research dollars.

From its headquarters in Denver, Colorado, the Foundation's staff directs and supports the efforts of more than 800 volunteers who serve on the board of trustees and various committees. These volunteers represent many facets of the water industry, and contribute their expertise to select and monitor research studies that benefit the entire drinking water community.

The results of research are disseminated through a number of channels, including reports, the Web site, conferences, and periodicals.

For subscribers, the Foundation serves as a cooperative program in which water suppliers unite to pool their resources. By applying Foundation research findings, these water suppliers can save substantial costs and stay on the leading edge of drinking water science and technology. Since its inception, AwwaRF has supplied the water community with more than $300 million in applied research.

More information about the Foundation and how to become a subscriber is available on the Web at **www.awwarf.org**.

**©2007 AwwaRF. All Rights Reserved.**

# Long-Term Variability of BDOM and NOM as Precursors in Watershed Sources

Prepared by:
**David A. Reckhow**, **Paula L. Rees**, **Klaus Nüsslein**, **Gladys Makdissy**, **Greg Devine**, **Teresa Conneely**, **Alison Boutin**, and **Darleen Bryan**
University of Massachusetts
Amherst, MA 01003

Sponsored by:
**Awwa Research Foundation**
6666 West Quincy Avenue, Denver, CO 80235-3098

and

**U.S. Environmental Protection Agency**
Washington, DC 20460

Published by:

Awwa Research Foundation

American Water Works Association

IWA Publishing

©2007 AwwaRF. All Rights Reserved.

## DISCLAIMER

This study was jointly funded by the Awwa Research Foundation (AwwaRF) and the U.S. Environmental Protection Agency (USEPA) under Cooperative Agreement No. R829409-01. AwwaRF and USEPA assume no responsibility for the content of the research study reported in this publication or for the opinions or statements of fact expressed in the report. The mention of trade names for commercial products does not represent or imply the approval or endorsement of AwwaRF or USEPA. This report is presented solely for informational purposes.

Copyright © 2007
by Awwa Research Foundation

All Rights Reserved

Printed in the U.S.A.

Printed on recycled paper

©2007 AwwaRF. All Rights Reserved.

# CONTENTS

LIST OF TABLES ..... xi

LIST OF FIGURES ..... xv

FOREWORD ..... xxv

ACKNOWLEDGMENTS ..... xxvii

EXECUTIVE SUMMARY ..... xxix

CHAPTER 1: INTRODUCTION ..... 1
- NOM Types and Genesis ..... 1
- Biodegradability Tests for NOM ..... 2
- NOM Modeling and Management ..... 4

CHAPTER 2: MATERIALS AND METHODS ..... 7
- Research Objectives ..... 7
- General Approach ..... 7
- Synopsis of Project ..... 7
  - Utility Survey ..... 7
  - Study of Selected Watersheds ..... 9
  - General Assessment and Modeling ..... 12
- Detailed Experimental Methods ..... 12
  - Laboratory Treatments ..... 12
  - Chemical and Biological Analyses ..... 13

CHAPTER 3: DATA COLLECTION FROM UTILITIES ..... 17
- Mail Survey ..... 17
  - Methods ..... 17
  - Response Rate ..... 17
  - Results ..... 18
- Data From Other Sources ..... 24
  - Length of Record ..... 24
  - SUVA Criteria ..... 25
  - Criteria Based on Ratios of Known to Unknown TOX ..... 25
  - Ecoregion ..... 27
- Primary List of Candidate Utilities ..... 28
- Secondary List of Candidate Utilities ..... 28
- Targeted Data Collection From Candidate Utilities ..... 28
  - Data Sources ..... 38
  - Full Database ..... 39
- Summary and Conclusions ..... 39



©2007 AwwaRF. All Rights Reserved.

CHAPTER 4: ANALYSIS OF PRECURSOR LEVELS ACROSS NORTH AMERICA ..... 43
Introduction ..... 43
DBP Precursor Tests and Models ..... 44
Approach to DBP Precursor Assessment ..... 44
Collecting New Data: Chlorination Conditions ..... 45
Standardizing Literature Data: The Hybrid Model ..... 45
Whole Waters Database ..... 47
Correlations From the Literature ..... 47
Full North American Database ..... 50
Field Studies in the Wachusett Watershed ..... 54
Database of NOM Fractions ..... 56
DBP Precursor Content in a Coherent Set of NOM Fractions ..... 60
Meta Analysis of Published Data on NOM Precursors ..... 63
Summary and Conclusions ..... 74

CHAPTER 5: ANALYSIS OF AQUATIC ORGANIC MATTER ..... 75
Review of Current Methods for AOC and BDOC ..... 75
The Connection Among BOM, DOM, and TOC ..... 75
Assimilable Organic Carbon ..... 76
Measures of Biomass ..... 77
Biodegradable Dissolved Organic Carbon ..... 84
BDOC Versus AOC ..... 91
Summary ..... 91
Methods Development ..... 93
NOM Isolation Protocol ..... 93
Lignin Testing With Alkaline CuO Method ..... 97
Evaluation of the LeChevallier ATP Rapid Method for Measuring AOC ..... 98
Isolation and Characterization of NOM Fractions ..... 101
Introduction ..... 101
Experimental Methods ..... 102
Characteristics of the Water Samples ..... 104
NOM Fractionation ..... 104
Bench-scale Treatment ..... 109
AOC and BDOC Tests ..... 110
Chlorination By-product Yields ..... 117
Chlorination By-product Precursor Recovery ..... 118
Summary and Conclusions ..... 124

CHAPTER 6: LEACHING STUDIES ..... 129
Introduction ..... 129
Background ..... 129
Leaching of DOC From Plant Material ..... 129
Biodegradation of Natural Organic Matter ..... 131



©2007 AwwaRF. All Rights Reserved.

Experimental Design ..... 132
Study 1 ..... 132
Study 2 ..... 133
Results of Study 1 ..... 135
Leaching ..... 135
Degradation ..... 145
Results of Study 2 ..... 151
Leaching ..... 151
Degradation ..... 151
Discussion ..... 156
Leaching of NOM ..... 156
Biodegradation of NOM ..... 157
Specific DBP Formation Potentials ..... 157
Relationship to Natural Processes ..... 158
Summary and Conclusions ..... 161

CHAPTER 7: ROLE OF LIGNIN ..... 163
Introduction ..... 163
Experimental Methods ..... 164
Chemicals and Standards ..... 164
Oxidation and Extraction ..... 166
HPLC Analysis ..... 166
Detection and Quantification ..... 167
Chlorination ..... 167
Analysis of DBPs and TOX ..... 167
Results and Discussion ..... 167
Lignin Phenolic Yields ..... 167
DBP Yields of Lignin Monomers ..... 168
Estimated DBP Yield ..... 169
Summary and Conclusions ..... 175

CHAPTER 8: CONCEPTUAL MODEL FORMULATION ..... 177
Introduction ..... 177
Empirical Observations From Field Studies ..... 179
Long-term Field Studies in the Wachusett Watershed ..... 179
Wet Weather Studies at Wachusett ..... 181
Cambridge: Another Intensive Regional Study ..... 183
Multicomponent Degradation of Allochthonous NOM ..... 183
Summary and Conclusions ..... 187

CHAPTER 9: NATIONAL TRENDS FOR ORGANIC CARBON ..... 191
Overview ..... 191
Select Results of Past Studies ..... 191
Available TOC Data ..... 193
Statistical Methods ..... 194
Background ..... 194
Model Selection ..... 195



©2007 AwwaRF. All Rights Reserved.

Spatial Trends ..... 197
Variability Between Ecoregions ..... 197
Variability Within Ecoregions ..... 200
Temporal Trends ..... 201
Mann–Kendall Technique ..... 201
Results ..... 206
Summary and Conclusions ..... 212

CHAPTER 10: UTILITY DATA ANALYSIS AND MODELING ..... 215
Overview ..... 215
Select Results of Past Studies ..... 215
Empirical Methods ..... 217
Nationwide Model for Organic Carbon ..... 219
Available Land Use and TOC Data ..... 219
National Period-of-Record Analysis ..... 220
Seasonal Analysis ..... 220
Site-Specific Studies for DBP Precursors and Organic Carbon ..... 225
Brick, New Jersey: Metedeconk River ..... 225
Cambridge System Analysis ..... 249
Site-Specific Analysis of Microbial Communities ..... 253
Introduction ..... 253
Results ..... 254
Summary and Conclusions ..... 258
Empirical Analyses ..... 258
Short-Term DBP Precursor and Microbial Analyses ..... 260
Future Work ..... 260

CHAPTER 11: MECHANISTIC MODELING ..... 261
Overview ..... 261
Select Results of Past Mechanistic Modeling Studies ..... 262
Methodology ..... 263
Fibonacci Series ..... 263
Golden Section ..... 263
Fibonacci Algorithm ..... 265
Model Development ..... 266
Fibonacci Optimization ..... 268
Data Requirements ..... 269
Results ..... 269
Calibration ..... 270
Summary and Conclusions ..... 274

CHAPTER 12: CONCLUSIONS ..... 275
Properties of NOM ..... 275
Watershed Processes ..... 276



©2007 AwwaRF. All Rights Reserved.

CHAPTER 13: RECOMMENDATIONS ..... 277
Practical Applications and Recommendations to Utilities..... 277
Recommendations for Future Research ..... 277

APPENDIX A: MAIL SURVEY ..... 279

APPENDIX B: TELEPHONE SURVEY QUESTIONS..... 283

REFERENCES ..... 289

ABBREVIATIONS ..... 305



©2007 AwwaRF. All Rights Reserved.

©2007 AwwaRF. All Rights Reserved.

# TABLES


2.1 NOM fractions: Operational versus conceptual nomenclature ..... 11

2.2 Testing of whole waters and NOM fractions ..... 11

3.1 Comparative raw water quality for high- and low-SUVA waters ..... 26

3.2 Selected utilities representing extremes in known to unknown TOX ratios ..... 27

3.3 Primary list of candidate utilities ..... 29

3.4 Secondary list of candidate utilities ..... 32

4.1 Model parameters used to adjust DBP yields ..... 46

4.2 Reference conditions for DBP formation tests ..... 47

4.3 Population characteristics for THMFP in surface waters ..... 55

4.4 Population characteristics for THM, DHAA, and THAA formation potentials in hydrophobic NOM fractions ..... 69

4.5 Population characteristics for THM, DHAA, and THAA formation potentials in transphilic NOM fractions ..... 69

4.6 Population characteristics for THM, DHAA, and THAA formation potentials in hydrophilic NOM fractions ..... 70

5.1 Summary of AOC methods ..... 78

5.2 Comparative summary of methods to measure BDOC ..... 86

5.3 Model compounds for fractionation tests ..... 94

5.4 UF fractionation of synthetic waters ..... 95

5.5 XAD-8 fractionation of synthetic waters ..... 95

5.6 Commercial lignin analysis by microwave and oven digestion ..... 99

5.7 Analysis of humic substances using the lignin phenol method with oven digestion ..... 100




©2007 AwwaRF. All Rights Reserved.

5.8 AOC values for the LeChevallier and the *Standard Methods* tests for a water sample series prepared from raw water from Tampa Reservoir waters ..... 100

5.9 Water samples ..... 102

5.10 General characteristics of water samples ..... 105

5.11 Chickahominy fractionation: Carbon mass balance and elemental analysis ..... 106

5.12 Assomption fractionation: Carbon mass balance and elemental analysis ..... 106

5.13 Shoal fractionation: Carbon mass balance and elemental analysis ..... 107

5.14 Hillsborough fractionation: Carbon mass balance and elemental analysis ..... 107

5.15 Quabbin fractionation: Carbon mass balance and elemental analysis ..... 108

5.16 Overall distribution of organic carbon ..... 108

5.17 Chlorine dose for low-dose conditions ..... 109

5.18 Summary of test results for AOC and BDOC assays of raw waters and corresponding NOM fractions ..... 111

5.19 The relative contribution of each fraction compared to the raw water sample ..... 117

6.1 Summary of literature on leaching studies with plant matter ..... 130

6.2 Analytical schedule for study 2 testing ..... 136

6.3 Leaching experiment DOC results for all samples ..... 138

6.4 Net carbon release rates ..... 140

6.5 $UV_{254}$, DOC, and SUVA on final day of leaching for all samples ..... 140

6.6 DBPFP data for last day of leaching ..... 142

6.7 Specific DBPFP data for last day of leaching ..... 144

6.8 Biodegradation experiment UV absorbance results for all samples ..... 147

6.9 Biodegradation experiment TOC results for all samples ..... 148

6.10 $UV_{254}$, TOC, and SUVA on final day of biodegradation for all samples ..... 149



©2007 AwwaRF. All Rights Reserved.

6.11 DBPFP data for last day of biodegradation ..... 150

6.12 Specific DBPFP data for last day of biodegradation ..... 151

6.13 Mass balance of leached material ..... 151

7.1 Characterization of key lignin monomers ..... 164

7.2 Characteristics of commercial lignins ..... 165

7.3 HPLC gradient program ..... 166

7.4 Phenolic composition of lignin ..... 169

7.5 DBP yields of lignin monomers following chlorination ..... 170

7.6 Estimated DBP yields of monomers within the organosolve lignin polymer ..... 171

7.7 Estimated DBP yields of monomers within the alkali lignin polymer ..... 172

8.1 Arbitrary model parameters used to create multicomponent model ..... 185

9.1 Descriptive statistics of data contained in the TOC database, by ecoregion ..... 198

9.2 Statistical comparison via the Mann–Whitney test of the median values, modern data ..... 199

9.3 Sample of seasonal Kendall results and Theil estimator of slope ..... 209

9.4 Results of the Mann–Whitney rank-sum test for modern data set ..... 210

9.5 Sample of results of covariance analysis within ecoregion 11.1, modern data set ..... 211

9.6 Results (*p*-values) of the Mann–Whitney rank-sum test for historical data set ..... 213

10.1 TOC data for each utility within the regression analyses ..... 219

10.2 Errors between measured TOC concentration and predicted TOC values for national regression ..... 226

10.3 Key characteristics of the Metedeconk River watershed ..... 227

10.4 Errors between measured TOC concentration and predicted TOC values for North and South branches ..... 234

10.5 Sampling locations and related information ..... 235



©2007 AwwaRF. All Rights Reserved.

10.6 General characteristics of water samples prior to chlorination ..... 235

10.7 Chlorine demand data ..... 236

10.8 THM concentrations ..... 237

10.9 THAA concentrations ..... 237

10.10 DHAA concentrations ..... 238

10.11 Bromine incorporation factor and substitution factor for THMs ..... 241

10.12 Tertiary LULC categories for eight sub-basins in the Brick watershed ..... 243

10.13 Secondary LULC categories for eight sub-basins in the Brick watershed ..... 244

10.14 Primary LULC categories for eight sub-basins in the Brick watershed ..... 245

10.15 Model summary for full watershed LULC ..... 246

10.16 Model summary for riparian zone LULC ..... 247

10.17 THMFP data for Cambridge water supply catchment ..... 253

11.1 Summary statistics for TOC data within the Metedeconk system ..... 269

11.2 Comparison between measured TOC concentration and predicted TOC values during the calibration period ..... 271

11.3 Comparison between measured TOC concentration and predicted TOC values during validation period ..... 273



©2007 AwwaRF. All Rights Reserved.

# FIGURES


1.1 Three major pools of aquatic NOM ..... 3

2.1 Experimental scheme for isolating the three major pools of NOM ..... 10

3.1 Issues of interest to water utilities ..... 23

3.2 Distribution of raw water NOM characteristics for 195 large U.S. plants (summarized from ICR data) and 1 plant in Winnipeg ..... 25

3.3 Known versus unknown TOX in selected ICR data ..... 26

3.4 Input form used for primary data collection: NOM-related data ..... 39

3.5 Input form used for primary data collection: water source information ..... 40

3.6 A selection of data fields ..... 40

4.1 Retesting of four databases used by Chapra, Canale, and Amy ..... 48

4.2 Re-evaluation of data from Figure 4.1 based on a common chlorination protocol.... 49

4.3 Standard specific THMFP versus TOC for the four data sets from Chapra, Canale, and Amy ..... 50

4.4 Standard specific THMFP versus TOC for all 27 data sets from the literature ..... 51

4.5 Standard specific THMFP versus TOC for surface waters based on all 27 data sets from the literature ..... 51

4.6 Standard specific THMFP versus SUVA for all data sets from the literature ..... 52

4.7 Standard specific dihaloacetic acid formation potential (DHAAFP) versus TOC for all data sets from the literature ..... 52

4.8 Standard specific trihaloacetic acid formation potential (THAAFP) versus TOC for all data sets from the literature ..... 53

4.9 Ratio of THAA precursors to THM precursors versus TOC for all data sets from the literature ..... 53

4.10 Cumulative frequency plot for specific THM precursor content of all surface and groundwaters ..... 54




©2007 AwwaRF. All Rights Reserved.

4.11 Formation of major DBPs from chlorination of samples collected from the Wachusett watershed ..... 55

4.12 Relationship between specific THM precursor levels and TOC in Wachusett Reservoir tributaries ..... 56

4.13 Relationship between THM precursors and SUVA for samples from Wachusett Reservoir watershed (part 1) ..... 57

4.14 Relationship between THM precursors and SUVA for samples from Wachusett Reservoir watershed (part 2) ..... 57

4.15 Specific THM precursor levels in Wachusett samples as compared to the North American surface water database, all plotted versus the TOC..... 58

4.16 Specific THM precursor levels in Wachusett samples as compared to the North American surface water database, all plotted versus the SUVA ..... 58

4.17 Specific DHAA precursor levels in Wachusett samples as compared to the North American surface water database, all plotted versus the TOC..... 59

4.18 Specific THAA precursor levels in Wachusett samples as compared to the North American surface water database, all plotted versus the TOC..... 59

4.19 Ratio of THAA to THM precursors in Wachusett samples as compared to the North American surface water database, all plotted versus the TOC ..... 60

4.20 Relationship between specific chlorine demand and SUVA for isolated NOM fractions from New England raw and treated waters..... 61

4.21 Relationship between specific THM precursor content and SUVA for isolated NOM fractions from New England raw and treated waters ..... 61

4.22 Relationship between specific THM precursor content and SUVA for isolated NOM fractions from New England raw and treated waters, including boxed NOM domains..... 62

4.23 Relationship between specific TCAA precursor content and SUVA for isolated NOM fractions from New England raw and treated waters ..... 63

4.24 Relationship between specific DHAA precursor content and SUVA for isolated NOM fractions from New England raw and treated waters ..... 64

4.25 Relationship between specific DHAN precursor content and SUVA for isolated NOM fractions from New England raw and treated waters ..... 64



©2007 AwwaRF. All Rights Reserved.

4.26 Relationship between TCAA/THM precursor ratio and SUVA for isolated NOM fractions from New England raw waters ..... 65

4.27 Relationship between specific THM precursor content and molecular size for isolated NOM fractions from New England (Forge Pond and Connecticut) waters ..... 65

4.28 Cumulative frequency plot for THM precursor content in major raw water fractions (untreated waters) ..... 67

4.29 Cumulative frequency plot for DHAA precursor content in major raw water fractions (untreated waters) ..... 67

4.30 Cumulative frequency plot for THAA precursor content in major raw water fractions (untreated waters) ..... 68

4.31 Cumulative frequency plot for THAA/THM precursor ratio in major raw water fractions (untreated waters) ..... 68

4.32 Cumulative frequency plot for DHAA/THM precursor ratio in major raw water fractions (untreated waters) ..... 69

4.33 Specific THM precursor content of all fractions plotted versus the SUVA (untreated waters) ..... 70

4.34 Relationship between specific THM precursor content of hydrophilic and transphilic fractions to the specific content in the corresponding hydrophobic fraction plotted versus the NOM fraction SUVA (untreated waters) ..... 71

4.35 Ratio of specific THM precursor content of NOM fractions to the specific content in the whole water from which they were derived plotted versus NOM fraction SUVA (untreated waters) ..... 72

4.36 Ratio of specific THM precursor content of NOM fractions to the specific content in the whole water from which they were derived (untreated waters) ..... 72

4.37 Relationship between specific THM precursor content of hydrophilic and transphilic fractions to the specific content in the corresponding hydrophobic fraction (untreated waters) ..... 73

4.38 Cumulative frequency plot for the ratio of the specific THM precursor content for a hydrophilic fraction to its corresponding hydrophobic faction ..... 73

5.1 Experimental schematic for fractionation tests ..... 94

5.2 Experimental scheme for isolating the three major pools of NOM ..... 96



©2007 AwwaRF. All Rights Reserved.

5.3 Digestion reactor used with classical oven method ..... 98

5.4 Correlation between LeChevallier test and *Standard Methods* test for AOC ..... 101

5.5 Method for isolating NOM fractions ..... 103

5.6 Comparison of TOC values with national ICR data ..... 105

5.7 AOC and BDOC of Chickahominy River water (Newport News, Va.) and corresponding NOM fractions ..... 112

5.8 AOC and BDOC of Assomption River (Repentigny, Que.) water and corresponding NOM fractions ..... 113

5.9 AOC and BDOC of Shoal Lake (Winnipeg, Man.) water and corresponding NOM fractions ..... 114

5.10 AOC and BDOC of Hillsborough River (Tampa, Fla.) water and corresponding NOM fractions ..... 115

5.11 AOC and BDOC of Quabbin Reservoir (Boston, Mass.) water and corresponding NOM fractions ..... 116

5.12 Chlorination by-products precursor content for the raw waters ..... 118

5.13 DBP precursor content for NOM fractions extracted from Chickahominy River water ..... 119

5.14 DBP precursor content for NOM fractions extracted from Assomption River water ..... 119

5.15 DBP precursor content for NOM fractions extracted from Shoal Lake water ..... 120

5.16 DBP precursor content for NOM fractions extracted from Hillsborough River water ..... 120

5.17 DBP precursor content for NOM fractions extracted from Quabbin Reservoir water ..... 121

5.18 Recovery of organic carbon versus recovery of chlorinated by-product precursor (Chickahominy River water) ..... 121

5.19 Recovery of organic carbon versus recovery of chlorinated by-product precursor (Assomption River water) ..... 122



©2007 AwwaRF. All Rights Reserved.

5.20 Recovery of organic carbon versus recovery of chlorinated by-product precursor (Shoal Lake water) ..... 122

5.21 Recovery of organic carbon versus recovery of chlorinated by-product precursor (Hillsborough River water) ..... 123

5.22 Recovery of organic carbon versus recovery of chlorinated by-product precursor (Quabbin Reservoir water) ..... 123

5.23 THM precursor content of watershed NOM shown against the backdrop of the full literature database (untreated waters) ..... 124

5.24 DHAA precursor content of watershed NOM shown against the backdrop of the full literature database (untreated waters) ..... 125

5.25 THAA precursor content of watershed NOM shown against the backdrop of the full literature database (untreated waters) ..... 125

5.26 DHAA/THM precursor ratio of watershed NOM shown against the backdrop of the full literature database (untreated waters) ..... 126

5.27 THAA/THM precursor ratio of watershed NOM shown against the backdrop of the full literature database (untreated waters) ..... 126

5.28 Chlorination by-product precursor content for NOM fractions extracted from Quabbin River water ..... 127

5.29 Chlorination by-product precursor content for NOM fractions extracted from Hillsborough River water ..... 127

6.1 Experimental design for leaching study 2 ..... 134

6.2 UV absorbance levels over time for leaching of red maple leaves ..... 136

6.3 UV absorbance levels over time for leaching of white oak leaves ..... 137

6.4 UV absorbance levels over time for leaching of white pine needles ..... 137

6.5 UV absorbance levels over time for leaching of maple, oak, and pine foliage ..... 139

6.6 Changes in total carbon over time for leaching of maple, oak, and pine samples ..... 139

6.7 TOC on final day of leaching for all samples ..... 141

6.8 Leaching of organic matter versus time for late fall leaf litter ..... 142



©2007 AwwaRF. All Rights Reserved.

6.9 DBPFPs for last day of leaching ..... 143

6.10 Specific DBPFPs for last day of leaching ..... 144

6.11 Specific THM formation as a function of leaf type and leaching conditions ..... 145

6.12 Specific THAA formation as a function of leaf type and leaching conditions ..... 146

6.13 Degradation of leached organic matter versus time ..... 146

6.14 Specific DBPFPs for last day of leaching ..... 147

6.15 UV absorbance over time for biodegradation of maple, oak, and pine leachate ..... 148

6.16 Changes in total carbon over time for biodegradation of maple, oak, and pine leachate ..... 149

6.17 DBPFPs for last day of biodegradation ..... 150

6.18 Degradation of DOC for tests with new maple leaves ..... 152

6.19 Degradation of DOC for tests with old maple leaves ..... 153

6.20 BDOC results for leachate tests with new maple leaves ..... 153

6.21 BDOC results for leachate tests with old maple leaves ..... 154

6.22 AOC results for leachate tests with new maple leaves ..... 155

6.23 AOC results for leachate tests with old maple leaves ..... 155

6.24 Specific THMFPs for leached and biodegraded organic matter ..... 158

6.25 DHAAFPs for leached and biodegraded organic matter ..... 159

6.26 Specific THAAFPs for leached and biodegraded organic matter ..... 159

6.27 Specific THM formation versus TOC for natural water samples and leaf leachate ..... 160

7.1 Structures identified from CuO degradation of NOM ..... 165

7.2 Comparison of measured and calculated DBP yields for the two commercial lignins ..... 173



©2007 AwwaRF. All Rights Reserved.

7.3 Contribution of lignin to DBP formation in NOM based on specific monomer reactivity ..... 174

7.4 Contribution of lignin to DBP formation in NOM based on total lignin assay ..... 175

8.1 Important watershed processes in the cycling of NOM ..... 178

8.2 Relationship between specific THM precursors and TOC for individual locations in the Wachusett Reservoir watershed (part 1) ..... 178

8.3 Relationship between specific THM precursors and TOC for individual locations in the Wachusett Reservoir watershed (part 2) ..... 179

8.4 Relationship between specific DHAA precursors and TOC for individual locations in the Wachusett Reservoir watershed (part 1) ..... 180

8.5 Relationship between specific DHAA precursors and TOC for individual locations in the Wachusett Reservoir watershed (part 2) ..... 181

8.6 DOC and specific HAAFP as a function of time for a rain event at the upstream, moderately impacted GBWD site ..... 182

8.7 DOC and specific HAAFP as a function of time for a rain event at the downstream, heavily residential GBBR site ..... 182

8.8 Relationship between specific THM precursor levels and TOC in Cambridge watershed..... 183

8.9 Multicomponent allochthonous NOM degradation: major bulk parameters ..... 185

8.10 Multicomponent allochthonous NOM degradation: biochemical components ..... 186

8.11 Multicomponent allochthonous NOM degradation: THM precursor composition ... 186

8.12 Conceptual model for NOM degradation: low-flow conditions ..... 188

8.13 Conceptual model for NOM degradation: high-flow conditions ..... 188

9.1 USEPA ecoregion level II designations ..... 192

9.2 Utility watershed locations and the HUC locations for which USEPA data were downloaded..... 195

9.3 Box-plot summary of data contained in the Legacy (1960–1998) database, organized by ecoregion ..... 198



©2007 AwwaRF. All Rights Reserved.

9.4 Decadal TOC variability by ecoregion ..... 200

9.5 Seasonal TOC variability by ecoregion ..... 201

9.6 Ecoregion 8.3 HUCs and locations ..... 202

9.7 Box-plot representation of interquartile range of data for each HUC within ecoregion 8.3 ..... 202

9.8 Seasonal trends by decade for ecoregion 8.3 ..... 203

9.9 Kendall's τ with ecoregions for modern data set ..... 207

9.10 Theil estimator of slope for modern data set ..... 208

9.11 Theil estimator of slope with ecoregions for modern data set ..... 208

9.12 Kendall's τ with ecoregions for historical data set ..... 211

9.13 Theil estimator of slope with ecoregions for historical data set ..... 212

10.1 National regression comparison between empirical TOC predictions and actual TOC concentrations ..... 221

10.2 Variability of available TOC data for each utility ..... 221

10.3 National winter regression comparison between empirical TOC predictions and actual TOC concentration ..... 223

10.4 National spring regression comparison between empirical TOC predictions and actual TOC concentration ..... 223

10.5 National summer regression comparison between empirical TOC predictions and actual TOC concentration ..... 224

10.6 National fall regression comparison between empirical TOC predictions and actual TOC concentration ..... 224

10.7 Full set of LULC categories across the Metedeconk watershed ..... 228

10.8 Metedeconk system regression comparison between empirical TOC predictions and actual TOC concentration ..... 228

10.9 North Branch regression comparison between empirical TOC predictions and actual TOC concentration ..... 230



©2007 AwwaRF. All Rights Reserved.

10.10 South Branch regression comparison between empirical TOC predictions and actual TOC concentration ..... 230

10.11 Winter MLR for North and South branch comparison between empirical TOC predictions and actual TOC concentration ..... 232

10.12 Summer MLR for North and South branch comparison between empirical TOC predictions and actual TOC concentration ..... 233

10.13 Specific THM4 yield following chlorination of water samples ..... 238

10.14 Standard specific THMFP versus TOC in surface waters ..... 239

10.15 DHAA and THAA yields following chlorination of water samples ..... 239

10.16 Ratio of THAA precursors to THM precursors in surface waters ..... 240

10.17 Ratio of DHAA precursors to THM precursors in surface waters ..... 240

10.18 Rainfall for the Brick region during the sampling period in May 2005 ..... 241

10.19 Streamflow for the North Branch of the Metedeconk River during the sampling period in May 2005 ..... 242

10.20 Relationship between wetland area and TOC concentration ..... 246

10.21 Full set of LULC categories in riparian zones across the Metedeconk watershed ..... 247

10.22 Cambridge regression comparison between empirical TOC predictions and actual TOC concentration ..... 250

10.23 Cambridge winter regression comparison between empirical TOC predictions and actual TOC concentration ..... 252

10.24 Cambridge summer regression comparison between empirical TOC predictions and actual TOC concentration ..... 253

10.25 Cambridge THMFP regression comparison between empirical TOC predictions and actual TOC concentration ..... 254

10.26 Phylogenetic analysis of bacterial 16S rRNA in raw water from the Newport News watershed (full chart) ..... 255

10.27 Phylogenetic analysis of bacterial 16S rRNA in raw water from the Newport News watershed (expansion of top half of Figure 10.26, for clarity) ..... 256



©2007 AwwaRF. All Rights Reserved.

10.28 Phylogenetic analysis of bacterial 16S rRNA in raw water from the Newport News watershed (expansion of bottom half of Figure 10.26, for clarity) ... 257

11.1 Golden section geometry ..... 264

11.2 Procedure of Fibonacci optimization technique ..... 266

11.3 Modeling simplification of Metedeconk watershed ..... 268

11.4 Metedeconk calibration results ..... 271

11.5 Metedeconk validation results ..... 272



©2007 AwwaRF. All Rights Reserved.

# FOREWORD

The Awwa Research Foundation (AwwaRF) is a nonprofit corporation that is dedicated to the implementation of a research effort to help utilities respond to regulatory requirements and traditional high-priority concerns of the industry. The research agenda is developed through a process of consultation with subscribers and drinking water professionals. Under the umbrella of a Strategic Research Plan, the Research Advisory Council prioritizes the suggested projects based upon current and future needs, applicability, and past work; the recommendations are forwarded to the Board of Trustees for final selection. The foundation also sponsors research projects through an unsolicited proposal process; the Collaborative Research, Research Applications, and Tailored Collaboration programs; and various joint research efforts with organizations such as the U.S. Environmental Protection Agency, the U.S. Bureau of Reclamation, and the Association of California Water Agencies.

This publication is a result of one of these sponsored studies, and it is hoped that its findings will be applied in communities throughout the world. The following report serves not only as a means of communicating the results of the water industry's centralized research program but also as a tool to enlist the further support of the nonmember utilities and individuals.

Projects are managed closely from their inception to the final report by the foundation's staff and large cadre of volunteers who willingly contribute their time and expertise. The foundation serves a planning and management function and awards contracts to other institutions such as water utilities, universities, and engineering firms. The funding for this research effort comes primarily from the Subscription Program, through which water utilities subscribe to the research program and make an annual payment proportionate to the volume of water they deliver and consultants and manufacturers subscribe based on their annual billings. The program offers a cost-effective and fair method for funding research in the public interest.

A broad spectrum of water supply issues is addressed by the foundation's research agenda: resources, treatment and operations, distribution and storage, water quality and analysis, toxicology, economics, and management. The ultimate purpose of the coordinated effort is to assist water suppliers to provide the highest possible quality of water economically and reliably. The true benefits are realized when the results are implemented at the utility level. The foundation's trustees are pleased to offer this publication as a contribution toward that end.

Walter J. Bishop
Chair, Board of Trustees
Awwa Research Foundation

Robert C. Renner, P.E.
Executive Director
Awwa Research Foundation



©2007 AwwaRF. All Rights Reserved.

©2007 AwwaRF. All Rights Reserved.

# ACKNOWLEDGMENTS

All large projects incorporating diverse research skills such as this one rely on many supporting individuals to be successful. This project has many individuals deserving of special recognition, especially David Loomis and his students who prepared and executed the mail survey, and Sarah Titus who assisted with the literature review and helped with some of the methods development work.

We acknowledge the assistance of the numerous utility personnel, engineers, and scientists with this project. More than 300 individuals had a part in this work, from those responding to the mail surveys to others who assisted in collecting water samples, answering questions, and generally passing on their accumulated wisdom to the project team.

We also thank the project officer, Djanette Khiari, and the project advisory team—Mel Suffet, Michael Elovitz, Anne Backes, and especially Jerry Leenheer—for their help at various stages of this work.



©2007 AwwaRF. All Rights Reserved.

©2007 AwwaRF. All Rights Reserved.

# EXECUTIVE SUMMARY

## INTRODUCTION

Most efforts at controlling disinfection by-product (DBP) formation and allied biodegradable compounds have focused on the treatment plant and distribution system. Drinking water providers now have an array of tools for addressing this problem through enhanced removal of precursors, better process sequencing to minimize formation, use of alternative disinfectants to reduce the formation of chlorine-based DBPs, and distribution system management to minimize dead ends. One area of drinking water system management that has not been widely discussed in the context of DBP control is management of the watershed and raw water quality.

Aquatic ecologists, limnologists, and geochemists have long studied the generation of dissolved organic carbon (DOC) and its cycling in the freshwater environment. Much has been learned about the sources, transport, and degradation of bulk organic matter and selected specific organic compounds in the aquatic environment. However, very few studies have been conducted on the fate and transport of DBP precursors. This research examines the relationship between DOC and dissolved DBP precursors in surface waters, and extends this information to the modeling of natural organic matter (NOM) and precursor organics in drinking water catchments. In addition, the highly biodegradable components of NOM and long-term trends in bulk NOM concentrations have been the focus of substudies within this project.

## DBP PRECURSOR DATA

To improve our current understanding of the full range of precursor levels in U.S. waters, three types of data collection efforts were undertaken. The first stage involved a survey mailed to 553 of the largest utilities in the United States using surface water (chapter 3). The second stage involved phone followups and site visits to selected utilities from the first list of survey recipients (chapter 3). To obtain the most data from this work, information was also collected on organic carbon levels, source types, land use characteristics, and utility interest in these issues. Finally, an extensive literature survey was conducted, spanning the period from 1975 to the present (chapter 4).

A substantial number of utilities have collected DBP data beyond that which has been mandated by regulatory agencies. About 17% of respondents had collected total organic halide (TOX) data in addition to any that were required under the Information Collection Rule (ICR). About 40% had collected data on DBP precursors in their raw water (including trihalomethane [THM], haloacetic acid [HAA], and TOX precursors). Of these, 12 utilities reported that they had collected precursor data prior to 1981. About 60% of the respondents reported that they had collected total organic carbon (TOC) or similar NOM-related data at locations in their watersheds. A slightly smaller number of utilities (~40%) reported that they had conducted watershed studies focused on storm events. The vast majority (90%) were willing to share their data. Data from the ICR revealed that some chlorinating utilities have high ratios of known to unknown TOX (>50%) and some have quite low ratios (<15%). There was no obvious connection between unknown TOX level and geographic location, raw water TOC, or specific ultraviolet absorbance (SUVA).

From these site-specific data, a set of about 50 primary utilities and about 120 secondary utilities were identified. Some utilities were the focus of this research project; however, time and resources did not permit exploration of all candidate utilities. Much unpublished data exist in the



©2007 AwwaRF. All Rights Reserved.

databases of these utilities. Given their expressed interest in watershed management for precursor control, it is recommended that additional studies be conducted to better develop an understanding of relevant watershed processes. Any future studies regarding NOM in watersheds should make use of these utility lists and the accompanying database to best capture the full range of organic water qualities.

## COMPARATIVE ANALYSIS OF PRECURSOR DATA

Efforts at modeling DBP precursors in watersheds can build on existing knowledge of DOC generation and cycling only if the relationship between DOC and DBP precursors is well defined. In chapter 4, attempts are made to clarify the relationship between DBP precursors and DOC concentrations and characteristics of the dissolved organic matter (DOM). New data were generated on DBP formation in a range of natural waters and extracted NOM fractions. These data were compared with similar data acquired from an extensive literature search. Appropriate comparisons required that differing conditions for precursor tests be accommodated. This was done with the help of DBP formation models, summarized from the literature. The following conclusions are based on an analysis of this large combined data set.

The formation of DBPs is heavily dependent on the characteristics of the precursor NOM. Although SUVA has been used successfully as a measure of NOM reactivity with oxidants and disinfectants, it falls short of capturing the full complexity of these reactions. Many other easily measured parameters can improve our ability to predict DBP formation. These include molecular size, partitioning behavior, and watershed characteristics.

The major reason for higher THM levels in surface waters than groundwaters is attributable to higher TOC levels. On a per-carbon basis, surface waters produce only about one-third more THMs than groundwaters. THM and HAA precursor levels are surprisingly uniform across the North American continent. The median THM precursor level in raw waters (based on the simulated distribution system protocol) is 23 µg/mg carbon (C), and the bulk of the data falls within 15 to 40 µg/mg C (10th–90th percentiles). The average THM precursor level in hydrophobic acids is nearly identical to the level for the raw water database. Similar conclusions and numerical assessments have been made for dihaloacetic acid (DHAA) and trihaloacetic acid (THAA) precursors.

Some general observations gleaned from this data set are useful in interpreting the role of watershed processes in precursor export. There are indications that specific THM formation decreases as the TOC of surface water increases. Highly colored (i.e., high SUVA) organic matter has a much higher tendency to form THAAs. Hydrophobic and hydrophilic bases produce a disproportional amount of dihaloacetonitriles. Hydrophilic acids produce a higher level of DHAAs as compared to the other DBPs.

Comparison of the North American database with samples from a single intensively studied watershed shows remarkable agreement in median values and spread of precursor densities (THM, THAA, DHAA) as well as SUVA. The fact that this study of narrow geographic and climatic scope can produce an array of data that nearly matches the full continental database has many implications. Perhaps the most important implication pertains to local versus regional factors. It seems that time of year as well as micro- and mesoscale processes (local precipitation, flow paths, riparian-zone topography, and sub-basin size) play a primary role in determining NOM character. Of lesser importance are the macroscale characteristics that only differ over broad geographic locations (e.g., ecoregion, dominant forest ecosystem, mean elevation, mean annual temperature, mean annual rainfall).



©2007 AwwaRF. All Rights Reserved.

## RELATION BETWEEN DBP FORMATION AND NOM CHARACTER

Interpreting the role of hydrologic and geographic factors in biodegradable organic matter (BOM) and DBP precursor loading requires a better understanding of NOM characteristics and its variability across different watersheds. In the case of BOM assessment, there is a need to clarify the methodologies and their implications to behavior in treatment and distribution systems. Chapter 5 begins with a review of BOM assay techniques as they pertain to drinking water. The proliferation of methods and lack of a validated biodegradable dissolved organic carbon (BDOC) protocol are perceived as challenging in the water field. Currently, assimilable organic carbon (AOC) methods appear to be more helpful in understanding BOM, even if they only assay a subset of the organic compounds. Some suggestions are offered for future research to improve the existing approaches.

A new set of data on NOM fractionation and characterization was generated with the objective of investigating some key properties germane to the general subject matter of this report. Five water samples were selected for this study, all of which were raw water supplies from major metropolitan areas: (1) Chickahominy River, (2) Assomption River, (3) Shoal Lake, (4) Hillsborough River, and (5) Quabbin Reservoir. Three of these waters are above the national average for TOC in surface waters, one is about average, and one (Quabbin) is well below average. Based on SUVA measurement, Assomption and Hillsborough waters are classified as high-SUVA water, suggesting high lignin content, which is typical for highly terrestrial and forested sources. Moderate- and low-SUVA sources are represented by the Chickahominy River, Shoal Lake, and Quabbin Reservoir waters.

Correlation between DBP precursor content and DOC is source specific, as represented by the specific DBP precursor level. The high-molecular-weight (HMW) fraction isolated from high-SUVA waters (Assomption and Hillsborough waters) shows high DBP precursor content compared to the low-molecular-weight fractions (hydrophilic [HPI] and hydrophobic [HPO]). The opposite is true for low- and moderate-SUVA waters (Quabbin and Chickahominy). Shoal Lake water shows a decreasing trend in DBP precursor content going from HPI to HMW to HPO fractions. DHAA and THM precursor levels decreased going from HPI to HPO fractions. However, the opposite is true for the fractions extracted from the more highly colored Assomption River water. THAA formation from the HPO fraction exceeded its formation from the HPI fraction. The reverse was observed for the fractions extracted from the autochthonous Shoal Lake water. In general, the HPO fractions showed low abundance of DBP precursors falling below the database median. However, the HPI fractions were quite divergent and dependent on the nature of the source water.

Both ozonation and chlorination resulted in AOC and BDOC increases of all waters and NOM fractions. Ozonation generally resulted in the highest levels, but in several cases, chlorination caused increases that were greater than those observed in the ozonated samples.

Organic fractionation methods based on size (ultrafiltration) and hydrophobicity (resin adsorption) resulted in losses at multiple stages of sample treatment. In addition, freeze-drying seemed to enhance the production of DBPs, suggesting a change in chemical structure of NOM. Decarboxylation (loss of carbon dioxide) and dehydration (loss of $H_2O$) probably occur during freeze-drying, and these are likely to increase reactivity with an electrophilic substance such as hypochlorous acid.



©2007 AwwaRF. All Rights Reserved.

## SOURCES OF PRECURSORS IN WATERSHEDS

In chapter 6, two sets of long-term experiments are presented, with the intention of quantifying the generation of DOM from leaf litter and its subsequent biodegradation. In both cases, leaves were collected and subjected to leaching tests of about 1 week in duration. Then the leachate was seeded and allowed to degrade under aerobic conditions. Samples were periodically removed for analysis of DOC, ultraviolet absorbance, biodegradability, and DBP precursor content.

These experiments showed that significant DOC concentrations were leached from maple, oak, and pine leaves within the first 24 hours. SUVA levels varied considerably among species for the initial leaching phase, suggesting that the nature of organic matter in forested watersheds is dependent on the dominant tree species. Leaching rates ranged from near zero to 24 mg C per gram of dry weight per day. Maple leaves had the highest yield, and newer leaves produced more DOC than older or previously weathered leaves.

A substantial fraction of the DOC was lost to biodegradation in the first 5 days of post-leaching aeration. Organic matter that had leached from pine needles degraded almost two times more quickly than organic matter from the other species. SUVA levels increased by two to three times over the original level after 81 days of biodegradation for all species. High levels of specific DBP formation potentials (DBPFPs) were formed for both the fresh and degraded organic matter. On average, the biodegraded organic matter formed DBP levels that were one-and-a-half to three times higher than the fresh organic matter on an equivalent carbon basis. Overall, specific THAA formation potentials were the highest and were almost two to three times higher than specific THM formation potentials. AOC and BDOC tests indicate reduced values of biologically available leachate from old leaves compared to new leaves. Although initial leaching tests offered the highest amount of assimilable carbon to the microorganisms, prolonged immersion of leaf litter produced more BDOC than short-term immersion.

Data presented in chapter 7 shed light on one prominent class of precursor molecules, lignin and its various substructures. A high-performance liquid-chromatography version of the classic cupric oxide (CuO) degradation method for lignin analysis was refined and worked well with small quantities of commercial lignin polymers and NOM isolates. Classic lignin parameters (e.g., syringyl-to-vanillyl ratios) were determined, and the values were commensurate with known origins of the commercial lignins. Total phenolic yields were 11 and 27 mg of monomer per 100 mg C. The same analysis was performed on a commercial terrestrial humic acid and an aquatic fulvic acid. The total lignin phenolic yields were 7 and 6 mg of monomer per 100 mg C, respectively. The syringyl-to-vanillyl ratios were quite high for both materials, suggesting large hardwood content with some degradative loss of vanillyl structures.

Lignin monomers were prepared in pure solution and treated in the laboratory with chlorine to determine DBP yields. All phenolic monomers were found to be reactive with aqueous chlorine, producing THMs and HAAs. Most of the syringyl, vanillyl, and *p*-hydroxy phenolics produced large amounts of trichloroacetic acid (TCAA) as compared to dichloroacetic acid and the THMs. They also tended to produce only minor amounts of other (termed “unknown”) halogenated by-products. Natural polymeric lignin products also produced THMs and HAAs following chlorination. The amount produced could be roughly estimated from the monomeric composition (as determined by the CuO degradation method) and the specific DBPFP of the monomers. In contrast, the TOX produced by lignin polymers was much greater than predicted from the analysis of monomers.



©2007 AwwaRF. All Rights Reserved.

DBP formation from lignin structures within NOM factions was estimated from two sets of model calculations, one termed "mechanistic" and one "empirical." The mechanistic approach used specific lignin monomer measurements for the NOM fractions and combined these with measured DBP yields from each monomer. The product of these values was summed over all monomers to reach an overall DBP formation from lignin-based structures. The empirical approach used two commercial lignins as models of isolated lignin-based precursors. The normalized DBP yields from these materials were adjusted for the lower lignin monomer content of the NOM fraction and compared with the measured (overall) DBP yields. Both calculations were made with and without corrections for incomplete recovery of vanillyl phenols. The median predictions for these four models have resulted in the best estimate at this time. They project a 9% role in THM formation, 12% for DHAA, and 25% for THAA based on the Suwannee River fulvic acid sample. The high THAA number suggests that watersheds dominated by vascular plants will give rise to higher levels of TCAA, such that the ratio of THAA/THM will be greater.

## MODELING OF DBP PRECURSORS AND LONG-TERM CHANGES

Short- and long-term temporal trends were used to infer some key fate and transport mechanisms (chapter 8). First, multiseason precursor data from two intensively studied locations were presented and analyzed for selected trends. THM, DHAA, and THAA data from the Wachusett Reservoir watershed and the Cambridge watershed showed consistently decreasing precursor levels with increasing TOC concentrations. Second, rain event data were analyzed for certain chemical parameters over the duration of the events. Although the initial TOC after the onset of rainfall is sometimes quite high, the first flush NOM has a low specific formation potential.

The temporal trends were interpreted within the context of the leaching study and other information in the literature. From this data, a conceptual model was proposed to describe the semi-empirical observations. First-flush NOM and NOM that dominates high-flow periods tends to be young or more recently leached from terrestrial plant material. This organic matter contains a suite of plant biochemicals that reflect availability, solubility, and rates of depolymerization or solubilization. As this material ages, it undergoes sequential biodegradation within the timescales of interest. For example, sugars are rapidly degraded, followed by cellulosic material, finally leaving the most recalcitrant biomolecules, such as lignins and tannins. These latter compounds tend to be the most productive DBP precursors. As a result, the normalized DBP formation (per unit DOC) from bulk NOM will increase as the surviving molecules are winnowed down to those most reactive with chlorine.

An illustrative quantitative model was calibrated and predictions made as a means of testing the order-of-magnitude validity of the conceptual model. This included class-specific biodegradation, DBP yield, and natural abundance in fresh DOM. Using plausible values for each, concentration estimates of major NOM classes (sugars, proteins, lignins, cellulose, hemicellulose, tannins, terpenoids, fats) versus water age were made. This allowed the calculation of DBP formation from each of the biochemical classes. The final calculations showed that increases in the specific DBP precursor levels with biodegradation time would be expected. This model also showed that lignins and especially tannins should be the major contributors to THM formation, especially after long periods of biodegradation. Proteins may be important contributors during the early stages of DOM formation, but these will be lost and terpenoids may take their place as the third most important source of THM formation. Other DBPs will follow different source patterns, which underscores the importance of considering all reactive biochemical classes.



©2007 AwwaRF. All Rights Reserved.

The large body of data on TOC in U.S. waters was mined for capturing information on the link between TOC levels and watershed characteristics, as well as to look for long-term changes (chapter 9). A database of TOC data for the United States was created, including modern (1995–2004) and historical (1965–2000) information from across the country, obtained from the U.S. Geological Survey (USGS) and U.S. Environmental Protection Agency (USEPA). Data were downloaded from 160 hydrologic unit codes that contained source watersheds for the 52 preselected utilities. A large body of additional data, scattered among numerous private utilities, local collection agencies, civic organizations, and individual researchers, is not readily available. A subset of this less readily available data was obtained through personal communication with the 52 selected utilities.

Variability between ecoregions, within ecoregions, and across seasons was also explored during the period of record. Inter-ecoregion variability was large and relatively constant on a decadal basis, with a few minor exceptions. No seasonal trends were evident, suggesting that inter-ecoregion variability is greater than interseasonal variability.

The seasonal Kendall test and Theil estimator of slope were applied to historical and modern data sets to capture both long-term and modern trends in TOC. An almost equal number of increasing and decreasing trends in TOC were observed in the modern data set. Decreases typically occurred in less populated regions and increases occurred in more populated regions of the country. Temporal trends were not consistent either within or between ecoregions, suggesting that factors other than variability in native vegetation and soils are important determinants of both TOC levels and trends.

Empirical models were developed for individual watersheds and for nationwide assessments using recently collected data (chapter 10). Most of these models were based on TOC data, as DBP precursor data are far less common. However, when available, the precursor data showed similar trends in the TOC. Regardless of analysis method, wetlands persisted across 13 watersheds (11 national and 2 individual) as a strong indicator of TOC contributions within water supply catchments. Examples were presented from eastern Massachusetts and coastal New Jersey illustrating this general trend. The empirical models from the national survey also described the strong influence of wetlands on TOC concentration. This land use type was dominant in each series of relationships between land use/land cover (LULC) and TOC.

Anthropogenic activity was retained as a significant predictor within the national empirical models and was represented by such LULC characteristics as urban development, industry, and residential areas. The national seasonal empirical models incorporated anthropogenic activity during the summer and fall seasons as well as within the overall model. The illustrative single watershed models also reflected seasonality to anthropogenic contributions. This was accompanied by a more generalized seasonal affect on TOC in the models from the 13 watersheds. When integrated across the United States, seasonal effects seemed to diminish.

A mechanistic model was developed for one of the test watersheds, the Metedeconk River system (chapter 11). It was a simple input–output model accounting for travel time and first-order reaction kinetics. The governing equations of the model were mass and flow balances, and did not explicitly include the effects of seasonality, precipitation, topography, and groundwater flow. Precipitation and groundwater flow indirectly affected the river discharge included in the model, whereas topography indirectly affected time of transport, also included in the model. During validation, a 20% total error associated with the Metedeconk River system was found. This error was attributed to underlying hydrologic and TOC variability not captured by the model. Loss of variability was due in part to temporal smoothing (monthly averages) and lack of data (only biweekly



©2007 AwwaRF. All Rights Reserved.

measurements), resulting in a poor representation of certain characteristics and governing processes within the watershed, such as precipitation and snowmelt. It is evident that a more mechanistic model that incorporates these parameters may be necessary to improve model performance.

## RECOMMENDATIONS

Wetlands and urban areas contribute a disproportionately large amount of NOM and DBP precursors to surface waters. Hydraulic connectivity between forested wetlands and raw water sources should be minimized to the extent possible.

Future watershed studies should focus on microscale processes. It is important to understand and document changes in NOM and precursor flux with higher spatial and temporal resolution. This must include careful analysis of hydrologic and precipitation data.

Utilities in North America should plan for the possibility that TOC and DBP precursor levels in raw waters may increase because of long-term changes in land use and climate. Collection of all historical and modern TOC data should be undertaken for more extensive investigation of long-term trends. This should include the full USGS and USEPA databases, as well as the remaining data scattered among numerous local agencies, civic organizations, drinking water utilities, and research projects (published and gray literature).

The role of tannins in THM and HAA formation should be explored. These compounds are expected to be major contributors in NOM from forested watersheds. Attempts should be made to identify plant-based sources, ultimately linking dominant flora to DBP export.



©2007 AwwaRF. All Rights Reserved.

©2007 AwwaRF. All Rights Reserved.

# CHAPTER 1
# INTRODUCTION

All drinking water supplies contain some background organic matter. Usually these organic compounds are from naturally occurring plant products or their derivatives. Thus, this material is called natural organic matter (NOM). Until the early 1970s, NOM was considered a benign nuisance. NOM was known to consume chemical additives or foul granular activated carbon (GAC) adsorbents, but it was not associated with any known threats to human health.

In the early 1970s, NOM became a focus of concern when its central role in the formation of potentially carcinogenic disinfection by-products (DBPs) was recognized. DBPs, including trihalomethanes (THMs) and haloacetic acids (HAAs), resulted from the reaction of chlorine ($Cl_2$) with NOM. Alternative disinfectants were explored, but these also produced potentially harmful DBPs.

Starting in the late 1970s, regulations were established by the US Environmental Protection Agency (USEPA) with the intention of minimizing DBP exposure. These regulations accompanied extensive occurrence studies (e.g., the Information Collection Rule, or ICR) and stimulated the development of in-plant control strategies. Though somewhat successful, these strategies have not completely solved the DBP problem. In addition, it is recognized that NOM contributes to the growth of bacteria in distribution systems. Biodegradable dissolved organic matter (BDOM) in raw waters is supplemented by newly formed BDOM in treated waters to provide stimulus for microbial growth. Different watersheds are known to result in waters with different levels of BDOM; however, the factors that lead to these differences are not understood (e.g., Volk and LeChevallier 2002). Clearly, more attention must be directed to the upstream processes that give rise to NOM in raw waters.

## NOM TYPES AND GENESIS

NOM differs markedly in concentration and quality from one raw water source to another. Specific ultraviolet absorbance (SUVA) is a qualitative parameter that has been correlated with chemical reactivity and the tendency to form DBPs. Total organic carbon (TOC) is one of the most widely used measures for quantifying the amount of NOM in water. These two parameters have been observed to be positively correlated, although the correlation is not extremely strong. Looking at just two examples, Tulsa (Oklahoma) and Winnipeg (Canada), the following questions can be asked: Why are the TOC and SUVA characteristics so variable? What differences in the watersheds lead to these contrasting characteristics, and can such effects be predicted? Unfortunately, definitive answers to these questions have not been ascertained.

One of the key issues in watershed management for NOM control is an understanding of the origin and genesis of aquatic organic matter. In many cases the discrimination between those organics termed “allochthonous” versus those called “autochthonous” is the goal. Some have framed this in the context of aquagenic versus pedogenic organic matter. In either case, the underlying goal is worth pursuing, although the articulation may be overly simplistic. Researchers agree that humic substances are a complex mixture of plant and bacterial products which have been partly degraded and then recombined through a mixture of covalent bonding (e.g., by nearly random, nonenzymatic reactions; MacCarthy 2001) and weaker associations supported by hydrophobic forces and hydrogen bonding (e.g., Sutton and Sposito 2005). On average, these humic



©2007 AwwaRF. All Rights Reserved.

substances are quite refractory, a property that has been attributed to nonenzymatic processes. They also tend to be the single most important reservoirs of the THM and HAA precursors.

In contrast, the nonhumics are largely natural plant products that may have undergone some degree of biodegradation or biosynthesis. One newly recognized and prominent group of nonhumics is the acylheteropolysaccharides (APS; Benner et al. 1992). These are high-molecular-weight (>1,000 D), algal-derived polymers of simple sugars, acetate, and lipids. They are recalcitrant, much like humic substances. Recent reports have placed this material at 10%–35% of the dissolved organic carbon (DOC) in lakes and rivers (Repeta et al. 2002). The APS pool probably constitutes the majority of what has often been characterized as the hydrophilic neutral fraction of NOM. This fraction is abundant in large reservoirs, and it is known to have a substantial THM and HAA precursor content, albeit somewhat lower than the humic fractions (Reckhow et al. 1999). This fraction has also been shown to be high in biodegradable dissolved organic carbon (BDOC; Reckhow, Tobiason, and Edzwald 1999).

Much of the remaining NOM is thought to be among the simple, identifiable plant products. Included in these compounds are carbohydrates, nucleic acids, proteins, lipids, and their various building blocks (e.g., sugars, nucleotides, amino acids, fatty acids), as well as some other prominent compounds (e.g., phenolics, terpenoids, alkaloids). This diverse group includes rapidly biodegradable substances, as well as those requiring a relatively slow prehydrolysis step. The labile nature of the SPP (simple plant products) fraction is most easily seen in organic material leached from leaf litter. This material will rapidly biodegrade in aerobic systems, whether in a laboratory or in a natural surface water. If given enough time, the carbon in these compounds will become completely mineralized, having incorporated into new cell biomass or abiotically transformed to humic substances (Figure 1.1). This helps to explain the decreasing biodegradability of NOM, sometimes in the face of increasing DOC, as water moves downstream in many fluvial systems (e.g., Leff and Meyer 1991). The SPP fraction is probably responsible for the majority of the BDOC measured in raw drinking waters. Some researchers have divided the BDOC into subfractions that are immediately biodegradable (S) and those requiring prior hydrolysis at varying rates (H1 and H2) (SANCHO model; Servais, Billen, and Bouillot 1994).

Numerous studies of organic carbon in surface waters have demonstrated the importance of algal sources (autochthonous) as well as those from terrestrial plants (e.g., Wetzel 2001). Walker (1983) illustrated the strong empirical relationship between reservoir phosphorus level and TOC. He proposed that THM precursors would behave similarly. However, others have disputed the wisdom of directly linking TOC with DBP precursors (Stepczuk et al. 1998b). Nevertheless, a body of evidence exists that supports algae as important sources of precursors in some reservoirs (e.g., Karimi and Singer 1991, Plummer and Edzwald 2001).

## BIODEGRADABILITY TESTS FOR NOM

The development of tests for assessing biodegradability in raw and treated drinking waters has been a slow and continually evolving process (Huck, Fedorak, and Anderson 1990; Frias, Ribas, and Lucena 1995). The most common type of test is probably the family of protocols that is referred to as the BDOC test. Samples are inoculated with what is usually a natural consortium of bacteria. The sample is then allowed to sit under aerobic conditions for a period of hours to weeks. DOC is measured before and after incubation. The difference represents the amount of



©2007 AwwaRF. All Rights Reserved.

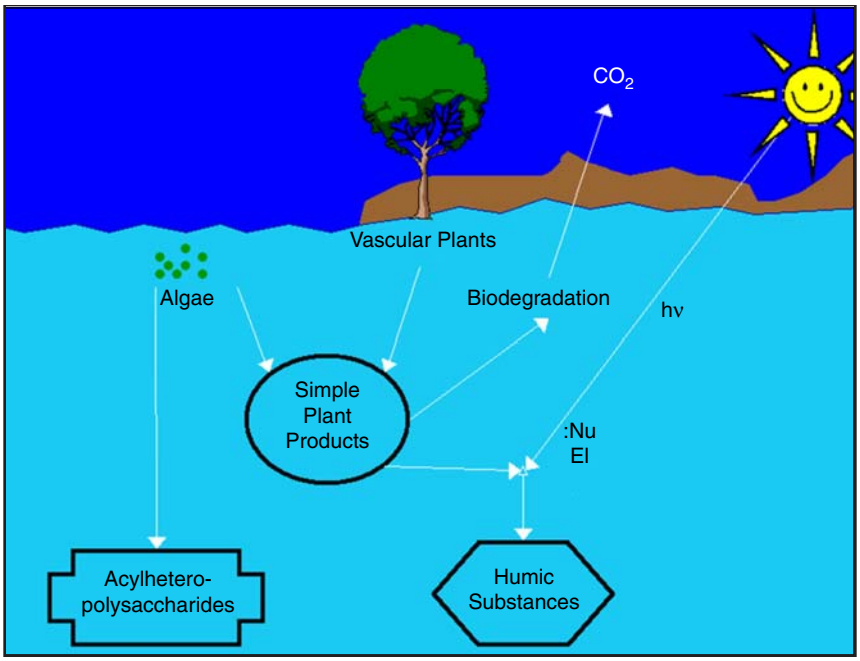



**Figure 1.1 Three major pools of aquatic NOM**

carbon that was mineralized or incorporated into the cells. This amount of carbon (calculated in milligrams per liter) is called the BDOC.

Many variations of this test exist, and most involve the nature of the inoculum (attached or suspended growth), the type of solid support if used (sand, sintered glass), the length of time, and the type of sample/bacterial contact (shaking bottle, recirculating column) (e.g., Mogren, Scarpino, and Summers 1990; Volk et al. 1994; Allgeier et al. 1996).

Another widely used test is the assimilable organic carbon (AOC) assay (Jago 1993; Kaplan, Bott, and Reasoner 1993; LeChevallier et al. 1993; Frias, Ribas, and Lucena 1994). This test is described in the widely accepted *Standard Methods for the Examination of Water and Wastewater* (APHA, AWWA, and WEF 1998). Here a sample is inoculated with one or two pure strains of bacteria. Bacterial growth is then monitored over a period of several days by enumeration techniques or adenosine triphosphate (ATP) analysis. The growth of bacteria is the critical measurement, and this is related to the growth expected for a standard acetate or oxalate solution. Therefore, the AOC values are given in micrograms of acetate carbon per liter.

Similar approaches have been taken by research on an ad hoc basis. For example, some have developed composite natural consortia from sites of interest and used these as inocula (Leff and Meyer 1991). These authors then measured the peak biomass, which was usually observed on day 3.

Other approaches to assessing biodegradability include the use of chemical surrogates. Because ozonation is known to enhance the biodegradability of a water as well as form a wide array of carbonyl compounds (e.g., aldehydes, ketones, keto-acids—compounds that are themselves highly biodegradable), analysis of the carbonyl by-products is sometimes used as a gauge of biodegradability. This is usually done only for ozonated waters.



©2007 AwwaRF. All Rights Reserved.

## NOM MODELING AND MANAGEMENT

Management of watersheds and surface waters for the purpose of NOM control requires some tangible knowledge of the impacts of human-made changes in the watershed on NOM quantity and quality. This knowledge can come from experience in similar watersheds, and the tools to translate this knowledge are the watershed models.

Watershed models may range from the highly empirical to the boldly mechanistic. Many limnologists, aquatic biologists, and geochemists have studied the origins and export of organic carbon from watersheds. Some excellent mechanistic research on TOC or DOC in catchments, rivers, and lakes has been published by large research groups (e.g., by G.E. Likens from the Institute of Ecosystem Studies in Millbrook, N.Y., and H.F. Hemond from the Massachusetts Institute of Technology). However, comparatively little has been done on specific subfractions of the NOM, such as the DBP precursors.

Dissolved organic matter (DOM) is primarily generated from terrestrial and aquatic plants. This material is transported through the upper soil horizons, shallow groundwater, deep groundwater, and by overland flow. During transport, portions of the DOM are attenuated because of biodegradation, sorption, photolysis, and chemical precipitation. Other processes can be important but were not explicitly studied in this work. Nevertheless, during watershed-scale modeling, DOM can rapidly adsorb to fine suspended sediments such as clay, organic, carbonate, or hydroxide metal oxide particles, resulting in the sedimentation of DOC to littoral zones and stream beds (Wetzel 2001, Steinberg and Muenster 1985, Meyer 1990). All of these processes have been studied by limnologists and geochemists, and emerging from this work is a good, generic understanding of DOC cycling in the environment. Far less work has been performed on understanding the movement of specific components of the DOC, and in particular, very little has been done on DBP precursors.

More than a decade ago, Cooke and Carlson (1989) prepared an Awwa Research Foundation (AwwaRF)-sponsored manual to help water quality managers control THM precursor levels in source waters. At the time, very little was known about the origin, transformation, and transport of these compounds. As a result, their work reads more like a general discussion of “good” watershed and reservoir management practices as they pertain to erosion and eutrophication control.

Even now, knowledge of watershed precursor control remains rudimentary. NOM management is currently viewed as an exercise in discouraging the release of plant products into the aquatic environment and encouraging biodegradation of that matter when it does enter the water. To achieve the former, it is critical to know the origin of the plant products. How much comes from terrestrial plants, and how much comes from aquatic plants (especially algae)? The distinction is important because of the management approaches that must be taken in dealing with each condition. Many have attempted to use existing analytical tools to help discriminate between organics from these two sources. Most NOM characterization methods have not been very successful in this regard (e.g., Christman et al. 1998). Those showing promise focus on certain elemental or structural signatures that are enriched in one source versus another.

High nitrogen (N) or amine content seems to be one indicator of an algal source. Formation of acetonitriles following chlorination has proven to be an especially sensitive means of detecting these N-compounds. In contrast to the algal organic matter, terrestrial organics seem to be enriched in phenolic structures, reminiscent of lignin monomers. These structures have been demonstrated from cupric oxide (CuO) oxidation studies by the research groups of Ertel and Christman (e.g., Ertel, Hedges, and Perdue 1984; Liao et al. 1982).



©2007 AwwaRF. All Rights Reserved.

Stringent DBP regulations in the United States are driving utilities to investigate ways of minimizing by-product formation starting from the intake and ending at the consumer's tap. Concerns about bacterial regrowth have extended this to biodegradable organic constituents (e.g., BDOM). Comparatively little effort has been made in minimizing precursor or BDOM formation in the watershed and transport to the intake. At the same time, many communities are faced with population growth and economic pressures that can profoundly impact land use and other activities in watersheds. A careful reexamination of watershed processes in DBP precursor and BDOM control is long overdue.

The purpose of this research was to explore the relationship between the character of NOM, its tendency to form DBPs, and the characteristics of the watersheds giving rise to that organic matter. Existing data from utilities on their source water NOM are discussed in chapter 3. This is followed by a quantitative review of precursor content in natural waters and NOM fractions (chapter 4). Additional data on NOM fractions, treatment-related characterizations, and impacts on precursor content are presented in chapter 5. DOM generation and biodegradation are discussed at length in chapter 6. This focuses on natural plant material with its full spectrum of biochemical residues. Chapter 7 exclusively focuses on lignin, with an orientation toward chemical models. These findings are then summarized in the first chapter on modeling, in the form of a semiquantitative conceptual model (chapter 8). The question of long-term changes in DOC is handled in chapter 9. The second modeling chapter takes a rather empirical approach (chapter 10), and this is followed by a more mechanistic treatment (chapter 11). The report ends with brief chapters on conclusions (chapter 12) and recommendations (chapter 13).



©2007 AwwaRF. All Rights Reserved.

©2007 AwwaRF. All Rights Reserved.

# CHAPTER 2
# MATERIALS AND METHODS

## RESEARCH OBJECTIVES

Simply stated, the research objectives of this study are to investigate sources, nature, and long-term variability of NOM and BDOM in source waters and to evaluate their impact on the formation of DBPs.

## GENERAL APPROACH

This work encompassed data collection, data generation, and a review of the literature. It builds on some of the latest fundamental advancements in NOM characterization and origins. In the first phase of this study, an extensive literature review was conducted on DBP precursors and BDOM in source waters. In addition, a survey of larger utilities was undertaken with the objective of finding long-term data sets on precursors and biodegradables in the source water.

Based on this early work, about a dozen utilities were identified for inclusion in an in-depth watershed analysis. These utilities were asked to share their raw water data, as well as a broad range of ancillary information. Bulk raw water samples were shipped to the University of Massachusetts (UMass) for advanced characterization tests. In some cases, many samples were collected at different points within a watershed. At UMass, a suite of organic analyses were conducted (e.g., TOC, DOC, ultraviolet [UV] absorbance, carbon-hydrogen-oxygen-nitrogen [CHON] and precursor analyses), along with biodegradability tests, and a detailed set of NOM fractionation and characterization tests. This information, along with the utility-supplied data, was used to help develop a set of NOM watershed models. These models make use of modern geographic information system (GIS)-based data, and span the range from empirical to semi-mechanistic. The goal of the modeling effort was to provide insight into NOM sources, degradation processes, NOM characteristics, and impacts of land use and management practices on NOM.

## SYNOPSIS OF PROJECT

### Utility Survey

The utility survey (appendix A) concerns the collection of data regarding long-term trends in DBP precursor concentrations in U.S. drinking water supplies. Long-term trends are very poorly understood. They may occur because of gradual changes in land use, point discharges, climate, or flora. Although rudimentary mathematical models exist that are intended to describe DBP levels as a function of changing watershed characteristics, this field is still in its infancy. None of these models has been adequately tested. As a result, very little is known regarding this issue. To understand how changes in a watershed can translate into changes in DBP levels requires the existence of long-term data on those concentrations.

A large group of candidate utilities was identified for dissemination of a preliminary written questionnaire. These utilities included all those currently serving more than 50,000 people that also use a surface water source. Some smaller utilities were added to this list based on the existence of published DBP data prior to 1985. It is expected that the utilities which published early data of this



©2007 AwwaRF. All Rights Reserved.

type were either strongly interested and concerned about their DBP levels or had collaborated with a researcher (e.g., someone from a local university) with long-term interests in the DBP issue.

Concurrent with the identification of candidate utilities, development of a brief written survey instrument was begun. The instrument was designed to acquire information about the following items:

- Drinking water sources
- Current DBP precursor levels in the source(s)
- Length of time during which DBP precursor measurements or surrogate measurements have been collected
- Types of previously-mentioned measurements and approximate frequencies
- Possible studies that have been or are being conducted to examine DBP and NOM sources in the watershed(s)
- Interest in issues addressed by DBP and NOM studies of this type

The survey instrument was formed into a small booklet (i.e., four half-size pages) and formatted to be inviting and user friendly. A single-page cover letter that described the survey's purpose and a stamped, self-addressed business-reply envelope were included. The recipient was allowed to respond via the Internet, making use of e-mail and a project Web address. Based on initial testing with a few utilities, the survey instrument was refined and finalized for mailing to all remaining candidate utilities. After 2 weeks, those recipients who had not responded were sent a followup postcard reminder. Those who still had not responded after 4 weeks were sent a second set of survey material with a new cover letter. These methods, which were intended to enhance the survey response rate, are based on the Dillman (1978) method. In the past, this procedure has routinely yielded response rates of 60% to 80% (Dillman 1978; Ditton, Loomis, and Choi 1992; Loomis 1996).

Next, a smaller group of utilities was selected from the initial group for followup activities based on the potential to acquire long-term DBP precursor data. These utilities were identified from respondents to the mailed survey, and all nonrespondents who were identified from the literature survey were included. The followup survey (appendix B) was largely conducted by telephone using UMass environmental engineering graduate students, faculty, and UMass staff. In this phase of the survey, attempts were made to acquire followup information about

- Size and nature of DBP precursor (or surrogate) database
- Format and accessibility of that database
- Types of ancillary data on raw water quality
- Special studies conducted in the watershed that influence NOM
- General information on watershed and known sources of NOM
- Interest in participating in this project

At this point, a selected group of utilities was asked to submit a synoptic survey of its watershed DBP precursors. Each utility was provided with a written set of directions that described collection of about a dozen samples from key points in the utility's watershed. Then the utilities were asked to analyze data at their own expense and provide the data to UMass. In return, the project investigators would offer their interpretations of these data to the utilities.



©2007 AwwaRF. All Rights Reserved.

## Study of Selected Watersheds

In the course of some in-depth study of the selected watersheds for DBP precursors, some research questions evolved:

- What is the nature of the NOM origin? Algal versus terrestrial?
- How much geographical variability is there?
- How does the DBP precursor and BDOM (or BDOM precursor) level depend on land use, best management practices, general watershed management, climate, and so forth?
- What type of watershed models are most useful for predicting or describing the DBP precursor or BDOM level in a source water?
- Are there simple surrogates that can be used instead?

### *Final Selection of Utilities*

Based on a careful analysis of the survey data, a group of utilities was selected to advance for subsequent studies. This group was expected to encompass about a dozen systems. The criteria used to make this selection included

- Long-term data set
- Good data quality
- Representative data
- Ancillary data
- Observable trends
- Utility's willingness to participate

### *Collection of Relevant Watershed/Utility Data*

An extensive set of data was collected from each of the selected utilities. Every effort was made to collect data by telephone, mail, fax, and e-mail to avoid costly and time-consuming site visits. The types of data requested include key water quality parameters relating to NOM, physical data on the watershed, and important hydrologic data. This information was requested for the current year, and for as many prior years as possible or as deemed to be helpful.

### *Site-Specific Data Analysis and Modeling*

After the data were assembled at UMass, the process of data analysis began. Data analysis was used to support empirical correlation models. These models involved the analysis of statistical correlations between the parameters of interest (e.g., precursor concentrations) and other more easily measured parameters. Although these models are not designed to explore cause and effect, they can be quite useful. Most of the modeling effort concerned export coefficient-type formulations based on land use. Some more mechanistic long-term models were explored as constructed from the literature, and from experience and data acquired over the course of this study. Such long-term modeling would be helpful from both a management point of view as well as for advancing the basic understanding of NOM processes.



©2007 AwwaRF. All Rights Reserved.

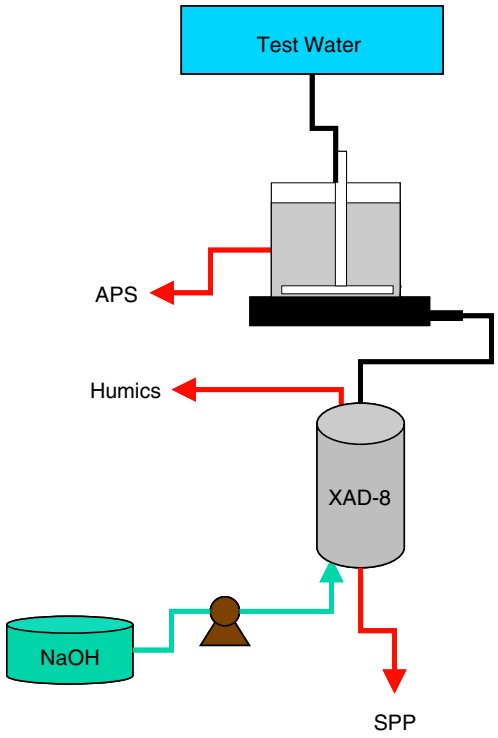



**Figure 2.1 Experimental scheme for isolating the three major pools of NOM**

### *Sample Analysis*

This work was proposed for filling in key gaps in understanding these systems regarding NOM origin and reactivity. As presented in the introduction, recent advances have helped to clarify the nature of NOM as it pertains to origin, biodegradability, and DBP formation. As a part of this effort, a new approach to NOM classification was proposed and then applied to a broad range of waters.

Bulk samples of about 40 L each were collected from each of the study sites and shipped overnight to the UMass laboratories. Collection sites were selected to obtain a representative sample of the raw water and, if appropriate to the modeling effort, one or more of the major tributaries or point discharges. These waters were immediately disbursed into a set of representative subsamples and kept refrigerated until needed. The largest portion of the sample volume was devoted to a preparative fractionation procedure intended to isolate the three major NOM pools. Two pools were brought to dryness (hydrophobics or "humics," and high molecular weight [MW] or "APS"). The third pool (hydrophilics or "SPP") was kept in its neutral dissolved form and modestly concentrated by vacuum rotary evaporation.

The fractionation scheme is illustrated in Figure 2.1. Samples were acidified and applied to an XAD-8 (Rohm and Haas Company, Philadelphia, Pa.) resin column. After passing the entire volume through the column, the retained organic matter was eluted in sodium hydroxide (NaOH). This is the classical method for isolating aquatic humic and fulvic acids.

Three NOM pools exist in the conceptual model, and it is proposed that these may be largely separated by the proposed resin and membrane method (Table 2.1). The two sets of terms (conceptual and operational) are roughly equivalent but are not exactly the same.



©2007 AwwaRF. All Rights Reserved.

**Table 2.1**
**NOM fractions: Operational versus conceptual nomenclature**

| Conceptual fraction | Operational fraction |
|---|---|
| APS | High molecular weight |
| Humics | Hydrophobic |
| SPP | Hydrophilic |

**Table 2.2**
**Testing of whole waters and NOM fractions**

| Pretreatments | Tests | Notes | Samples |
|---|---|---|---|
| None | DOC/TOC | NA* | All |
| | UV-visible absorbance | NA | All |
| | CHON analysis | NA | Humics and APS |
| | CuO degradation and gas chromatography/mass spectrometry | Lignin signatures | Humics and APS |
| | BDOC and AOC | Raw water biodegradability | All |
| Chlorination (medium dose); short term | THMs, HAAs, haloacetonitriles (HANs), and total organic halide (TOX) | Fast raw water precursors | All |
| Chlorination (medium dose); long term | THMs, HAAs, HANs, and TOX | Slow raw water precursors | All |
| Chlorination (low dose) | BDOC and AOC | Treated water ($Cl_2$) biodegradability | All |
| Ozonation | BDOC and AOC | Treated water (ozone, or $O_3$) biodegradability | All |

*NA = not applicable.

When fractionation was completed, a large battery of analytical tests was conducted on each of the isolates. Table 2.2 summarizes this testing. Important features included analysis of lignin signatures, which was performed in part to verify the effectiveness of the fractionation scheme in separating these compounds and to provide some additional characterization. In particular, the strength of the lignin signal in the humics fraction was intended to help with assessing the importance of nonalgal sources to this material. Qualitative characteristics from these tests may also help in discriminating the relative importance of gymnosperm versus angiosperm in its genesis.

Additional testing included biodegradability assays, where measurements of BDOC and AOC were used, which are the main indicators of biological stability and BDOM of raw and treated drinking waters. The biological potential for degradation of DOC was determined by analyzing AOC and BDOC for all untreated test waters, as well as for each of the three major



©2007 AwwaRF. All Rights Reserved.

NOM pools after their fractionation (high molecular weight [HMW], hydrophobic [HPO], hydrophilic [HPI]). This testing was extended to fractions treated with chlorine and ozone ($O_3$). The purpose of ozone treatment was to provide an indication of the presence of compounds that may not in and of themselves be biodegradable but could react in treatment systems to form biodegradable substances. This is analogous to the precursor tests for DBPs.

Finally, all samples were analyzed for a general set of bulk organic measurements (e.g., DOC and UV-visible absorbance), as well as a battery of comprehensive precursor tests. The latter included four pretreatment scenarios followed by analysis of THMs, HAAs, haloacetonitriles (HANs), and total organic halide (TOX). In an effort to ascertain some kinetic information, two identical aliquots of these samples—one set after a reaction time of 4 hours and the other after 3 days—were chlorinated and analyzed. These are referred to as short-term and long-term precursor tests.

### General Assessment and Modeling

#### *Development of a Database for NOM and DBP Precursors*

This research project offers a prime opportunity to initiate the development of a comprehensive North American database on NOM in raw waters. Such a database should include information on watershed characteristics, NOM quantity, and NOM quality. Project data were entered into a Microsoft Access database (Microsoft Corporation, Redmond, Wash.). The database design was kept simple so that someone with only minimal knowledge of the software could use it. It is also hoped that utilities and researchers will be encouraged by AwwaRF to submit future data so that this database remains up to date.

#### *Empirical Models for NOM and Precursors*

Individual site-specific models were developed during this research. On this foundation, attempts were made to develop some more-general tools (i.e., models) for use by water quality managers concerned with other watersheds. Most of the individual short-term models are intended to provide data on the relationship between land use and management practices with precursor and BDOM export.

## DETAILED EXPERIMENTAL METHODS

### Laboratory Treatments

#### *Chlorination Procedures*

Samples intended for precursor analysis were treated with chlorine in the UMass laboratory. Briefly, samples were buffered at pH 7 and then tested for chlorine demand. Buffered samples were then treated with a chlorine dose that leaves a residual of at least 1 mg/L at the end of the incubation period. Medium dose tests, intended to assess precursor levels, used a fixed dose of 20 mg/L. Reactions were conducted in headspace-free glass vessels (chlorine demand free) in the dark at a fixed temperature (20°C). For the chlorination studies, two reaction times were used: 4 hours and 3 days. This was intended to simulate the residence time in a finished water clearwell



©2007 AwwaRF. All Rights Reserved.

and a distant point in a large distribution system, respectively. In addition, these times were carefully selected to capture some of the most important kinetic information on DBP reactions while avoiding the expense of a full kinetic study (usually 10 or more reaction times).

### *Ozonation Procedures*

When required, samples were ozonated in a semibatch system. Ozone was generated from pure oxygen by means of a laboratory corona discharge generator. The ozone/oxygen product gas was introduced into a 2-L glass reaction vessel containing the water to be treated. Flow was controlled with an electronic flow controller, and the ozone content was monitored by direct UV absorbance spectrophotometry. The gas was mixed with the sample by a porous quartz frit. Off-gas was redirected through a spectrophotometer for determination of ozone content. A membrane ozone electrode (Orbisphere) was fitted into the side of the glass reactor so that aqueous ozone concentration could be continuously monitored. Ozone transferred was determined from the flow rates and the differences in ozone content in the applied gas versus the off-gas. For the purpose of this study, a fixed ozone/TOC ratio was used that resulted in an ozone exposure (concentration times time) representative of a system using ozone for primary disinfection.

### *Ultrafiltration*

Ultrafiltration (UF) was used for isolating the APS (HMW) fraction. All samples were also prefiltered with a 0.2-µm polycarbonate cartridge for removing bacterial and small particles. Membranes rated at 3 kD (kilodaltons) were used. This approach has been used to fully isolate the APS fraction of NOM.

## Chemical and Biological Analyses

### *Total Organic Carbon*

TOC was measured on nearly all samples in this research by the high-temperature combustion method (APHA, AWWA, and WEF 1998). A Shimadzu 5000 instrument (Shimadzu Scientific Instruments, Columbia, Md.) was used for these measurements at UMass.

### *Residual Chlorine (Free and Combined)*

Residual chlorine was measured by titrimetric *N*,*N*-diethyl-*p*-phenylenediamine (DPD) methodology (4500-Cl, D, and F; APHA, AWWA, and WEF 1998). Residual chlorine species was measured on all samples collected for DBP analysis.

### *THMs and Other Neutral Extractables*

The liquid–liquid extraction, gas chromatography procedure for the analysis of THM in this research is closely aligned with USEPA method 551.1. The THM standard stock solution was prepared from USEPA halogenated volatiles mix stock solution (Supelco Inc., Bellefonte, Pa.). Pentane was used as the extracting solvent and 1,2-dibromopropane was used as an internal standard. This method was also used for measurement of HANs (useful indicators of algal origins of NOM).



©2007 AwwaRF. All Rights Reserved.

### *Haloacetic Acids*

HAA analysis was performed in accordance with USEPA method 552.2. The HAA standard stock solution was prepared in MTBE (methyl-*tert*-butyl-ether) from USEPA Halogenated Acetic Acids Mix (Supelco Inc., Bellefonte, Pa.), which consisted of monochloroacetic acid (MCAA), dichloroacetic acid (DCAA), trichloroacetic acid (TCAA), monobromoacetic acid (MBAA), dibromoacetic acid (DBAA), and bromochloroacetic acid (BCAA), and from separate standards of bromodichloroacetic acid (BDCAA), chlorodibromoacetic acid (CDBAA), and tribromoacetic acid (TBAA) (Supelco). MTBE was used as the extracting solvent and 1,2,3-trichloropropane was used as an internal standard. All stock solutions were stored at –4°C. THM and HAA analyses were performed on an HP 5890A series II gas chromatograph (Agilent Corporation, Palo Alto, Calif.) with an electron capture detector. At the beginning of each analytical run, solvent blanks and solvent samples containing the internal standard were injected to condition the gas chromatograph and to verify that interferences were absent.

### *Total Organic Halide*

TOX was analyzed by the adsorption-pyrolysis-titrimetric method using a TOX analyzer (Euroglas EFU 1700; Delft, the Netherlands). Methodology generally follows that established in *Standard Methods* (APHA, AWWA, and WEF 1998). Analyses were conducted on 50 mL of solution that passed through two carbon columns to separate the dissolved organic material by adsorption onto the carbon. The carbon columns were then washed with 15 mL of 1,000 mg/L nitrate solution. Inorganic halides present on the activated carbon were removed by competitive displacement by nitrate ions. The activated carbon with organic halides was then pyrolyzed in a furnace at 1,000°C, and the formed hydrogen halides were transported to a microcoulometric titration cell containing 75% acetic acid and silver electrodes, using ultrahigh-purity oxygen as the carrier gas. The amount of hydrogen halides was quantified by measuring the current produced by silver-ion precipitation of the halides. The results were expressed as micrograms of chlorine per liter.

### *Specific and Delta UV Absorbance*

The full UV-visible absorbance spectrum was measured for all waters prior to treatment with disinfectants. UV spectroscopy has been extensively used in studying humic substances. SUVA at 254 nm is widely used to assess the humic content of NOM. All absorbance measurements were made at UMass on a Hewlett-Packard 8452 diode array spectrophotometer or a ThermoElectron Genesys 10-UV spectrophotometer (ThermoElectron Corporation, Waltham, Mass.).

### *CuO Oxidation and Product Analysis*

Oxidative degradation methods have been used along with gas chromatography–mass spectrometry (GC–MS) for the characterization of NOM since the early 1970s. Though many different oxidants have proven successful in preserving structural features in degraded NOM, CuO oxidation has probably been the most useful (Christman et al. 1983; Ertel, Hedges, and Perdue 1984; Hautala, Peuravuori, and Pihlaja 1997, 1998; Hyotylainen et al. 1998; Liao et al. 1982; Louchouarn, Opsahl, and Benner 2000). Using this technique, researchers from both Ertel's and Christman's laboratories have clearly identified a range of lignin-based structures in aquatic



©2007 AwwaRF. All Rights Reserved.

NOM. Cupric oxide methods are mild and have been reported to preserve 25%–75% of such lignin structures in environmental samples.

The alkaline CuO conditions employed by both of these research groups were used. The degradation products were then analyzed by liquid chromatography–mass spectrometry (LC–MS) in accordance with a subsequent modification of the standard protocols (Hedges and Ertel 1982).

### *Analysis of BDOM*

In this research, BDOC was determined using a modification of the method described by Servais, Anzil, and Ventresque (1989), and AOC was determined using a method based on that described in *Standard Methods* (APHA, AWWA, and WEF 1998) with slight modifications introduced from the method proposed by van der Kooij, Visser, and Hijnen (1982).

The biological potential for degradation of DOC was determined by analyzing bacterial numbers and bacterial viability for all untreated test waters, as well as for each of the three major NOM pools after their fractionation (HMW, HPO, and HPI, reflecting APS, humics, and SPP, respectively), and the remaining organic matter. The number of bacterial cells was monitored by both direct counting techniques and plating techniques. Direct counting was combined with the determination of the ratio of viable versus dead cells, which is based on epifluorescence microscopy and membrane-selective differential fluorescent dyes. For plating tests, media selective for oligotrophic microorganisms were used, as electron-donating substrates in biologically stable drinking waters are characterized to be present at low concentrations. Incubation conditions were aerobic, which mimics the common concentration of dissolved oxygen in drinking water.

Both BDOC and AOC measurements provide trends about changes in biological stability of drinking water (Escobar and Randall 2001), but they are not perfectly quantitative for determining total BDOC (Woolschlager and Rittmann 1995). During incubation, cells form soluble microbial products and bacteria might undergo endogenous decay. Thus, internal control experiments were added in which the use of a complex mixture of added DOC of known concentration was tested.



©2007 AwwaRF. All Rights Reserved.

©2007 AwwaRF. All Rights Reserved.

# CHAPTER 3
# DATA COLLECTION FROM UTILITIES

This chapter discusses the initial efforts to learn about the availability of data on NOM that might not be found in the open literature but is available from utilities and associated groups. As part of this effort, data were collected about general interest in NOM and DBP precursor assessment and modeling on the part of the utilities.

## MAIL SURVEY

### Methods

A two-page questionnaire was designed to gather information on the types of data that utilities had on BDOM and DBP precursors in their source waters, how many years they had been collecting those data, and whether they would be willing to share this information with the university. On Oct. 3, 2003, the questionnaire was mailed to 553 utilities throughout the United States. These facilities were known to use a surface water source and to serve more than 50,000 people. A cover letter accompanied the questionnaire, which explained the purpose of the survey, ensured the confidentiality of the respondent, and encouraged respondents to contact the principal investigator with questions. A second survey and cover letter was mailed to nonrespondents on Oct. 29, 2003. On November 18, a third survey was mailed to those who had not responded, with a cover letter that further encouraged their participation.

### Response Rate

Formal written surveys were mailed to 553 of the largest U.S. utilities that use surface water. The mailing list was obtained from USEPA, which was based on the following selection criteria:

- Serving a population of 50,000 or greater
- Surface water source

Completed surveys were returned by 249 respondents, representing a return rate of 45%. Many who did not return the surveys telephoned or e-mailed with explanations as to why they were not responding. The stated reasons for not responding, in order of decreasing frequency, were as follows:

- We don’t use surface water.
- We purchase all of our water from another supplier, and therefore don’t have our own watershed.
- We don’t respond to surveys.

When contacted directly, none of the utilities expressed concern about confidentiality of the data. As a result, specific utilities are referenced in a few isolated cases.



©2007 AwwaRF. All Rights Reserved.

## Results

Results of a question-by-question statistical summary are included in this section. Short discussions accompany some of the results. The responses to question 3 were often complicated and site-specific in nature. As a result, question 3 is omitted in the analysis that follows.

**Question 1. The approximate total water production rate at your utility is _____ mgd.**

| mgd | Number of respondents | Percent |
|---|---|---|
| <10 | 35 | 14.2 |
| 10 to 19 | 66 | 26.7 |
| 20 to 29 | 39 | 15.8 |
| 30 to 39 | 21 | 8.5 |
| 40 to 49 | 16 | 6.5 |
| 50 to 59 | 9 | 3.6 |
| 60 to 69 | 7 | 2.8 |
| 70 to 79 | 8 | 3.2 |
| 80 to 89 | 13 | 5.3 |
| 90 to 99 | 3 | 1.2 |
| 100 to 200 | 23 | 9.3 |
| >200 | 7 | 2.8 |
| Total | 247 | 100.0* |

*Rounded.

**Question 2. The fraction of water coming from surface supplies at your utility is _____%.**

| Fraction | Number of respondents | Percent |
|---|---|---|
| <25 | 13 | 5.3 |
| 25 to 49 | 4 | 1.6 |
| 50 to 74 | 21 | 8.5 |
| 75 to 99 | 42 | 17.1 |
| 100 | 166 | 67.5 |
| Total | 246 | 100.0 |

**Question 4. Aside from data reported to USEPA as part of the ICR, have any total organic halide (TOX) measurements been made on your *treated* water?**

| Response | Number of respondents | Percent |
|---|---|---|
| Yes | 42 | 17.4 |
| No | 200 | 82.6 |
| Total | 242 | 100.0 |



©2007 AwwaRF. All Rights Reserved.

**Question 5. Does your utility have data on disinfection by-product (DBP) precursor levels for any of your *raw waters*?**

| Response | Number of respondents | Percent |
|---|---|---|
| Yes | 100 | 41.8 |
| No | 139 | 58.2 |
| Total | 239 | 100.0 |

**Question 6. Which DBPs were measured as part of a precursor test?**
Results for question 5 are shown *only* for those responding from the 100 utilities answering "yes" to question 4.

Trihalomethane (THM) precursors:

| Response | Number of respondents | Percent |
|---|---|---|
| Yes | 82 | 93.2 |
| No | 6 | 6.8 |
| Total | 88 | 100.0 |

Haloacetic acid (HAA) precursors:

| Response | Number of respondents | Percent |
|---|---|---|
| Yes | 55 | 72.4 |
| No | 21 | 27.6 |
| Total | 76 | 100.0 |

Total organic halide (TOX) precursors:

| Response | Number of respondents | Percent |
|---|---|---|
| Yes | 21 | 32.8 |
| No | 43 | 67.2 |
| Total | 64 | 100.0 |

Twenty-two respondents listed one of five other DBPs that were measured as part of a precursor test.



©2007 AwwaRF. All Rights Reserved.

**Question 7. What year were *any* of these DBP data *first* collected on raw water at your utility?** Results for question 6 are shown *only* for those responding from the 100 utilities answering "yes" to question 4.

| Year | Number of respondents | Percent |
|---|---|---|
| Before 1975 | 3 | 3.4 |
| 1976–1980 | 9 | 10.2 |
| 1981–1985 | 6 | 6.8 |
| 1986–1990 | 11 | 12.5 |
| 1991–1995 | 18 | 20.5 |
| 1996–2000 | 29 | 33.0 |
| After 2001 | 12 | 13.6 |
| Total | 88 | 100.0 |

Most utilities did not answer this question, which has been interpreted as uncertainty on the part of the respondent regarding the exact date and the perceived difficulty in obtaining this information. Of those who did respond, it is suspected that the data are biased toward dates that are more recent. Eight utilities claimed to have collected such data before 1980 (1974 was the earliest date cited):

1. North Marin Water District
2. Brick Township Municipal Utilities Authority (MUA)
3. Dallas Water Utility
4. Manatee County Public Works
5. Syracuse Water Department
6. Cincinnati Public Water System
7. Upper Mohawk Valley Region Water B
8. Harnett County Department of Public Utilities

Nine utilities indicated that they started collecting such data between 1980 and 1985:

1. Manchester Waterworks
2. City of Tampa Water Department
3. Eugene Water and Electric Board
4. East Jefferson Waterworks District 1
5. Kansas City
6. West Jefferson Waterworks District 2
7. Phoenix Municipal Water System
8. Passaic Valley Water Commission
9. City of Vallejo



©2007 AwwaRF. All Rights Reserved.

**Question 8. Has *your utility* collected DBP precursor or other natural organic matter (NOM)-related data (e.g., TOC, DOC, UV absorbance) in the surface water samples collected within your watershed(s)?**

| Response | Number of respondents | Percent |
|---|---|---|
| Yes | 144 | 59.8 |
| No | 97 | 40.2 |
| Total | 241 | 100.0 |

**Question 9. Have any studies been done by *other groups* on DBP precursors or other NOM-related data in surface water samples collected within your watershed(s)?**

| Response | Number of respondents | Percent |
|---|---|---|
| Yes | 80 | 35.4 |
| No | 146 | 64.6 |
| Total | 226 | 100.0 |

Questions 8 and 9 were intended to assess the likelihood of a useful historical record for NOM-related parameters (e.g., TOC, DOC, UV absorbance, DBP precursors) in the utility's watershed. Question 10 was intended to assess the same type of historical data for general hydrological parameters. Utilities that responded unambiguously to any of these questions in the affirmative were judged to be of greater interest concerning the project's objectives. In total, 93 of 249 respondents were included in this category.

**Question 10. Have hydrologic studies been carried out in your watershed(s)?**

| Response | Number of respondents | Percent |
|---|---|---|
| Yes, by my utility | 30 | 14.0 |
| Yes, by other groups | 139 | 65.0 |
| No | 45 | 21.0 |
| Total | 214 | 100.0 |

**Question 11. Have storm events in your watershed(s) ever been specially monitored *by your utility* in terms of the following?**

| | Water quality | | Flow | | System response | |
|---|---|---|---|---|---|---|
| Response | Number of respondents | Percent | Number of respondents | Percent | Number of respondents | Percent |
| Yes | 95 | 39.4 | 80 | 33.8 | 39 | 17.6 |
| No | 146 | 60.6 | 157 | 66.2 | 183 | 82.4 |
| Total | 241 | 100.0 | 237 | 100.0 | 222 | 100.0 |



©2007 AwwaRF. All Rights Reserved.

**Question 12. Have storm events in your watershed(s) ever been specially monitored by *other groups* in terms of the following?**

| | Water quality | | Flow | | System response | |
|---|---|---|---|---|---|---|
| Response | Number of respondents | Percent | Number of respondents | Percent | Number of respondents | Percent |
| Yes | 89 | 40.3 | 86 | 39.6 | 46 | 24.0 |
| No | 132 | 59.7 | 131 | 60.4 | 146 | 76.0 |
| Total | 221 | 100.0 | 217 | 100.0 | 192 | 100.0 |

**Question 13. To what extent would the following issues be of interest to you?**

DBP control:

| Response | Number of respondents | Percent |
|---|---|---|
| Not at all interested | 11 | 4.5 |
| Somewhat interested | 21 | 8.6 |
| Modestly interested | 35 | 14.3 |
| Very interested | 92 | 37.7 |
| Extremely interested | 85 | 34.8 |
| Total | 244 | 100.0* |

*Rounded.

Watershed management for control of organics:

| Response | Number of respondents | Percent |
|---|---|---|
| Not at all interested | 12 | 4.9 |
| Somewhat interested | 35 | 14.4 |
| Modestly interested | 48 | 19.8 |
| Very interested | 89 | 36.6 |
| Extremely interested | 59 | 24.3 |
| Total | 243 | 100.0 |



©2007 AwwaRF. All Rights Reserved.

Watershed management for control of pathogens:

| Response | Number of respondents | Percent |
|---|---|---|
| Not at all interested | 11 | 4.5 |
| Somewhat interested | 32 | 13.2 |
| Modestly interested | 56 | 23.1 |
| Very interested | 85 | 35.1 |
| Extremely interested | 58 | 24.0 |
| Total | 242 | 100.0* |

*Rounded.

Role of storm events in water quality:

| Response | Number of respondents | Percent |
|---|---|---|
| Not at all interested | 11 | 4.5 |
| Somewhat interested | 32 | 13.1 |
| Modestly interested | 69 | 28.2 |
| Very interested | 82 | 33.5 |
| Extremely interested | 51 | 20.8 |
| Total | 245 | 100.0* |

*Rounded.

Figure 3.1 shows responses to question 13.

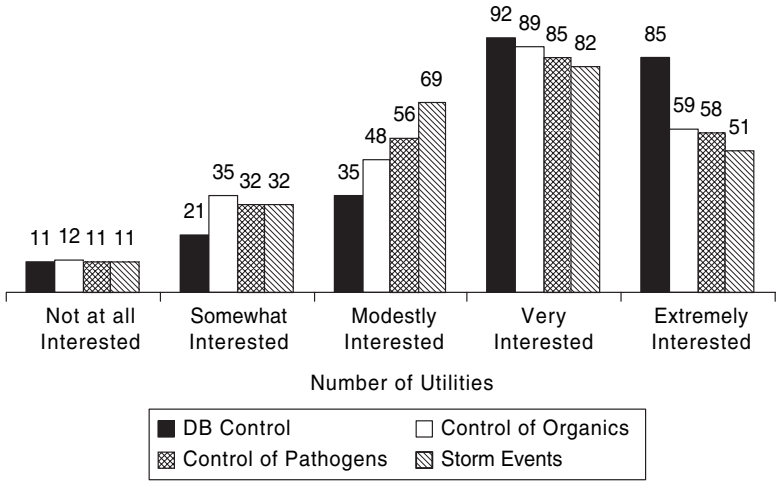



**Figure 3.1 Issues of interest to water utilities**



©2007 AwwaRF. All Rights Reserved.

**Question 14. Would your utility be willing to share existing water quality data with the University of Massachusetts?**

| Response | Number of respondents | Percent |
|---|---|---|
| Yes | 219 | 90.5 |
| No | 23 | 9.5 |
| Total | 242 | 100.0 |

**Question 15. Would your utility be willing to collect additional water quality data as part of an AwwaRF-sponsored project?**

| Response | Number of respondents | Percent |
|---|---|---|
| Yes | 197 | 84.2 |
| No | 37 | 15.8 |
| Total | 234 | 100.0 |

## DATA FROM OTHER SOURCES

This section pertains to use of additional survey data and other types of information in the selection of utilities for studies of NOM and DBP precursors in watersheds. Selection of utilities should be based on factors such as the existence of a long historical record of raw water NOM characteristics (e.g., TOC, UV absorbance, DBP precursors), availability of land-use data, and a utility's willingness to collect data and participate in the project. Utilities drawing water from large and difficult-to-characterize watersheds should be avoided.

In addition, because of an interest in known characteristics that would help to identify waters of contrasting behavior, the intent was to capture the full range of relevant characteristics found within North American waters. Given the types of information available from many utilities, the following selection criteria were of greatest importance:

- Contrasting watershed characteristics and watershed ecoregion locations
- High and low SUVA at a given TOC
- High and low ratios of known by-products to other TOX (e.g., total trihalomethanes [TTHM] plus the sum of nine HAAs [HAA9] divided by unknown TOX)
- High and low nitrogen contents or HAN/TOX ratios

### Length of Record

Eleven utilities cited dates from 1974 to 1980 as the period when their first raw water NOM measurements were made. Another six utilities cited dates between 1981 and 1985. Most respondents did not fill out this particular portion of the survey, presumably because the information was not readily available. For this reason, it is suspected that this information is unreliable.



©2007 AwwaRF. All Rights Reserved.

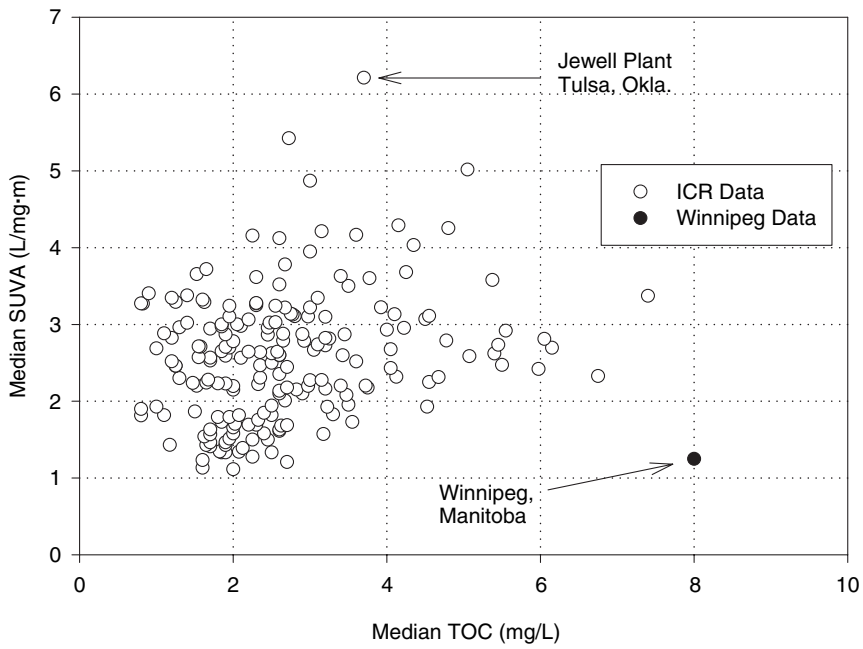



**Figure 3.2 Distribution of raw water NOM characteristics for 195 large U.S. plants (summarized from ICR data) and 1 plant in Winnipeg**

### SUVA Criteria

A set of high- and low-SUVA waters were identified for use in a recent AwwaRF project (Reckhow et al. 2007). Winnipeg was selected as a low-SUVA site. The Winnipeg location was particularly interesting because it has abnormally high TOC given its SUVA (Figure 3.2). This utility is also quite interested in and knowledgeable about water quality research and was recently a participant in a tailored collaboration on impacts of UV treatment.

Tulsa's Jewell Plant was identified as having high-SUVA waters. This plant is also located in a region of the United States that is not well represented in national studies of NOM and DBPs. For this reason, it will help to capture some of the geographical variability that is the subject of a subsequent criterion. Table 3.1 presents a comparison of the Tulsa and Winnipeg waters.

### Criteria Based on Ratios of Known to Unknown TOX

Another set of criteria is related to the distribution or speciation of DBPs. Some waters produce higher levels of the regulated DBPs than others when normalized to the TOX. By selecting utilities with raw waters that have contrasting ratios of known TOX (e.g., the THMs and HAAs) to unknown (the remaining) TOX, one ensures a broad range of precursor types being captured.

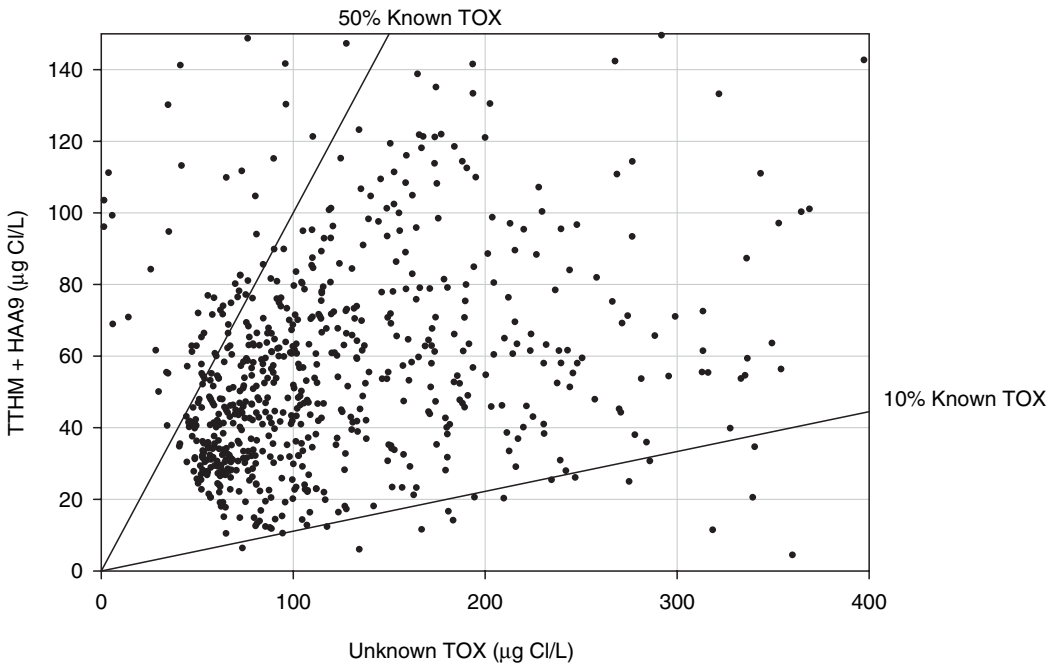
Based on averaged data from the ICR, about 30% of the measured TOX is accounted for in the major by-products (THMs and HAAs). However, individual systems show a broad range in the ratio of known TOX (essentially TTHM + HAA9) to unknown TOX (i.e., the TOX not accounted for in these compounds). To determine which systems represent extremes, data from the ICR were carefully selected. To avoid systems that were not practicing free chlorination, only samples exhibiting 75% of total residual chlorine in the free form were included. Furthermore,



©2007 AwwaRF. All Rights Reserved.

**Table 3.1**
**Comparative raw water quality for high- and low-SUVA waters**

| Parameter | Tulsa, Okla. (Jewell plant), raw water | Winnipeg, Man., raw water |
|---|---|---|
| TOC (mg/L) | 3.8 | 8.0 |
| SUVA (L/mgm) | 5.5 | 1.3 |
| UV absorbance ($cm^{-1}$) | 0.21 | 0.10 |
| Hardness (mg/L) | 142 | 83 |
| Alkalinity (mg/L) | 113 | 81 |
| Bromide (mg/L) | 0.065 | Low |
| pH | 7.9 | 8.2 |
| Turbidity (ntu) | 22 | 1.0 |




NOTES: Selected SDS data (see text). HAA9 concentrations were either measured or estimated using the method of Roberts, Singer, and Obolensky (2002).

**Figure 3.3 Known versus unknown TOX in selected ICR data**

systems with TOX values below 75 µg/L were excluded in the interest of avoiding high relative uncertainties in the TOX value. Finally, the simulated distribution system (SDS) data were extracted so that confounding factors such as losses to biodegradation in distribution systems might be avoided. The remaining data, representing nearly 500 systems, are shown in Figure 3.3. Nearly all of the data are bracketed by the two lines, representing 10% and 50% known TOX. Utilities that correspond to the data near the two extreme lines are good candidates for capturing variability in TOX speciation. To avoid the occasional false data point, only those utilities that appeared more than once in the extreme zones were selected. These utilities are listed in Table 3.2.



©2007 AwwaRF. All Rights Reserved.

**Table 3.2**
**Selected utilities representing extremes in known to unknown TOX ratios**

| City | State | Plant identification numbers, PWSID* (ICR) | Flow (mgd) | Raw water quality TOC | Raw water quality SUVA | Other information |
|---|---|---|---|---|---|---|
| **Low Known/Unknown TOX** | | | | | | |
| Brockton | MA | 4044000(402) | 11.5 | 3.8 | 3.1 | Trident 210A (Siemens, Warrendale, Pa.) |
| Layton | UT | 4900512(667) | 26 | 3.3 | 3.3 | Weber Basin 2; conventional with GAC caps |
| Rochester | NY | 0004518(514) | 38 | 2.6 | 1.9 | Hemlock plant; direct filtration |
| Salt Lake City | UT | 4900390(662) | 100 | 1.3 | 1.7 | City Creek WTP†: several rivers; conventional |
| Clayton Co. | GA | 0630000(324) | 27 | 3.4 | 2.2 | Hooper WTP |
| New York | NY | 0003493(721) | 1,400 | 1.5 | 2.6 | Catskill system; unfiltered |
| **High Known/Unknown TOX** | | | | | | |
| Shrewsbury | NJ | 1345001(473) | 38 | 3.0 | 6.9 | Swimming River WTP |
| Glendale | AZ | 0407093(126) | 42 | 6.8 | 7.3 | Cholla WTP |
| Sacramento | CA | 3410020(205) | 122 | 1.3 | 2.6 | Sacramento River WTP; Fairbairn WTP (ICR 204) may be similar |
| Durham | NC | 0332010(450) | 35 | 4.4 | 4.8 | Brown WTP |

* PWSID = public water system identification.
† WTP = water treatment plant.

Based on well-controlled laboratory data using aquatic humic substances (i.e., Black Lake fulvic acid [FA]; Reckhow et al. 2007), the fraction of known to unknown TOX would be expected to range from 0.3 (low pH, reaction time, and dose) to about 1.0 (high pH, reaction time, and dose) under the full range of conditions one might expect to find in a water treatment plant. This is equivalent to about 25%–50% known TOX accountable in the major by-products.

## Ecoregion

Climate, terrain, mean rainfall, geochemistry, and flora are among the characteristics that affect export of carbon from pristine watersheds. Much of this natural variability is captured in the delineation of ecoregions. USEPA (2006) has described these ecoregions as follows.

> Ecoregions are defined as areas of relative homogeneity in ecological systems and their components. Factors associated with spatial differences in the quality and quantity of ecosystem components, including soils, vegetation, climate, geology, and physiography, are relatively homogeneous within an ecoregion. Ecoregions



©2007 AwwaRF. All Rights Reserved.

separate different patterns of human stresses on the environment and different patterns in the existing and attainable quality of environmental resources. They have proven to be an effective aid for inventorying and assessing national and regional environmental resources, for setting regional resource management goals, and for developing biological criteria and water quality standard.

North American ecoregions have been delineated in as many as three levels of resolution. Maps and descriptive material may be found at several Internet sites (e.g., <http://www.epa.gov/wed/pages/ecoregions/ecoregions.htm>).

In addition to the characteristics previously mentioned, the ecosystem designation for the watershed of each candidate utility was noted. Attempts were made to include samples from all of the major ecosystems, and from as many secondary ones as could be accommodated given the other characteristics and the limitations on site numbers.

## PRIMARY LIST OF CANDIDATE UTILITIES

A primary group of approximately 50 utilities was extracted from the original 553 and supplemented with a few smaller utilities with key attributes (Table 3.3). Heavy emphasis was placed on the existence of a long record of NOM data, coordination with other studies, and an interest in having a broad geographical representation (e.g., ecoregions).

## SECONDARY LIST OF CANDIDATE UTILITIES

Approximately 120 additional utilities (Table 3.4) were identified as being good alternatives to the primary list of utilities. This secondary list includes candidates for further study from most of the ecoregions represented in the primary list. This group also contains all remaining utilities who indicated that there were substantial existing data about their watersheds (questions 8–10 of the survey).

## TARGETED DATA COLLECTION FROM CANDIDATE UTILITIES

An extensive set of data was collected from as many of the selected utilities as possible. Every effort was made to collect data by telephone, mail, fax, and e-mail to avoid costly and time-consuming site visits. A special effort was made to avoid placing unnecessary or unwanted burdens on utility personnel. The types of data requested were as follows:

1. Raw water quality data (TOC, DOC, UV absorbance, trihalomethane formation potential [THMFP], haloacetic acid formation potential [HAAFP], haloacetonitrile formation potential, total organic halide formation potential [TOXFP], inorganic anions and cations, alkalinity, pH, conductivity, total Kjeldahl nitrogen, nutrients, algal enumerations, chlorophyll a, etc.)
2. Water quality from other locations in the watershed (same as item 1)
3. Watershed description (size, land use, topography, flora, hydrology)
4. Industrial or municipal discharges in the watershed
5. Hydrologic and meteorologic data

This information was generally requested for the current year, and for as many prior years as possible or as deemed to be helpful. Much of the watershed data was available in GIS format.



©2007 AwwaRF. All Rights Reserved.

**Table 3.3**
**Primary list of candidate utilities**

| State | PWSID | Public water system | Population | Ecoregions I | Ecoregions II | Ecoregions III | Old data* | Select Q8–10† | Unusual DBPs‡ | AwwaRF studies§ | TPL study** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| AK | AK2320078 | Barrow Utilities and Electric Cooperative | 4,550 | 2 | 2 | 1 | | | | | |
| AK | AK2280155 | McGrath Water System (Koyukuk) | 666 | 3 | 1 | | | | | | |
| CT | CT0150011 | Bridgeport Main System | 324,859 | 6 | 2 | 14 | | | X | | X |
| UT | UT4903012 | Central Utah Water Conservancy District–Utah Valley | 150,000 | 6 | 2 | 13 | | X | | T | |
| AK | AK2210906 | Municipality of Anchorage | 114,909 | 7 | 1 | 3 | | | | | X |
| OR | OR4100657 | Portland Bureau of Waterworks | 460,000 | 7 | 1 | 9 | | X | | | X |
| WA | WA5324050 | Everett Public Works Department | 86,730 | 7 | 1 | 7 | | | | | X |
| CT | CT0930011 | South Central Connecticut Regional Water Authority | 383,623 | 8 | 1 | 7 | | X | X | | X |
| MA | MA6000000 | Massachusetts Water Resources Authority | 2,000,000 | 8 | 1 | 7 | | X | X | T,W,R | |
| MI | MI0000220 | Ann Arbor | 109,592 | 8 | 1 | 4 | | X | | W | |
| NY | NY3202411 | Upper Mohawk Valley Region Water B | 120,000 | 8 | 1 | | 1978 | | | | |
| NH | NH1471010 | Manchester Waterworks | 128,000 | 8 | 1 | | 1980 | | X | | |
| MA | MA3049000 | Cambridge Water Department | 95,802 | 8 | 1 | | | | | T | |
| NY | NY0301651 | Binghamton | 53,008 | 8 | 1 | 3 | | X | | | |
| NY | NY3304334 | Syracuse Water Department | 192,000 | 8 | 1 | 1 | 1977 | X | | | |
| IL | IL1671200 | Springfield | 146,000 | 8 | 2 | 3 | | | | | X |
| OH | OH0200811 | City of Lima | 74,750 | 8 | 2 | 4 | | | | | X |
| NC | NC0343045 | Harnett County Department of Public Utilities | 50,000 | 8 | 3 | 4 | 1979 | X | | | |
| GA | GA2150000 | Columbus | 185,000 | 8 | 3 | 5 | | X | | B | |
| IL | IL1635040 | Illinois American Water Company–East Saint Louis | 139,200 | 8 | 3 | 2 | | | | W | |

(continued)



©2007 AwwaRF. All Rights Reserved.

**Table 3.3 (Continued)**

| State | PWSID | Public water system | Population | Ecoregions I | Ecoregions II | Ecoregions III | Old data* | Select Q8–10† | Unusual DBPs‡ | AwwaRF studies§ | TPL study** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| NJ | NJ0714001 | Newark Water Department | 275,221 | 8 | 3 | 1 | | | | | |
| NC | NC0368010 | Orange Water and Sewer Authority | 55,000 | 8 | 3 | 4 | | X | | | X |
| TN | TN0000494 | Nashville Water Department 1 | 412,067 | 8 | 3 | 3 | | X | X | | |
| AR | AR0000465 | Central Arkansas Water | 204,543 | 8 | 4 | 8 | | X | | | |
| AR | AR0000507 | Fort Smith Waterworks | 72,798 | 8 | 4 | 7 | | X | | | |
| MO | MO5010754 | Springfield | 149,237 | 8 | 4 | 5 | | X | | | |
| PA | PA3390024 | Allentown City Bureau of Water | 105,200 | 8 | 4 | 1 | | X | | | |
| PA | PA5650032 | Westmoreland Municipal Authority–Sweeney Plant | 139,335 | 8 | 4 | 3 | | X | | | |
| NJ | NJ1506001 | Brick Township MUA | 75,561 | 8 | 5 | 1 | 1975 | X | X | | |
| FL | FL6290327 | City of Tampa Water Department | 475,000 | 8 | 5 | | 1980 | | | | |
| FL | FL6411132 | Manatee County Public Works | 186,000 | 8 | 5 | 3 | 1976 | X | | | |
| NJ | NJ1605002 | Passaic Valley Water Commission | 275,000 | 8 | 5 | 3 | X | X | | | X |
| VA | VA3700500 | City of Newport News | 400,000 | 8 | 5 | 1 | | | X | T,W | X |
| IA | IA7727031 | Des Moines Waterworks | 193,189 | 9 | 2 | 3 | | X | | | X |
| OK | OK1021307 | Tulsa Metropolitan Utility Board | 403,000 | 9 | 2 | 4 | | | | R | |
| TX | TX0570004 | Dallas Water Utility | 1,068,800 | 9 | 4 | | 1975 | X | | | |
| CO | CO0116001 | Denver Water Board | 1,000,000 | 9 | 4 | 1 | | | | | X |
| TX | TX1550008 | City of Waco | 110,000 | 9 | 4 | | | | | | |
| TX | TX2200012 | City of Fort Worth | 479,884 | 9 | 4 | 5 | | X | | W | |
| UT | UT4900391 | Jordan Valley Water Conservancy District | 229,000 | 10 | 1 | 4 | | X | | | |

(continued)



©2007 AwwaRF. All Rights Reserved.

**Table 3.3 (Continued)**

| State | PWSID | Public water system | Population | Ecoregions I | Ecoregions II | Ecoregions III | Old data* | Select Q8–10† | Unusual DBPs‡ | AwwaRF studies§ | TPL study** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| UT | UT4900392 | Metropolitan Water District of Salt Lake City | 700,000 | 10 | 1 | 5 | | | | | |
| NV | NV0000158 | Incline Village General Improvement District | 9,313 | 10 | 2 | 1 | | | | | X |
| CA | CA1510040 | Kern County Water Agency | 100,000 | 11 | 1 | 2 | | X | | | |
| CA | CA2110003 | North Marin Water District | 53,000 | 11 | 1 | 1 | 1974 | X | | | |
| CA | CA0710003 | Contra Costa Water District | 200,000 | 11 | 1 | 1 | | X | | T,W | |
| CA | CA3710025 | Sweetwater Authority | 163,000 | 11 | 1 | 1 | | X | | | X |
| CA | CA4210010 | Santa Barbara Water District | 95,064 | 11 | 1 | 3 | | | | | |
| AZ | AZ0405005 | Gila Resources (Safford) | 18,900 | 12 | 1 | | | | | | |
| AZ | AZ0403008 | Flagstaff Municipal Water | 53,000 | 13 | 1 | | | | | | |
| FL | FL4501559 | City of West Palm Beach | 87,769 | 15 | 4 | 1 | | X | | | |
| MB | NA†† | Winnipeg | 650,000 | 5 | 2 | | | | | | |
| PQ | NA | Repentigny | 75,000 | 8 | 1 | | | | | | |

* Utility indicated that raw water NOM data had been collected prior to 1985.
† Utility answered "yes" to questions 8, 9, or 10, indicating NOM or hydrological data already exists for the watershed.
‡ Utilities with either high or low unknown TOX/TOX ratios, or unusual ratios of dihaloacetonitriles as indicated in the ICR data.
§ Utilities where relevant NOM or DBP-related data have recently been collected (T = partner for this study; W = Arizona State University study on total organic nitrogen; R = UMass TOX study; B = East Bay Municipal Utility District DBP study).
** Utility selected for Trust for Public Land (TPL) study.
†† NA = not applicable for plants outside of the United States.



©2007 AwwaRF. All Rights Reserved.

**Table 3.4**
**Secondary list of candidate utilities**

| State | PWSID | Public water system | Population | Ecoregions I | Ecoregions II | Ecoregions III | Old data* | Select Q8–10† | Unusual DBPs‡ | AwwaRF studies§ | TPL study** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| UT | UT4900332 | Orem | 84,000 | 6 | 2 | 13 | | | | | |
| CO | CO0107152 | City of Boulder | 105,740 | 6 | 2 | 14 | | X | | | |
| WA | WA5305600 | City of Bellingham Water Division | 67,971 | 7 | 1 | 7 | | X | | | |
| WA | WA5308200 | City of Bremerton | 52,500 | 7 | 1 | 7 | | X | | | |
| WA | WA5386800 | City of Tacoma Water Division | 306,000 | 7 | 1 | 7 | | X | X | | |
| OR | OR4100225 | City of Corvallis | 50,101 | 7 | 1 | 9 | | X | | | X |
| OR | OR4100731 | Salem Public Works | 155,000 | 7 | 1 | 9 | | | | | X |
| OR | OR4100287 | Eugene Water and Electric Board | 161,000 | 7 | 1 | | 1981 | | | | |
| NY | NY2704518 | Rochester | 220,000 | 8 | 1 | 1 | | | X | | |
| NY | NY1400422 | Buffalo Water Authority | 345,974 | 8 | 1 | 1 | | X | | | |
| NY | NY2701047 | Monroe County Water Authority—Shoremont Water Treatment Plant | 225,370 | 8 | 1 | 1 | | X | | | |
| OH | OH7700011 | City of Akron | 308,720 | 8 | 1 | 1 | | X | | | |
| MN | MN1620026 | Saint Paul Regional Water Services | 414,735 | 8 | 1 | 4 | | X | | | |
| MA | MA1281000 | Springfield Water and Sewer Commission | 250,000 | 8 | 1 | 7 | | X | | | |
| MA | MA3336000 | Weymouth Water Department | 52,140 | 8 | 1 | 7 | | X | | | |
| ME | ME0091300 | Portland Water District | 113,560 | 8 | 1 | 7 | | X | | | |
| NJ | NJ0238001 | United Water New Jersey | 713,737 | 8 | 1 | 7 | | X | | | |
| OH | OH4801411 | City of Toledo | 380,000 | 8 | 2 | 2 | | X | | | |

(continued)



©2007 AwwaRF. All Rights Reserved.

**Table 3.4 (Continued)**

| State | PWSID | Public water system | Population | Ecoregions I | Ecoregions II | Ecoregions III | Old data* | Select Q8–10† | Unusual DBPs‡ | AwwaRF studies§ | TPL study** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| IL | IL1835120 | Aqua Illinois Water—Vermilion County Division | 38,000 | 8 | 2 | 3 | | | | | X |
| IL | IL1150150 | Decatur | 83,885 | 8 | 2 | 3 | | | | | X |
| IL | IL0316000 | Chicago | 2,783,726 | 8 | 2 | 3 | | X | | | |
| IL | IL0894380 | Elgin | 77,010 | 8 | 2 | 3 | | X | | | |
| IL | IL0310810 | Evanston | 73,233 | 8 | 2 | 3 | | X | | | |
| WI | WI2520062 | Racine Waterworks | 99,100 | 8 | 2 | 3 | | X | | | |
| OH | OH2504412 | Columbus Public Water System | 897,213 | 8 | 2 | 4 | | X | | | |
| MD | MD0300002 | City of Baltimore | 1,600,000 | 8 | 3 | 1 | | X | | | |
| DE | DE0000564 | United Water Delaware | 93,000 | 8 | 3 | 1 | | X | | | |
| OH | OH3102612 | Cincinnati Public Water System | 780,000 | 8 | 3 | 3 | 1978 | X | | | |
| KY | KY0590220 | Northern Kentucky Water Service District | 197,749 | 8 | 3 | 3 | | X | | | |
| TN | TN0000745 | White House Utility District | 57,137 | 8 | 3 | 3 | | X | | | |
| TX | TX1010013 | City of Houston | 2,354,040 | 8 | 3 | | | | | | |
| VA | VA4041845 | Swift Creek Waterworks | 253,751 | 8 | 3 | 4 | | X | | | X |
| GA | GA0450002 | Carrollton Water Works | 19,000 | 8 | 3 | 4 | | | | | X |
| SC | SC4220006 | SJWD Water District | 40,000 | 8 | 3 | 4 | | | | | X |
| GA | GA0590000 | Athens–Clarke County Water System | 80,000 | 8 | 3 | 4 | | X | | | |
| GA | GA0970000 | Douglasville–Douglas County Authority | 73,000 | 8 | 3 | 4 | | X | | | |
| GA | GA0890001 | Dekalb County | 594,400 | 8 | 3 | 4 | | X | | | |
| NC | NC0160010 | Charlotte–Mecklenburg Utility Department | 641,000 | 8 | 3 | 4 | | X | | | |
| NC | NC0332010 | City of Durham | 153,000 | 8 | 3 | 4 | | X | X | B | |

(continued)



©2007 AwwaRF. All Rights Reserved.

**Table 3.4 (Continued)**

| State | PWSID | Public water system | Population | Ecoregions I | Ecoregions II | Ecoregions III | Old data* | Select Q8–10† | Unusual DBPs‡ | AwwaRF studies§ | TPL study** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| NC | NC0241010 | City of Greensboro | 205,000 | 8 | 3 | 4 | | X | | | |
| NC | NC0234010 | City of Winston-Salem | 236,214 | 8 | 3 | 4 | | X | X | | |
| VA | VA4041035 | Appomattox River Water Authority | | 8 | 3 | 5 | | | | | X |
| DC | DC0000001 | Washington Aqueduct Division | 927,055 | 8 | 3 | 5 | | | | | X |
| NC | NC0474010 | Greenville Utilities Commission | 60,928 | 8 | 3 | 5 | | X | | | |
| SC | SC4010001 | City of Columbia | 259,293 | 8 | 3 | 5 | | X | | | |
| VA | VA4760100 | City of Richmond | 209,000 | 8 | 3 | 5 | | X | | | |
| PA | PA5260036 | Westmoreland County Municipal Authority—Yough Plant | 130,000 | 8 | 4 | 3 | | | | | |
| AR | AR0000871 | Benton–Washington County Water | | 8 | 4 | 5 | | X | | | |
| LA | LA1051001 | East Jefferson Waterworks District 1 | 308,362 | 8 | 5 | | X | | | | |
| NC | NC0465010 | City of Wilmington | 77,390 | 8 | 5 | 1 | | X | | | X |
| NJ | NJ0327001 | New Jersey American Water Company—Western Division | 304,772 | 8 | 5 | 1 | | | | | X |
| NJ | NJ1345001 | New Jersey American Water Company—Monmouth | 302,491 | 8 | 5 | 1 | | | X | | X |
| NJ | NJ1111001 | Trenton Water Department | 225,000 | 8 | 5 | 1 | | X | | | |
| NY | NY7003666 | New York City—Croton System | 1,000,000 | 8 | 5 | 1 | | X | | | |
| PA | PA1510001 | Philadelphia Water Department | 1,755,000 | 8 | 5 | 1 | | X | | W | |
| VA | VA3710100 | City of Norfolk | 261,250 | 8 | 5 | 1 | | X | | | |
| AL | AL0001005 | Mobile Area Water Service System | 279,000 | 8 | 5 | 3 | | X | | | |
| LA | LA1051004 | West Jefferson Waterworks District 2 | 186,079 | 8 | 5 | | X | | | | |

(continued)



©2007 AwwaRF. All Rights Reserved.

**Table 3.4 (Continued)**

| State | PWSID | Public water system | Population | Ecoregions I | Ecoregions II | Ecoregions III | Old data* | Select Q8–10† | Unusual DBPs‡ | AwwaRF studies§ | TPL study** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| IA | IA5715093 | Cedar Rapids Water Plant | 108,772 | 9 | 2 | 3 | | X | | | |
| IA | IA5225079 | Iowa City Water Department | 54,231 | 9 | 2 | 3 | | X | | | |
| KS | KS2009110 | Water District 1 of Johnson County | 313,434 | 9 | 2 | 4 | | X | | | |
| OK | OK1020418 | Tulsa | 197,000 | 9 | 2 | 4 | | X | | | |
| NE | NE3105507 | Metropolitan Utilities District | 506,420 | 9 | 2 | | | | X | | |
| MO | MO1010415 | Kansas City | 450,000 | 9 | 2 | 4 | X | X | | | |
| CO | CO0130001 | City of Arvada | 101,000 | 9 | 4 | 1 | | X | | | |
| CO | CO0101170 | City of Westminster | 100,000 | 9 | 4 | 1 | | X | | | |
| OK | OK1020902 | Oklahoma City—Hefner | 276,000 | 9 | 4 | 5 | | | | | |
| TX | TX1080029 | North Alamo Water Supply Corporation | 68,949 | 9 | 4 | 5 | | | | | |
| OK | OK1020723 | Edmond Public Works Authority | 53,000 | 9 | 4 | 5 | | X | | | |
| TX | TX0610002 | City of Denton | 73,050 | 9 | 4 | 5 | | X | | | |
| OK | OK1020802 | Oklahoma City—Draper | 276,000 | 9 | 4 | 5 | | | | | |
| UT | UT4900463 | West Jordan Water System | 68,000 | 10 | 1 | 4 | | | | | |
| UT | UT4900390 | Salt Lake City Water System | 305,835 | 10 | 1 | 5 | | | X | | X |
| AZ | AZ0407025 | Phoenix Municipal Water System | 1,200,000 | 10 | 2 | 1 | X | X | X | W | |
| AZ | AZ0407095 | City of Mesa | 350,000 | 10 | 2 | 1 | | X | | | |
| AZ | AZ0407098 | Scottsdale Municipal Water | 180,000 | 10 | 2 | 1 | | X | | | |
| AZ | AZ0407100 | Tempe Municipal Water Department | 160,000 | 10 | 2 | 1 | | X | | | |
| NV | NV0000076 | Henderson Water Company | 57,000 | 10 | 2 | 1 | | X | | | |
| CA | CA0110001 | Alameda County Water District | 292,307 | 11 | 1 | 1 | | | | | X |

(continued)



©2007 AwwaRF. All Rights Reserved.

**Table 3.4 (Continued)**

| State | PWSID | Public water system | Population | Ecoregions I | Ecoregions II | Ecoregions III | Old data* | Select Q8–10† | Unusual DBPs‡ | AwwaRF studies§ | TPL study** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| CA | CA4810007 | City of Vallejo | 130,308 | 11 | 1 | 1 | X | X | | | |
| CA | CA4810003 | City of Fairfield | 80,000 | 11 | 1 | 1 | | X | | | |
| CA | CA1910065 | Long Beach Water Department | 425,807 | 11 | 1 | 1 | | X | | | |
| CA | CA0110005 | East Bay Municipal Utility District | 1,300,000 | 11 | 1 | 1 | | X | X | | |
| CA | CA4010022 | Lopez Project–San Luis Obispo County Public Works | | 11 | 1 | 1 | | X | | | |
| IL | IL1435030 | Illinois American Water Company—Peoria | 158,564 | | | | | | | W | X |
| NJ | | New Jersey Water Supply Authority | 41,500 | | | | | | | | X |
| CA | CA4210004 | Goleta Water District | 75,000 | | | | | X | | | |
| CA | CA3710014 | City of Oceanside Water Utilities | 149,600 | | | | | X | | | |
| CO | CO0101150 | City of Thornton | 85,000 | | | | | X | | | |
| OK | OK1020806 | Midwest City | 50,000 | | | | | X | | | |
| AZ | AZ0407090 | City of Chandler | 180,000 | | | | | | | W | |
| AZ | AZ0407092 | Town of Gilbert | 100,000 | | | | | | | W | |
| AZ | AZ0407093 | City of Glendale | 180,000 | | | | | | X | | |
| AZ | AZ0407096 | Peoria | 86,000 | | | | | | | W | |
| CA | CA1910067 | Los Angeles Department of Water and Power | 3,800,000 | | | | | | | | |
| CA | CA3410020 | City of Sacramento | 396,200 | | | | | | X | | |
| FL | FL2161328 | City of Jacksonville (SG) Arlington Water Treatment Plant | | | | | | | X | | |
| FL | FL6290787 | Hillsborough County Water Department | | | | | | | X | | |
| GA | GA0630000 | Clayton County Water Authority | 134,693 | | | | | | X | | |

(continued)



©2007 AwwaRF. All Rights Reserved.

**Table 3.4 (Continued)**

| State | PWSID | Public water system | Population | Ecoregions I | Ecoregions II | Ecoregions III | Old data* | Select Q8–10† | Unusual DBPs‡ | AwwaRF studies§ | TPL study** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| IL | IL0971070 | Central Lake County Joint Action Water Agency | 161,560 | | | | | | | W | |
| IN | IN5218012 | Indiana American Water Company—Muncie | 100,000 | | | | | | | W | |
| IN | IN5234007 | Indiana American Water Company—Kokomo | 52,002 | | | | | | | W | |
| IN | IN5249004 | Indianapolis Water Company | 839,000 | | | | | | | W | |
| KY | KY0560258 | Louisville Water Company | 719,235 | | | | | | | W | |
| MA | MA4044000 | Brockton Water Commission | 97,547 | | | | | | X | | |
| MO | MO5010413 | Joplin | 50,700 | | | | | | | W | |
| NV | NV0000289 | Southern Nevada Water System | 500,000 | | | | | | | W | |
| NY | NY0701008 | Elmira Water Board | 70,000 | | | | | | | | |
| RI | RI1592021 | City of Pawtucket | 110,000 | | | | | | X | | |
| RI | RI1592024 | City of Providence | 261,845 | | | | | | | | |
| TX | TX1520002 | Lubbock Public Water System | 199,445 | | | | | | X | | |
| UT | UT4900512 | Weber Basin Water Conservancy District/Weber County | 95,000 | | | | | | X | | |
| VA | VA3700100 | Fort Eustis | 16,900 | | | | | | | B | |
| WA | WA5377050 | Seattle Public Utilities | 595,430 | | | | | | | | |
| WI | WI2410100 | Milwaukee Waterworks | 682,332 | | | | | | | W | |
| WV | WV3302016 | West Virginia–American Water Company—Kanawha Valley | 173,005 | | | | | | X | | |

*Utility indicated that raw water NOM data had been collected prior to 1985.
†Utility answered "yes" to questions 8, 9, or 10, indicating NOM or hydrological data already exists for the watershed.
‡Utilities with either high or low unknown TOX/TOX ratios, or unusual ratios of dihaloacetonitriles as indicated in the ICR data.
§Utilities where relevant NOM or DBP-related data have recently been collected (T = partner for this study; W = Arizona State University study on total organic nitrogen; R = UMass TOX study; B = East Bay Municipal Utility District DBP study).
**Utility selected for Trust for Public Land (TPL) study.



©2007 AwwaRF. All Rights Reserved.

## Data Sources

### *Open Literature*

Some data about watersheds and raw waters could be collected from the refereed and open, nonrefereed literature. The refereed literature was searched using electronic databases such as the Web of Science (<http://scientific.thomson.com/webofknowledge>). Additional literature was identified through specialized indices (e.g., U.S. Geological Survey [USGS]), literature citations, or with the help of a Web search engine.

### *Public or Quasi-Public Databases*

Some data could be obtained from electronic databases maintained by the U.S. government, private foundations, or other private entities. Examples include

- USEPA's ICR database
- The American Water Works Association Water:\Stats database
- Penwell Corporation's (Tulsa, Okla.) water utility database
- USGS National Water-Quality Assessment Program
- USGS National Land Cover Dataset

### *Telephone Surveys*

Telephone surveys were conducted as part of this research. These were intended as followups to the written survey for selected utilities. The generic questions that were used as a guide to these telephone inquiries are presented in appendix B. In most cases, several followup calls were made to accommodate time constraints or availability on the part of utility personnel.

### *Site Visits*

In many cases, site visits were made by one of the project investigators. For practical and economic reasons, most of these were in the northeastern United States. The utility locations that were visited in connection with this project are

- Utica, N.Y.
- Rochester, N.Y.
- Binghamton, N.Y.
- Springfield, Mass.
- Cambridge, Mass.
- Sweetwater, Calif.
- Celina, Ohio
- Bridgeport, Conn.
- Lima, Ohio
- Repentigny, Que.
- Logan, Utah
- Weber Basin, Utah
- Elgin, Ill.



©2007 AwwaRF. All Rights Reserved.

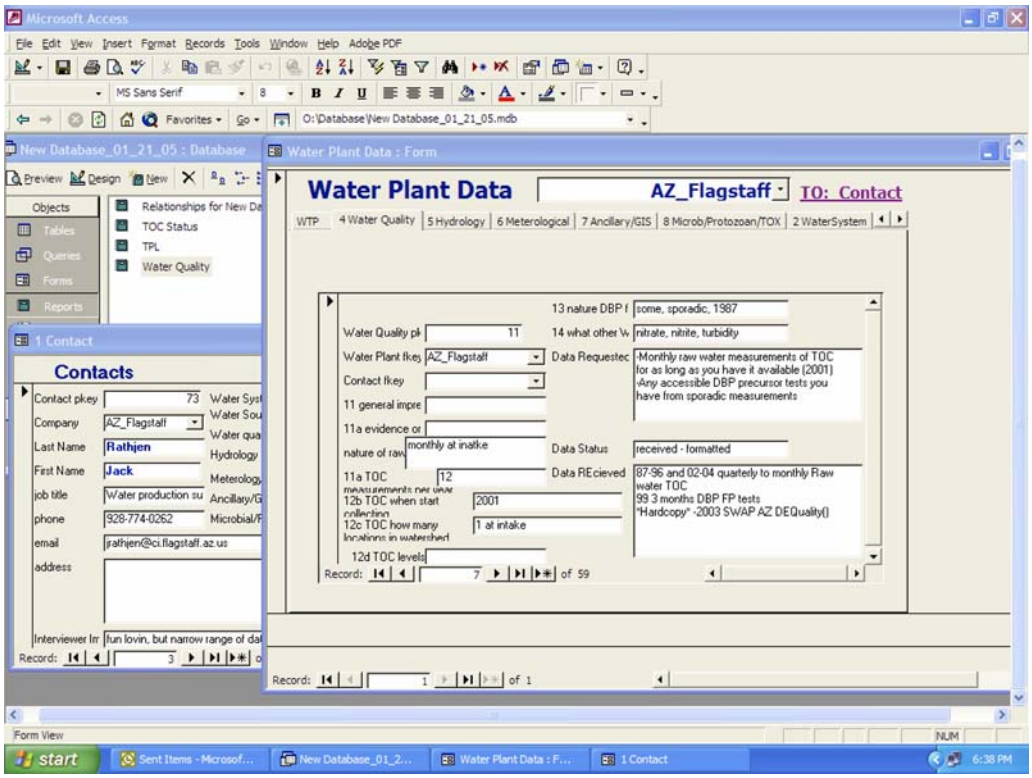



**Figure 3.4 Input form used for primary data collection: NOM-related data**

## Full Database

The data were collected and compiled in various forms. However, much of the critical quantitative and qualitative data were entered into an electronic relational database (Microsoft Access). The format of this database was developed by the UMass research group for the current project. It includes tables with separate records for individual plants. These are cross-linked to similar tables for utilities and contacts, and finally to tables with sources and associated data. Figure 3.4 shows the basic input screen with contact information in the background. Figure 3.5 shows another part of the input fields, and Figure 3.6 lists the various headings on many of these fields.

## SUMMARY AND CONCLUSIONS

To help improve current understanding of the full range of precursor levels in U.S. waters, three types of data collection efforts were undertaken. The first stage of this data collection involved mailing a survey to 553 of the largest U.S. utilities that use surface water. The second stage added telephone followups and site visits to selected utilities from the first list. To obtain the most data from this work, information was also collected on organic carbon levels, source types, land use characteristics, and utility interest in these issues. Finally, an extensive literature survey was conducted, spanning the period from 1975 to the present. Results from this work are presented in chapter 4.



©2007 AwwaRF. All Rights Reserved.

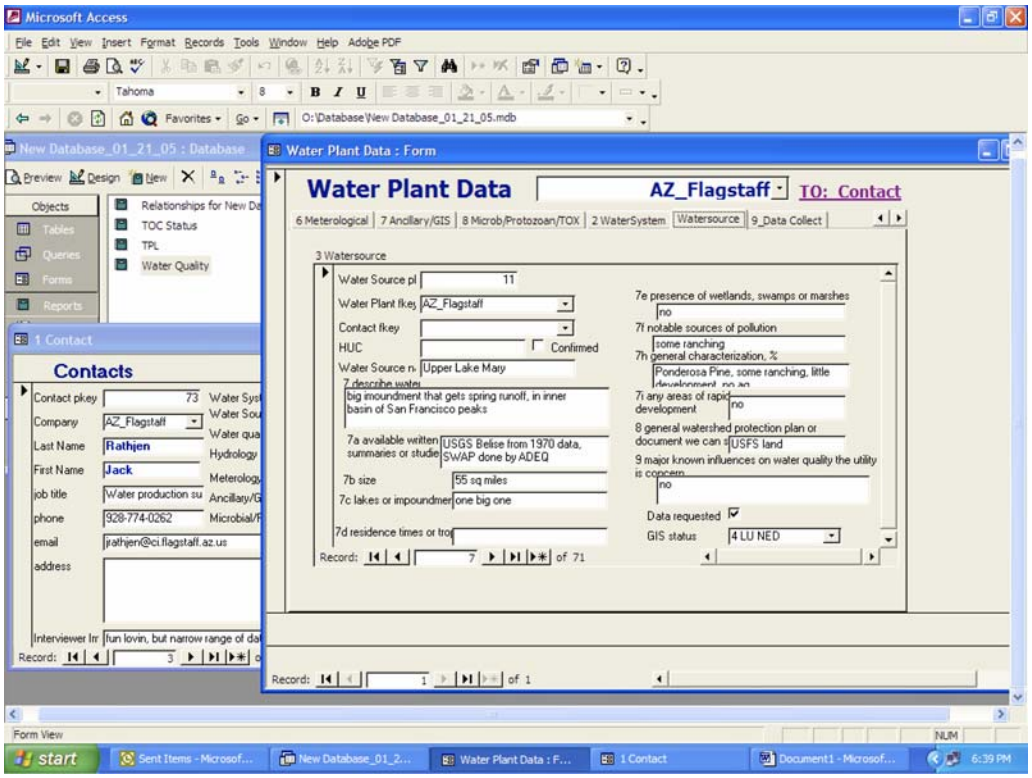



**Figure 3.5 Input form used for primary data collection: water source information**

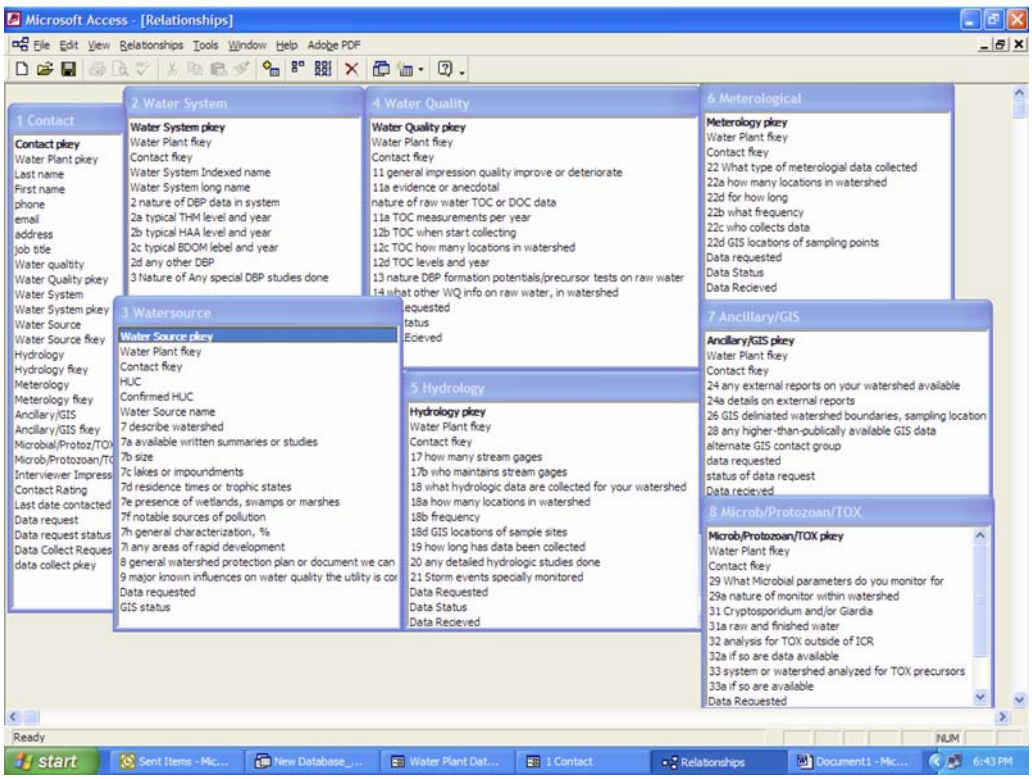



**Figure 3.6 A selection of data fields**



©2007 AwwaRF. All Rights Reserved.

A substantial number of utilities have collected DBP data beyond what has been mandated by regulatory agencies. About 17% of respondents had collected TOX data in addition to any that were required under the ICR. About 40% had collected data on DBP precursors in their raw water (including THM, HAA, and TOX precursors). Of these, 12 reported that they had collected precursor data prior to 1981. About 60% of the respondents reported that they had collected TOC or similar NOM-related data at locations in their watersheds. A slightly smaller number (~40%) reported that they had conducted watershed studies focused on storm events. The vast majority (90%) were willing to share their data. ICR data revealed that some chlorinating utilities had high ratios of known to unknown TOX (>50%) and some had quite low ratios (<15%). There was no obvious connection between unknown TOX level and geographic location, raw water TOC, or SUVA.

From these site-specific data, a set of about 50 primary utilities and about 120 additional secondary utilities were identified. Some of these were the focus of this research project; however, time and resources did not permit exploration of all candidate utilities. Much unpublished data exist in the databases of these utilities. Given their expressed interest in watershed management for precursor control, it is recommended that additional studies be conducted to better develop an understanding of relevant watershed processes. Any future studies regarding NOM in watersheds should make use of these utility lists and the accompanying database to best capture the full range of organic water qualities.



©2007 AwwaRF. All Rights Reserved.

©2007 AwwaRF. All Rights Reserved.

# CHAPTER 4
# ANALYSIS OF PRECURSOR LEVELS ACROSS NORTH AMERICA

## INTRODUCTION

Natural organic matter reacts quickly with aqueous chlorine to form an extensive mixture of halogenated and nonhalogenated by-products. These reactions occur in most U.S. drinking water treatment plants where chlorine is still the disinfectant of choice. Toxicologists and epidemiologists have slowly amassed a convincing body of evidence supporting chlorination by-products of NOM as a cause of cancer and other serious health problems in humans (e.g., see Xie 2003).

Efforts by water treatment scientists and engineers to minimize DBP formation have been partly successful. Advancements in process technology and control have reduced DBP formation by 50% or more in treatment facilities across the country. As these efforts begin to reach their practical limit, the field is now looking toward new approaches. Among these is the use of innovative treatment technologies intended to reduce TOC levels prior to chlorination and watershed management intended to reduce organic DBP precursors in source waters.

Testing and assessment of new technologies or management practices will be most effective if there is a good quantitative understanding of the DBP problem. This requires some insight into the chemical nature of DBP precursors and an understanding as to why certain waters tend to favor production of one DBP over another. It would also be helpful if the mean and typical range of DBP precursor levels were better known. Such knowledge would be useful in identifying outliers and unusual sources of DBP precursors. In this chapter, the current state of knowledge of DBP precursors is critically assessed and summarized. This includes tests from a broad geographical range and intensive studies of NOM fractions from a more limited set of field sites.

DBP formation is explored through analysis of existing data on whole waters (raw and treated) and supplemented with data on NOM extracts. The former exists in the literature in sufficient abundance to make a broad-scale statistical summary possible. By far, the majority of these data focus on THM formation, and for this reason, this analysis begins with that particular class of DBP. Other DBPs that will be examined are the HAAs. These are grouped into those having two halogens (the dihaloacetic acids, or DHAAs) and those having three halogens (the trihaloacetic acids, or THAAs). This is primarily in recognition of the differing precursor types that give rise to these groups (e.g., Reckhow and Singer 1985). The monohaloacetic acids (MHAAs) are not considered here because of the generally poor quality of MHAA data. In a few cases, other precursors are considered (e.g., dihaloacetonitriles, halonitromethanes), but the scarcity of these data makes the analysis less helpful.

The NOM extracts proved to be a bit more complicated than the whole waters. This is because NOM fractions can be extracted in different ways depending on the research group conducting the work. As a result, the few data sets on DBPs and NOM fractions that do exist in the literature are all based on slightly different methodologies. For this reason, the focus is given to the most detailed part of this analysis on NOM fractions prepared and treated only at UMass. This results in fewer data points. However, it gives a complete data set that includes THMs, HAAs, and minor by-products. The other part of this analysis considers all data pertaining to the NOM fractions based on partitioning with XAD resins. This particular fractionation protocol is judged uniform to yield useful results.



©2007 AwwaRF. All Rights Reserved.

## DBP PRECURSOR TESTS AND MODELS

### Approach to DBP Precursor Assessment

The quantitative measurement of DBP precursors must be done by means of an empirical precursor test. A wide range of conditions have been employed for precursor tests since their earliest use in the mid-1970s. All involve some bench-scale treatment of the sample with chlorine, a fixed reaction period, and then reduction of the residual chlorine followed by analysis of the DBPs, usually by gas chromatography. Researchers have not been in agreement about the exact conditions to be used (chlorine dose, pH, reaction time, temperature, etc.). Some of this disagreement occurs because of legitimate differences in the objectives for measuring precursor levels.

Most researchers would agree that the ideal precursor test should be capable of accurately measuring the total amount of precursor sites that are potentially accessible to reaction with the oxidant (usually free chlorine) under conditions typical of drinking water treatment. These tests should be completely independent of the inorganic matrix, and they should be reproducible and robust. Because organic matter in natural waters is quite dilute, a substantial amount of reaction between organic molecules would be unlikely. Consequently, each organic molecule will react independently of the others and in accordance with the kinetic laws that govern their elementary reactions. These kinetics are primarily second order (first in precursor organic, and first in chlorine species), or first order overall (in precursor organic). In reactions with model compounds (known simple organic molecules of high purity), a complex web of reactions that drifts from second order to first order can be seen as rapid bimolecular halogenation reactions subside and slower pathways involving some other rate-limiting step (e.g., hydrolysis) predominate.

These reactions always show direct proportionality between initial model compound concentration and by-product formation as long as all other conditions are held constant. Nonreactive solutes (e.g., $Na^+$, $Ca^{+2}$, $NO_3^-$) will naturally remain constant through a test of this sort, and pH can be held fixed with an appropriate buffer. Reactive inorganic species are more difficult to control. Ideally, one would like to remove all such species (except, of course, free chlorine) so that their concentrations always remain near zero. This is not practical, so it is important to measure their initial concentrations (e.g., bromide, iodide, ammonia, sulfide). Ideally, one would also like to have a "chlorine buffer" or means of holding the chlorine residual constant instead of allowing it to become depleted on reaction. At low chlorine levels, this requires too much intervention and potential loss of volatiles or introduction of contaminants to be practical. At very high chlorine levels (e.g., >10 times the demand), the relative depletion of chlorine is insignificant. However, such high levels are likely to access pathways that are not normally seen in typical drinking water systems. The approach has been to use a compromise value of 20 mg/L for waters with 3-day demands of 10 mg/L or less. This keeps the chlorine residual between 10 and 20 mg/L at all times, and thereby maintains a relatively constant chemical driving force that is not strongly influenced by the organic matter.

In contrast, the practice of using a fixed chlorine-to-carbon ratio typically results in a high initial residual that drops to some very low level at the end of the incubation time. Waters with a higher DOC and higher chlorine demand will require that the chlorine dose be higher. This is ideal when assessing the tendency of a particular water, with its unique DOC level to produce DBPs, but it is not best if one is interested in the inherent reactivity of the organic compounds making up the DOM. The latter is of greatest importance when considering the transport and inevitable dilution and degradation of NOM in watersheds.



©2007 AwwaRF. All Rights Reserved.

## Collecting New Data: Chlorination Conditions

For this work, NOM solutions and raw waters were chlorinated in the laboratory under the standard UMass conditions (20 mg/L dose, 72 hours, 20°C) as used with prior NOM and model compound studies. As already mentioned, the high-level fixed-dose testing protocol was adopted to minimize the chances that the residual would dip down to levels below 1 mg/L. Beyond this, the protocol imposes a chlorine residual profile that is relatively constant from one sample to the next (i.e., residual is initially 20 mg/L and rarely drops below 10 mg/L). This was done to avoid variations between tests in the reactive chemical environment (i.e., chlorine residual). At the end of the 72-hour incubation period, these samples were partitioned into a set of identical aliquots and analyzed for residual chlorine, HAAs, THMs, and other neutral extractable by-products. Controls were prepared by chlorinating high-purity water as previously described.

The high chlorine dose causes predictable shifts in the DBP profiles, the same as differences caused by pH, bromide, temperature, and reaction time. These can be accounted for by using the accumulated data on DBP response to varying reaction conditions. In this way, data collected by different groups using diverse protocols (or even data collected by the same group, but using different doses because of diverse TOC levels) can be compared on a near-equal basis. The following section explains the approach used for this correction or "standardization" process.

## Standardizing Literature Data: The Hybrid Model

One of the challenges of working with large, diverse DBP precursor data sets is the varying conditions that the organic matter is subject to during chlorination. The variables include those that are easily controlled such as temperature, pH, reaction time, and chlorine dose. However, differences in bromide level and chlorine residual not only change from test to test but will also change during the course of each test.

To control for such differences, all measured DBP levels were adjusted to that which would be expected from a standard chemical environment, where only the organic precursor content (quality or characteristics) affects the specific yields. This requires the use of a mathematical model for DBP formation. Models employing chemical kinetic rate laws and accounting for instantaneous concentrations of reactive species are preferred. However, these have not been widely used or calibrated. Although they do exist for a few waters (e.g., McClellan et al. 2000), they are not easily translated in their present form across a broad database such as the one used in this study.

Therefore, the more widely tested power function modeling approach was used. This approach is based on several assumptions, and it is important to be aware of the various limitations. First, the power function models (Equation 4.1) are generally developed with terms that contain a single independent variable each, with few or no interaction terms. This means that the model assumes affects due to changing pH, chlorine dose, time, temperature, and so forth, are uniform and unaffected by changes in other variables:

$$\text{DBP} = a(\text{Br} + b)^c(\text{pH} + d)^e(\text{temperature} + f)^g(\text{Cl}_2\ \text{dose})^h(\text{time})^i \tag{4.1}$$

The model parameters shown in Table 4.1 were arrived at in various ways. The literature was surveyed and models were critically assessed. The TTHM model was supported by the largest number of studies that used the biggest databases. In addition, all TTHM models encompassed the



©2007 AwwaRF. All Rights Reserved.

**Table 4.1**
**Model parameters used to adjust DBP yields**

| | Bromide (µg/L) | | pH | | Temperature (°C) | | Chlorine dose (mg/L) | Time (hr) |
|---|---|---|---|---|---|---|---|---|
| DBP group | Offset *b* | Exponent *c* | Offset *d* | Exponent *e* | Offset *f* | Exponent *g* | Exponent *h* | Exponent *i* |
| TTHM | 1 | 0.036 | –2.6 | 0.715 | 0 | 0.791 | 0.272 | 0.261 |
| THAA | 1 | 0.036 | 0 | –1.495 | 0 | 0.307 | 0.378 | 0.188 |
| DHAA | 1 | 0.036 | 0 | 0.259 | 0 | 0.377 | 0.346 | 0.186 |

full set of four THM species. Several published models were identified whose prediction space encompassed the range of conditions in the NOM fraction database (Amy, Chadik, and Chowdhury 1987; Owen, Amy, and Chowdhury 1993; Rodriguez and Serodes 2001; Rathbun 1996; Amy et al. 1998; USEPA 2005). In addition, one unpublished model based on an aquatic fulvic acid database (Reckhow and Makdissy 2007) was used. Because all were derived from a broad set of conditions, these models could be applied to the NOM fraction data. To avoid issues of bias, coefficients from all seven models were averaged when they could be compared on an equivalent basis (i.e., when there were no "offsets" or when the offsets were identical). Because of the low bromide levels in most hydrophobic NOM fractions, it was necessary to have a model with a positive bromide offset (to avoid the problem of bromide levels near zero). Thus, the bromide coefficients from Amy, Chadik, and Chowdhury's (1987) model were adopted without attempting to blend in the other model predictions.

A different approach had to be taken for the HAAs, because these models were less numerous and generally not inclusive of all brominated forms. In addition, there was some concern about the accuracy of the earliest studies. In general, only four models were combined to obtain the final set of coefficients (Amy et al. 1998; Reckhow, Tobiason, and Edzwald 1999; Malafronte 2003; Reckhow and Makdissy 2007). However, none of these models were based on a large data set, including the full suite of brominated forms. Instead, the bromide model for the THMs was incorporated without modification. Roberts and colleagues have shown that bromine incorporation into the HAAs follows a quantitative pattern quite similar to that of the THMs (Roberts, Singer, and Obolensky 2002). This was recently verified using data from the National Occurrence Survey (Weinberg et al. 2002) and was used as justification for adopting bromine incorporation factors from the THMs for estimating HAA speciation.

In many of the graphical presentations that follow, DBP precursor levels are shown based on the standard conditions used at UMass. To help put these numbers into context, a left offset scale is included that allows translation of these numbers into uniform formation conditions (UFC) values, sometimes used as standard SDS conditions. The particular conditions that these tests refer to are summarized in Table 4.2. In many of these figures there is also a right-hand vertical scale with the "pre-exponential value." This is the *a* value in the power function model (Equation 4.1).



©2007 AwwaRF. All Rights Reserved.

**Table 4.2**
**Reference conditions for DBP formation tests**

| Condition | Units | Formation potential | UFC (SDS) |
|---|---|---|---|
| Chlorine dose | mg/L | 20 | 1.5 |
| pH | | 7.0 | 8.0 |
| Temperature | °C | 20 | 20 |
| Reaction time | hr | 72 | 24 |
| Bromide | µg/L | 0 | 50 |

## WHOLE WATERS DATABASE

### Correlations From the Literature

Several researchers have examined THM formation in samples from a given water source over extended periods of time. In one of the most complete studies of this type, Edzwald and co-workers examined raw and treated waters from the Canton and Oneida, N.Y., water treatment plants for a period of more than 2 years (Edzwald, Becker, and Wattier 1985). In both cases, they found strong linear correlations between THMFP and TOC (Equations 4.2 and 4.3). The slopes reflect the intrinsic THM precursor content (micrograms per milligram of carbon), but in both cases there was a positive intercept.

$$\text{THMFP} = 75.8\,(\text{TOC}) + 13.6 \quad \text{(Grasse River)} \tag{4.2}$$

$$\text{THMFP} = 45.1(\text{TOC}) + 113.4 \quad \text{(Glenmore Reservoir)} \tag{4.3}$$

Several researchers have also examined THM precursor levels across a broad geographic distribution of sites. One of the largest of these studies encompassed 133 data points from five literature sources (Chapra, Canale, and Amy 1997). They concluded that the THMFP of this broad spectrum of natural waters increases with TOC raised to the power of about 1.25 (Equation 4.4). This implies that low-TOC waters have a specific THMFP (THMFP/TOC) of about 45 µg/mg C (carbon; at ~1 mg/L TOC), whereas higher TOC waters (~10 mg/L TOC) would be closer to 85 µg/mg C.

$$\text{THMFP} = 43.78(\text{TOC})^{1.248} \tag{4.4}$$

which can be rearranged to

$$\frac{\text{THMFP}}{\text{TOC}} = 43.78(\text{TOC})^{0.248} \tag{4.5}$$

This suggests that the specific THMFP increases with increasing TOC. However, these data were not corrected in any way for differences in chlorination testing conditions. Of the five



©2007 AwwaRF. All Rights Reserved.

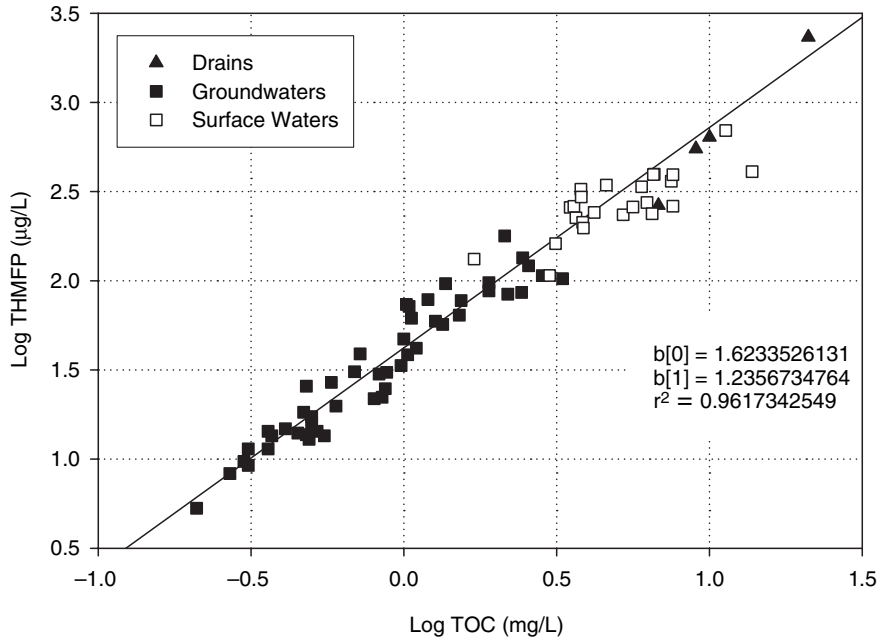



**Figure 4.1 Retesting of four databases used by Chapra, Canale, and Amy (1997)**

databases used by these authors, four were readily available in the refereed literature (Amy, Chadik, and Chowdhury 1987; Amy et al. 1990; Edzwald, Becker, and Wattier 1985; Miller et al. 1990). Figure 4.1 shows the data from these four sources, and a linear correlation of the log transform gives a model (Equation 4.6) that is quite close to the one reported by Chapra, Canale, and Amy (1997; Equation 4.4), despite the missing fifth data set.

$$\text{THMFP} = 42.01(\text{TOC})^{1.236} \tag{4.6}$$

Direct comparison of DBP formation potential (DBPFP) data across research groups is risky when multiple sets of chlorination conditions are used in the compiled data set. In addition, raw DBPFP data can be heavily influenced by the TOC of the water being tested, when chlorine doses are positively correlated with TOC. This is a problem even when comparing data obtained from the same DBPFP testing protocol. To a lesser extent, varying raw water bromide levels can create apparent differences in the reactivity of organic matter. For this reason, it is helpful to normalize data based on treatment conditions, TOC, and perhaps bromide levels. As a first approximation, the power function model developed by Amy, Chadik, and Chowdhury (1987) was selected to achieve some level of normalization for these data (Equation 4.7). This particular model was selected because much of the data that is in the Chapra, Canale, and Amy (1997) database is identical or similar to the data used to develop this model. Subsequent literature analysis will make use of the hybrid model previously discussed. The normalized conditions adopted for this analysis were standard UMass conditions: 20 mg/L chlorine dose, 72-hour reaction time, pH 7, no bromide, and 20°C. These values, as well as the model coefficients and offsets, are represented in Equation 4.7:



©2007 AwwaRF. All Rights Reserved.

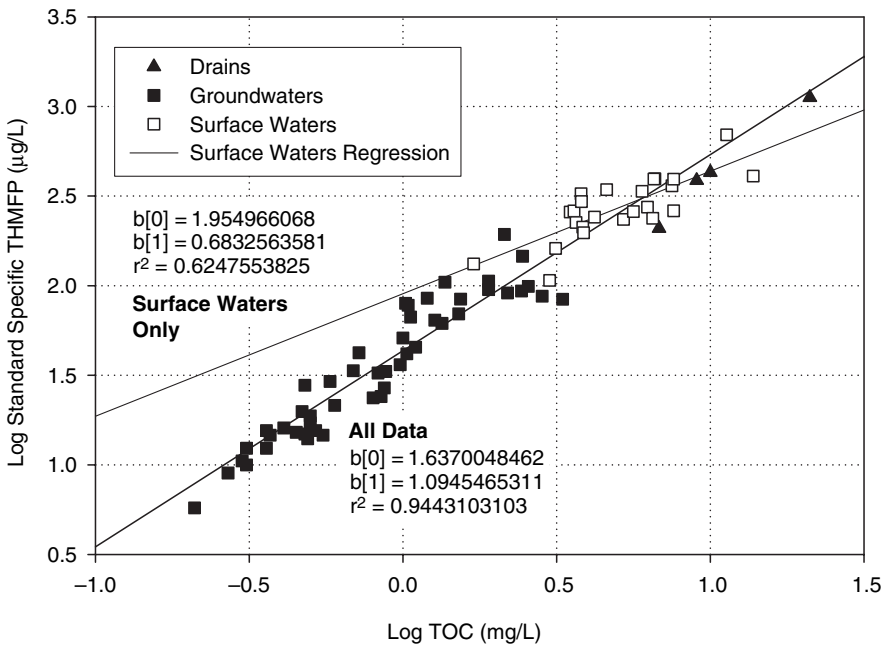



**Figure 4.2 Re-evaluation of data from Figure 4.1 based on a common chlorination protocol**

$$\text{std.THMFP} = \text{THMFP}\left(\frac{20}{\text{Cl}_2\text{ dose}}\right)^{0.409}\left(\frac{72}{\text{time}}\right)^{0.265}\left(\frac{7-2.6}{\text{pH}-2.6}\right)^{0.715}\left(\frac{1}{\text{Br}+1}\right)^{0.036}\left(\frac{20}{\text{temperature}}\right)^{1.06} \tag{4.7}$$

Figure 4.2 shows some of these data after manipulation to bring them all to the same set of formation potential conditions. The THMFPs for the higher TOC waters were readjusted downward more than the others, because of the higher chlorine doses necessary for testing these waters and, therefore, the artificially high reactive environment experienced by the organics in these waters. As a result, the slope of the overall regression line dropped to 1.09 and the standard specific (std.sp.) THMFP is very weakly correlated to TOC, as in Equation 4.8.

$$\text{std. sp. THMFP} \equiv \frac{\text{THMFP}}{\text{TOC}} = 43.4(\text{TOC})^{0.095} \tag{4.8}$$

Because all of the earlier data sets were exclusively populated with surface waters, the most appropriate comparison for the data in Figure 4.2 should also involve just the surface waters. Excluding all groundwaters and agricultural drains, the regression slope drops to 0.68. When plotted as specific THMFP on a linear scale, the surface water data resemble much more closely those correlations developed for the two Massachusetts watersheds (Figure 4.3). Again, the groundwaters and agricultural drains do not fit this relationship, and possible reasons for this will be discussed in the following paragraphs.



©2007 AwwaRF. All Rights Reserved.

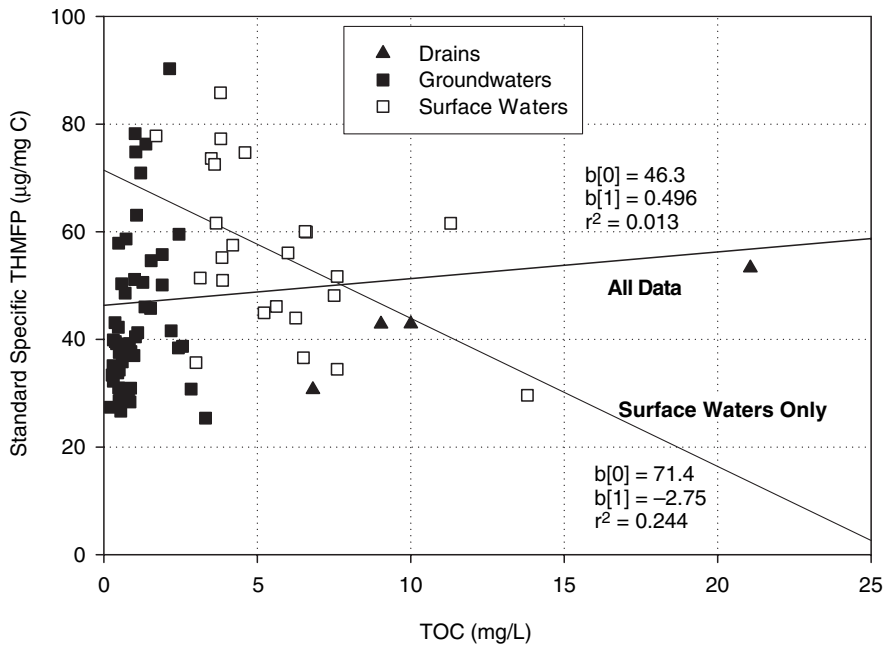



**Figure 4.3 Standard specific THMFP versus TOC for the four data sets from Chapra, Canale, and Amy (1997)**

## Full North American Database

Many other good databases can be incorporated into this analysis. Figure 4.4 shows the results of a compilation of 26 data sets that are accessible in the open literature (Reckhow and Singer 1990; Amy et al. 1990; Summers et al. 1996; Krasner et al. 1996; Symons et al. 1994; Sinsabaugh et al. 1986; Garvey and Tobiason 2003a, b; Edzwald, Becker, and Wattier 1985; Amy, Chadik, and Chowdhury 1987; Amy et al. 1987, 1992; Chadik and Amy 1983; Glaze and Wallace 1984; Randtke et al. 1988; Fujii et al. 1998; Waldron and Bent 2001; Miller et al. 1990; Oliver and Lawrence 1979; Edzwald, Gong, and McGowan 1981; Childress et al. 1999; Chang and Singer 1991; Thompson et al. 1988; Shukairy, Miltner, and Summers 1994; Siddiqui et al. 1994; McClellan et al. 2000). Again, the regression based on these data indicates that specific THMFP drops with increasing TOC. For clarity, Figure 4.5 shows only the surface water data.

Figures 4.4 and 4.5 show a cluster of data almost entirely residing between 20 and 120 µg/mg C under formation potential conditions or between 10 and 60 µg/mg C under SDS conditions. These diverse data fail to show any real correlation between the natural water TOC level and the specific THM precursor level of that NOM. When comparing with SUVA, there is a slight positive correlation (Figure 4.6), but the slope and goodness of fit are weak.

Although the North American database has fewer HAA data, sufficient data exist to make some useful generalizations. Figures 4.7 and 4.8 suggest that there is a negative correlation between specific HAA precursor content and TOC. This is especially convincing for the DHAAs. Possible reasons for this will be discussed later.

One final comparison using the North American database is shown in Figure 4.9. Here the ratio of THAA to THM from the various precursor tests is plotted versus the particular water's TOC value. This is believed to reflect water's tendency to form HAAs over THMs from the



©2007 AwwaRF. All Rights Reserved.

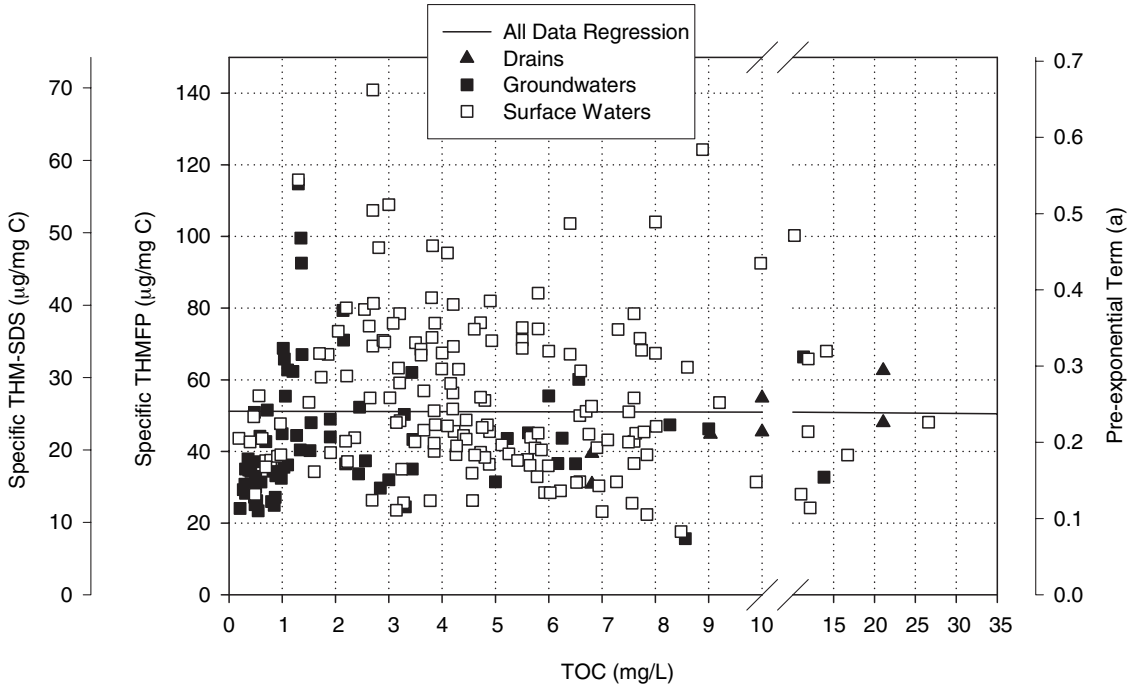



**Figure 4.4 Standard specific THMFP versus TOC for all 27 data sets from the literature**

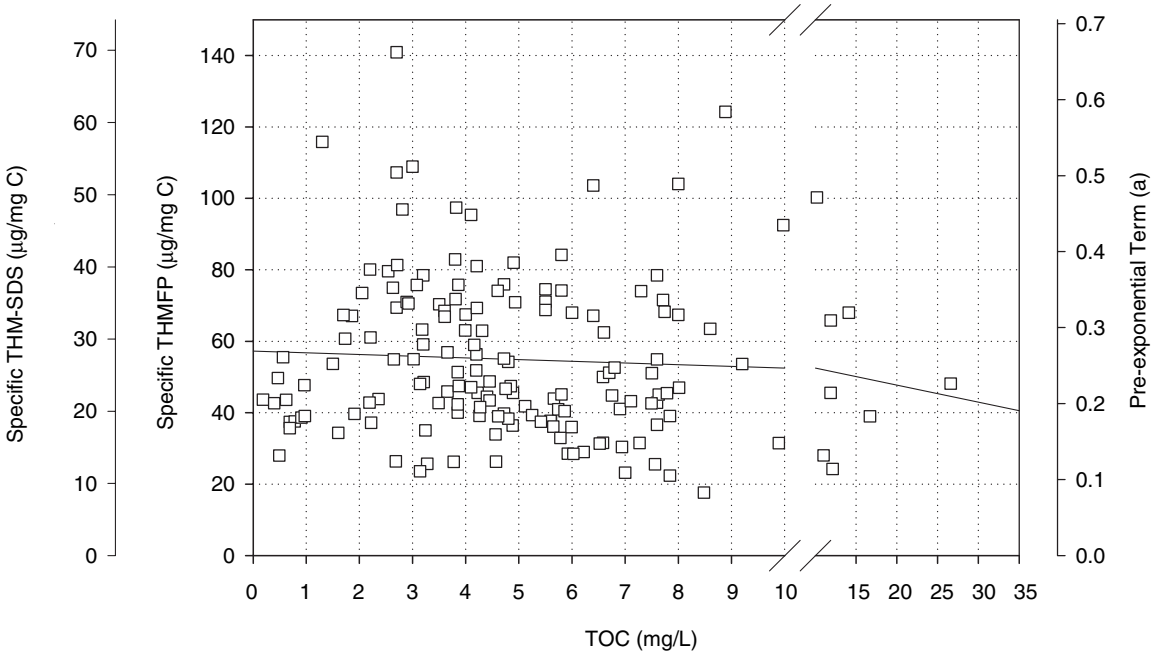



**Figure 4.5 Standard specific THMFP versus TOC for surface waters based on all 27 data sets from the literature**



©2007 AwwaRF. All Rights Reserved.

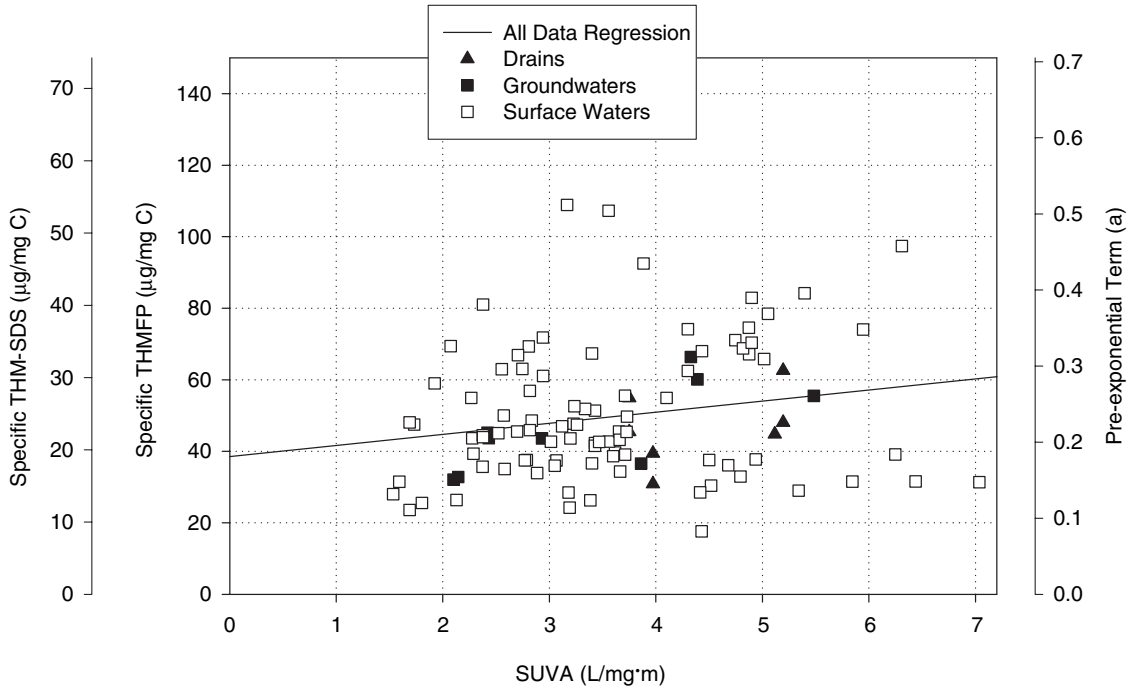



**Figure 4.6 Standard specific THMFP versus SUVA for all data sets from the literature**

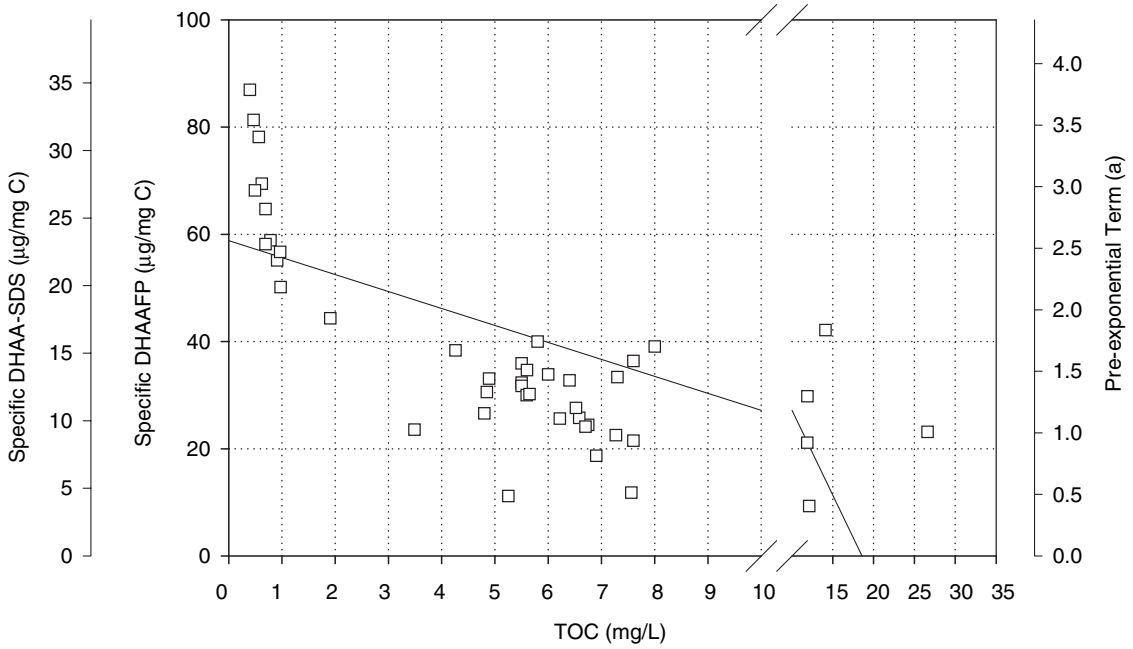



**Figure 4.7 Standard specific dihaloacetic acid formation potential (DHAAFP) versus TOC for all data sets from the literature**



©2007 AwwaRF. All Rights Reserved.

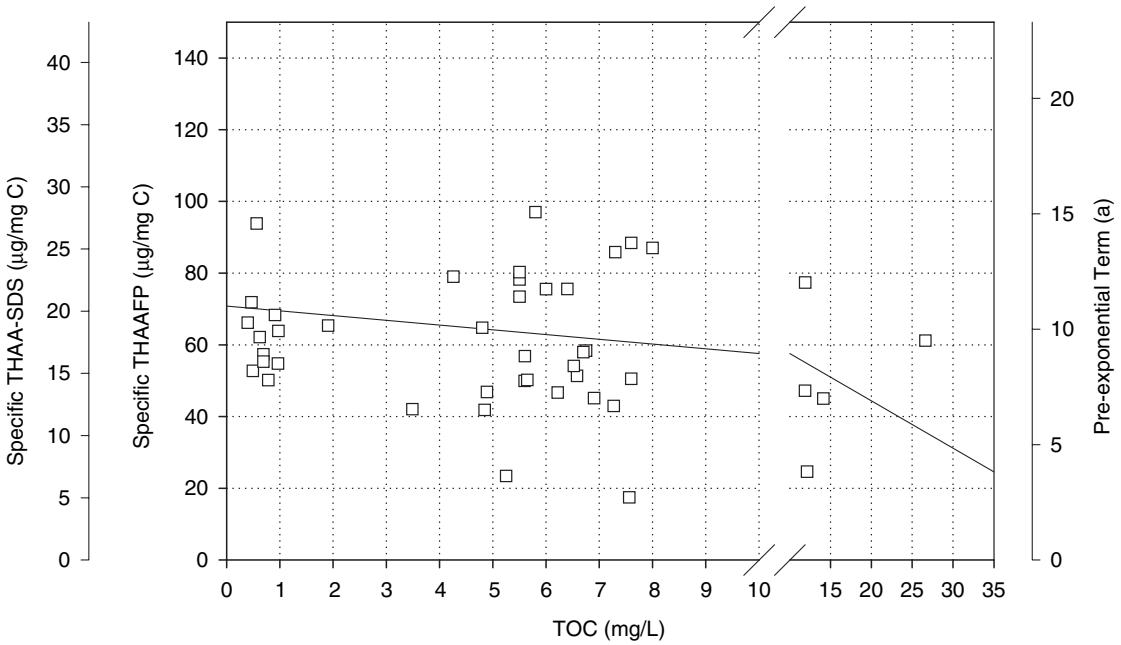
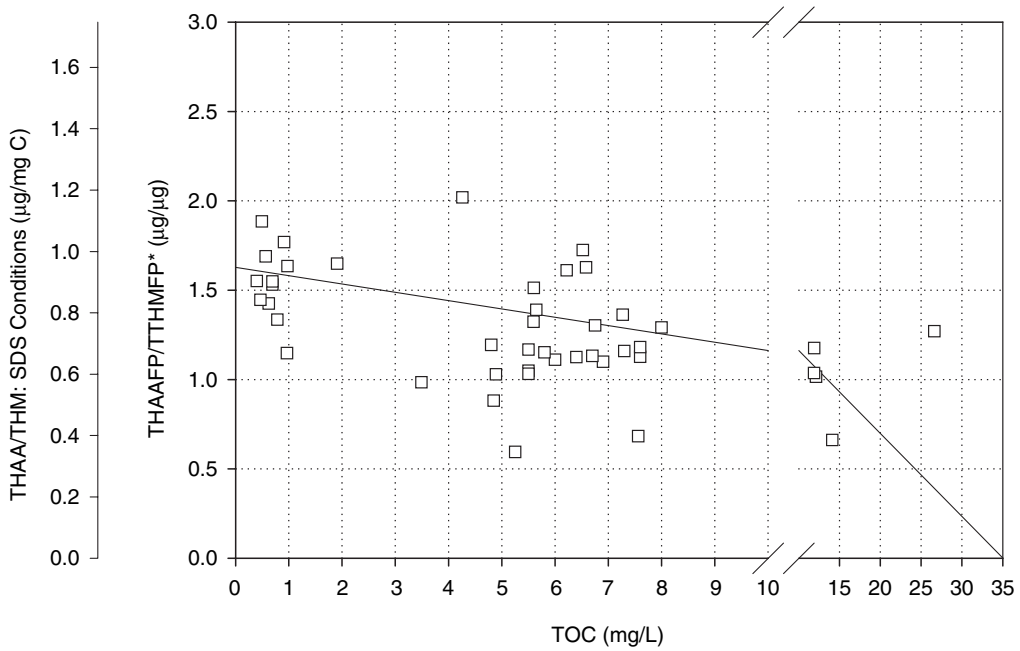



**Figure 4.8 Standard specific trihaloacetic acid formation potential (THAAFP) versus TOC for all data sets from the literature**




* TTHMFP = total trihalomethane formation potential

**Figure 4.9 Ratio of THAA precursors to THM precursors versus TOC for all data sets from the literature**



©2007 AwwaRF. All Rights Reserved.

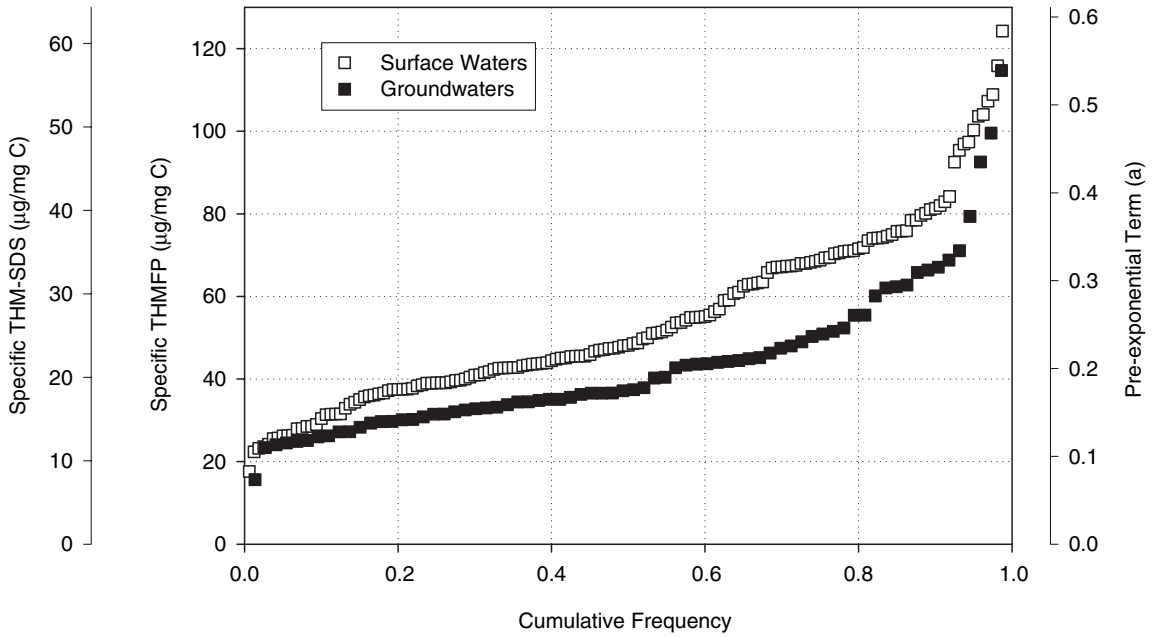



**Figure 4.10 Cumulative frequency plot for specific THM precursor content of all surface and groundwaters**

common trichloromethyl intermediate (e.g., see Reckhow and Singer 1985). Figure 4.9 implies that the average nature of NOM in high TOC waters is more likely to favor the THM pathway than the THAA route.

The slope in Figure 4.4 is not statistically different from zero, so that entire population should be considered independent of the TOC concentration. Using the Kolmogorov–Smirnov test, it can be shown that the standard specific THMFP for the groundwaters failed to resemble a normal distribution. However, the log-transformed data passed the test. Although the slope for the surface waters was significantly different from zero, these data were characterized as a single population as was done for the groundwaters. Again, the untransformed surface water data failed the test for a normal distribution, but it did prove to be log normal. The two population means are statistically different as determined by the Mann–Whitney rank-sum test.

A more direct and intuitive way of presenting the data is through a cumulative frequency plot as found in Figure 4.10. This reinforces the observation that groundwaters generally have lower specific formation potentials, though not much lower. In addition, one can extract medians and population percentiles from these data (Table 4.3). The median value for surface waters is about 48 µg/mg C.

### Field Studies in the Wachusett Watershed

Wachusett Reservoir is the primary source of drinking water for Boston and 44 eastern Massachusetts communities (2.2 million consumers). Created in 1897 by impounding the Nashua River, Wachusett Reservoir and its surrounding watershed have seen steady encroachment by development in the Worcester area. The two major tributaries feeding Wachusett are the Stillwater and Quinapoxet rivers. These rivers and several other minor rivers have been monitored by the



©2007 AwwaRF. All Rights Reserved.

**Table 4.3**
**Population characteristics for THMFP in surface waters**

| Cumulative frequency | THMFP (µg/mg C) |
|---|---|
| 10th percentile | 30.4 |
| 25th percentile | 39.1 |
| 50th percentile | 48.2 |
| 75th percentile | 68.8 |
| 90th percentile | 81.3 |

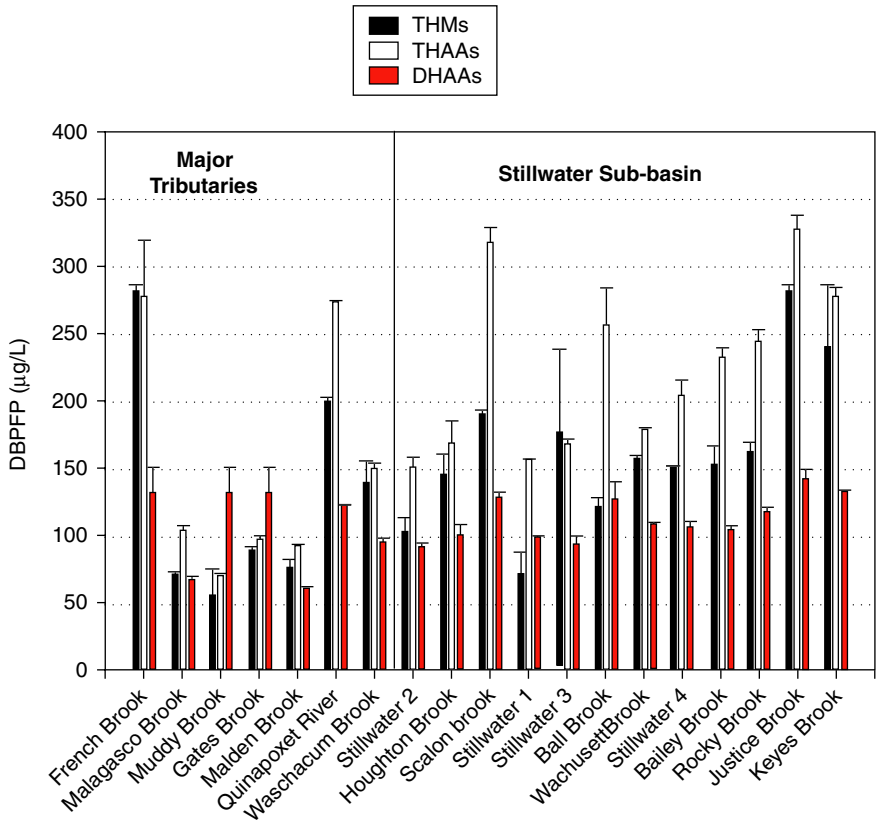



**Figure 4.11 Formation of major DBPs from chlorination of samples collected from the Wachusett watershed**

UMass research team for several years. Through this effort, a long-term database was created for DBP precursors in this central Massachusetts watershed.

Figure 4.11 shows a typical data set from a single dry-weather sampling event. There is substantial variability in the level of DBPs formed from the major tributaries as well as in the channel locations and small feeder streams to the Stillwater River. Much of this variability is quite clearly due to differences in the TOC level. However, analysis of specific DBP precursor



©2007 AwwaRF. All Rights Reserved.

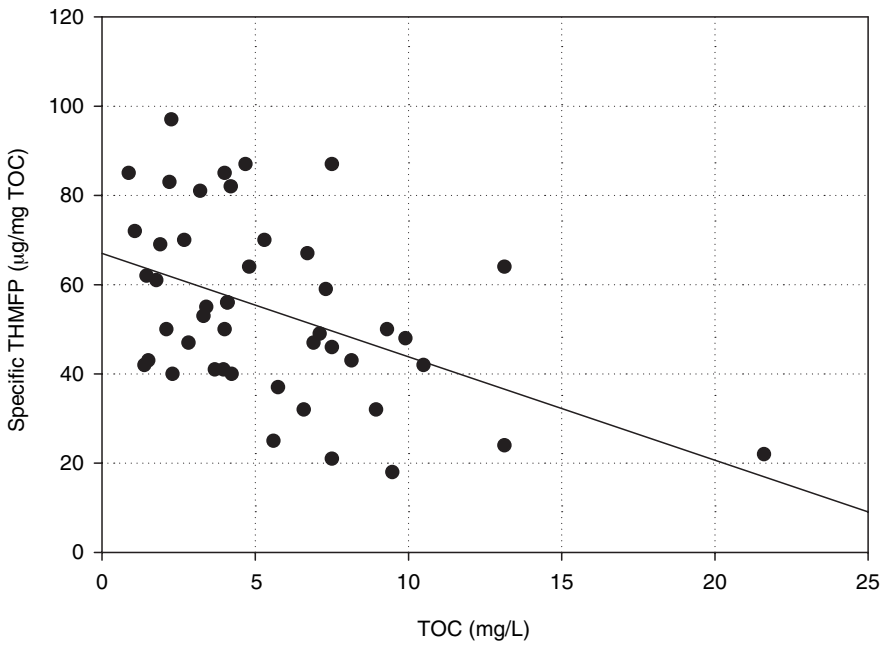



**Figure 4.12 Relationship between specific THM precursor levels and TOC in Wachusett Reservoir tributaries**

concentrations (precursor levels in micrograms per liter divided by the TOC in milligrams per liter) indicates there is still a substantial variability that cannot be attributed to TOC alone.

Figure 4.12 presents the specific THM precursor levels versus TOC for more than a half-dozen sampling campaigns spanning almost 3 years. First, the values run from as low as 20 µg/mg C to almost 100 µg/mg C. Second, there is a distinct decreasing trend with increasing TOC.

SUVA is not a very good indicator of THM precursor content in these waters (Figures 4.13 and 4.14).

When the Wachusett watershed data are superimposed over the entire North American database, there appears to be extensive overlap (Figures 4.15 and 4.16). This suggests that the Wachusett NOM is not uniquely different from NOM found elsewhere in the United States and Canada.

Another interpretation from Figures 4.15 and 4.16 is that the Wachusett watershed is capable by itself of producing NOM with nearly the full range of character (as measured by specific THMFP) as the entire North America continent. This surprising result implies that geography and ecoregion distinctions are not primary determinants of NOM character. Rather, hydrology and microenvironments are the overwhelming factors.

Figures 4.17–4.19 present evidence for the same conclusion based on DHAA and THAA precursors.

## DATABASE OF NOM FRACTIONS

This section presents a summary of data on NOM fractions collected from the open literature and from some previously unpublished UMass studies. The purpose is to gain insight into the reactivity of NOM from different locations and to provide population statistics that can be used to place new NOM data into context. The first part concerns a wide range of NOM fractions from a



©2007 AwwaRF. All Rights Reserved.

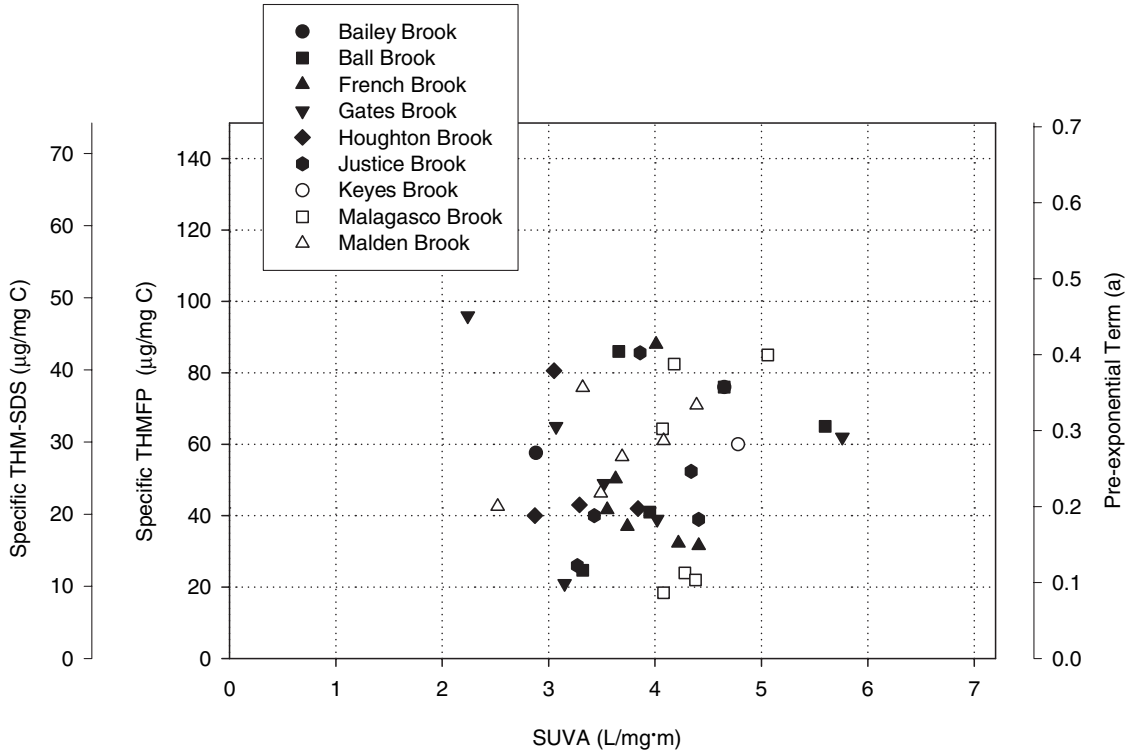



**Figure 4.13 Relationship between THM precursors and SUVA for samples from Wachusett Reservoir watershed (part 1)**

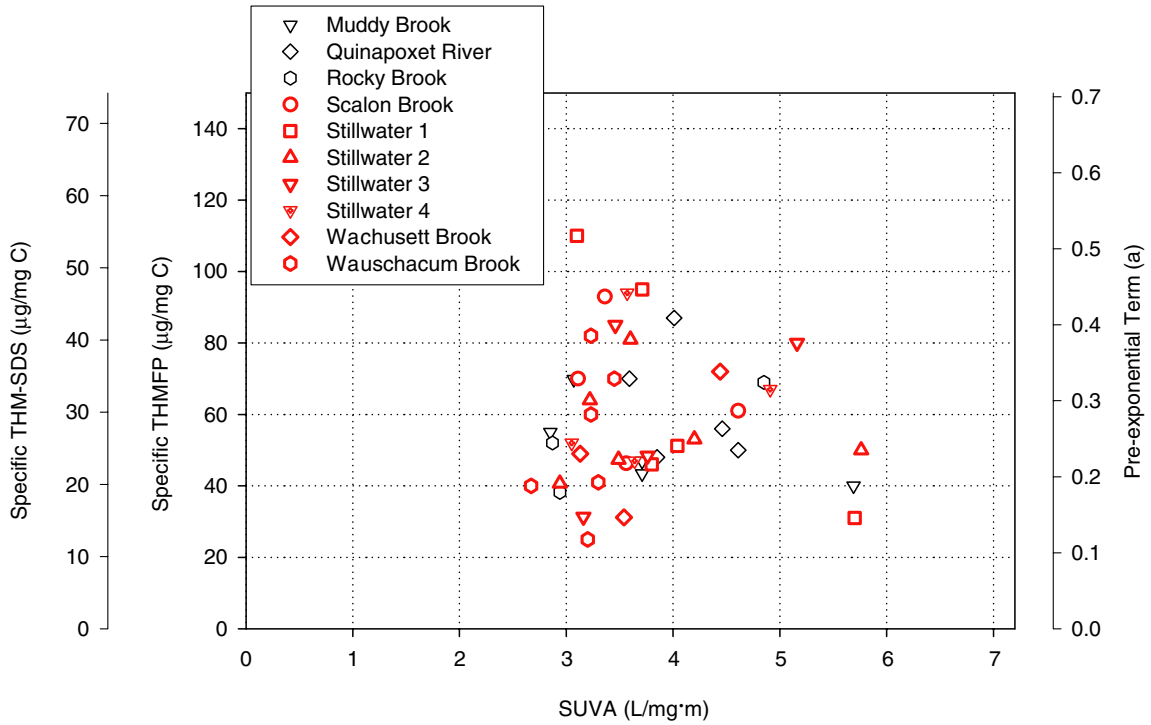



**Figure 4.14 Relationship between THM precursors and SUVA for samples from Wachusett Reservoir watershed (part 2)**



©2007 AwwaRF. All Rights Reserved.

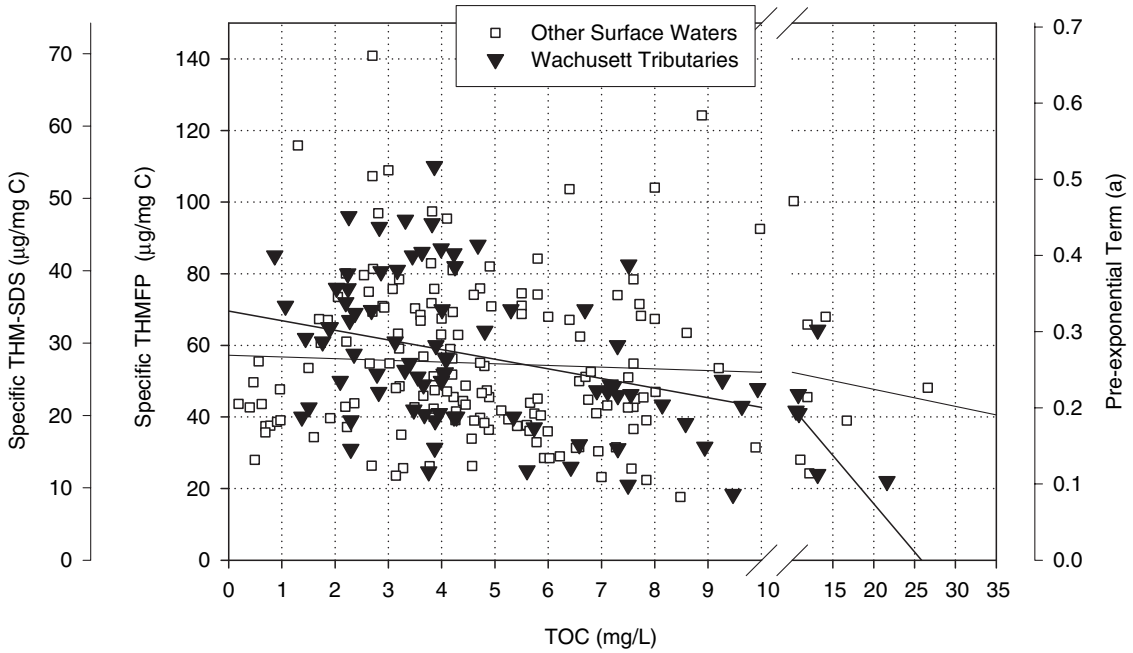



**Figure 4.15 Specific THM precursor levels in Wachusett samples as compared to the North American surface water database, all plotted versus the TOC**

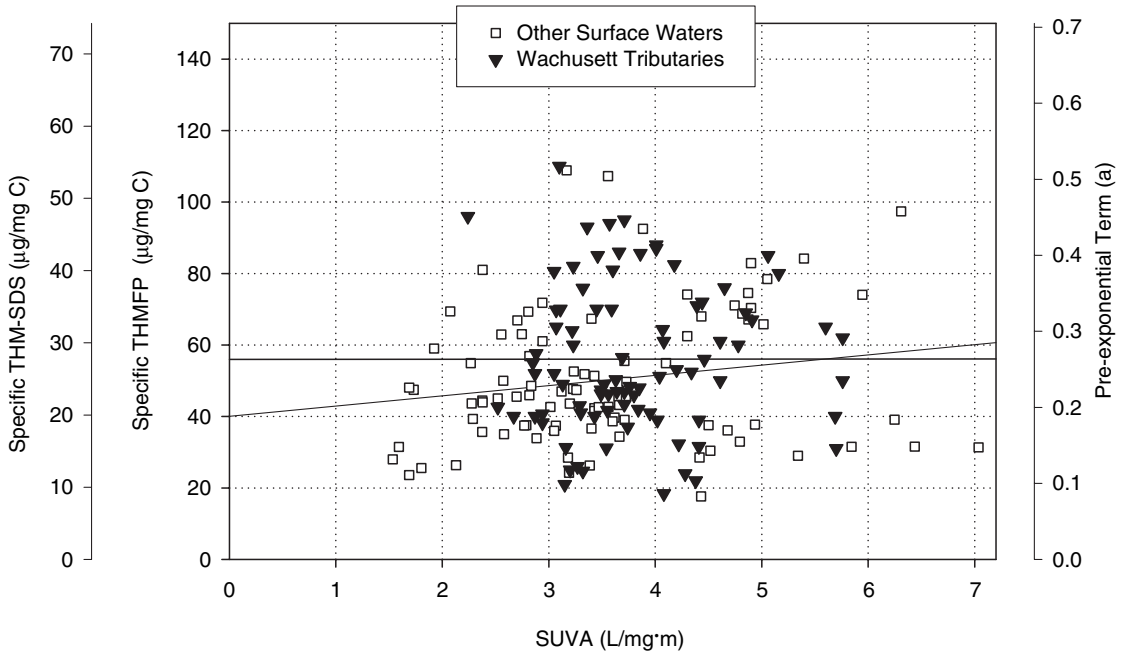



**Figure 4.16 Specific THM precursor levels in Wachusett samples as compared to the North American surface water database, all plotted versus the SUVA**



©2007 AwwaRF. All Rights Reserved.

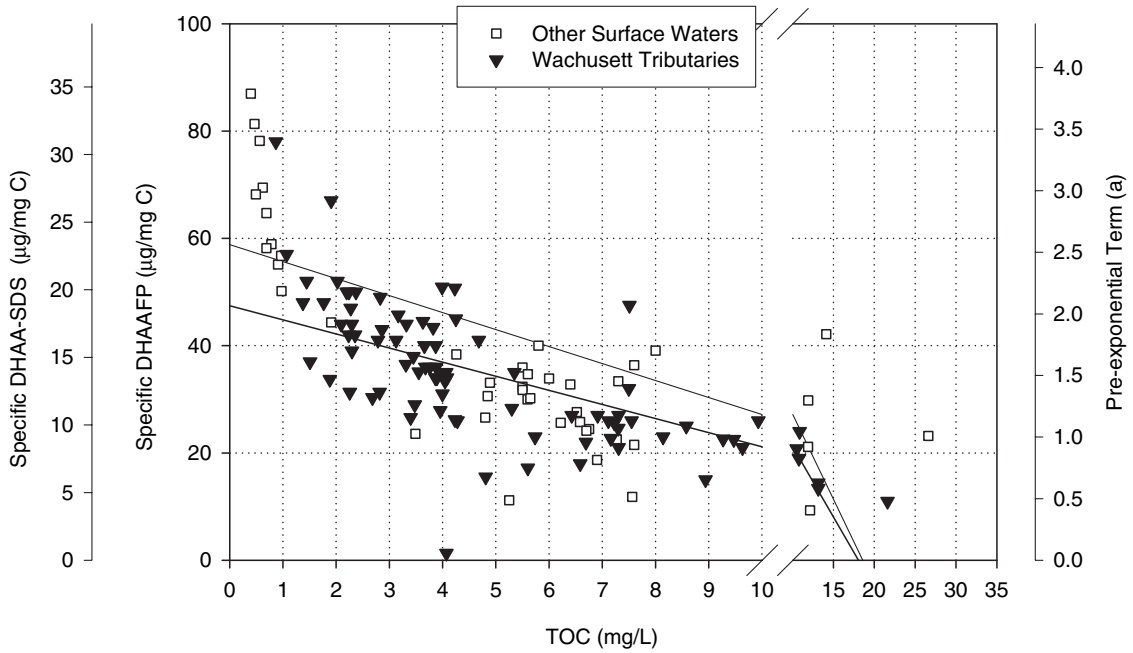



**Figure 4.17 Specific DHAA precursor levels in Wachusett samples as compared to the North American surface water database, all plotted versus the TOC**

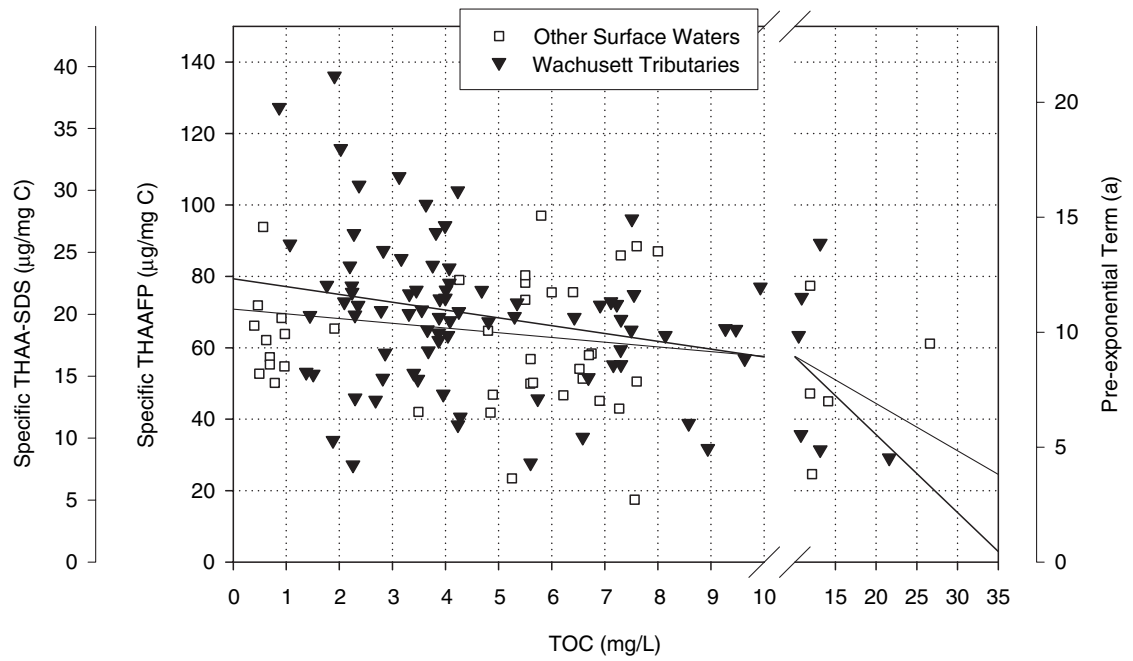



**Figure 4.18 Specific THAA precursor levels in Wachusett samples as compared to the North American surface water database, all plotted versus the TOC**



©2007 AwwaRF. All Rights Reserved.

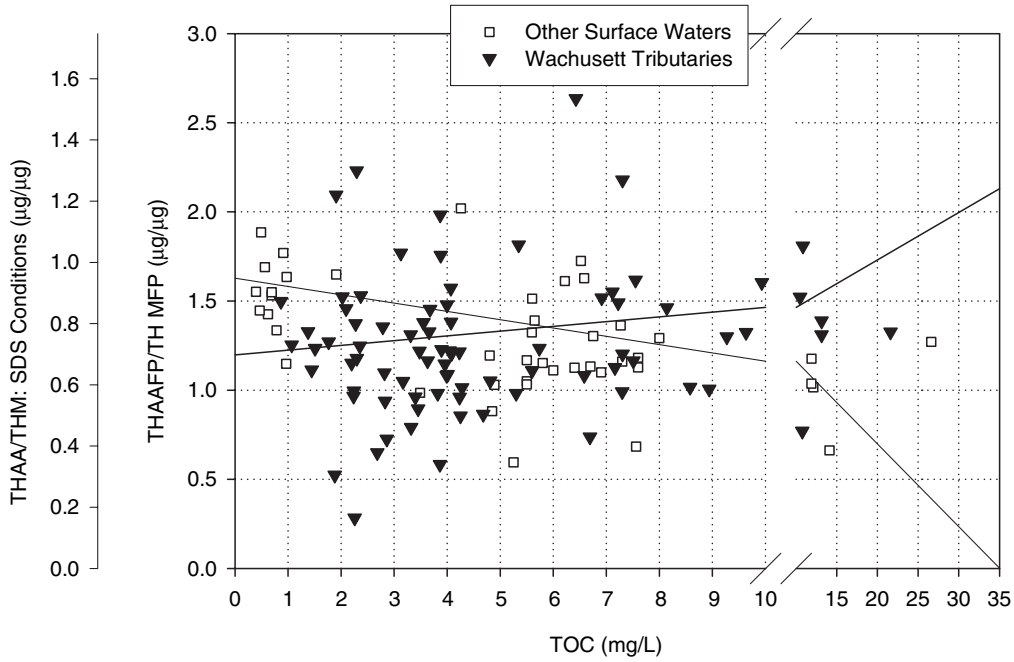



**Figure 4.19 Ratio of THAA to THM precursors in Wachusett samples as compared to the North American surface water database, all plotted versus the TOC**

limited number of sites and collected at only one laboratory. The second section attempts to include discussion about all literature on this subject but narrows the scope to include only the major solubility-based fractions.

### DBP Precursor Content in a Coherent Set of NOM Fractions

In this section, a single set of internally consistent NOM fractions are compared. All were isolated at UMass using very similar protocols and then were tested for DBP formation using identical protocols. These NOM fractions were all derived from raw and treated waters collected at various locations in Massachusetts and Connecticut. All fractions were dissolved in buffered water at DOC concentrations between 3 and 5 mg/L. They were then treated under the standard UMass chlorination conditions as previously described.

SUVA is the absorbance of a sample measured at 254 nm and normalized to a concentration of 1 mg/L DOC. This is becoming a widely used measure of NOM quality. It is easy and inexpensive to measure, and it is well correlated to various treatment-related parameters. For this reason, nearly all of the data in this section will be presented as a function of SUVA.

Specific chlorine demand ranges from less than 0.5 mg/mg C to as much as 7.8 mg/mg C (Figure 4.20). There is a weak correlation with SUVA, which is largely dependent on the weight of the humic acid fraction. Some very weak absorbers are also strongly reactive with chlorine. The possibility of a confounding inorganic demand cannot be ruled out in these cases.

When looking at the THM formation, there appears to be a stronger positive correlation with SUVA (Figure 4.21). Clear clusters of NOM fractions also exist, suggesting a common reactivity for the same fractions from different sources (Figure 4.22). In general, the neutrals and bases are quite



©2007 AwwaRF. All Rights Reserved.

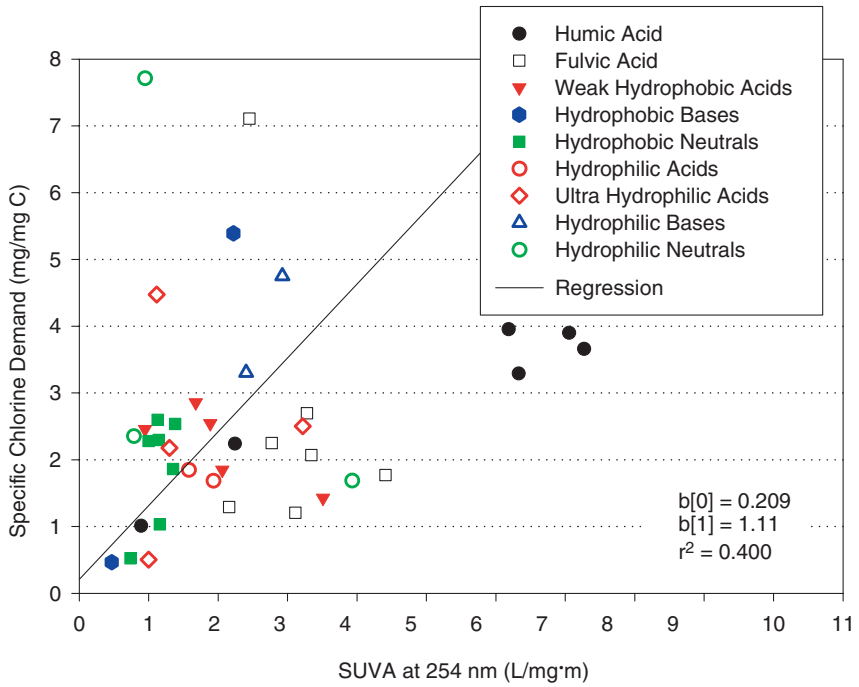



**Figure 4.20 Relationship between specific chlorine demand and SUVA for isolated NOM fractions from New England raw and treated waters**

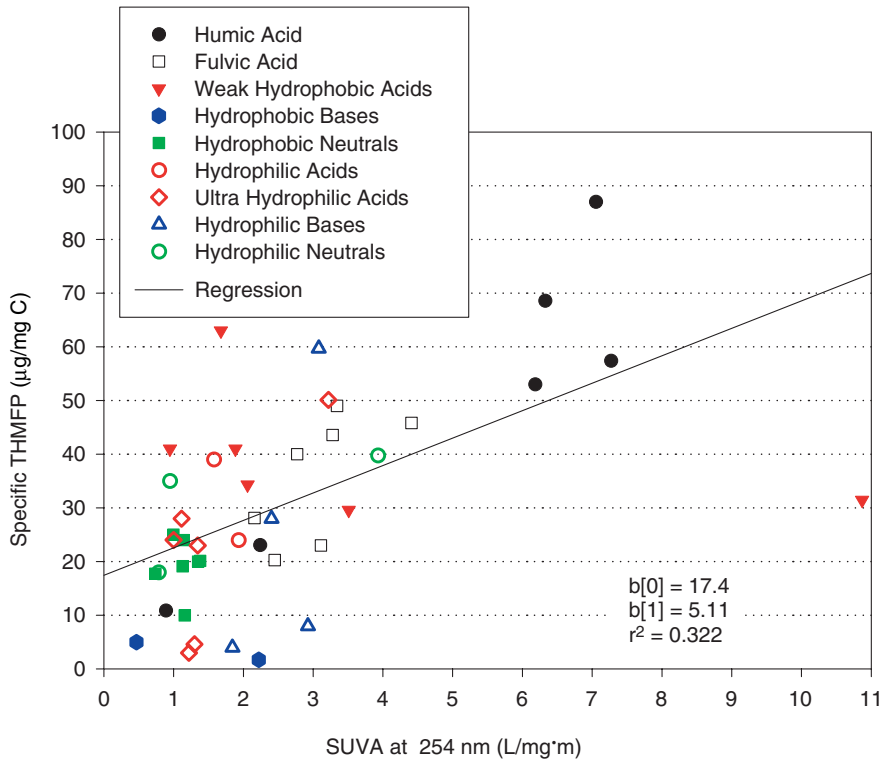



**Figure 4.21 Relationship between specific THM precursor content and SUVA for isolated NOM fractions from New England raw and treated waters**



©2007 AwwaRF. All Rights Reserved.

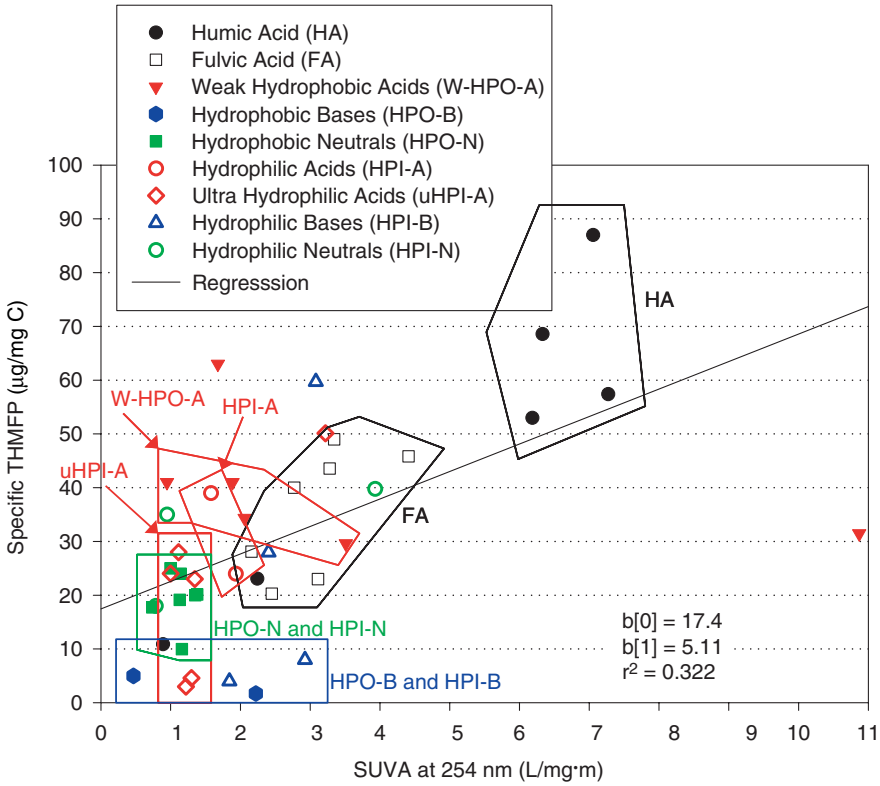



**Figure 4.22 Relationship between specific THM precursor content and SUVA for isolated NOM fractions from New England raw and treated waters, including boxed NOM domains**

unreactive. Most do not exceed 25 µg/mg C, which places them at the lowest end of the surface water data (compare Figures 4.22 and 4.10). Most of the hydrophilic acids and weak hydrophobic acids are only slightly more reactive. In terms of mean reactivity, it is the fulvic acids that most resemble the average whole surface waters. As expected, the humic acids are even more reactive; however, their abundance in surface waters is generally low.

Of all the DBP precursors, the THAAs exhibit the largest range in concentration (Figure 4.23). They also show a strong correlation with SUVA. The mechanistic explanation for this can be found in the literature (e.g., Reckhow and Singer 1985). A high degree of conjugation facilitates electron transfer from the site of THAA cleavage to the reactive sites undergoing oxidative attack. This leads to faster THAA formation and, therefore, less time for hydrolysis (which yields THMs). Of course, the conjugated bonds also result in higher absorbance in the UV range (SUVA).

The DHAA precursors also show a strong correlation with SUVA (Figure 4.24). This is more surprising based on observations with treated drinking waters and is not easily explained with mechanistic reasoning. Nevertheless, for the DHAAs there is almost no difference between the yields for fulvic acid as compared to the hydrophilic and weak hydrophobic acids. This is in contrast to the THAAs and seems to explain why coagulated waters (i.e., waters with the fulvic acid largely removed) produce higher ratios of DHAA/THAA than do raw waters.

Dihaloacetonitriles (DHANs) are necessarily produced by chlorination of nitrogen-containing organic compounds. Figure 4.25 supports this as the base fractions show higher levels of DHAN formation potential (DHANFP) as a group than any other. Humic acids are still relatively



©2007 AwwaRF. All Rights Reserved.

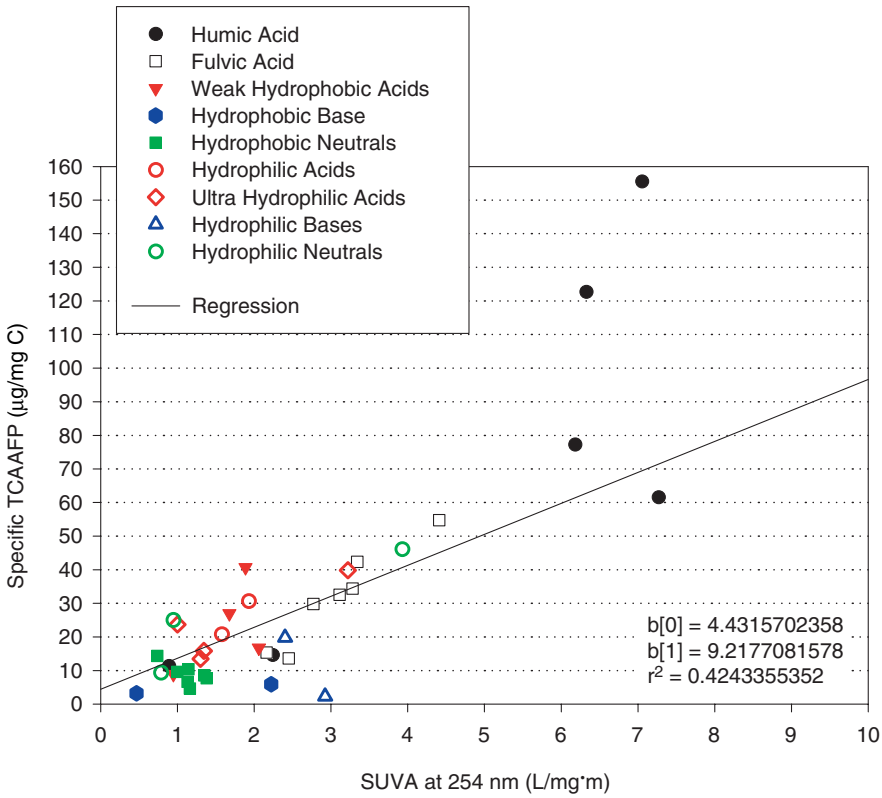



**Figure 4.23 Relationship between specific TCAA precursor content and SUVA for isolated NOM fractions from New England raw and treated waters**

high, as their nitrogen content is known to be substantially higher on average than the fulvic acids. Correlation with SUVA across all fractions is very weak.

The ratio of TCAA to THM formation most directly addresses the dual pathway to these two trihalomethyl species, as presented in Reckhow and Singer (1985). Figure 4.26 shows a strong positive correlation between SUVA and the TCAA/THM ratio. If a highly conjugated system encourages THAA (TCAA) formation, as previously discussed, higher UV absorbances should accompany a shift from THM to THAA. This is observed in Figure 4.26.

It has often been proposed that THM precursor content is positively correlated with molecular size. Although size data were not collected on all of the NOM fractions in this study, UF tests were run for a subset. Figure 4.27 shows how the THMFP compared with the percentage retention by a 1-kD ultrafilter. Though the data are limited, there is a weak correlation. Molecular size may not directly affect DBP formation. However, it might be correlated with other chemical properties that could play a role. For example, high-MW lignin structures and condensed polyphenols would be expected to be better THM formers than low-MW biomolecules such as simple fatty acids, terpenes, and so forth.

### Meta Analysis of Published Data on NOM Precursors

All studies included in this analysis used XAD-8 to selectively remove the hydrophobic acid. Nearly all made this separation at pH 2 and using a k' of about 50. Most of the comparisons



©2007 AwwaRF. All Rights Reserved.

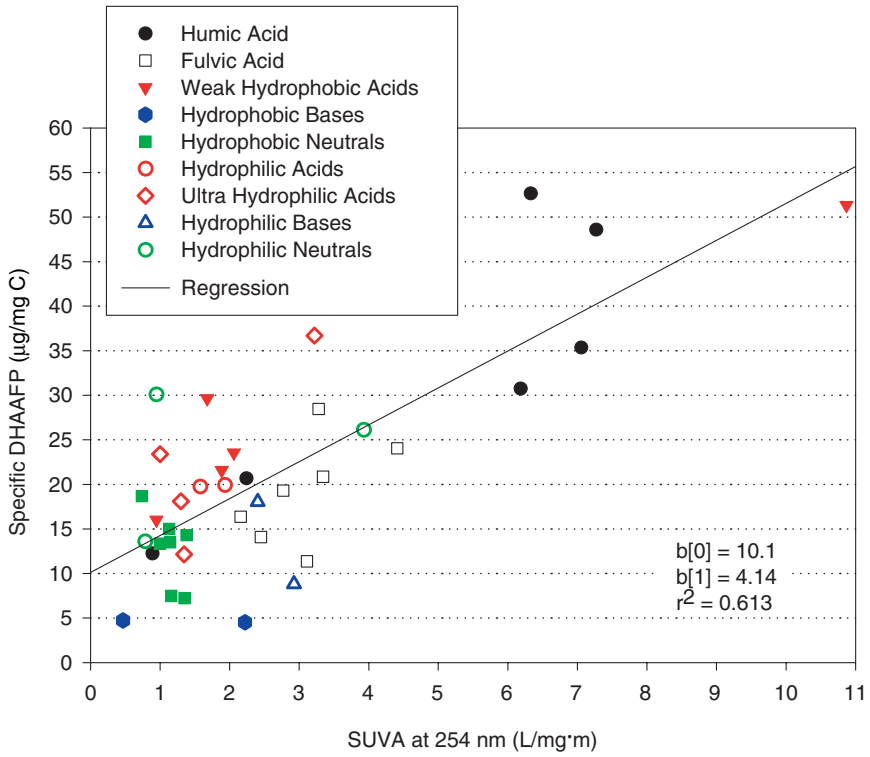



**Figure 4.24 Relationship between specific DHAA precursor content and SUVA for isolated NOM fractions from New England raw and treated waters**

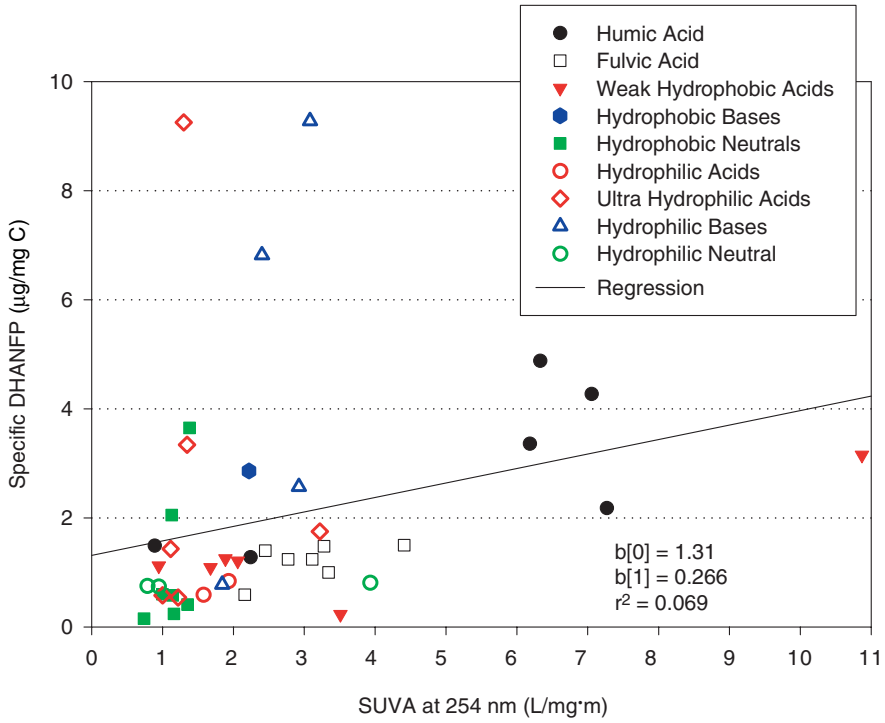



**Figure 4.25 Relationship between specific DHAN precursor content and SUVA for isolated NOM fractions from New England raw and treated waters**



©2007 AwwaRF. All Rights Reserved.

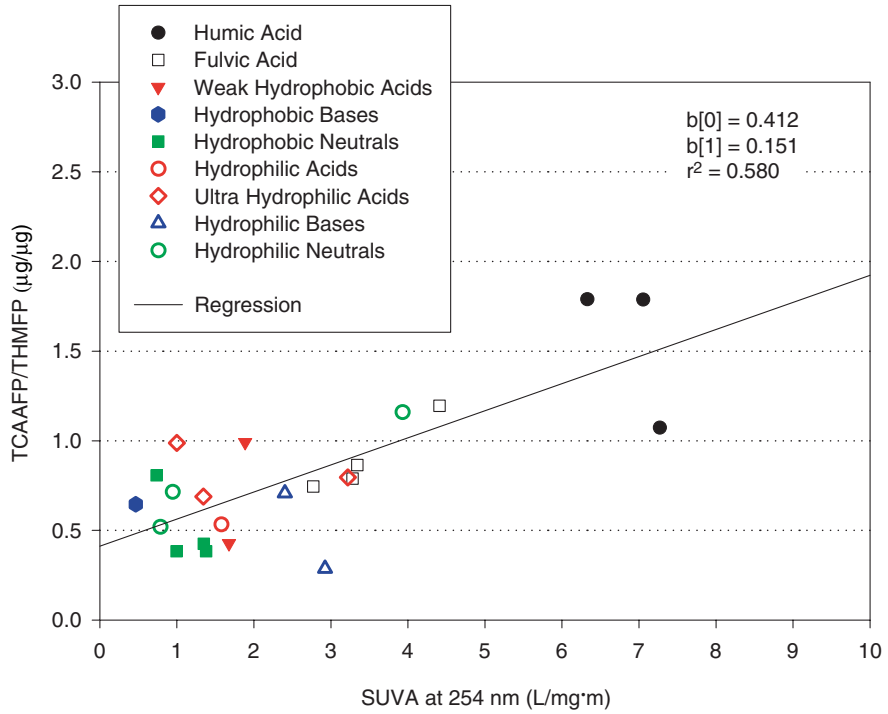



**Figure 4.26 Relationship between TCAA/THM precursor ratio and SUVA for isolated NOM fractions from New England raw waters**

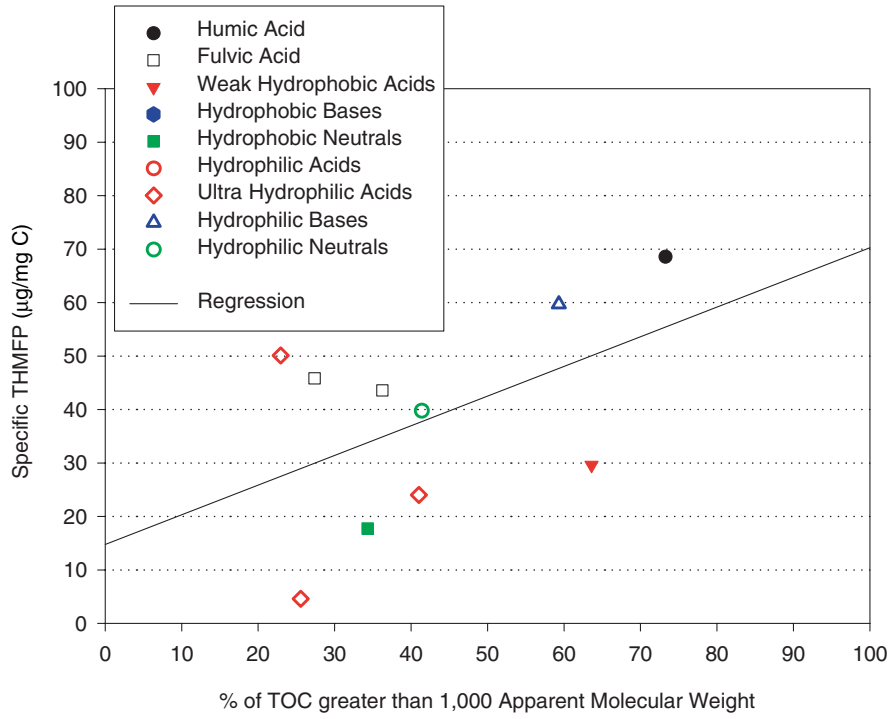



**Figure 4.27 Relationship between specific THM precursor content and molecular size for isolated NOM fractions from New England (Forge Pond and Connecticut) waters**



©2007 AwwaRF. All Rights Reserved.

were done with the major acid fractions (e.g., hydrophobic acids, hydrophilic acids, and transphilic acids). When hydrophobic acids were separated into humic and fulvic acids (or weak hydrophobic acids), only the fulvic acids were retained for this analysis. The others are generally of much lower abundance in natural waters and therefore not representative of the bulk hydrophobic acids. In most cases, the hydrophilic fraction was left whole and not further fractionated into acids, bases, or neutrals. These were retained in this analysis, based on the objective of comparing hydrophobic organics with hydrophilic organics.

Comparisons were drawn between the major phobic and philic components of raw waters. The full database was sampled and only the major acid-bearing fractions were selected. These included hydrophilic NOM, hydrophobic acids, fulvic acids, hydrophilic acids, hydrophilic NOM, transphilic acids, and transphilic NOM. These are believed to be dominated by the acidic compounds and are primarily distinguished by hydrophobic behavior. Excluded from this are fractions containing only bases, neutrals, or humic acids. Also excluded are other fractions that are separated based on size or other properties (HMW, colloids, etc.).

In addition, fractions isolated from treated waters were excluded. Finally, whenever the same fraction was treated with chlorine under two or more different treatment conditions, only the one subject to those most closely resembling the standard formation potential conditions was retained for further analysis.

### *DBP Yields From Analysis of Full Isolate Database*

Figure 4.28 shows the summary of the full database for specific THM precursors. This type of presentation allows one to see the full range of values recorded and corresponding percentiles. The median value (50th percentile) for hydrophobic NOM is about 50 µg/mg C; the hydrophilic is about 40 µg/mg C. Figure 4.28 also clearly shows that the hydrophobic fractions are shifted toward the higher THM precursor levels as compared to the hydrophilic ones. The transphilic fractions are shown as being low compared to the others; however, this might be attributed to a higher degree of uncertainty (i.e., fewer data) for this less commonly isolated fraction.

In contrast to the THMs, the DHAA precursors show very little systematic difference among the three NOM fractions (Figure 4.29). All three show median values of about 20 µg/mg C (formation potential conditions) with little separation at other percentiles. The THAAs (Figure 4.30) exhibited at least as much displacement between the fractions as the THMs, if not more. Once again, it is the hydrophobic fraction that has the higher precursor density.

A more direct set of comparisons between complementary NOM isolates is shown in Figures 4.31 and 4.32. These use the THMs as a benchmark and present the HAA values as ratios to the THMs. Figure 4.31 shows that although the THMs are favored in hydrophobic fractions, the THAAs are favored more. This is true even when comparing complementary fractions from the same water. However, the reverse is true for the DHAAs (Figure 4.32). In this graph, the hydrophobic data fall below the hydrophilic.

Population statistics based on Figures 4.28–4.32 are compiled in Table 4.4 for the hydrophobic fraction, Table 4.5 for the transphilic fraction, and Table 4.6 for the hydrophilic fraction.

### *Analysis of Selected Properties and Inter-relationships*

The national database shows a substantial positive correlation between specific THMFP and SUVA for all NOM fractions (Figure 4.33). This correlation exists within each of the three



©2007 AwwaRF. All Rights Reserved.

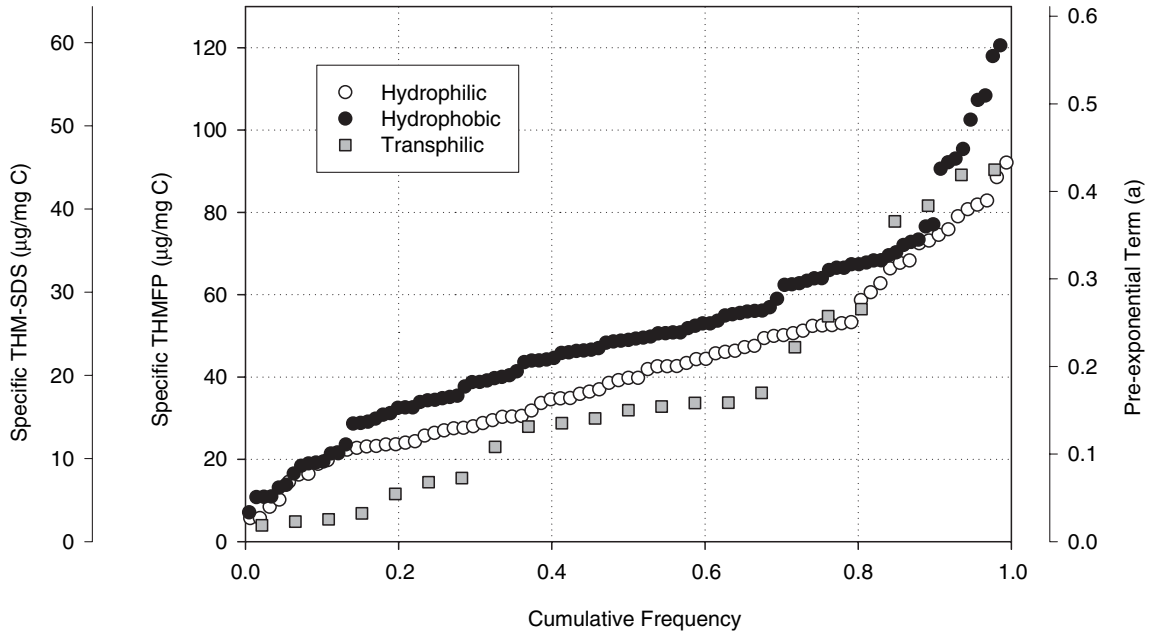



**Figure 4.28 Cumulative frequency plot for THM precursor content in major raw water fractions (untreated waters)**

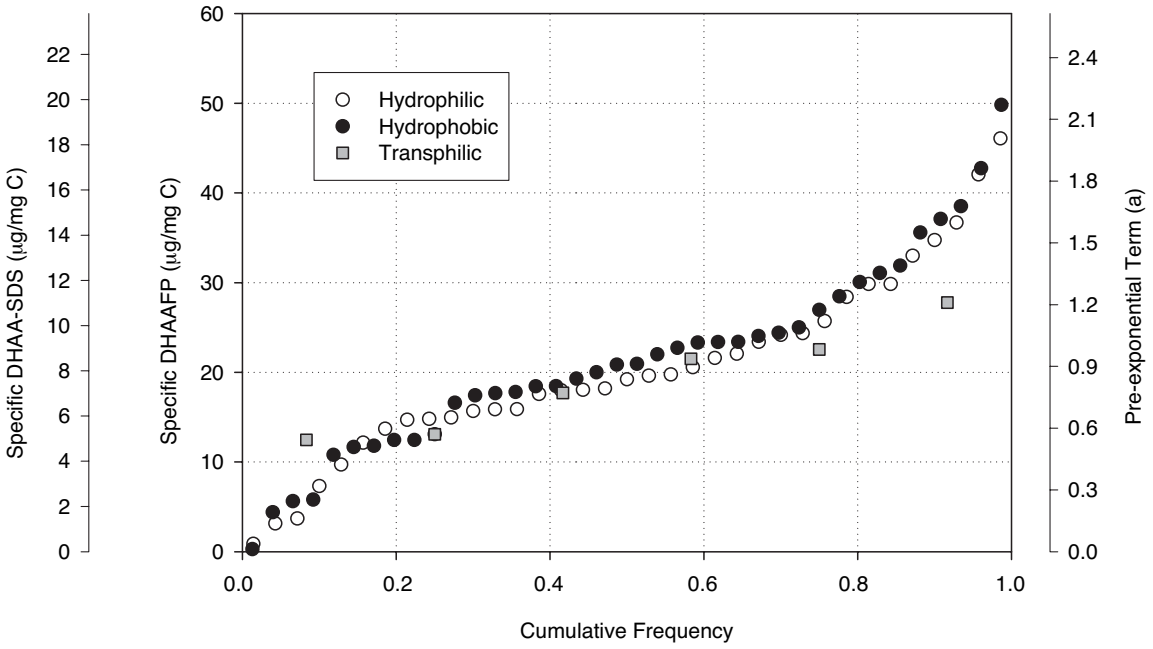



**Figure 4.29 Cumulative frequency plot for DHAA precursor content in major raw water fractions (untreated waters)**



©2007 AwwaRF. All Rights Reserved.

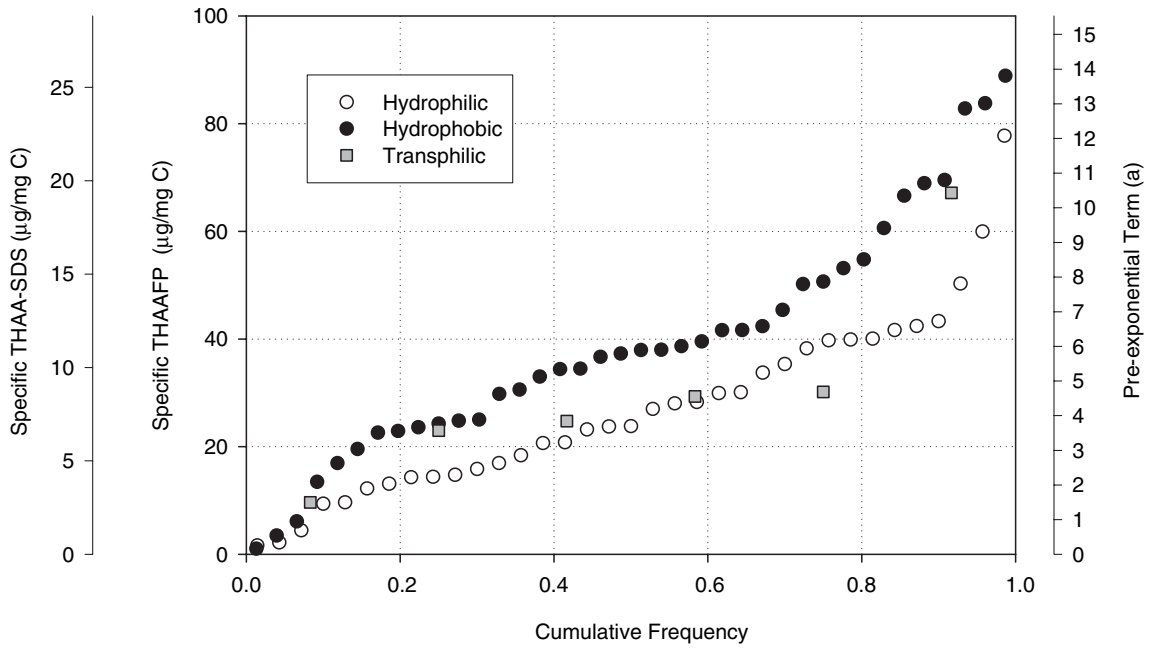



**Figure 4.30 Cumulative frequency plot for THAA precursor content in major raw water fractions (untreated waters)**

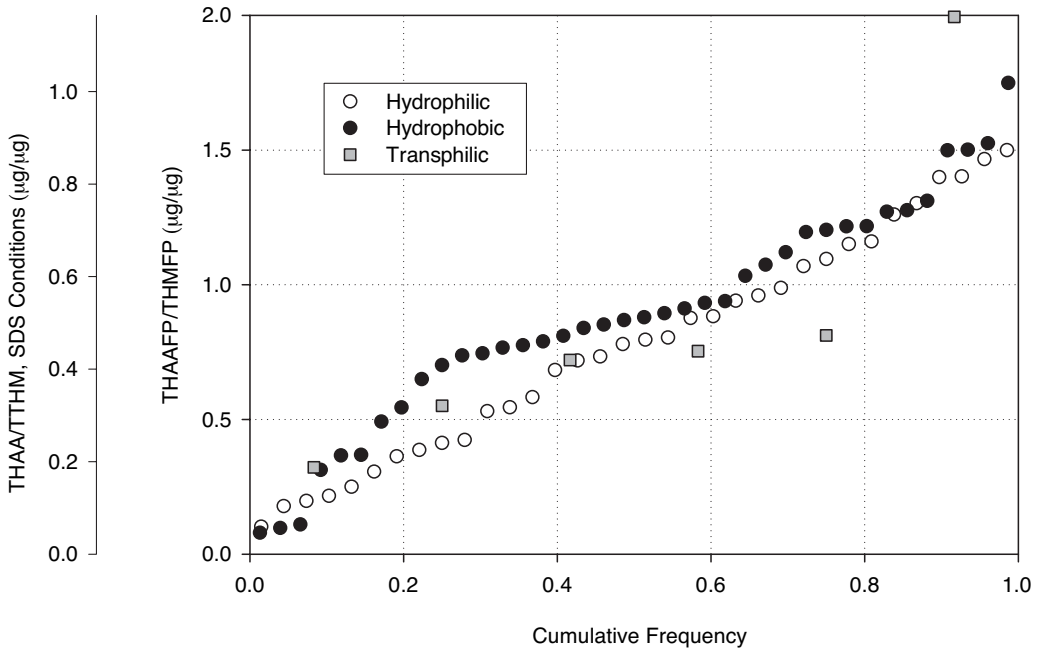



**Figure 4.31 Cumulative frequency plot for THAA/THM precursor ratio in major raw water fractions (untreated waters)**



©2007 AwwaRF. All Rights Reserved.

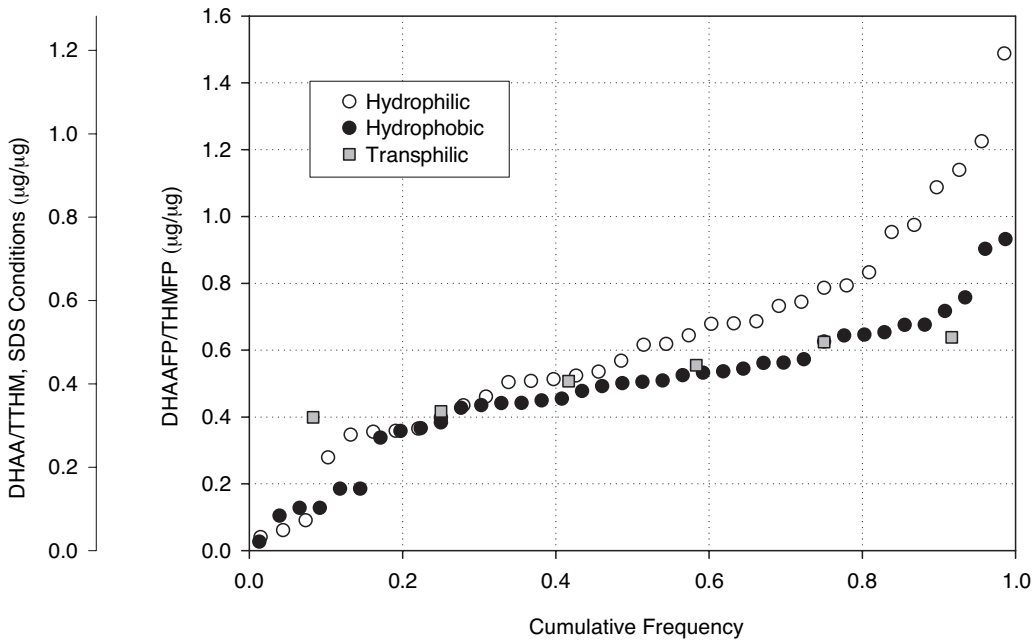



**Figure 4.32 Cumulative frequency plot for DHAA/THM precursor ratio in major raw water fractions (untreated waters)**

**Table 4.4**
**Population characteristics for THM, DHAA, and THAA formation potentials in hydrophobic NOM fractions**

| Cumulative frequency | TTHM (µg/mg C) | DHAA (µg/mg C) | THAA (µg/mg C) | THAA/TTHM (µg/µg) |
|---|---|---|---|---|
| 10 percentile | 19.4 | 7.2 | 9.4 | 0.33 |
| 25 percentile | 34.5 | 13.1 | 14.4 | 0.70 |
| 50 percentile | 49.0 | 20.9 | 23.8 | 0.87 |
| 75 percentile | 64.0 | 27.0 | 39.7 | 1.20 |
| 90 percentile | 80.0 | 36.7 | 43.3 | 1.44 |

**Table 4.5**
**Population characteristics for THM, DHAA, and THAA formation potentials in transphilic NOM fractions***

| Cumulative frequency | TTHM (µg/mg C) | DHAA (µg/mg C) | THAA (µg/mg C) | THAA/TTHM (µg/µg) |
|---|---|---|---|---|
| 25 percentile | 14.6 | | | |
| 50 percentile | 31.9 | 19.6 | 27.0 | 0.74 |
| 75 percentile | 53.0 | | | |

*An insufficient amount of data were available for DHAA and THAA to allow meaningful assessments of the 25th and 75th percentiles.



©2007 AwwaRF. All Rights Reserved.

**Table 4.6**
**Population characteristics for THM, DHAA, and THAA formation potentials in hydrophilic NOM fractions**

| Cumulative frequency | TTHM (µg/mg C) | DHAA (µg/mg C) | THAA (µg/mg C) | THAA/TTHM (µg/µg) |
|---|---|---|---|---|
| 10 percentile | 19.3 | 7.3 | 14.5 | 0.21 |
| 25 percentile | 26.5 | 24.8 | 24.3 | 0.41 |
| 50 percentile | 39.7 | 19.2 | 37.6 | 0.78 |
| 75 percentile | 52.4 | 25.6 | 50.6 | 1.10 |
| 90 percentile | 73.8 | 34.7 | 69.4 | 1.40 |

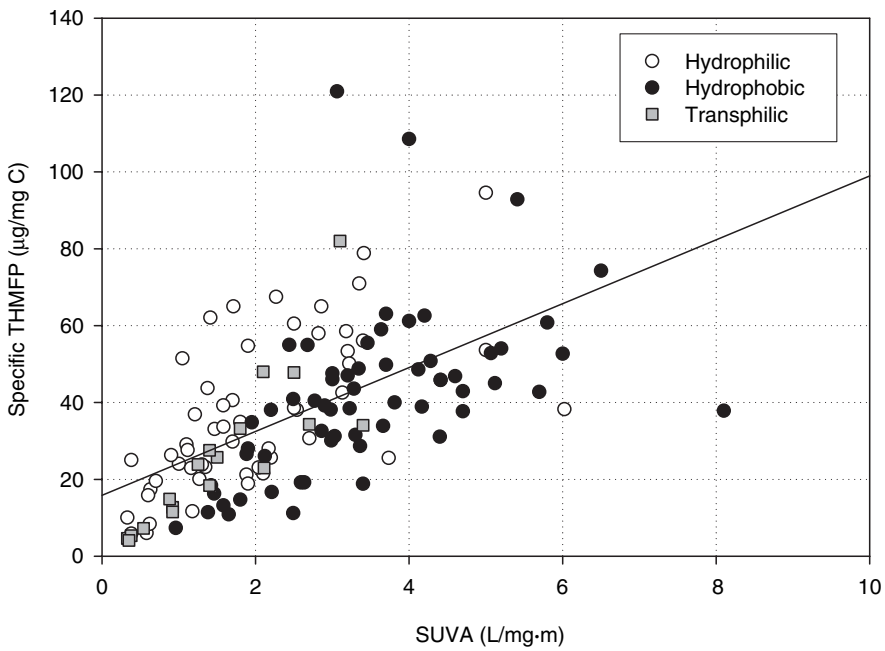



**Figure 4.33 Specific THM precursor content of all fractions plotted versus the SUVA (untreated waters)**

major NOM fractions as well (i.e., philic, phobic, and trans). Although the tendency of hydrophobic fractions to have higher SUVA values is evident, differences in the relative precursor densities between these solubility-based fractions is not obvious. Figure 4.34 shows the specific THMFP ratios for each of the paired NOM fractions derived from the same raw water. The median value for these ratios is less than unity, although a substantial number of data with ratios greater than 1 occur. This indicates that hydrophilic fractions tend to have slightly lower THM precursor densities than their corresponding hydrophobic fractions, but for any given water, the reverse might be true. The trend is similar for the transphilics, although less well defined, probably because of the smaller data set.



©2007 AwwaRF. All Rights Reserved.

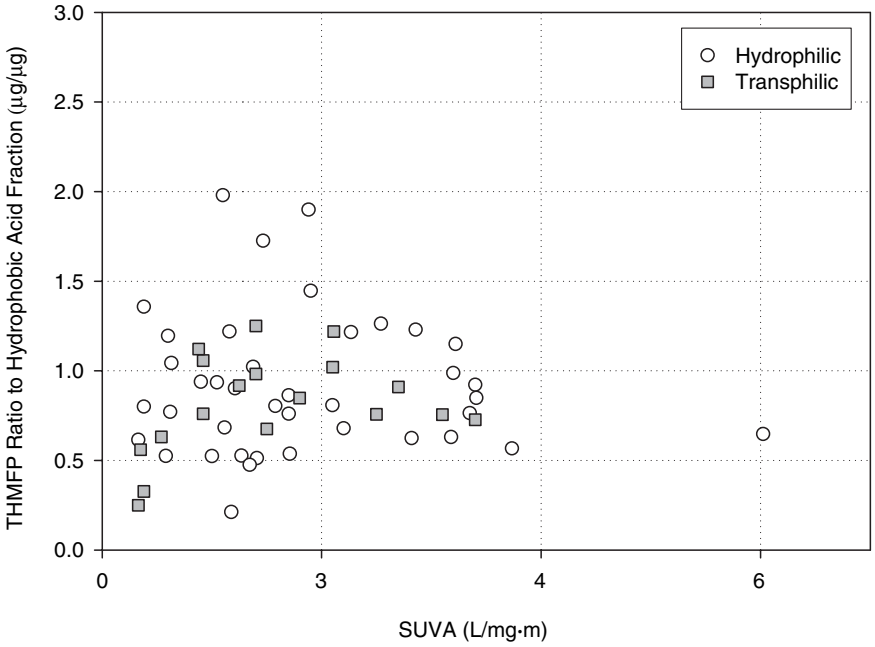



**Figure 4.34 Relationship between specific THM precursor content of hydrophilic and transphilic fractions to the specific content in the corresponding hydrophobic fraction plotted versus the NOM fraction SUVA (untreated waters)**

Many of the NOM fractions included in this analysis were extracted from whole waters that were subject to precursor testing. The existence of complementary data of this type permits comparison of the complete and unfractionated NOM with its resin-based isolates. Figure 4.35 shows such a comparison for the THM precursors. First, all of the data are clustered around a value of about 1. Second, there does seem to be a positive correlation with the particular fraction's SUVA, although this parameter can only describe some of the variability. Clearly, this presentation shows that the hydrophobic (solid circles) differ from the other fractions in their higher SUVA and not as much in their tendency to form THMs.

Figure 4.36 shows the same comparison but plotted as a function of the SUVA of the whole water from which the fractions were derived. In this case, there does not appear to be any trend as to increasing or decreasing ratio with changes in the raw (whole) water SUVA. It would appear that although SUVA can serve as an excellent surrogate for bulk THM precursor content, it fails to provide useful information on the THM-forming properties of the NOM fractions within a whole water.

The direct comparison of hydrophilic fractions to their respective hydrophobic fractions shows that most are less heavily populated with precursors (i.e., ratio below 1) when considering THM precursor content (Figure 4.37). Of course, this supports the earlier discussion where the unpaired data were graphed separately in frequency plots. Figure 4.38 shows this ratio in the form of a frequency plot. It appears that about 70% of the data have higher THM yields in the hydrophobic fractions than in their corresponding hydrophilic fractions. This tendency also carries out to the fulvic and humic acids. Aquatic humic acids, which are excluded from this analysis because of their low natural abundance, are considered even more hydrophobic than fulvic acids or bulk hydrophobic acids. When these two have been analyzed for precursors from the same source, humic acids usually have the higher specific THM precursor content (e.g., Reckhow and Singer 1990).



©2007 AwwaRF. All Rights Reserved.

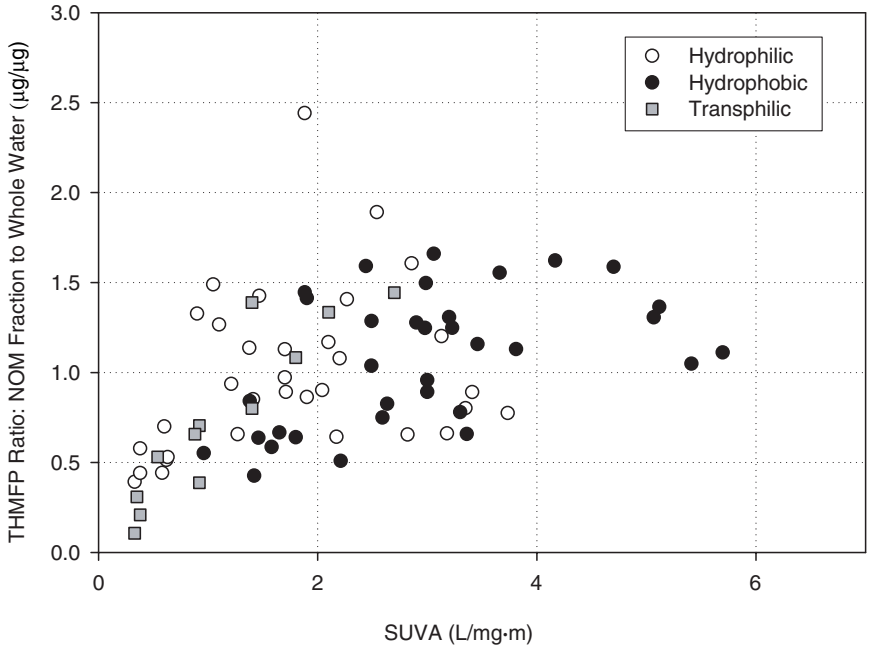



**Figure 4.35 Ratio of specific THM precursor content of NOM fractions to the specific content in the whole water from which they were derived plotted versus NOM fraction SUVA (untreated waters)**

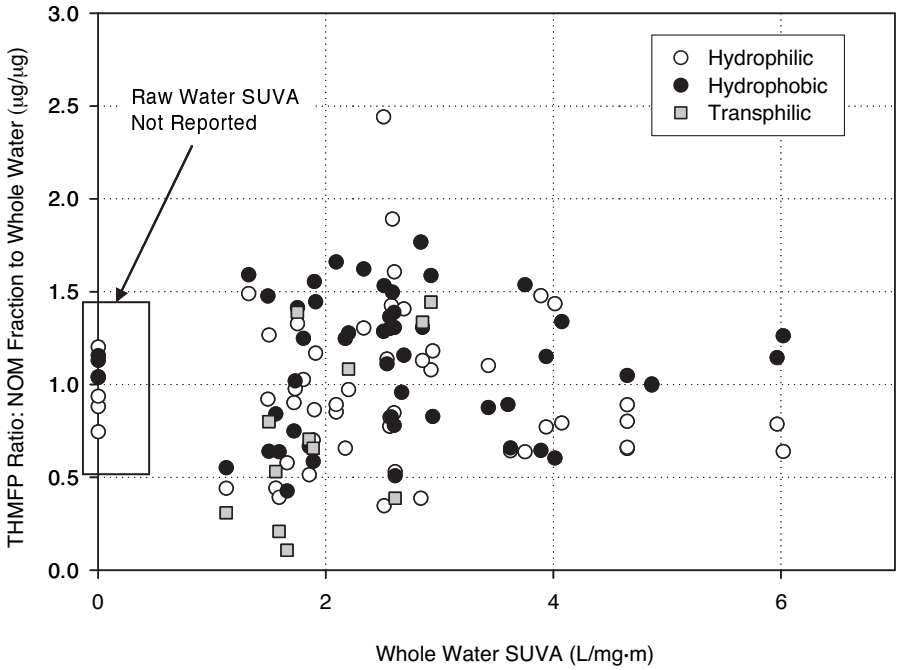



**Figure 4.36 Ratio of specific THM precursor content of NOM fractions to the specific content in the whole water from which they were derived (untreated waters)**



©2007 AwwaRF. All Rights Reserved.

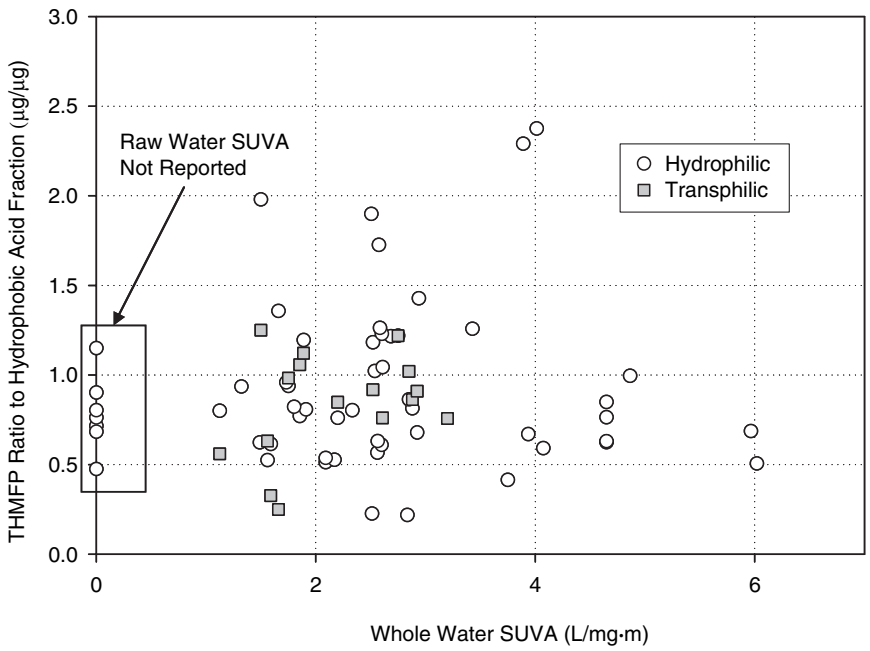



**Figure 4.37 Relationship between specific THM precursor content of hydrophilic and transphilic fractions to the specific content in the corresponding hydrophobic fraction (untreated waters)**

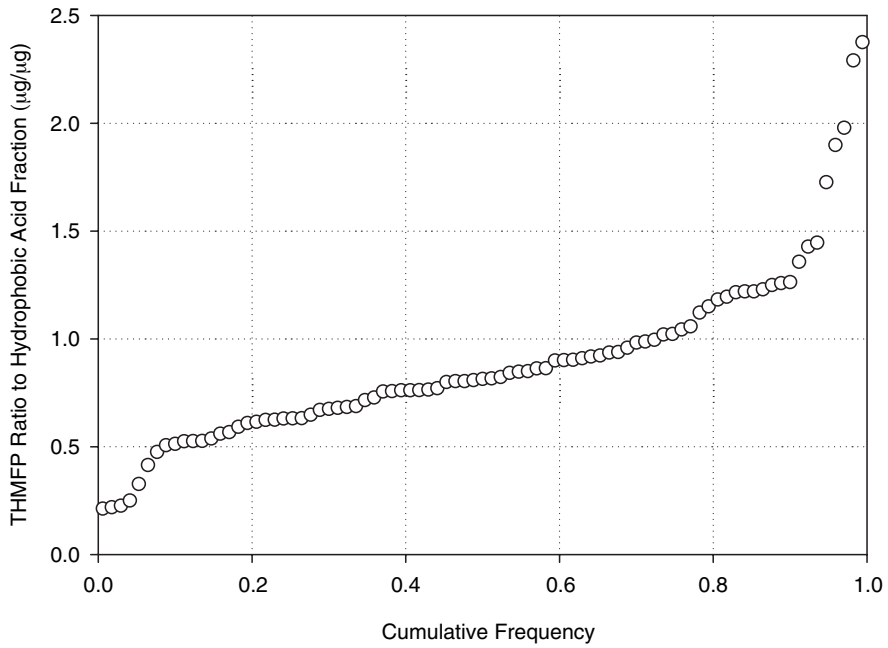



**Figure 4.38 Cumulative frequency plot for the ratio of the specific THM precursor content for a hydrophilic fraction to its corresponding hydrophobic faction**



©2007 AwwaRF. All Rights Reserved.

## SUMMARY AND CONCLUSIONS

New data were generated on DBP formation in a range of natural waters and extracted NOM fractions. These data were compared with similar data acquired from an extensive literature search. Appropriate comparisons required that differing conditions for precursor test be accommodated. This was done with the help of DBP formation models, summarized from the literature. The following conclusions are based on an analysis of this large combined data set.

THM precursor content of natural waters can vary sharply with the source type, location, and time of year. Although much of this variability can be attributed to differences in TOC, some cannot. By calculating the specific THMFP, one can explore the nature of the variability in this intensive property of NOM.

The formation of DBPs is heavily dependent on the characteristics of the precursor NOM. Although SUVA has been used successfully as a measure of NOM reactivity with oxidants/disinfectants, it falls short of capturing the full complexity of these reactions. This study documents the range of variability in DBP yields from a broad spectrum of waters and NOM types. Specifically, the follow conclusions can be drawn:

1. The major reason for higher THM levels in surface waters than groundwaters is attributable to higher TOC levels. On a per-carbon basis, surface waters produce only about one-third more THMs than groundwaters.
2. THM precursor levels are surprisingly uniform across the North American continent. The median THM precursor level in raw waters (based on the UMass protocol) is 48 µg/mg C, and the bulk of the population falls within 30–80 µg/mg C (10th–90th percentiles). This translates to about 23 µg/mg C under SDS conditions with a range of about 15–40 µg/mg C.
3. The average THM precursor level in hydrophobic acids closely mirrors the raw water database. The median THM precursor level (based on the UMass protocol) is 49 µg/mg C, and the bulk of the population falls within 20–80 µg/mg C (10th–90th percentiles). Once again, these are surprisingly uniform.
4. Indications suggest that specific THM formation goes down as the TOC of a surface water increases.
5. Highly colored (i.e., high-SUVA) organic matter has a much greater tendency to form THAAs.
6. Hydrophobic and hydrophilic bases produce a disproportional amount of DHANs.
7. Hydrophilic acids produce a higher level of DHAAs as compared to the other DBPs.

Comparison of the North American database with samples from a single intensively studied watershed show remarkable agreement in median values and spread of precursor densities (THM, THAA, DHAA) as well as SUVA. The fact that this study of narrow geographic and climatic scope can produce an array of data that nearly matches the full continental database has many implications. It seems that time of year, as well as micro- and mesoscale factors (local precipitation, flow paths, riparian-zone topography, sub-basin size) play a primary role in determining NOM character. Of lesser importance are the macroscale characteristics that only differ across broad geographic locations (e.g., ecoregion, dominant forest ecosystem, mean elevation, mean annual temperature, mean annual rainfall).



©2007 AwwaRF. All Rights Reserved.

# CHAPTER 5
# ANALYSIS OF AQUATIC ORGANIC MATTER

The intent of this chapter is to evaluate selected properties of aquatic organic matter in raw drinking waters. The main focus is reactivity with chlorine, formation of halogenated organic by-products, and biodegradability. Methods have been examined, selected, and refined as needed. These methods were then applied to the analysis of some contrasting raw waters from across North America.

## REVIEW OF CURRENT METHODS FOR AOC AND BDOC

Increasing demands are being placed on the water treatment industry to reduce the risks of illness from both chemicals and microorganisms (Nwachcuku and Gerba 2004). Growth of microorganisms in drinking water during distribution can lead to deterioration of water quality, violation of quality standards, and increased operating costs. Bacterial regrowth in drinking water distribution systems is the multiplication of viable bacteria downstream from the treatment plant (Charnock and Kjønno 2000). For bacterial growth to occur, various nutrients must be present. In particular, organic compounds, either dissolved or particulate, provide energy and carbon sources for heterotrophic bacteria to produce new cellular materials. The organic carbon in water supplies is mainly composed of humic and fulvic acids, carbohydrates, proteins, and carboxylic acids (e.g., hydrophilic materials). In many distribution systems it is assumed that the level of biodegradable organic matter (BOM) is the limiting nutrient for bacterial growth. Thus, BOM concentrations often measured as AOC and BDOC provide a good indicator for the bacterial regrowth potential of the treated water. Further, AOC has actually been correlated with bacterial counts in water distribution systems (Escobar, Hong, and Randall 2000). The last comprehensive, peer-reviewed, summary of methods to measure the biodegradability of aquatic NOM was published by Huck in 1990. Since then, new tests have been developed and old methods have been revised. Comparisons of some of these revised methods have been undertaken and the findings published. However, as of yet, no single test has been developed that fulfills the requirements of the water industry (i.e., real-time results and cost-effectiveness).

This AwwaRF review discusses the methods currently in use and under investigation for use in drinking water plants. It also attempts to predict the direction in which the development of new methods is headed. For example, pour plate counts, according to German drinking water regulations, are not sensitive enough to detect regrowth in distribution systems in sufficient time to enable the purveyor to react (Uhl and Schaule 2004). This review emphasizes the direction that regulatory bodies will be taking and the need for alternative, standardized tests. AOC and BDOC tests can be used to determine the level of treatment needed to minimize regrowth of bacteria in the water distribution system.

### The Connection Among BOM, DOM, and TOC

The amount of NOM in a system is usually expressed in terms of the concentration of TOC or DOC due to the ill-defined character of organic matter (Yavich et al. 2004). TOC is often synonymous with NOM, as organic contaminants in natural systems generally represent an insignificant fraction of the TOC. TOC can be subdivided into operationally defined fractions such as DOC and



©2007 AwwaRF. All Rights Reserved.

particulate organic carbon (POC). DOC is the organic carbon smaller than 0.45 µm in diameter and represents about half the DOM in natural waters (most organic matter is about 50% carbon by weight). POC is that fraction of the TOC retained on a 0.45-µm-pore size filter (Leenheer and Croué 2003). USEPA mandates utilities to remove certain concentrations of TOC to reduce DBPs.

Most NOM is considered refractory to rapid biodegradation (Leenheer and Croué 2003). BOM is that portion of NOM that is biodegradable and is available as a source of carbon and electrons for bacteria. BOM leads to undesirable water quality—for example, tastes, odors, color, nitrite, consumption of dissolved oxygen, accelerated corrosion of pipes, growth of biofilms (pipe surfaces in distribution systems can contain biofilms with bacterial densities up to $10^9$ cfu/cm$^2$)—and constitutes a possible source of pathogens.

Two types of bioassays are commonly used to measure BOM: the AOC assay and the BDOC assay. After the disinfection process, the rate of bacteria regrowth in water is mainly influenced by the concentration of AOC or BDOC present in the water. Both parameters represent the fraction of DOC that can be metabolized by bacteria in a given period of time.

### Assimilable Organic Carbon

AOC is that portion of BDOC which can be converted to cell mass—thus, it can be seen as a measure for bacterial (re-)growth potential—and is related to that portion of NOM which has a molecular weight <1,000. The removal of this portion of NOM results in reduced regrowth, reduced biofilm formation potential and biofilm growth, and reduced recovery of cells injured by disinfection. These biomass-based measurements are predicated on BDOC being a limiting nutrient for bacterial growth in drinking water; therefore, growth of bacteria is directly related to the amount of AOC in the water (Gibbs, Scutt, and Croll 1993). AOC promotes bacterial (re-) growth in finished water by supplying nutrients and may be related to the occurrence of coliforms in distribution systems. “Regrowth” is defined as the recovery of disinfectant-injured cells that have entered the distribution system from the water source or treatment plant. “Aftergrowth” is defined as the growth of microorganisms native to the water distribution system. “Breakthrough” is defined as the increase in bacterial numbers in the distribution system resulting from viable bacteria passing unharmed through the disinfection process (Escobar and Randall 2001).

Several water disinfection methods increase AOC; for example: (a) prechlorination in combination with anthracite–sand filtration, (b) prechlorination preceding GAC filtration, (c) ozonation (particularly for *Spirillum NOX*, or NOX), and (d) postchlorination. Water treatment methods that are known to decrease AOC include (a) filtration, with GAC as a filter matrix yielding lower AOC levels than sand or anthracite–sand; (b) iron-based coagulants, which show superior TOC removal to polymeric coagulants, and sometimes alum (Volk and LeChevallier 1999); (c) preozonation combined with sand or GAC filtration; and (d) the combination of coagulation, flocculation, and sedimentation (Hem and Efraimsen 2001).

AOC is expressed as an equivalent carbon concentration by means of a conversion factor or calibration. Water is considered biologically stable if AOC <10 µg C/L. Biological stability indicates that chlorine does not have to be added and bacteria levels will not increase during the test period. Regrowth is limited by AOC levels of <50 to 100 µg C/L in the presence of a chlorine residual or ozonation treatment (van der Kooij 1992). These numbers have been established for the Netherlands and may differ from region to region.



©2007 AwwaRF. All Rights Reserved.

## Measures of Biomass

In this section, some key elements among the existing AOC tests will be described. These will be discussed in the same order as presented in Table 5.1 The tests are numbered to facilitate cross-referencing between the text and table.

**1. AOC standard test (van der Kooij method).** This assay is based on determining the maximum levels of growth of two pure bacterial cultures inoculated into pasteurized samples of the water to be tested. The strains are *Pseudomonas fluorescens* P17, which is capable of using a wide range of compounds at low concentrations (1–10 µg C/L; van der Kooij, Visser, and Hijnen 1982) and *Spirillum* sp. NOX, which only utilizes carboxylic acids (van der Kooij 1984). AOC concentrations are calculated from the maximum colony counts ($N_{max}$) of these two strains grown as a mixed culture in the water samples using the yield values of these organisms for acetate. It may take 5–25 days to achieve $N_{max}$, including a lag phase of up to 24 hours. The length of run time may lead to differences in population succession over time, which will affect results; therefore, only pure cultures can be used (Arana et al. 2000). However, the use of pure culture is an advantage as it remains a predictable source; therefore, the results can be compared with those of other researchers. In some cases, inferences may be made about the type of substrate present or absent based on the observed growth of the pure cultures (Huck 1990). Bacterial growth in the water samples is monitored by colony counts, and the average growth ($N_{avg}$) observed during the incubation is converted into AOC units (µg/L acetate carbon) by using the growth yield of the bacteria derived from calibration curves obtained using standard concentrations of organic compounds (e.g., acetate or oxalate). The carbon conversion factors needed by this method to measure the AOC have been worked out for acetate and oxalate. Measurements of a few selected compounds for determination of growth kinetics and yield factors were performed by van der Kooij (1990). Compounds such as acetate and oxalate are ubiquitously present in treated drinking waters and are favorable sources of energy and carbon for a wide variety of bacteria. Both strain P17 and NOX grow on acetate, whereas NOX does not grow on glucose and P17 cannot grow on oxalate. The modification of the van der Kooij method to use 40-mL aliquots, compared to 1-L, by Kaplan (Kaplan, Bott, and Reasoner 1993), increased the ratio of surface area to volume in the glassware used and resulted in higher AOC values. The growth of heterotrophic microbes correlates with the increase in concentration of AOC, particularly in samples without chlorination (Miettinen et al. 1998). However, this method may only measure a small fraction of AOC because of the use of only two strains of bacteria and the underestimation of growth of these bacteria by the plate count method. This is especially relevant with raw water samples. Another limitation of this assay is that values obtained may be considerably lower than those obtained with other methods, especially for those assays that use a longer incubation time and a wider variation of microorganisms in the inoculum (Huck 1990). Additional factors constitute a potential source of error, including long incubation time, trace carbon-free glassware, proper dilutions, media variability, plate counts, and turbidity in water, which could interfere with accurate plate counts leading to an over- or underestimation of AOC (LeChevallier et al. 1993).

Multiple variations of this standard method exist:

**A. AOC standard test run with four characterized bacterial species (Kemmy and Fry method).** *P. fluorescens*, *Curtobacterium* sp., *Corynebacterium* sp., and an unidentified coryneform are used as this mix of cultures and will provide greater breadth in the utilization of organic compounds. Triplicate experimental parallels are prepared in



©2007 AwwaRF. All Rights Reserved.

## Table 5.1
## Summary of AOC methods

| Biomass-based method | Inoculum | Sample treatment | Measurement parameter | Convesion factor | Time effort | Advantages | |
|---|---|---|---|---|---|---|---|
| 1. Current standard method (van der Kooij 1990) | Known organisms: *Pseudomonas fluorescens* P17*, *Spirillum* sp. NOX†, *Flavobacterium*‡ | Pasteurized sample | Plate count (cfu) | Yes | Incubation at 20°C rather than 15°C reduces incubation time from 9 to 5 days. This assay can be run up to 20 days, but 9 days is standard. | • Inoculum does not change over time.<br>• Results can be compared with data of other workers (or studies).<br>• Inferences may be made about type of substrate.<br>• A standard method. | • Possible changes in AOC due to pasteurization.<br>• Added labor required to process plate counts. |
| 1a. Kemmy and Fry method (Kemmy, Fry, and Breach 1989) | Known organisms: *P. fluorescens*, *Curtobacterium* sp., *Coryne-bacterium* sp., unidentified coryneform | Filtered sample | Plate count (cfu) | Yes | 6 days | • Inoculum does not change over time.<br>• Results can be compared with data of other workers.<br>• Inferences may be made about type of substrate. | • Possible leaching of AOC from filter.<br>• Difficult to filter turbid samples.<br>• Added labor required to process plate counts. |
| 1f. USEPA CGR (coliform growth response) test (Rice et al. 1990) | Known organisms: *Enterobacter cloacae*, *Escherichia coli* | Filtered sample | Plate count (cfu) | No | 5 days | • Inoculum does not change over time.<br>• Results can be compared with data of other workers.<br>• Inferences may be made about type of substrate. | • Possible leaching of AOC from filter.<br>• Difficult to filter turbid samples.<br>• Added labor required to process plate counts. |
| 2. Rapid AOC method (LeChevallier et al. 1993) | Known organisms: *P. fluorescens* P17*, *Spirillum* sp. NOX† | Heat treated and filtered | Adenotri-phosphate (ATP) | No | 3 days | • Rapid assay.<br>• Can be amended to a high throughput format. | • Unacceptable variation found between sample readings. |
| 2a. Water Research Centre (United Kingdom) method (Jago 1993) | Indigenous bacteria | Filtered sample | ATP | Yes | Until maximum ATP concentration is reached. | • More representative of exist-ing conditions. | • Variability of inoculum among water sources.<br>• Added instrumentation. |
| 3. Turbidity test (Jago 1993) | Indigenous bacteria | Filtered sample | Turbidity | No | 2.5 or 5 days | • More representative of exist-ing conditions. | • Variability of inoculum among water sources.<br>• Added instrumentation. |

(continued)



©2007 AwwaRF. All Rights Reserved.

## Table 5.1 (Continued)

| Biomass-based method | Inoculum | Sample treatment | Measurement parameter | Convesion factor | Time effort | Advantages | Disadvantages |
|---|---|---|---|---|---|---|---|
| 4. Bioluminescent method (Haddix, Shaw, and LeChevallier 2004) | *P. fluorescens* P17*, Spirillum sp. NOX,† *+luxCDABE* variants | Pasteurized sample | Luminescence | Yes | 2–3 days | • Inoculum does not change over time.<br>• Results can be compared with data of other workers. | • Added instrumentation.<br>• Constant measurement. |
| 5. CheckLight method (CheckLight Company 2007 | Indigenous bacteria | | Luminescence | No | 2 hours | • More representative of existing conditions.<br>• Rapid (2–3 hours), inexpensive and sensitive enough to detect sub-ppm concentrations of diverse groups of utilizable organic compounds in water.<br>• Multiple tests can be run in a 96-well microplate format. | • Highly sensitive—cannot reuse glassware.<br>• Variability of inoculum among water sources. |
| 6. Olson method (Bradford, Palmer, and Olson 1994) | *P. fluorescens* P17*, *Spirillum* sp. NOX† | Pasteurized sample | Cell elongation | Yes | 18 hours | • Inoculum does not change over time<br>• Rapid assay | • Using a single pure cell culture may underestimate AOC in some water samples.<br>• Antibiotic concentration is ill-defined as concentration may need to be changed with different cultures. |
| 7. Long-term test (Servais, Billen, and Hascoet 1987) | Indigenous bacteria | Filtered sample | Cell numbers and volume | Yes | 10–30 days | • More representative of existing conditions. | • More laborious.<br>• Variability of inoculum among water sources.<br>• Added instrumentation. |

*Source:* Adapted from Huck 1990.

*P17 can grow on a wide variety of compounds at low concentrations but cannot utilize carboxylic acids including oxalate. *Pseudomonas fluorescens* is ubiquitous in surface, ground-, and drinking waters.

† NOX can utilize a wide range of carboxylic acids but not carbohydrates, alcohols, or aromatic acids, nor amino acids when grown on mixtures of compounds.

‡ *Flavobacterium* S12 can utilize glycerol and several oligo- and polysaccharides, and starch.



©2007 AwwaRF. All Rights Reserved.

which approximately 100 cfu/mL of each of the four species are inoculated into a 50-mL filter-sterilized, dechlorinated water sample. The samples are incubated without agitation at 20°C for the relatively short period of 6 days. At the end of the incubation period, the bacterial content is measured with a drop-counting method such as that published by Hoben (Hoben and Somasegaran 1982) using nutrient agar. The results are compared to calibration curves established with a known mixture of organic compounds (e.g., acetate, glucose, peptones, and yeast extract). This calibration provides a linear correlation from 0 to $10^6$ µg C/L. However, with this technique, the AOC value may be overestimated when reading from the calibration curve. This may be due to the incomplete degradation of the complex carbon used in preparing the calibration standard, or it may be that organic carbon is not the sole factor limiting bacterial growth. Further studies were to be conducted by the authors (Kemmy, Fry, and Breach 1989).

**B. AOC successional P17 and NOX test.** Water samples are pasteurized for 30 minutes at 70°C. The strain P17 is added and samples are incubated for 2 days at 25°C. After determining the plate count, the sample is repasteurized to kill P17, after which the strain NOX is added. After an incubation period of 3 days at 25°C, another plate count is established and AOC is determined by comparing the colony-forming units and yield coefficient. The yield coefficient of P17 and NOX were measured to be $1.4 \times 10^7$ cfu/µg acetate carbon and $1.8 \times 10^7$ cfu/µg acetate carbon, respectively (Liu et al. 2002). In this situation, strain NOX utilizes compounds that had not been utilized by strain P17 (e.g., oxalate, formate, and glyoxylate). An advantage to using the different carbon-utilizing strains is to indicate the types of organic compounds available to bacteria in the water source.

**C. AOC with mixtures of bacteria that utilize groups of specific compounds.** *Flavobacterium* (strain S12) can be added to the standard assay for waters that have been subjected to water treatment where starch was used as a coagulant aid, as P17 is not able to fully utilize the polysaccharides present in this water type (Jago 1993).

**D. AOC standard test—*P. fluorescens* only.** In this variation of the AOC test, the metabolic capability of a single microorganism is measured, as this bacterial strain grows on a wide variety of compounds at low concentrations but cannot utilize carboxylic acids including oxalate. An advantage of this adaptation of the AOC assay is that there is only one bacterial species to control, and this species should give a sufficient indication of the amount of organic carbon in the water system due to the wide range of organic compounds it assimilates. This is beneficial because it does not require as high a skill level as other variations. A disadvantage is the potential to underestimate the AOC in areas where there may be a large amount of carboxylic acids in the waters.

**E. AOC with mixtures of environmental bacteria.** This method is characterized by a great nutritional versatility.

**F. Coliform growth response (CGR) method.** With this method the water sample is filter-sterilized and inoculated with a pure culture of *Enterobacter cloacae*. The inoculated flask is incubated in the dark at 20°C for 5 days, and samples are taken to determine the starting and final concentrations of the inoculum. The CGR value is determined as the logarithm of the ratio of the finished concentration to that at the time of inoculation (Jago 1993). The results are compared with those obtained from replicate determinations with the van der Kooij method using P17 as the inoculum. Results obtained with this technique suggest a positive correlation between CGR and AOC. Higher CGR



©2007 AwwaRF. All Rights Reserved.

values are obtained from ozonated samples as compared to those from unozonated samples, which suggest that coliform growth responds to the increase in organic carbon availability known to follow ozonation.

It is significant that *Pseudomonadaceae* regrowth is favored by the presence of low concentrations of BDOC. The species *Aeromonas*, a member of the group *Pseudomonadaceae*, has been used as an indicator of potential regrowth in drinking water systems in the Netherlands. As a general standard, Dutch water regulations permit a plate count of 20 cfu of *Aeromonas* per 100 mL in finished water and a tenfold higher count in distribution system waters. In the drinking water of the Barcelona area, *Pseudomonas* spp. have been found at four times these levels. This result suggests that *Pseudomonas* could be used as a better indicator of potential bacterial regrowth than *Aeromonas* (Ribas et al. 2000).

**2. AOC rapid ATP procedure (rapid AOC method).** This assay uses enumeration of the test organism by measurement of cell ATP levels to increase the speed and simplicity of the AOC assay. Water samples are quenched, if chlorinated, by a sodium thiosulfate solution (0.1 mL of a 13.2-g/L solution in a 40-mL sample), heat treated at 70°C for 30 minutes, and then three aliquots of 40 mL are inoculated in triplicate with ~$10^4$ cfu/mL of low-nutrient adapted cultures of either P17 or NOX. The inoculum is prepared by inoculating an isolated colony of each strain into 100 mL of sterile, chlorine-neutralized tap water for 7 days. An aliquot (0.1 mL) of this culture is used to inoculate 100 mL of sodium acetate and mineral salts solution, and this culture is incubated for 7 days. A typical growth level of $8.2 \times 10^6$ cfu/mL for P17 and $2.4 \times 10^7$ cfu/mL for NOX is obtained. Three additional aliquots are kept at 4°C as uninoculated controls. The inoculated aliquots are incubated for 1 to 3 days at room temperature (~22°C). Samples are filtered through a polyvinylidine-difluoride membrane filter (25-mm pore size, 0.22-µm, Durapore filter; Millipore Corp., Billerica, Mass.) under a 152.4-mm mercury vacuum. The filters are placed in Hepes buffer with 0.5 mL of ATP releasing agent with phosphatase inhibitor. The samples are vortexed for 10 seconds. After 20 minutes, 150 µL of the sample is placed in a polypropylene test tube and the ATP level is measured (LeChevallier et al. 1993). After 3, 4, and 5 days of incubation (this time can be reduced with increased inoculum and temperature) at 22°C, ATP levels per cell are measured using a luminometer, and measurements are converted to AOC units (Volk and LeChevallier 2002, LeChevallier et al. 1993). Because there is no plating, there is no necessity for long incubation times for plate counts, media preparation, dilutions bottles, nor cleanup of glassware.

Results from ATP measurements are equivalent to plate count results as they are both indirect measures of active bacterial metabolism. The ATP technique may be more accurate as higher AOC results are attributed to an underestimation of bacterial levels by the plate count method. Underestimation can occur in plate counts because of the attachment of cells to particulates. There is also more opportunity for errors such as in the preparation of the dilutions. This method also uses averages of the stationary phase levels rather than $N_{max}$; reducing the variability of the assay as the time to achieve $N_{max}$ can vary from 5 to 25 days. The strains NOX and P17 are measured in separate experiments because of differences in ATP levels between these microorganisms. This separation prevents competition for nutrients between the two species, but it also means that mixed indigenous populations cannot be used because they cannot be easily separated into single-species treatments. Turbidity in the water sample can also interfere with ATP determination requiring filtration of the water sample. This might constitute a bias as some of the AOC might be trapped in the particulate organic matter causing the turbidity. Levels of free ATP may also



©2007 AwwaRF. All Rights Reserved.

obscure results. However, concentrating the cells on a filter reduces the background levels of non-cell associated ATP (LeChevallier et al. 1993).

The detection limit of this method for strain P17 is $5 \times 10^3$ cfu/mL or AOC of 1.5 μg/L, and for strain NOX the limit is $5 \times 10^4$ cfu/mL or an AOC of 4 μg/L.

In practice, this method had unacceptable variation and was too technically cumbersome for use at water utility laboratories (Haddix, Shaw, and LeChevallier 2004). Another negative comment about this method was that there appeared to be no relationship between AOC concentrations and the spatial and seasonal variation in plate counts. The lack of correlation may have been caused by a lack of sensitivity in the AOC technique or changes in the culturability of bacterial inocula rather than an increase in bacteria numbers leading to no change in the AOC used. The sensitivity of the AOC technique may be affected by the origin of the inoculum when using indigenous populations (Gibbs, Scutt, and Croll 1993).

A. **Jago method (Water Research Centre method).** This method is based on the same principle as the van der Kooij method, where the AOC value is directly related to the growth response of the microbial population. However, several significant differences exist and are outlined in the text that follows. The indigenous bacteria are removed from the water sample by filtration rather than pasteurization. Filtration is used because heat treatment may have an influence on the AOC concentration because of hydrolysis of refractory compounds to simpler biodegradable compounds (Jago 1993). The filtered sample is inoculated with a natural heterotrophic community of bacteria. The mixed culture is chosen because it would be more likely to utilize the wide range of organic compounds that would be found in a water sample. The inoculated sample is incubated at 20°C, which should decrease the length of time for maximum growth. Samples are taken on a daily basis and analyzed for their ATP content until a maximum is achieved. The ATP measurement is chosen because plate count techniques can give an underestimation of populations of unknown mixtures of bacteria as a result of the diverse nutritional requirements of the bacterial populations. The ATP yield can be directly converted to an AOC concentration by using a conversion factor. This conversion factor is derived from the defined ratio of cell ATP content to cellular carbon and from the calculated conversion efficiency of substrate carbon to cellular carbon. This method has been used in several water treatment plants in England (Jago 1993).

**3. Turbidity (Werner method).** This method is based on the measurement of the growth response of mixed cultures as recorded by turbidity (Jago 1993). Results are obtained showing both the magnitude and the rate of production of biomass as a result of the continuous turbidity monitoring performed on each sample (Jago 1993). The test water is filtered, supplemented with a carbon-free mineral salts solution, and inoculated with a natural heterotrophic community of bacteria. The inoculum is obtained from washings of the filter used in the initial sterilization of the sample. The inoculated sample is then incubated in a turbidimeter at 20°C. Turbidity readings are taken automatically every 30 minutes, and a plot of the growth response is produced. Measurements are also made of the DOC and the total bacteria count. The procedure can be calibrated from the growth response to known concentrations of carbon compounds. The results obtained give valuable information not only on the total regrowth potential of the water sample, but a measurement of the ease of biodegradability of the organic compounds in the water can be deduced from the slope of the growth curve. This is an important factor in assessing how a water



©2007 AwwaRF. All Rights Reserved.

sample is going to react during treatment, and when it enters the distribution system. An advantage of this method is that because the samples are supplemented with mineral salts, carbon should be the only growth-limiting factor. The result obtained is therefore the maximum regrowth potential of the water sample. Most other AOC procedures do not supplement the water sample with mineral salts. Therefore, if the water sample tested in the other procedures is naturally deficient in an essential growth requirement other than carbon, then this would limit growth, yet the result would be expressed as AOC, which constitutes an underestimation of the true AOC value.

**4. AOC measurement with bioluminescent derivatives of test bacteria.** The principle of this technique is based on the observation that bacterial bioluminescence in microbial batch cultures peaks at full cell yield just prior to the onset of the stationary phase. The test strains P17 and NOX are mutagenized with a *luxCDABE*-inducible transposon, and although P17 is inducible with *p*-aminosalicylate, NOX does not need an inducer. The AOC values determined with peak bioluminescence and those based on plate counts are linearly related.

The procedure for this AOC measurement requires the water sample to be pasteurized at 65°C for 30 minutes. Each vial is inoculated with a final concentration of $10^4$ cfu/mL. Luminescence is measured twice a day for 3 days, making this one of the shorter tests of AOC. To induce bioluminescence, 1 mL of culture is measured for 90 seconds with 0.1 mL of injected 0.2% *n*-decanal in denatured ethanol. It is possible to adapt this test to a microtiter plate format (Haddix, Shaw, and LeChevallier 2004). The unit of measurement is relative light units. The modified strain of NOX has a tenfold greater sensitivity than the P17 derivative (0.012 µg acetate carbon/L compared to 0.67 µg acetate carbon/L, respectively). These values are determined without a luminescence inducer. Addition of inducer may be necessary to promote luminescence in the P17 genetic background. Other AOC methods, such as the standard method, are not seen to show differences in AOC sensitivities for these two strains.

**5. CheckLight AOC test.** This test is based on the effect of AOCs on the development of luminescence in *Vibrio fischeri*. The test is intended to be simple with a relatively quick result while giving a high correlation with the capacity of bacteria to divide in the water sample. The luminescent bacterium *V. fischeri* is given all conditions necessary for light production with the exception of an organic carbon source that is provided by the water sample. The bacteria are provided freeze-dried, and on hydration the bacterial luminescence system is induced if the sample contains AOC. Bioluminescence increases with time with an intensity dependent on the concentration of the organic compound. Concentrations below the parts-per-minute level can be determined within 2–3 hours. Significant correlation was reported by the manufacturers of the commercial CheckLight kit (CheckLight Ltd., Qiryat-Tiv'on, Israel) compared to the standard AOC test.

CheckLight has developed an automatic version of its novel AOC test that enables the detection of 5–1,000 µg/L of AOC in water (either surface, brackish, or seawater) within 10 minutes. The monitor may be programmed to alert the user whenever the AOC level exceeds a given threshold. The analyzer has automatic safeguards to ensure reagent and data quality; it is equipped with positive and negative controls and programmed flushing and cleaning cycles to maintain consistency of the tests (CheckLight Company 2007).

**6. Rapid microscopic method (Olson method).** This method, based on the van der Kooij method, is adapted to give a measurement of AOC within a day without intensive plate counting. Duplicates of 50-mL volumes of water are treated with 20 µg/mL of lomefloxacin, an antibiotic that inhibits cell division, and inoculated with either NOX or P17. Nutrients in the water sample cause viable cells to grow, but because their division is inhibited, the cells merely elongate (Jago



©2007 AwwaRF. All Rights Reserved.

1993). The principle of this method is that the number of cells which elongate is proportional to the nutrient concentration of the water sample. The elongated cells are detected by staining and observation using epifluorescence microscopy and manual measurement of the degree of cell elongation that may take place. The length of each cell is measured and the number of elongated cells expressed as a percentage of the total number of cells measured. A standard curve of the percentage of the elongated cells against carbon concentration (acetate) is constructed and used to calculate the actual AOC in the test water samples. The bioassay is conducted within a 24-hour time period. Bradford, Palmer, and Olson (1994) showed that the correlation between this method and method 1B is 0.91 for strain P17 between AOC concentrations of 0 and 150 µg C/L, and 0.91 for NOX between 0 and 50 µg C/L. Jago (1993) suggested there may be problems in defining the concentration of antibiotic sufficient to prevent cell division but not inhibit cell growth. It was also thought that the use of a single pure culture would produce an underestimation of AOC in some water samples. Also, adapting this method to an indigenous bacterial population is likely to complicate the choice and concentration of the antibiotic used to inhibit cell division (Jago 1993).

As reported in the literature, the AOC method currently used most often is the modified van der Kooij assay as described in *Standard Methods* (APHA, AWWA, and WEF 1998). Many of the previously described methods have been researched to develop a shorter, less labor-intensive assay. The two most promising methods utilize luminescence properties of bacteria to measure AOC (e.g., the bioluminescence derivatives assay and the CheckLight AOC assay).

**Biodegradable Dissolved Organic Carbon**

BDOC is that portion of the dissolved organic carbon in water which can be mineralized by indigenous, heterotrophic microorganisms. It is measured by the DOC fraction assimilated and mineralized by heterotrophic bacteria (i.e., the measure of the DOC before and after incubation with natural flora). BDOC values represent 10%–30% of the total DOC. For lakes, the global average of BDOC is 14% of the total DOC pool. Most of the ambient DOC is refractory to bacterial decomposition or can only be decomposed slowly in combination with photochemical reactions (Søndergaard and Worm 2001). BDOC is a regrowth indicator because it may be used as a nutrient source for bacterial regrowth in water treatment and distribution systems, and the removal of BDOC is essential in minimizing biofilm formation (Park et al. 2004). The most common ozonation by-products, aldehydes and simple carboxylic acids, are highly biodegradable. These two groups of compounds constitute a minor fraction of BDOC, and the structures of the majority of the compounds are unknown (Swietlik et al. 2004).

"Biological stability" is defined as the absence of BDOC consumption, which is when BDOC ≤0.16 mg/L (Servais, Laurent, and Randon 1993). When the validity of this value was further analyzed with respect to temperature, it was found that a better criterion for biological stability is when BDOC ≤0.15 mg/L at 20°C and ≤0.30 mg/L at 15°C. Coliform occurrence is related to the range of BDOC >0.1 to 0.15 mg/L (Volk et al. 1994). Coliforms regrow in water that has a BDOC concentration of 1.4 mg/L (Joret, Levi, and Volk 1991).

The approach of DOC measurements via BDOC is different from AOC in two main features: BDOC determines mineralization of DOC for energy gain rather than assimilation for cell growth, and it uses indigenous microbial communities rather than amended pure cultures.

Many benefits of using indigenous bacteria exist. Indigenous bacteria are likely to be more representative of the responses of treated water and have broader biodegradation capacities than the



©2007 AwwaRF. All Rights Reserved.

few species in mixed pure cultures (Servais, Anzil, and Ventresque 1989). Furthermore, their current biochemical cell composition is already adapted to have either no lag phase or a lag phase shorter than 6 hours. The growth of indigenous bacteria is less sensitive to concentrations of DOC than that of a pure culture such as *E. coli*. Coliforms cannot grow at lower concentrations of organic carbon (Arana et al. 2000). The origin of the bacterial inoculum has no effect on the BDOC value (Servais, Anzil, and Ventresque 1989; Block et al. 1992). There is no need for pretreatment of water samples, such as sterilization, filtration, or pasteurization (Block et al. 1992).

Raw water filtration is preferred to autoclaving as autoclaving can modify the biodegradability of DOC, thus contributing to an increased BDOC concentration. The percentage of BDOC reduction decreases with increasing filtration velocity and thus with decreasing contact time (Servais, Anzil, and Ventresque 1989). The main distinguishing feature within the available suite of BDOC methods that have been developed is whether bacterial cells are immobilized or remain in suspension. Finally, none of the BDOC methods has been officially accepted or standardized, which makes it difficult to draw comparisons among published values and obtain accurate BDOC values (Park et al. 2004).

Table 5.2 is a summary of the most frequently used BDOC tests. As with Table 5.1, these tests are grouped according to key characteristics and are numbered to facilitate cross-referencing between the text and table.

**1. Bacterial inoculum attached to inert material—plug-flow bioreactor (Frias method).** The following descriptions of Frias methods are dynamic procedures that measure the decrease of DOC in a sample of water (Frias, Ribas, and Lucena 1992). The bacterial inoculum is attached to an artificial matrix of porous glass beads over which the water is guided in a recirculation path. The bioreactors are not site specific and require no acclimatization to measure BDOC of samples from different sources (Søndergaard and Worm 2001).

In a single darkened-glass column that is filled with acid-washed and muffled porous glass beads (60-µm pore; Siran, Schott GmbH, Germany), test water is recirculated by a peristaltic pump. Colonization of the column with a microbial community is achieved at a flow rate of 1 mL/min using 1.0-µm filtered water from the water source and stabilized over a period of about 1 month. The biofilm is maintained by feeding with filtered sample water. Using a TOC analyzer, DOC is measured in the original water sample and on a series of daily samples taken from the recirculating water until a stable minimum DOC value is achieved (up to 5 days). Samples for DOC analysis are filtered through prerinsed Millipore GVPWP membranes (0.22-µm pore size). As a series of samples are taken for analysis over a period of days, the rate of DOC reduction within the reactor is monitored. This provides valuable information on the degree of biodegradability of the organic carbon compounds present in the sample. This information can be used to tailor biologically based water treatment processes for optimum reduction of BDOC. This information will also give an indication of how the water will react in a distribution system (Jago 1993).

An alternative version of this method involves two glass columns in series, where a single water sample passes through the columns. The two glass columns are filled with 300 g of an inert support to which the biofilm responsible for DOC removal is attached. The inert support can be biologically active sand (BAS) or Siran, sinterized porous glass beads that need to be colonized for about 2 weeks to 2 months depending on the water sample. The glass beads were the better medium. The distance between the inlet and outlet of each column is 620 mm, and the inner diameter of the columns is 28 mm. A peristaltic pump continuously pumps the water sample upward, across the columns. The optimal flow is 3.5 mL/min and the required retention time is



©2007 AwwaRF. All Rights Reserved.

## Table 5.2
## Comparative summary of methods to measure BDOC

| DOC-based methods | Inoculum | Support medium | Parameter | Concentration detection range | Time | Advantages | Disadvantages |
|---|---|---|---|---|---|---|---|
| 1. Spanish method (Ribas, Frias, and Lucena 1991) | Biofilm on biologically active sand (BAS) or colonized glass beads | BAS or glass beads | Δ DOC | | 2 hours | • Biofilm more representative of treatment process. | • Possible desorption of organic matter from sand or adsorption to sand. |
| 2. Billen–Servais method (Servais, Billen, and Hascoet 1987) | Indigenous bacteria | None | Δ DOC | >0.3 mg BDOC/L | 28 days | | • In this method the bacterial community may change over time to a community that is unrepresentative of the steady-state community. |
| 3. Joret–Levi method (Jago 1993) | Biofilm on BAS | BAS | Δ DOC | | 3–5 days | • Biofilm more representative of treatment process. | • Possible desorption of organic matter from sand or adsorption to sand. |
| 4. Mortality flux method (Servais, Billen, and Hascoet 1987) | Indigenous bacteria | None | Mortality flux (Fmort)/ biomass carbon yield (Y) | | 13 days | • Good agreement between DOC method and this method. Can be used when DOC is low. | • More time consuming than other methods. |
| 5. Combined BDOC and biochemical oxygen demand method (Khan et al. 1998) | | None | | DOC 4–15 mg/L | 28 to 5 days | • Using a larger volume and/or more concentrated inoculum reduces time. | |
| 6. Immobilized cell BDOC determination (Khan et al. 2003) | Indigenous bacteria | Cells immobilized in cellulose triacetate | Δ DOC | Detection limit is 150 µg/L | >3 hours | • Life expectancy reported to be 10 years. Cells do not require acclimatization or startup period. Method comparable to batch test. | • BDOC can be underestimated if the hydraulic residence time <3 hours. DOC can be released from carrier. |
| 7. Morgen method (Jago 1993). | Indigenous bacteria | Acclimatized sand | Δ DOC | | 5 days | | • Limited by the detection limits and sensitivity of DOC analysis. |

*Source:* Adapted from Huck 1990.



©2007 AwwaRF. All Rights Reserved.

1 hr/column. Samples are taken from the inlet and the outlet of both columns at 1-hour intervals. BDOC equals the difference between the inlet and the second column outlet. The columns are kept at a temperature of 21 ± 2°C. Samples are filtered for DOC analysis through prerinsed Millipore GSWP membranes (0.22-µm pore size). A significant removal of BDOC was found in the second column for all water types tested (Jago 1993). The Siran glass beads can be colonized according to one of the following methods:

**A.** Continuous circulation of water, making periodic DOC determinations at inlet and outlet.

**B.** Continuous circulation with a closed circuit at the same working flow of a special mixture of water (one-third river water [1.2 µm filtered] and two-thirds GAC-filtered water) that ensures a diverse population in the biofilm. The water mixture is changed daily. This takes ~5 days. There is then an adaptation period of 5–8 days with the sample water (Frias, Ribas, and Lucena 1992).

The second method was reported by Ribas, Frias, and Lucena (1991) to be superior.

A third version of this assay uses a single-column, single-pass system, where the darkened-glass column is also filled with acid-washed and muffled porous glass beads (Siran 60-µm pore). BAS can also be used as a support material. The column is attached to a peristaltic pump. Colonization of the column with a microbial community is achieved at a flow rate of 1 mL/min using 1.0 µm filtered water from the water source and stabilized over time (~1 month). The biofilm is maintained by feeding it with filtered sample water. DOC is measured in water from the inlet and outlet. The residence time of the sample water in the reactor is ~2 to 3 hours. BDOC is calculated from the difference in DOC between the inlet and outlet. Carbon needs to be limiting. A good relationship exists between batch and bioreactor—a bioreactor tends to utilize about 20% higher DOC (Søndergaard and Worm 2001). BDOC may be underestimated by this method if a retention time of less than 2 hours is used, as the reactor column may not allow full degradation of samples with high concentrations of biodegradable compounds. A major advantage of this method is that test time takes only a few hours (2–3 hours). This advantage is contrasted by the time it takes to set up a stable biofilm.

**2. Batch experiment (Billen–Servais method).** The Billen–Servais method is based on the removal of DOC that occurs during sample incubation with a nonattached, indigenous bacterial inoculum. The DOC is measured with a suspended bacterial inoculum for a period of 10–30 days during which time the DOC, and quantity and size of bacteria are measured daily, or on day 1 and at ~28 days. Samples are filtered through a 0.2-µm filter. The filters are prerinsed with 300 mL distilled water and 300 mL sample water, and indigenous microbes (5 mL [Servais, Billen, and Hascoet 1987]; 2 mL [Servais, Anzil, and Ventresque 1989]) are added to the sample after being filtered through a 2.0-µm filter. After incubation in the dark at 20°C, BDOC concentrations are derived from the difference between the initial and the resulting minimum DOC levels at the end of the incubation period. These samples can be preserved with a few drops of a saturated mercuric chloride ($HgCl_2$) or sodium azide solution and stored in sealed glass ampoules. DOC analyses are performed using a total carbon analyzer. Duplicate analyses are performed on each sample. In the range of 0.2 to 1.5 mg C/L, the agreement between added substrate concentration and BDOC is good and validates this method (Servais, Anzil, and Ventresque 1989). These authors also claim that the accuracy of DOC determinations is about 2% in the 0–10 mg C/L range (Servais, Billen, and Hascoet 1987).



©2007 AwwaRF. All Rights Reserved.

A disadvantage of this method is that the best recovery of BDOC requires a long incubation period and difficulties arise in choosing a clear-cut minimum DOC concentration. The incubation period can be shortened either by increasing the inoculum size or by adding a more highly concentrated inoculum. However, this leads to potential carbon contamination of low-carbon drinking water samples.

**3. Bacterial inoculum attached to sand particles (Joret–Levi method).** In the Joret–Levi approach, the main principle is to evaluate the BDOC of water by DOC reduction in a sample inoculated with a natural, acclimated biomass, fixed on sand particles, within a short period of time (Joret, Levi, and Volk 1991). In brief, water samples are inoculated 3:1, in duplicate, with prewashed (until no detectable DOC is released) BAS (e.g., 300 mL water:100 g sand and incubated at 20°C under a 4-L/hr aeration). The BAS is obtained from a sand filter at a water treatment plant that does not use prechlorination. A sample fraction, in duplicate, is collected daily and filtered through a 0.45-µm membrane. Using a TOC analyzer, DOC is measured until a minimum DOC is reached (~1 week). The initial TOC is measured and labeled $DOC_0$, and the samples are incubated at room temperature (22°–25°C). On days 3 through 6, DOC is measured until a DOC minimum is reached (Escobar and Randall 2001). BDOC concentrations are derived from the difference between the initial and minimum DOC levels (Volk and LeChevallier 2002, Volk et al. 1994).

Advantages of this method are that it is accurate and precise, gives reproducible results, is sensitive even for distributed waters, gives a good relationship between BDOC and AOC (Joret, Levi, and Volk 1991), and is good for biofilm situations. It is thus superior to the Billen–Servais method or to the use of inert material in terms of incubation period, recovery, and reproducibility, indicating a higher degree of precision. Park et al. (2004) recommends that this assay be used for analysis of drinking water samples to determine BDOC due to the benefits previously listed. Another advantage is that biodegradation of samples by recirculating the water through columns containing an acclimated community of microorganisms attached to a fixed medium results in reproducible measurements of BDOC. As others have discovered, this BDOC method is a promising technique and could be standardized by selecting a given column size, media size and type, recirculation flow rate, and seeding period. For repeated tests, columns could be maintained for long periods at treatment plants, thereby reducing startup costs (Cipparone, Diehl, and Speitel 1997). Because a series of samples are taken during each assay, it is also possible to determine the rate at which the DOC is consumed related to the biodegradability of the organic compounds present in the water sample (Jago 1993).

Disadvantages of this method are that the kinetics of DOC reduction depend on the origin of the bacterial inoculum. Conditions in a bottle exclude development of complex bacterial communities with diverse decomposition capability that could bias results (Søndergaard and Worm 2001). The release of organic carbon also necessitates daily analysis (Block et al. 1992). The biofilm column requires constant attention and maintenance, including periodic inspection of pumping equipment for a continuous supply of filtered and/or dechlorinated water to the column. Anaerobic conditions are likely to occur within the column, especially during extended running times or at elevated DOC concentrations. This assay requires an extended preincubation period for the bacterial attachment to sand particles (Park et al. 2004).

**4. Bacterial biomass determination (mortality flux)** (Servais, Billen, and Hascoet 1987). This assay estimates the flux of organic matter flowing through bacteria during the batch experiment previously described in method 2. Mortality can have a profound effect on the dynamics of bacterial populations and must be taken into account for correctly interpreting the



©2007 AwwaRF. All Rights Reserved.

time variations of bacterial numbers. Bacterial biomass reaches a maximum after 75 to 100 hours. A low level of biomass results after 200 to 300 hours, and the total integrated flux of bacterial mortality is equal to the total flux of bacterial production. Enumeration of bacteria is performed by epifluorescence microscopy after acridine orange staining (Kaplan and LeChevallier 1993). Cell sizes are visually estimated by reference to a calibrated grid. Biovolumes are converted to biomass carbon yield using a conversion factor of $1.2 \times 10^{-13}$ g C/$\mu m^3$. The determination of the bacterial mortality flux is performed on a 100-mL sample taken from the incubation mixture. To this subsample a small quantity of radioisotope-labeled thymidine is added. This reaction mixture is incubated in the dark, and within 20 hours the thymidine is assumed to have been completely incorporated into the bacterial DNA (deoxyribonucleic acid). The radioactivity of the bacterial DNA is then monitored every 10–20 hours during a period of several (~12) days. The decay observed is taken as an estimate of bacterial mortality. This value is then used to derive the integrated flux of bacterial mortality value. An estimate of the BDOC is obtained from the relationship between the observed biomass carbon yield and the derived mortality flux by using the following formula: BDOC = mortality flux (Fmort)/biomass carbon yield (Y) (Jago 1993). There is good agreement between the DOC method and this method.

An advantage of this method is that it can be used when DOC is low, as it is more accurate and more sensitive than other methods. A disadvantage is that it is more time consuming than other methods, requiring daily analysis of the bacterial biomass and at least one determination of the bacterial mortality rate.

**5. Combined BDOC and biochemical oxygen demand method.** Khan (Khan et al. 2003) developed a modified BDOC method by combining a batch BDOC protocol with biochemical oxygen demand (BOD) techniques. The water sample is filtered through a 0.7-µm glass-fiber filter prerinsed with 300 mL of deionized (DI) water (TOC <0.20 mg/L). The TOC of the filtrate is analyzed and reported as DOC. A dilution factor, F, is calculated to ensure that adequate dissolved oxygen (DO) of ≥1.0 mg/L remains at the end of the test. The dilutions are made with DI water to produce at least 320 mL of combined volume and placed in a washed container with at least 20% gas volume. The mixture is saturated with DO by shaking. After shaking, a 20-mL sample is collected, measured for TOC, and recorded as DOC*i*. The mixture is than placed in a washed BOD bottle. The DO is measured with a washed probe and recorded as DO*i*. A 2-mL inoculum of unfiltered water sample containing indigenous bacteria is then added. The bottle is water-sealed and incubated in the dark without shaking at 20° ±0.5°C for 28 days. At the end of the incubation period, the DO is measured (DO*f*). Then 20 mL of the supernatant is collected and directly measured for TOC (DOC*f*). A seed control is prepared in the same way except that the 2-mL seed is added to 300 mL of dilution water with no sample and the values are recorded as DO*bi*, DOC*bi*, DO*bf*, and DOC*bf*. The BDOC and ultimate soluble biochemical oxygen demand (SBOD*u*) are calculated using the following equations:

$$\text{BDOC (mg/L)} = [(\text{DOC}i - \text{DOC}f) - (\text{DOC}bi - \text{DOC}bf)]\text{F} \tag{5.1}$$

$$\text{SBOD}u\ \text{(mg/L)} = [(\text{DO}i - \text{DO}f) - (\text{DO}bi - \text{DO}bf)]\text{F} \tag{5.2}$$

where F = (mL of dilution water + mL of sample)/mL of sample, observing the criteria of (DO*i* – DO*f*) ≥2 mg/L and DO*f* ≥1 mg/L. This method can be used to characterize DOC in the range of 4–15 mg/L (Khan et al. 1998). Using a larger volume and/or more concentrated inoculum



©2007 AwwaRF. All Rights Reserved.

reduces incubation time from 28 to 5 days. The precision of the protocol ranges from 0.05 to 0.18 mg/L. The BDOC measurement is not sensitive to inoculum origin and size.

**6. Immobilized cell BDOC determination** (Khan et al. 2003). This method adapts a cell immobilization technique proven to work for wastewater treatment. Cell immobilization permits the use of high cell concentrations and better microbial utilization of substrate even at low concentrations. The cells are immobilized by mixing 20-g cells and 100 mL 10% w/v cellulose triacetate in methylene chloride and 20 mL water. Toluene is added as a hardening agent. When hardened, the mixture is cut into $5 \times 5 \times 5$ mm$^3$ cubes. The cubes are flushed with tap water until the toluene is washed out (12–24 hours). The reactor has a volume of 3.11 L and is constructed from a 50 cm $\times$ 8.9 cm polymethyl methacrylate column. Immobilized cells equal 1.49 L of the volume, inlet and outlet zones equal 0.62 L of the volume, and a void volume of 1 L air provided into the feedwater tank at an air flow rate of 16.86 mL/min water sample was fed through the bottom of the reactor. The air is kept at 4°–10°C to reduce bacterial regrowth. Effluent is discharged through the top of the column. The reactor is operated at 20 ± 5°C. A blank containing <200 µg C/L is also run through the system.

$$\text{BDOC} = (\text{DOC in} - \text{DOC out}) - (\text{DOC blank in} - \text{DOC blank out}) \tag{5.3}$$

The advantage of this assay is an extensive life expectancy of the embedded cells, potentially as long as 10 years, an estimate based on a 146-day measurement. Cells do not require acclimatization or a startup period because they have been incubated with the test water only. This method is comparable to a batch test. Disadvantages of this method are that BDOC can be underestimated if the hydraulic residence time is less than 3 hours, DOC can be released from the carrier, and the detection limit is 150 µg/L.

**7. Morgen method.** This method makes use of the advantages of high microbial biomass by using bacteria attached to sand grains. The use of a recirculation system ensures contact between the microbial community and the nutrients in the water sample. The sample water is recirculated through a column containing sand from a water treatment sand filter. The reactor column contains between 100 and 125 mL of acclimatized sand. A vacuum forces the sample water from a reservoir to the top of the sand column to recirculate the water sample. The water then percolates through the column at a rate of 25 to 30 mL/min and feeds back to the reservoir. The action of the vacuum also pulls air into the reservoir, thereby aerating the water. The reactor is protected from light to prevent algae growth. At the onset of the testing phase, the reactor is conditioned with the test water sample by rinsing the column through three times. This uses a total of 200 mL of the sample that is discarded. As soon as the rinsing is complete, an additional 300 mL of the sample is added to the reservoir. Daily samples are taken for DOC analysis. Prior to analysis, the DOC samples are filtered through 0.45-µm filters and acidified with phosphoric acid. DOC analyses are performed on each sample using a TOC analyzer. The resulting BDOC value is calculated as the difference between the lowest DOC concentration (typically day 5) and that present at the start of the test. The source of sand in the reactor made no difference in the results. The method is limited by the detection limits and sensitivity of DOC analysis, which only becomes a problem in waters with extremely low BDOC (Jago 1993).



©2007 AwwaRF. All Rights Reserved.

## BDOC Versus AOC

The expression of results for BDOC methods is in milligrams of carbon per liter rather than acetate carbon equivalents per liter. Comparisons of AOC and BDOC have confirmed that BDOC values are always higher than AOC values for a given sample type (Jago 1993).

A disadvantage of the AOC methods is that they do not provide information about the effects or the rate of uptake of nonorganic growth-promoting compounds (e.g., ammonia, sulfides). On the other hand, BDOC cannot be used to predict the level of regrowth because no significant correlation has been reported between this parameter and the counts of heterotrophic bacteria. This is probably due to the detection limit of 0.1–0.2 mg/L for the BDOC bioassay and because BDOC is both mineralized (to carbon dioxide, or $CO_2$) and assimilated (to biomass) carbon. The higher detection limit means significant AOC changes on the order of up to 100 µg/L as acetate carbon are usually undetectable by BDOC analysis (Escobar and Randall 2001).

## Summary

Growth-based methods include measurements of AOC based on distinct bacterial strains. The readout of these methods varies from standard plate counts, to ATP measurements, to the measurement of bioluminescence in more recent methods. The van der Kooij method (using P17 and NOX) is accepted as a typical bioassay method for measuring AOC by the Standard Methods Committee in *Standard Methods* (APHA, AWWA, and WEF 1998). However, this method takes, at best, 11 days to complete and collect results, in which time the results of the test cannot be directly related to the current conditions of the water source.

Organic carbon–based methods include measurements of BDOC using indigenous populations of bacteria in suspension or attached to a support medium. The time required to run these tests has been shortened from nearly 28 days to 2 hours. The main drawback of BDOC procedures is the poor sensitivity of the TOC analyzers. Improvements in TOC analysis have broadened the application of BDOC, with TOC analyzers now reaching sensitivities of 4 µg/mL, thereby enabling accurate measurements of drinking water sources with lower DOC levels. No method has either been officially accepted or standardized for BDOC, which makes it difficult to compare published values.

Methods for the measurement of BDOC and AOC are continuously being modified with emphasis on shortening the length of time necessary to run these tests. Current efforts are focusing on developing real-time measurements (e.g., the CheckLight method for measuring AOC). AOC and BDOC levels for the same raw water source are seen to fluctuate widely over time (Kemmy, Fry, and Breach 1989). CheckLight Company reports that their CheckLight system has a good correlation with the standard AOC method. This correlation has yet to be independently confirmed by peer-reviewed literature. Comparisons of BDOC bioassays for analyzing BDOC in drinking water are still being studied (Park et al. 2004). These comparisons can lead to suggestions for a standardized BDOC assay.

AOC and BDOC measurements should be viewed on a case-by-case (watershed-by-watershed) basis to determine which has the most significant meaning. Values may vary with geographic location and water source. Experiments have demonstrated (Yavich et al. 2004) that AOC and BDOC vary in source waters from region to region and from day to day. It is recommended that drinking water treatment plants should extensively evaluate the nature of their source waters before initiating changes in blending of their raw waters. Consideration should



©2007 AwwaRF. All Rights Reserved.

also be given to the inorganic matrix. In some regions of the world (Finland and Japan; Lehtola et al. 1999), phosphorus instead of AOC or BDOC limits the heterotrophic microbial growth in drinking water (Lehtola et al. 2003). Finnish surface waters contain a high content of humic substances that are derived from forests and peatlands typical for northern latitudes (boreal region). When inorganic nutrients are added to an AOC experiment, the AOC values are higher than without this addition (Miettinen, Vartiainen, and Martikainen 1999); here only a potential AOC can be measured rather than an actual AOC.

Seasonal changes in the nature of DOC can have significant impact on water treatment processes. There is, therefore, a need to characterize the nature of the organic matter entering a treatment system. DOC should be considered a complex mixture of organics rather than a bulk parameter (Goslan et al. 2004). BOM levels appear to vary with both source location and season. Routine monitoring of raw and treated water must be conducted on a site-specific basis to determine bacterial regrowth potential or biofilm formation.

The percentage of DOC that is actually available to microorganisms as AOC reflects the portion of organic carbon available for bacterial regrowth or a measure of the stability of organic matter. In a broad spectrum of studies, 0.5%–31% of the DOC for waters tested was assimilable, which is higher than the 0%–4% found for the Netherlands, as reported by van der Kooij, Hijnen, and Kruithof (1989). Surface waters have higher AOC concentrations than groundwaters. It is suggested that water utilities concerned with issues of biofilm growth measure AOC on a regular basis. This would help determine the reasonable AOC threshold for a given water supply below which biofilm-related problems are minimized. It is very difficult to generalize a "safe" AOC value across different systems (i.e., differing geographic regions, mean temperatures, pipe materials, raw water quality, treatment processes, etc.). Any treatment process such as ozonation that increases AOC must be followed by additional treatment to remove as much AOC as possible (Bradford, Palmer, and Olson 1994).

The lack of consistent correlation of AOC and/or BDOC with regrowth and biological stability is due to several issues. These include an inadequate number of BOM measurements to truly represent the level of organic carbon available for growth, a significant interaction of other factors (disinfectant, distribution system materials, etc.) that govern microbial growth, and/or the presence of organic carbon promoting biological growth not measured by these tests. BOM tests can be difficult for the average utility to perform, and high costs result in a limited number of data points for a particular system (Camper 2004). AOC and BDOC provide complementary information and it is advisable to measure both, if possible. There is concern that measuring only BDOC can lead to an overestimation of the biological stability in some cases. Measuring only AOC can, however, lead to an underestimation of biostability (the level of organic carbon present at which bacterial regrowth will not occur) (Escobar and Randall 2001).

A standardized AOC and BDOC method should be used in measuring the regrowth capabilities of drinking water systems until evaluations of real-time methods give a consistent and accurate result. Further investigations of alternative methods to measure AOC and BDOC should also be conducted. Possible directions for research are the use of, for example, fluorescent in situ hybridization (FISH) probes and flow cytometry (FCM) with the use of dyes. FISH probes use ribosomal ribonucleic acid (rRNA)-targeted oligonucleotide probes to directly probe nucleic acids. FISH probes in combination with staining techniques can be used to enumerate the numbers of bacterial cells associated with oligotrophic environments (Langmark et al. 2004). The number of bacteria can be related to the amount of BOM in a water system. It is hypothesized that creating probes to detect and count several known bacterial species (e.g., *Pseudomonas*,



©2007 AwwaRF. All Rights Reserved.

*Spirillum*, and *Aeromonas*) could give an estimate of AOC. FCM uses laser light to measure the number and types of microbes (Gruden, Skerlos, and Adriaens 2004). This method offers high speed, sensitivity, and reproducibility over current microbe detection methods; however, it is underutilized in environmental applications. A method such as FCM could replace plate counting in the standard AOC method, thereby reducing labor time and costs. FCM can also be used in conjunction with an esterified fluorogenic substrate that will measure intracellular esterase activity in physiologically active bacteria (Hoefel et al. 2003). The number of cells detected by FCM to be active can be related to the amount of AOC in the water system by comparing numbers to a standard curve where bacteria have been exposed to varying amounts of a carbon cocktail. This method could be used to measure the indigenous bacterial community, including chlorine-injured cells that are capable of regrowth in a water distribution system. However, this method will give variable results due to changes in the bacterial populations.

## METHODS DEVELOPMENT

### NOM Isolation Protocol

The original plan was to use XAD-8 resin adsorption (removal of humics) followed by UF (removal of APS) for the general fractionation procedure. However, concerns were expressed (J.A. Leenheer, personal communication) that APS could irreversibly bind with XAD-8. At the time the plan was to test this procedure along with one or more alternatives using extracted, purified marine APS and aquatic fulvic acid (International Humic Substances Society [IHSS] Suwannee River FA). After receiving a small sample of marine APS from the laboratory of Daniel Repeta, it was determined that this material was not retained on XAD-8 to an appreciable extent. However, further discussions led to the suspicion that the marine APS may be lower in molecular weight and substantially more hydrophilic than freshwater APS. The final judgment was to apply XAD-8 adsorption after UF in accordance with the original recommendation. Tests with Suwannee River FA also confirmed that this material would pass through the 3-kD ultrafilter and be almost completely retained on the XAD-8 resin.

#### *Preliminary Tests*

Three model organic substances (APS, FA, and SPP) were selected for the proposed fractionation tests and are listed in Table 5.3. APS and FA were dissolved in DI water at 2 mg C/L and 4 mg C/L, respectively. SPP solution is a mixture of amino acids (serine, glycine), sugars (fructose, sucrose), malonic acid, and phenol at 2 mg C/L each. A phosphate buffer ($NaH_2PO_4 \cdot H_2O$) at 1 m*M* was added to each solution to adjust the pH to 7. All samples were prefiltered through glass microfiber filters (GF/F) to remove large particles prior to either UF or XAD-8 adsorption.

An experimental schematic for fractionation is shown in Figure 5.1. UF experiments were performed at room temperature in 200-mL stirred cells (cell model 8200; Amicon, Beverly, Mass.) pressurized to 50 psi using nitrogen gas. The nominal molecular weight cutoff (MWCO) of the regenerated cellulose membrane, as specified by the manufacturer, was YM1 = 1,000 AMU (atomic mass units). After each batch analysis, the UF membrane was soaked in NaOH (0.1*N*) for a half hour and then stored in a refrigerator in 10% ethanol solution.

Source (feed) waters were adjusted to pH ~2 (with phosphoric acid, or $H_3PO_4$) before being passed through a nonionic XAD-8 resin (acrylic ester polymer). Each sample was loaded



©2007 AwwaRF. All Rights Reserved.

**Table 5.3**
**Model compounds for fractionation tests**

| Substance* | Supplier | Source | % C |
|---|---|---|---|
| APS | Marine Chemistry and Geochemistry Department, Woods Hole Oceanographic Institution (Woods Hole, Mass.) | Surface seawater sample from Kona, Hawaii, NELHA H02OOC | 36† |
| FA | IHSS (Denver, Colo.) | Suwannee River (reference sample) | 53.04‡ |
| SPP | | | |
| Amino acids: | | | |
| L-(-)-serine | Kodak (Rochester, N.Y.) | | 34.25 |
| Glycine | Aldrich (St. Louis, Mo.) | | 31.97 |
| Sugars: | | | |
| D-fructose | Aldrich | | 39.96 |
| Sucrose | Fisher Scientific (Pittsburgh, Pa.) | | 42.07 |
| Fatty acids: | | | |
| Malonic acid | Aldrich | | 34.59 |
| Phenolic compounds: | | | |
| Phenol | Aldrich | | 76.51 |

*APS = acylheteropolysaccharide; FA = fulvic acid, SPP = simple plant products.
†Elemental analysis (performed by Marine Chemistry and Geochemistry Department, Woods Hole Oceanographic Institution).
‡Elemental analysis (provided by supplier).

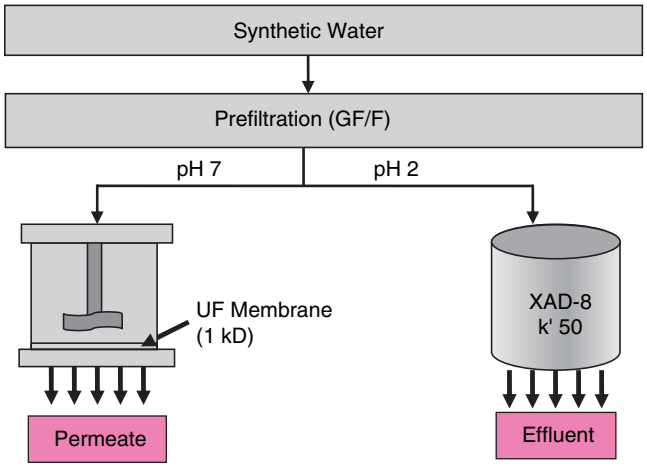



**Figure 5.1 Experimental schematic for fractionation tests**



©2007 AwwaRF. All Rights Reserved.

**Table 5.4**
**UF fractionation of synthetic waters**

| Synthetic waters | Feed (mg C/L) | Permeate (mg C/L) | Concentrate (mg C/L) | Concentration Factor* | Rejection* (%) |
|---|---|---|---|---|---|
| APS (2 mg C/L) | 1.99 | 0.23 | 4.37 | 2 | 88.42 |
| FA (4 mg C/L) | 3.72 | 0.70 | 7.12 | 2 | 81.18 |
| SPP (2 mg C/L each) | 11.54 | 10.70 | 11.68 | 1 | 7.24 |

*Concentration factor = Cc/Cf, and Rejection = (1 – Cp/Cf) × 100, where Cc = concentrate concentration; Cf = feed concentration; Cp = permeate concentration.

**Table 5.5**
**XAD-8 fractionation of synthetic waters**

| Synthetic waters | Feed (mg C/L) | Effluent (mg C/L) | Adsorption (%) |
|---|---|---|---|
| APS (2mg C/L) | 2.21 | 2.02 | 8.56 |
| FA (2 mg C/L) | 2.99 | 0.98 | 67.10 |
| SPP (2 mg C/L each) | 9.12 | 8.48 | 7.05 |

onto the XAD-8 resin (bed volume 8 mL, porosity 0.65) at a flow rate of 2 mL/min. The column capacity factor (grams of solute on resin per gram of solute in void volume) was fixed at 50.

Tables 5.4 and 5.5 summarize the results of UF and XAD-8 fractionation of each synthetic compound.

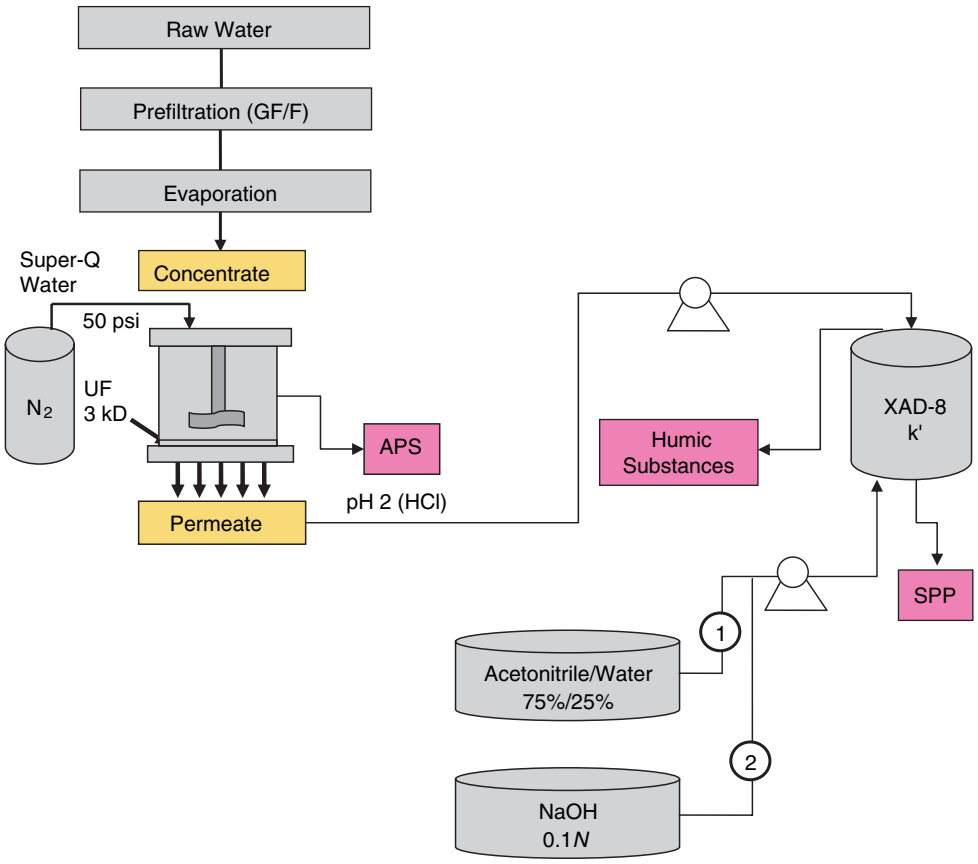
Results show that the SPP fraction is not retained by the YM1 membrane (Amicon) and does not adsorb onto XAD-8 resin. In contrast, the FA fraction is highly retained by the YM1 membrane and XAD-8 resin. Finally, the APS fraction is only retained by the YM1 membrane. Therefore, the initial conclusion was that the best experimental design would be adsorption through XAD-8 resin followed by UF. However, because of suspicion that the marine APS was older and more depolymerized than the freshwater APS fraction—which means that freshwater APS is more likely to coil up inside XAD-8 resin beads and be irreversibly retained by the UF membrane with higher MWCO—the fractionation protocol was set up as UF followed by XAD-8 resin adsorption.

### *Final Protocol*

The isolation protocol combines vacuum evaporation and UF before passage through nonionic XAD-8 resin (acrylic ester polymer). The experimental scheme for isolation of the three major pools of NOM is shown in Figure 5.2. Initially, the adopted protocol called for elution of the XAD-8 resin column—first with 75% acetonitrile and 25% water, followed by a second elution with sodium hydroxide 0.1 N, as shown in Figure 5.2. However, subsequent tests revealed that the second elution did not yield a substantial amount of additional material. This information,



©2007 AwwaRF. All Rights Reserved.




NOTES: $N_2$ = nitrogen gas; HCl = hydrochloric acid.

**Figure 5.2 Experimental scheme for isolating the three major pools of NOM**

along with concerns about the high salt content, resulted in removal of the NaOH elution step from the final protocol.

The stepwise comprehensive NOM isolation procedure is as follows.

**Step 1:** Acidify 30 L of sample to pH 4 using hydrochloric acid (HCl). Vacuum-evaporate a 30-L sample to 140 mL. Rinse the rotary evaporation flask with weak HCl (pH 5–6) to dissolve any salt (calcium carbonate) that adheres to the flask. Combine this rinse with the 140-mL sample. Keep the pH in the range of 4–5. Place the sample in three Amicon stirred cells (model 8200) pressurized to 50 psi using nitrogen gas ($N_2$) and rinse continuously with Super-Q water (Millipore) until permeate is colorless. (The $UV_{254}$ [ultraviolet absorbance at 254 nm] value will decrease and become stable when no color passes through the membrane.) (In this study, the nominal MWCO of the regenerated cellulose membrane, as specified by the manufacturer, was YM3 = 3,000 AMU.) Freeze-dry the concentrate to isolate the APS fraction. After each batch analysis, clean the UF membrane by soaking it in 0.1*N* NaOH for a half hour, then store in a refrigerator in a 10% ethanol solution.

**Step 2:** Adjust the pH of the combined permeate sample from UF cells to 2, then pump the sample through the XAD-8 resin. (The column capacity factor, k' value, is fixed at 50.) Follow the sample with a pH 2 formic acid (0.1 N), and rinse until the conductivity of the effluent is the same as the conductivity of the column influent in order to remove the residual chloride. Elute the XAD-8 resin column with 75% acetonitrile/25% water (the $UV_{254}$ value will decrease and become



©2007 AwwaRF. All Rights Reserved.

stable). Evaporate to remove elution solvent (acetonitrile) and formic acid, then freeze-dry to isolate the HPO fraction—also called the humic substances fraction. Adjust the pH of the fraction eluted with NaOH to 7 before the evaporation and freeze-drying processes.

**Step 3:** Evaporate the effluent of XAD-8 resin column, then freeze-dry to isolate the HPI fraction—also called SPP.

**Lignin Testing With Alkaline CuO Method**

A programmable microwave digester was purchased for use with the rapid lignin method. Unfortunately, the high NaOH concentrations used in this method resulted in arcing, charring, and focusing of the microwave energy, ultimately leading to a catastrophic failure of the digester. This required the consideration of other alternatives to the previously developed method for lignin analysis. Toward this end, the more classical oven method was applied, using polytetrafluoroethylene (PTFE)-lined stainless-steel reactors (Figure 5.3). Although the sample concentration, high-performance liquid chromatography (HPLC) separation, and quantification steps did not change, the digestion and chemical degradation steps are different. The current protocol is as follows:

A. Classical Oven Digestion

1. Add 50 mg of the dried NOM sample to a PTFE-lined mini-bomb.
2. Add:
   - 2 g of fine CuO powder
   - 0.20 g $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ (ferrous ammonium sulfate hexahydrate)
   - 14 mL of boiled and/or $N_2$-sparged 2*M* NaOH
   - 13.3 µL of the 5.49-m*M* stock solution of ethyl vanillin
   - Stir bar
3. Heat vessel for 3 hours at 180°C in a preheated oven.
4. Cool down with iced water.

B. Postdigestion Treatment

1. Add known amount of secondary recovery standard to each reaction vessel (40 µL *trans*-cinnamic acid [40 µL of 6.75-m*M* stock]).
2. Transfer contents to a 50-mL centrifuge tube, and rinse with a small amount of 1*N* NaOH.
3. Centrifuge (3,000 rpm for 10 minutes) samples to remove solids.
4. Transfer supernatant to extraction vial.
   A. Decant supernatant from reaction vessel into an appropriate vial.
      - For classical method, use 50-mL borosilicate glass tubes fitted with PTFE-coated caps, or larger heat-resistant glass and PTFE vessels, as needed.
      - For microwave digestion, use 50-mL heat-resistant glass tubes fitted with PTFE-coated caps.
   B. Add additional 1*N* NaOH to each tube to help with quantitative transfer.
      - For classical method, use about 20 mL of 0.1*N* NaOH.
      - For microwave digestion, use about 5 mL of 0.1*N* NaOH.
   C. Repeat centrifuge step (3,000 rpm for 10 minutes).



©2007 AwwaRF. All Rights Reserved.

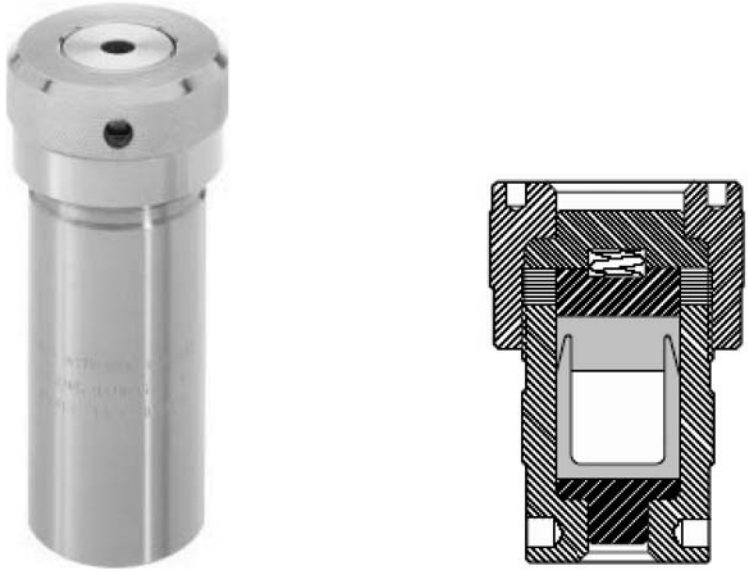


**Figure 5.3 Digestion reactor used with classical oven method**

5. Acidify solution to about pH 1.
   - For classical method, add about 4 mL of concentrated HCl.
   - For microwave digestion, add between 4 and 40 mL of concentrated HCl, depending on total volume.

Concern about the use of alkaline CuO for characterization of NOM in this research led to conducting some limited preliminary investigations with this method. Aside from questions about reproducibility, there was concern that the recovery of lignin monomers would be so low as to render this form of characterization nearly meaningless. For this reason, the decision was made to conduct a series of tests with commercial lignin material and to compare these in parallel tests with commercial aquatic humic substances. Results showed recoveries of up to 36% based on total mass with the oven method (Table 5.6).

Analysis of humic substances produced overall mass yields of 0.5% to 1% (Table 5.7). Based on the pure lignin yields, this suggests that the aquatic humic substances have no more than 10% lignin content, and possibly much lower. The significance of this finding will be discussed in the next chapter.

### Evaluation of the LeChevallier ATP Rapid Method for Measuring AOC

In this study, a determination was made as to whether this new method that claims to be "rapid" should replace the AOC standard method listed in *Standard Methods* EWW 9217 B (APHA, AWWA, and WEF 1998).

The method was followed exactly as described by LeChevallier et al. (1993) in the article titled "Development of a Rapid Assimilable Organic Carbon Method for Water." Briefly, water samples were prefiltered through GF/F and then pasteurized. Vials containing 40 mL of water sample were inoculated in triplicate with $10^4$ cfu/mL of pure culture *P. fluorescens* P17 or *Spirillum* strain NOX. The vials were incubated at room temperature for 1 to 3 days.



©2007 AwwaRF. All Rights Reserved.

**Table 5.6**
**Commercial lignin analysis by microwave and oven digestion (mg/100 mg C)**

| | Microwave digestion | | Oven digestion | |
|---|---|---|---|---|
| Structure/group (abbreviation) | Organosolve lignin | Alkali lignin | Organosolve lignin | Alkali lignin |
| *p*-Hydroxybenzoic acid (PAD) | ND* | ND | ND | 0.08 |
| Vanillic acid (VAD) | 0.11 | 1.28 | ND | 0.06 |
| *p*-Hydroxybenzaldehyde (PAL) | 0.07 | ND | 0.83 | ND |
| Syringic acid (SAD) | 0.17 | ND | ND | ND |
| Vanillin (VAL) | 3.69 | 7.33 | 11.85 | 7.78 |
| *p*-Hydroxyacetophenone (PON) | ND | ND | ND | ND |
| Syringaldehyde (SAL) | 11.55 | ND | 11.15 | ND |
| Acetovanillone (VON) | 1.13 | 2.38 | 5.01 | 3.31 |
| *p*-Coumaric acid (CAD) | ND | ND | ND | ND |
| Acetosyringone (SON) | 5.74 | ND | 7.12 | ND |
| Ferulic acid (FAD) | 0.00 | ND | ND | ND |
| Vanillyl phenols† | 4.94 | 10.99 | 16.86 | 11.14 |
| Syringyl phenols‡ | 17.46 | ND | 18.27 | ND |
| Cinnamyl phenols§ | ND | ND | ND | ND |
| *p*-Hydroxyl phenols** | 0.07 | ND | 0.83 | 0.08 |
| Total phenols†† | 22.47 | 10.99 | 35.96 | 11.22 |

*ND = not determined.
†Vanillyl phenols (VAD + VON + VAL).
‡Syringyl phenols (SAD + SON + SAL).
§Cinnamyl phenols (CAD + FAD).
***p*-Hydroxyl phenols (PAD + PON + PAL).
††Sum of all lignin phenols.

On days 1 through 3, after the incubation process was started, one sample vial inoculated with either P17 or NOX sample was filtered through a 0.22-µm filter. The filter was placed in 1 mL of a 1:1 mixture of HEPES (N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid) buffer and briefly vortexed. After 20 minutes, 150 µL was placed into a clean tube for light production measurement in a luminometer. One hundred microliters of luciferase was added, and the ATP released was indirectly measured as light output. One luminescence unit equals 1.3 µg AOC/L for P17 and 3.7 µg AOC/L for NOX. AOC is reported as the sum of micrograms of AOC per liter for P17 and for NOX. The results of this evaluation are shown in Table 5.8.

For the entire treatment series based on Tampa Reservoir water, a weak correlation ($r^2$ = 0.53) was found between the two methods (Figure 5.4). However, in a side-by-side comparison of individual treatment pairs, both methods indicated similar trends and thus showed a higher correlation than indicated in the averaging regression line of Figure 5.4. The standard method was the preferred method based on smaller standard deviations. Similar results were achieved with water samples from another watershed (Winnipeg; data not shown). The advantage of time savings with the ATP assay was outweighed by the greater costs for equipment and materials. Different users in the laboratory found that the standard method gave results that were more consistent.



©2007 AwwaRF. All Rights Reserved.

**Table 5.7**
**Analysis of humic substances using the lignin phenol method with oven digestion (mg/100 mg C)**

| Structure/group | Humic acid | Fulvic acid |
|---|---|---|
| PAD | 0.15 | 0.19 |
| VAD | 0.31 | ND* |
| PAL | ND | 0.17 |
| SAD | ND | ND |
| VAL | 0.21 | 0.04 |
| PON | ND | 0.17 |
| SAL | ND | ND |
| VON | 0.28 | ND |
| CAD | ND | ND |
| SON | ND | ND |
| FAD | ND | ND |
| Vanillyl phenols† | 0.79 | 0.04 |
| Syringyl phenols‡ | ND | ND |
| Cinnamyl phenols§ | ND | ND |
| *p*-Hydroxyl phenols** | 0.15 | 0.53 |
| Total phenols†† | 0.94 | 0.57 |

* ND = no data.
† Vanillyl phenols (VAD + VON + VAL).
‡ Syringyl phenols (SAD + SON + SAL).
§ Cinnamyl phenols (CAD + FAD).
***p*-Hydroxyl phenols (PAD + PON + PAL).
†† Sum of all lignin phenols.

**Table 5.8**
**AOC values for the LeChevallier and the *Standard Methods* tests for a water sample series prepared from raw water from Tampa Reservoir waters**

| Sample series (Tampa Reservoir watershed) | LeChevallier test AOC (µg/L) | Standard test AOC (µg C/L) |
|---|---|---|
| Raw | 13 ± 10 | 87 ± 23 |
| Raw ozonated | 98 ± 40 | 203 ± 27 |
| Raw chlorinated | 189 ± 99 | 163 ± 25 |
| HMW fraction | 2.3 ± 1.7 | 17 ± 22 |
| HPO fraction | 7.6 ± 6.0 | 8.5 ± 9.4 |
| HPO fraction, ozonated | 3.4 ± 4.2 | 103 ± 28 |
| HPI fraction | 33 ± 43 | 20 ± 8 |
| HPI fraction, ozonated | 186 ± 33 | 148 ± 28 |



©2007 AwwaRF. All Rights Reserved.

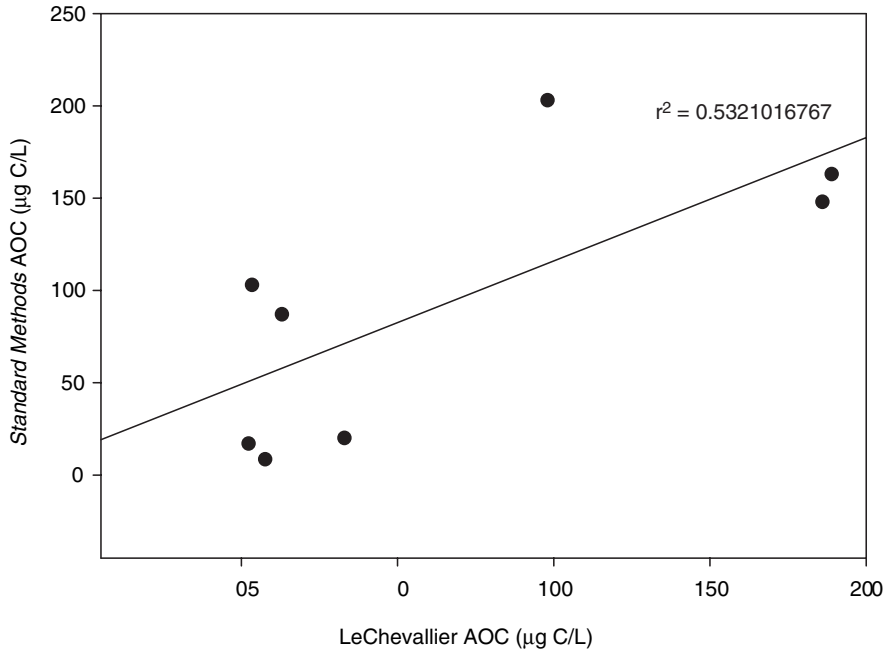



**Figure 5.4 Correlation between LeChevallier test and *Standard Methods* test for AOC. Both tests were applied to aliquots from the same water sample series.**

## ISOLATION AND CHARACTERIZATION OF NOM FRACTIONS

### Introduction

NOM is a heterogeneous mixture of organic molecules ranging from macromolecular humic substances to small-molecular-weight hydrophilic acids and various hydrocarbons (Thurman 1985). Many researchers believe that isolation of NOM into more homogeneous components is critical to better understanding its DBP reactivity. A wide range of methods (based on size or hydrophobic/hydrophilic character) have been developed to isolate NOM from water. Researchers have used UF to fractionate NOM into components of differing molecular weight. UF is a pressure-driven membrane process for separating molecules in solution on the basis of size, often quantified as an MWCO. MWCO is established by the manufacturers, who calibrate membranes by measuring membrane rejection of macromolecules having known molecular weights. However, these molecules have structural characteristics significantly different than NOM. Thus, it is not possible to relate the molecular weight of NOM to an exact MWCO. In addition, the accumulation of retained molecules may form a concentrated gel layer that can significantly alter the performance characteristics of the membrane. Some studies (Amy et al. 1987, Chadik and Amy 1987, Schnoor et al. 1979) have compared the reactivity of various molecular weight NOM fractions for DBP formation. In general, they don't find any significant trend between NOM molecular weight and DBP formation. However, molecular weight fractions greater than 5 kD generally exhibit higher reactivity. It seems that size is not a key parameter for predicting DBP formation, but some weak relationships occur. Others have used XAD-8 resin adsorption for isolating NOM components depending on their affinities to the resin and their



©2007 AwwaRF. All Rights Reserved.

**Table 5.9**
**Water samples**

| Water source | City using water source | Sample code | Sampling date |
|---|---|---|---|
| Chickahominy River | Newport News, Va. | 1 | Mar. 30, 2004 |
| Assomption River | Repentigny, Que. | 2 | Apr. 14, 2004 |
| Shoal Lake | Winnipeg, Man. | 3 | Aug. 10, 2004 |
| Hillsborough River | Tampa, Fla. | 4 | Dec. 1, 2004 |
| Quabbin Reservoir | Boston, Mass.* | 5 | July 6, 2005 |

*Includes Boston plus 47 other cities and towns, most in eastern Massachusetts, for a total of 2.2 million customers.

elution efficiencies. Aquatic NOM consists of both hydrophobic and hydrophilic organic material. The nature and the distribution of these materials differ depending on the water source. It is widely accepted in the form of an operational definition that hydrophobic NOM is retained on XAD-8 resin whereas the hydrophilic NOM is not.

During isolation processes, the original water's organic composition may change as a result of concentration, adsorption, chemical exposure to acids or bases, and freeze-drying. One issue of concern is whether these changes alter the reactivity of NOM in the water source. Many researchers have studied the reactivity of NOM fractions; however, few have examined the impact of fractionation on NOM reactivity. This work is aimed at investigating the formation of DBPs from chlorination of NOM originating from five water samples: (1) Chickahominy River, (2) Assomption River, (3) Shoal Lake, (4) Hillsborough River, and (5) Quabbin Reservoir.

## Experimental Methods

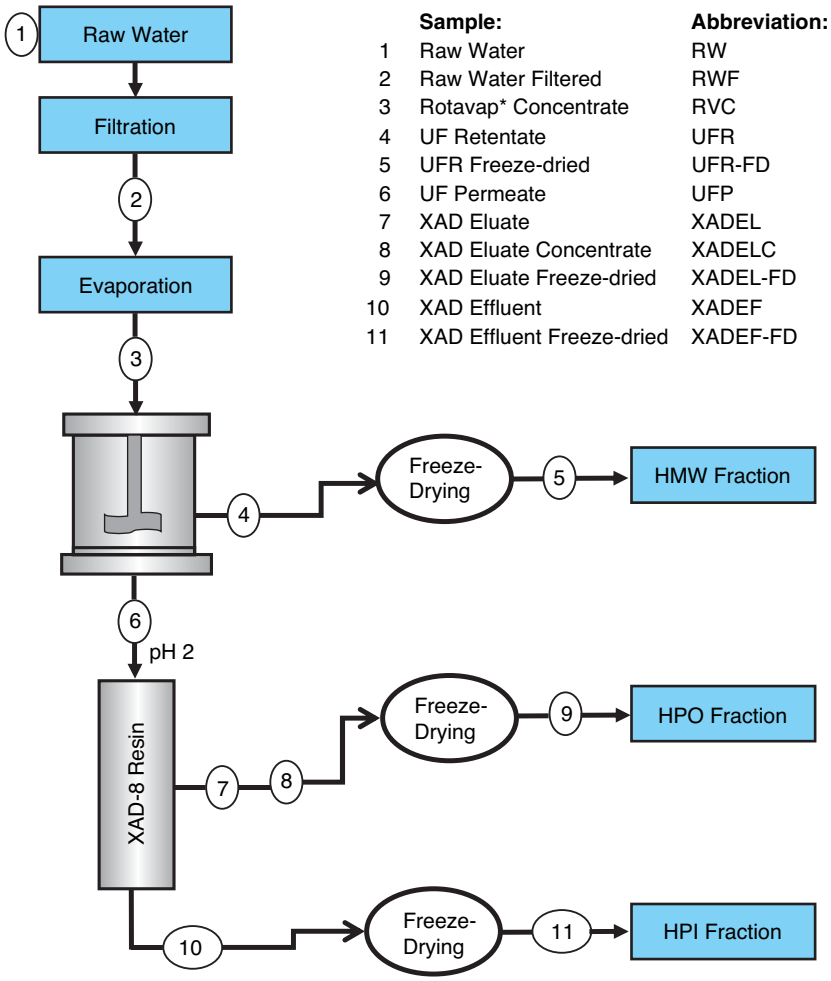
### *Sample Collection*

Five water samples were collected for this study. Sampling locations and sample codes are summarized in Table 5.9. Sampling date is also included in this table.

The Chickahominy River is the principal raw-water source for the city of Newport News in Virginia. It is characterized by low TOC and low SUVA. The Chickahominy is largely a coastal river with urbanized areas in the upper watershed (e.g., Richmond, Va.), and mostly forested areas in the lower reaches. The Assomption River is the raw water source of the city of Repentigny in Quebec. It is heavily impacted by agricultural runoff. The Assomption River contains higher amounts of organic matter than the Chickahominy with a moderate SUVA. Shoal Lake is a part of the larger water body from which Winnipeg draws its water. It is characterized by high TOC levels, yet uncharacteristically low-SUVA values given the TOC. Some have speculated that this water is heavily influenced by autochthonous material. The Hillsborough River serves the city of Tampa, Fla. It has a high and variable TOC, and it is thought to be largely allochthonous. Finally, the Quabbin Reservoir is an oligotrophic impoundment that feeds the Massachusetts Water Resources Authority system responsible for supplying Boston and much of eastern Massachusetts. It is in a well-protected watershed that receives substantial NOM loading from higher terrestrial plants. However, because of its long residence time (~7 years), much of the NOM is attenuated by bacterial degradation and photolysis.



©2007 AwwaRF. All Rights Reserved.




| | Sample: | Abbreviation: |
|---|---|---|
| 1 | Raw Water | RW |
| 2 | Raw Water Filtered | RWF |
| 3 | Rotavap* Concentrate | RVC |
| 4 | UF Retentate | UFR |
| 5 | UFR Freeze-dried | UFR-FD |
| 6 | UF Permeate | UFP |
| 7 | XAD Eluate | XADEL |
| 8 | XAD Eluate Concentrate | XADELC |
| 9 | XAD Eluate Freeze-dried | XADEL-FD |
| 10 | XAD Effluent | XADEF |
| 11 | XAD Effluent Freeze-dried | XADEF-FD |

*Rotary evaporator.

**Figure 5.5 Method for isolating NOM fractions**

### *Fractionation Scheme*

Water samples were processed in accordance with the NOM fractionation scheme shown in Figure 5.5. The isolation protocol used in this study combines vacuum evaporation and UF before adsorption onto nonionic XAD-8 resin (acrylic ester polymer). Figure 5.5 shows the experimental scheme for isolating three major pools of NOM.

The stepwise comprehensive NOM isolation procedure is as follows.

**Step 1:** Pass 30 L of sample through a GF/F (Whatman Inc., Florham Park, N.J.) to remove particulate matter, then acidify to pH 4 using HCl. Vacuum-evaporate the sample to 140 mL. Rinse the rotary evaporation flask with weak HCl (pH 5–6) to remove any salt (calcium carbonate) that adheres to the flask. Combine this rinse with the 140-mL sample. Keep the pH in the range of 4–5. Place the sample in stirred Amicon cell (model 8200), pressurized to 50 psi using nitrogen gas and rinse continuously with DI water until permeate is colorless (the $UV_{254}$ value will decrease and become stable when no color passes through the membrane). (In this study, the nominal MWCO of the regenerated cellulose membrane used, as specified by the



©2007 AwwaRF. All Rights Reserved.

manufacturer, was YM3 = 3,000 AMU.) Freeze-dry the concentrate to isolate the HMW fraction. After each batch analysis, clean the UF membrane by soaking it in 0.1*N* NaOH for a half hour, then store in a refrigerator in a 10% ethanol solution.

**Step 2:** Adjust the pH of the combined permeate sample from the UF cells to 2, then pump the sample through the XAD-8 resin. (The column capacity factor, k' [grams of solute on resin per grams of solute in void volume] is fixed to 50 and the flow is fixed to 500 mL/hr.) Follow the sample with a pH 2 formic acid (0.1*N*), and rinse until the conductivity of the effluent is the same as the conductivity of the column influent in order to remove the residual chloride. Elute the HPO fraction that is adsorbed onto the XAD-8 resin column with 75% acetonitrile/25% water (the $UV_{254}$ value will decrease and become stable when no color passes through the membrane). Evaporate to remove elution solvent (acetonitrile) and formic acid, then freeze-dry to isolate the HPO fraction.

**Step 3:** Evaporate the effluent of XAD-8 resin column, then freeze-dry to isolate HPI fraction.

NOM fractions are isolated as a dry powder and individually redissolved in DI water as needed. When possible, fractions are isolated into aqueous form without any further lyophilization process.

### Characteristics of the Water Samples

All water samples were analyzed for basic organic parameters (TOC, DOC, and UV absorbance) as well as bromide content. Samples for DOC and $UV_{254}$ measurements were filtered through a prerinsed GF/F. Bromide ion concentration was determined using a Dionex Corp. (Sunnyvale, Calif.) ion chromatographic system.

Table 5.10 summarizes the basic organic/inorganic characteristics of the water samples. The five raw waters ranged from low to very high DOC. These are in the low and high percentiles nationwide for raw surface waters (Figure 5.6). "Specific ultraviolet absorbance" is defined as the UV absorbance at 254 nm, multiplied by 100, and divided by the DOC concentration. Hillsborough and Assomption water samples are characterized by relatively high SUVA, suggesting high lignin content, which is typical of highly terrestrial and forested sources. The Quabbin water sample is characterized by low SUVA, indicating low aromaticity. Chickahominy and Shoal water samples are characterized by moderate SUVA, indicating moderate aromaticity. Different bromide concentrations were observed among the various water samples with the Hillsborough water sample exhibiting the highest bromide level.

### NOM Fractionation

All waters were fractionated based on the procedure shown in Figure 5.5. Tables 5.11–5.15 show concentrations, volumes, and mass balances of TOC at several points along the fractionation process. Elemental composition analysis—carbon, hydrogen, and nitrogen for the fractions—is also reported.

The overall distribution of NOM in each water sample is shown in Table 5.16 as the collection of isolatable fractions (extract) from the freeze-drying (FD) step and the amount lost either during freeze-drying and sample transfer (lost FD) or freeze-drying and adsorption onto XAD-8 (lost FD and XAD). A category of "missing NOM" is also distinguished as organic material lost during preliminary evaporation and UF. This portion of NOM is probably a combination of HPI, HPO, and HMW fractions, but the exact breakdown cannot be determined.



©2007 AwwaRF. All Rights Reserved.

**Table 5.10**
**General characteristics of water samples**

| Water sample | TOC (mg C/L) | DOC (mg C/L) | $UV_{254}$ ($cm^{-1}$) | SUVA (L/m·mg C) | Bromide (µg/L) |
|---|---|---|---|---|---|
| Chickahominy River | 5.25 | 4.81 | 0.110 | 2.29 | 32 |
| Assomption River | 12.11 | 7.21 | 0.230 | 3.19 | 45 |
| Shoal Lake | 7.56 | 7.21 | 0.130 | 1.80 | <10 |
| Hillsborough River | 11.87 | 11.77 | 0.430 | 3.65 | 67 |
| Quabbin Reservoir | 2.07 | 2.10 | 0.028 | 1.33 | <10 |

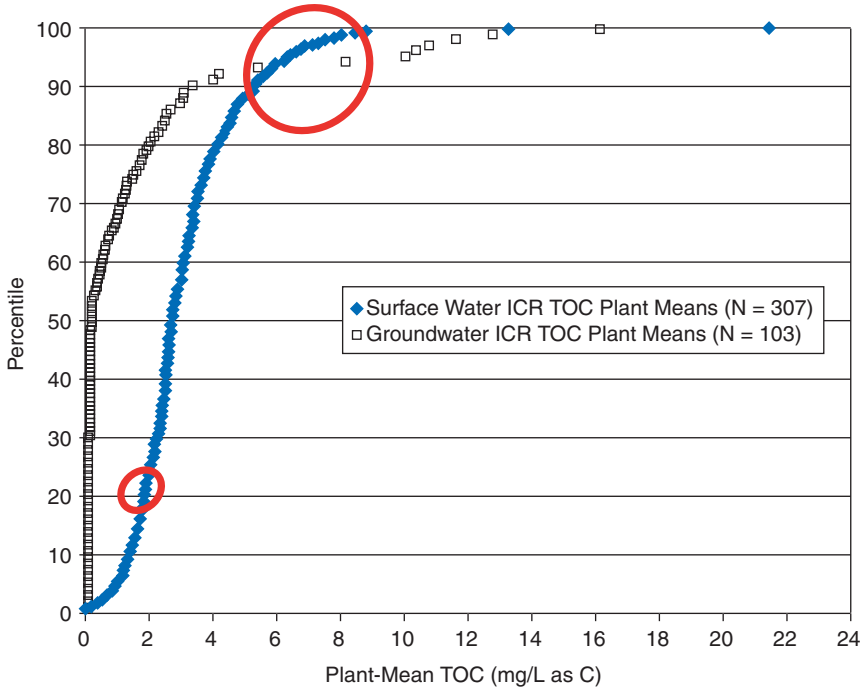



**Figure 5.6 Comparison of TOC values with national ICR data**

Table 5.16 shows that 34% and 23% of the DOC were lost during initial rotary evaporation (rotavap) of the Assomption and Quabbin samples, respectively. Shoal and Hillsborough waters did not show the same extent of loss. Loss in the Chickahominy sample amounted to only 3% of the DOC. NOM losses during this step can be attributed to loss of volatile NOM or sticky and/or insoluble organic and inorganic residues forming on the walls of evaporation vessels. During UF, NOM may be irreversibly adsorbed onto the surface or within the pores of the membrane, resulting in some extent of losses. Freeze-drying of NOM produces powders, which can be lost in the vapor stream of the freeze-dryer or on removal from the flask if their density is too low. For this reason, recovery based on the amount transferred to a final storage vessel can be low. The HMW fraction did not show any loss during this step.

The HMW fraction is substantial in Hillsborough and Quabbin waters, but rather minor in Chickahominy, Assomption, and Shoal waters. The breakdown between hydrophobic and



©2007 AwwaRF. All Rights Reserved.

**Table 5.11**
**Chickahominy fractionation: Carbon mass balance and elemental analysis**

| | | TOC | | Volume (mL) | | Total C | Elemental analysis | | |
|---|---|---|---|---|---|---|---|---|---|
| | Sample* | mg/L | mg | Used | Collected | (mg) | % C | % H | % N |
| 1 | RW | 5.25 | | | | | | | |
| 2 | RWF | 4.81 | | 30,000 | | 144 | | | |
| 3 | RVC | 1,002.6 | | | 140 | 140 | | | |
| 4 | UFR | 93.6 | | | 600 | 56 | | | |
| 5 | UFR-FD | | 10 | 600 | | 10 | 13.13 | 2.48 | 0.09 |
| 6 | UFP | 18.85 | | | 5,900 | 111 | | | |
| 7 | XADEL | | | | | | | | |
| 8 | XADELC | | | | | | | | |
| 9 | XADEL-FD | | 28 | | | 28 | 48.09 | 5.42 | ND† |
| 10 | XADEF | 9.57 | | | 5,900 | 56 | | | |
| 11 | XADEF-FD | | 29 | 5,750 | | 29 | 0.72 | 2.34 | ND |

*RVC = rotavap (rotary evaporation) concentrate; RW = raw water; RWF = raw water, filtered; UFP = ultrafiltration permeate; UFR = ultrafiltration retentate; UFR-FD = ultrafiltration retentate, freeze-dried; XADEF = XAD effluent; XADEF-FD = XAD effluent, freeze-dried; XADEL = XAD eluate; XADELC = XAD eluate concentrate; XADEL-FD = XAD eluate concentrate, freeze-dried.
†ND = not determined.

**Table 5.12**
**Assomption fractionation: Carbon mass balance and elemental analysis**

| | | TOC | | Volume (mL) | | Total C | Elemental analysis | | |
|---|---|---|---|---|---|---|---|---|---|
| | Sample* | mg/L | mg | Used | Collected | (mg) | % C | % H | % N |
| 1 | RW | 12.11 | | | | | | | |
| 2 | RWF | 7.21 | | 30,000 | | 216 | | | |
| 3 | RVC | 1,297 | | | 110 | 143 | | | |
| 4 | UFR | 36.6 | | | 600 | 22 | | | |
| 5 | UFR-FD | | 14 | 600 | | 14 | 17.61 | 4.16 | ND† |
| 6 | UFP | 22.28 | | | 5,000 | 111 | | | |
| 7 | XADEL | | | | | | | | |
| 8 | XADELC | | | | | | | | |
| 9 | XADEL-FD | | 40 | | | 40 | 48.56 | 4.99 | ND |
| 10 | XADEF | 7.88 | | | 5,000 | 39 | | | |
| 11 | XADEF-FD | | 45 | 5,000 | | 45 | 0.73 | 3.29 | 0.03 |

*RVC = rotavap (rotary evaporation) concentrate; RW = raw water; RWF = raw water, filtered; UFP = ultrafiltration permeate; UFR = ultrafiltration retentate; UFR-FD = ultrafiltration retentate, freeze-dried; XADEF = XAD effluent; XADEF-FD = XAD effluent, freeze-dried; XADEL = XAD eluate; XADELC = XAD eluate concentrate; XADEL-FD = XAD eluate concentrate, freeze-dried.
†ND = not determined.



©2007 AwwaRF. All Rights Reserved.

**Table 5.13**
**Shoal fractionation: Carbon mass balance and elemental analysis**

| | Sample* | TOC mg/L | TOC mg | Volume (mL) Used | Volume (mL) Collected | Total C (mg) | Elemental analysis % C | % H | % N |
|---|---|---|---|---|---|---|---|---|---|
| 1 | RW | 7.56 | | | | | | | |
| 2 | RWF | 7.21 | | 30,000 | | 216 | | | |
| 3 | RVC | 974 | | | 200 | 195 | | | |
| 4 | UFR | 24.16 | | | 600 | 14 | | | |
| 5 | UFR-FD | | 22 | 600 | | 22 | 36.31 | 5.97 | 3.03 |
| 6 | UFP | 34.14 | | | 5,000 | 171 | | | |
| 7 | XADEL | | | | | | | | |
| 8 | XADELC | | | | | | | | |
| 9 | XADEL-FD | | 87 | | | 87 | 53.55 | 6.02 | 1.52 |
| 10 | XADEF | 17.67 | | | 5,000 | 88 | | | |
| 11 | XADEF-FD | | 88 | 5,000 | | 88 | 3.08 | 4.08 | 0.22 |

*RVC = rotavap (rotary evaporation) concentrate; RW = raw water; RWF = raw water, filtered; UFP = ultrafiltration permeate; UFR = ultrafiltration retentate; UFR-FD = ultrafiltration retentate, freeze-dried; XADEF = XAD effluent; XADEF-FD = XAD effluent, freeze-dried; XADEL = XAD eluate; XADELC = XAD eluate concentrate; XADEL-FD = XAD eluate concentrate, freeze-dried.

**Table 5.14**
**Hillsborough fractionation: Carbon mass balance and elemental analysis**

| | Sample* | TOC mg/L | TOC mg | Volume (mL) Used | Volume (mL) Collected | Total C (mg) | Elemental analysis % C | % H | % N |
|---|---|---|---|---|---|---|---|---|---|
| 1 | RW | 11.87 | | | | | | | |
| 2 | RWF | 11.77 | | 30,000 | | 353 | | | |
| 3 | RVC | 1,946.6 | | | 167 | 324 | | | |
| 4 | UFR | 88.52 | | | 625 | 55 | | | |
| 5 | UFR-FD | | 78 | 600 | | 78 | 21.43 | 2.07 | 0.64 |
| 6 | UFP | 16.25 | | | 13,000 | 211 | | | |
| 7 | XADEL | | | | 1,800 | | | | |
| 8 | XADELC | 24,240 | | | 227 | | | | |
| 9 | XADEL-FD | | 117 | 217 | | 117 | 53.02 | 5.33 | 1.58 |
| 10 | XADEF | 6.58 | | | 12,900 | 85 | | | |
| 11 | XADEF-FD | | 66 | 12,550 | | 66 | 1.23 | 3.09 | 0.06 |

*RVC = rotavap (rotary evaporation) concentrate; RW = raw water; RWF = raw water, filtered; UFP = ultrafiltration permeate; UFR = ultrafiltration retentate; UFR-FD = ultrafiltration retentate, freeze-dried; XADEF = XAD effluent; XADEF-FD = XAD effluent, freeze-dried; XADEL = XAD eluate; XADELC = XAD eluate concentrate; XADEL-FD = XAD eluate concentrate, freeze-dried.



©2007 AwwaRF. All Rights Reserved.

**Table 5.15**
**Quabbin fractionation: Carbon mass balance and elemental analysis**

| | | TOC | | Volume (mL) | | Total C | Elemental analysis | | |
|---|---|---|---|---|---|---|---|---|---|
| | Sample* | mg/L | mg | Used | Collected | (mg) | % C | % H | % N |
| 1 | RW | 2.07 | | | | | | | |
| 2 | RWF | 2.10 | | 39,500 | | 83 | | | |
| 3 | RVC | 617.08 | | | 103 | 64 | | | |
| 4 | UFR | 28.24 | | | 645 | 18 | | | |
| 5 | UFR-FD | | 21 | 600 | | 21 | 29.60 | 5.43 | 2.20 |
| 6 | UFP | 9.26 | | | 4,165 | 39 | | | |
| 7 | XADEL | | | | 800 | | | | |
| 8 | XADELC | 12,925 | | | 150 | | | | |
| 9 | XADEL-FD | | 10 | 125 | | 10 | 51.99 | 5.96 | 1.83 |
| 10 | XADEF | 4.22 | | | 4,165 | 18 | | | |
| 11 | XADEF-FD | | 14 | 3,850 | | 14 | 1.65 | 2.20 | 0.17 |

*RVC = rotavap (rotary evaporation) concentrate; RW = raw water; RWF = raw water, filtered; UFP = ultrafiltration permeate; UFR = ultrafiltration retentate; UFR-FD = ultrafiltration retentate, freeze-dried; XADEF = XAD effluent; XADEF-FD = XAD effluent, freeze-dried; XADEL = XAD eluate; XADELC = XAD eluate concentrate; XADEL-FD = XAD eluate concentrate, freeze-dried.

**Table 5.16**
**Overall distribution of organic carbon (%)**

| | HMW | | HPO | | HPI | | Missing | |
|---|---|---|---|---|---|---|---|---|
| Water sample | Extract | Lost FD | Extract | Lost FD and XAD | Extract | Lost FD | Evaporation | UF |
| Chickahominy River | 7 | 0 | 19 | 19 | 20 | 19 | 3 | 14 |
| Assomption River | 7 | 0 | 19 | 15 | 21 | 0 | 34 | 8 |
| Shoal Lake | 10 | 0 | 40 | 0 | 40 | 0 | 10 | 1 |
| Hillsborough River | 22 | 0 | 33 | 3 | 19 | 5 | 8 | 10 |
| Quabbin Reservoir | 26 | 0 | 12 | 13 | 17 | 4 | 23 | 4 |

hydrophilic shows higher abundance of the hydrophobic fraction for low-SUVA (Quabbin) and high-SUVA (Assomption, Hillsborough) waters. Similar abundance was found in the case of moderate-SUVA waters (Chickahominy and Shoal). Recall that the missing NOM is probably a combination of HPI, HPO, and HMW fractions, but the exact breakdown cannot be determined.

Comparing the elemental analysis results, the major findings for the three NOM isolates were as follows:

- The order of abundance was generally HPO > HMW > HPI.
- The very low carbon content of the HPI fraction was due to its high salt content.



©2007 AwwaRF. All Rights Reserved.

**Table 5.17**
**Chlorine dose for low-dose conditions (mg/L)***

| Sample type | Chickahominy River | Assomption River | Shoal Lake | Hillsborough River | Quabbin Reservoir |
|---|---|---|---|---|---|
| Whole water | ND† | ND | 5.46 (0.6) | 14 (0.48) | 2.3 (1.15) |
| HMW fraction | ND | ND | ND | 4.31 (0.67) | ND |
| HPO fraction | 4 (1.4) | 4.97 (0.5) | 3.43 (1.3) | 3.76 (0.56) | ND |
| HPI fraction | 7.4 (2.5) | 5.88 (0.5) | 4.06 (1.0) | 4.85 (0.68) | ND |

*Values in parentheses are the residual chlorine values following chlorination at low dose.
†ND = not determined.

- The HPI fractions were characterized by very low C/H atomic ratios, indicating that this category of NOM is richer in aliphatic carbon than the two other fractions. It must also be acknowledged that hydrated salts were probably present in the HPI fractions (e.g., $MgCl_2 \cdot H_2O$), which would have depressed the C/H ratio as well. The HMW and HPO fractions were characterized by an intermediate and high C/H ratio, respectively. However, the C/H value of the HMW fraction isolated from Hillsborough River water was slightly higher than that of the HPO fraction, indicating similar aromatic contents.
- The HPI and HMW fractions of Quabbin Reservoir and Shoal Lake waters were characterized by a low C/N ratio compared to the HPO fraction, indicating the high amount of nitrogen compounds in these fractions. Evidence mentioned earlier suggests that these waters are heavily influenced by autochthonous organic matter, and it is suspected that this material ends up in the HMW and HPI fractions. In contrast, the HMW fraction isolated from Hillsborough River water did not contain as much nitrogenous material as the HPI fraction, perhaps reflecting the greater importance of lignin-type materials over proteins in that source.

## Bench-scale Treatment

All of the major NOM fractions were tested for impacts of ozonation. First, the pH of each sample was adjusted to 7. Samples were then ozonated in a semicontinuous system at a dose of 1 mg $O_3$/mg C. Ozone was generated from pure oxygen by means of a laboratory corona discharge generator (ozone generator; Welsbach—company no longer in existence). The ozone/oxygen product gas was introduced into the reactor containing the water to be treated. Flow was controlled with an electronic flow controller, and the ozone content was monitored by direct UV absorbance spectrophotometry. The gas was mixed with the sample by a porous quartz frit. Off-gas was redirected through a spectrophotometer (Lambda 3A UV-visible spectrophotometer; Perkin-Elmer, Wellesley, Mass.) for determination of ozone content.

These same NOM fractions were tested for impacts of low-dose chlorination. Samples were buffered at pH 7 (1 m*M* $Na_2HPO_4$ [sodium phosphate diabasic]) and then treated with a chlorine dose intended to be sufficient to maintain a residual of about 1 mg/L at the end of the 3-day reaction period (Table 5.17). Reactions were conducted in headspace-free glass vessels (chlorine demand free) in the dark at a fixed temperature of 20°C.



©2007 AwwaRF. All Rights Reserved.

## AOC and BDOC Tests

All water sampling sites were tested for AOC and BDOC values of their raw waters and the corresponding NOM fractions before and after treatment. Table 5.18 summarizes the raw data. Figures 5.7–5.11 show AOC and BDOC data in graphical form for each water sample and the NOM fractions extracted from the same water.

### *AOC Discussion*

For all samples tested, ozonation caused a substantial increase in the AOC value. In all but one of these samples (Assomption River HPO), ozonation caused a higher increase in AOC than chlorination. In almost every case, chlorination resulted in a substantial increase in AOC over the raw water, but that increase was less than the increase seen with ozonation.

The untreated, unfractionated waters (whole raw waters) exhibited higher AOC values than the untreated NOM fractions, except for the Quabbin Reservoir water sample in which the HPI fraction exhibited the highest AOC level. Among the NOM fractions, ozonated HPI isolates had the highest AOC levels as compared to all other HPI fractions, and all HMW and HPO fractions. However, there were not as many waters from which the HMW fraction was extracted with which to compare AOC levels. Similarly, chlorination produced the highest AOC values for most of the HPI fractions. The levels of AOC contributed by each fraction when untreated varied among the waters tested. AOC levels were below 150 µg acetate carbon/L for most samples. All samples were below 350 µg acetate carbon/L.

Ozonation is known to cause an increase in AOC compounds due to the breakdown of larger organic molecules. The effects of ozonation on BDOC are less clear (Escobar and Randall 2001). As previously noted in the literature, the strain *Spirillum* sp. NOX contributed the most to the AOC levels after ozonation and, to a lesser extent, after chlorination (Polanska, Huysman, and van Keer 2005). Chlorination is known to produce low-MW compounds and also to enhance the utilization of high-MW compounds by bacteria (van der Kooij 1990). The greatest contribution to AOC levels was from the HPI fraction. This can be expected as hydrophilic carbon compounds tend to have lower molecular weights than hydrophobic compounds and thus can be more easily assimilated by the bacteria inoculated in the AOC assay, particularly when further breakdown is contributed by the ozonation process (Hood and Williams 2001).

### *BDOC Discussion*

The variance in BDOC levels among the waters tested was more pronounced. For all raw waters, chlorination or ozonation increased BDOC levels. By comparison, ozonation increased the BDOC levels to a greater extent than chlorination in three of four water samples. Generally, ozonation or chlorination also increased BDOC levels in all fractions. For two water sources—Assomption River and Shoal Lake—ozonation increased BDOC levels to a greater extent than chlorination, whereas the opposite outcome resulted for the Chickahominy River water. As seen for the AOC values, the HPI fraction produced the greater increase in BDOC values compared to the HPO fraction. Again, there are not enough HMW fractions to permit a firm conclusion. Most



©2007 AwwaRF. All Rights Reserved.

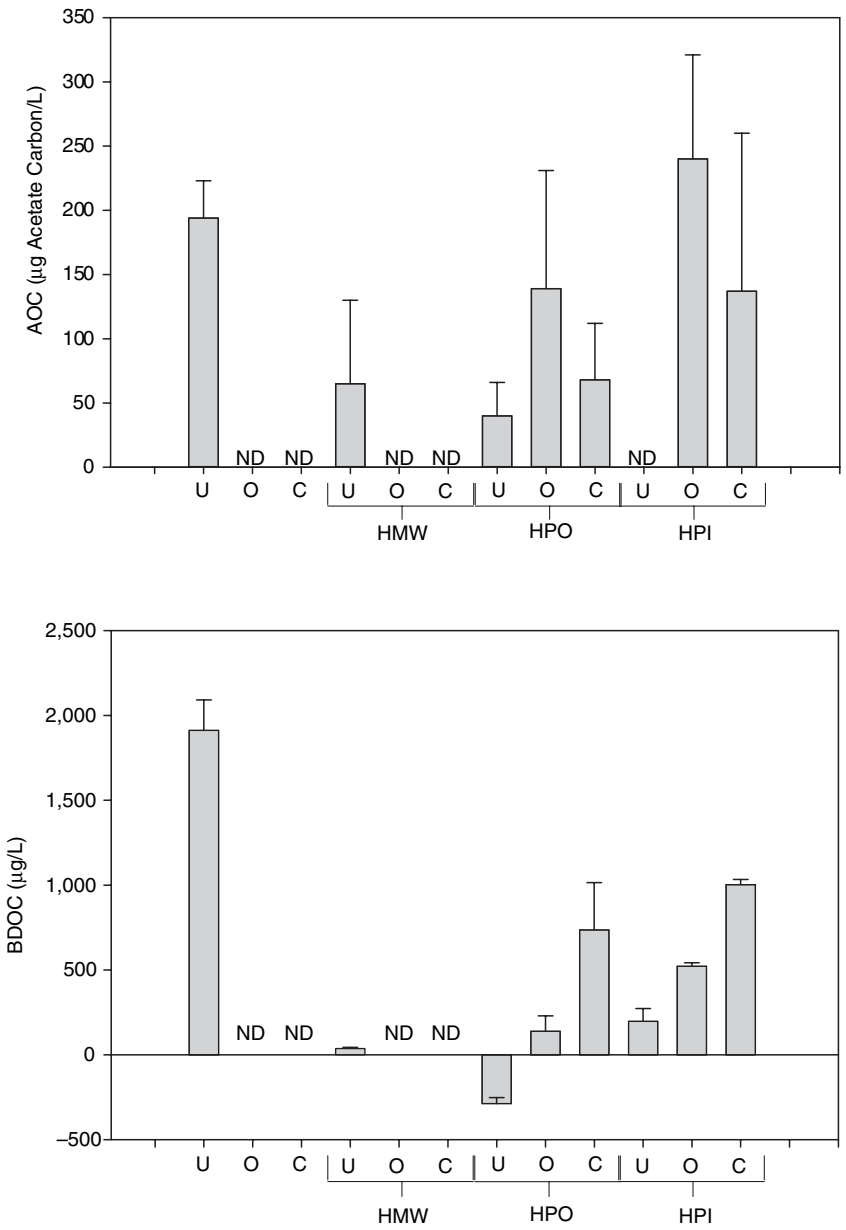
## Table 5.18
## Summary of test results for AOC and BDOC assays of raw waters and corresponding NOM fractions

| Sample analysis | Chickahominy River | Assomption River | Shoal Lake | Cobble Mountain Reservoir* | Hillsborough River | Quabbin Reservoir |
|---|---|---|---|---|---|---|
| Raw | | | | | | |
| AOC (µg C/L) | 194 ± 29 | 90 ±30 | 79 ± 39 | 55 ± 74 | 87 ± 23 | 22 ± 13 |
| BDOC (µg/L) | 1,912 ± 180 | 1,701 ± 127 | 909 ± 518 | 1,657 ± 63 | 373 ± 101 | |
| DOC (µg/L) | 3,551 | 7,998 | 26,322 | 4,610 | 10,781 | 4,926 |
| Raw + Ozonation | | | | | | |
| AOC (µg C/L) | ND† | ND | 218 ± 86 | 123 ± 41 | 203 ± 27 | 67 ± 24 |
| BDOC (µg/L) | | | 1,364 ± 602 | 1,769 ± 135 | 686 ± 50 | |
| DOC (µg/L) | | | 24,762 | 4,255 | 10,142 | 5,252 |
| Raw + Chlorination | | | | | | |
| AOC (µg C/L) | ND | ND | 86 ± 32 | 113 ± 18 | 163 ± 25 | 30 ± 10 |
| BDOC (µg/L) | | | 4,016 ± 1,899 | 428 ± 27 | 1,092 ± 632 | |
| DOC (µg/L) | | | 17,310 | 3,243 | 10,499 | 4,081 |
| HMW | | | | | | |
| AOC (µg C/L) | 65 ± 12 | 58 ± 19 | 32 ± 14 | ND | 17 ± 22 | 38 ± 16 |
| BDOC (µg/L) | 37 ± 7 | 260 ± 47 | 1,173 ± 359 | | 139 ± 57 | |
| DOC (µg/L) | 1,204 | 1,441 | 2,905 | | 2,804 | |
| HMW + Ozonation | | | | | | |
| AOC (µg C/L) | ND | 86 ± 10 | ND | ND | 94 ± 35 | ND |
| BDOC (µg/L) | | 429 ± 29 | | | 111 ± 43 | |
| DOC (µg/L) | | 1,569 | | | 2,393 | |
| HMW + Chlorination | | | | | | |
| AOC (µg C/L) | ND | ND | ND | ND | 75 ± 34 | ND |
| BDOC (µg/L) | | | | | 388 ± 30 | |
| DOC (µg/L) | | | | | 2,720 | |
| HPO | | | | | | |
| AOC (µg C/L) | 40 ± 26 | 20 ± 13 | 42 ± 11 | ND | 8.5 ± 9.4 | 19 ± 15 |
| BDOC (µg/L) | –288 ± 36 | 1,727 ± 24 | –67 ± 60 | | 263 ± 133 | |
| DOC (µg/L) | 4,841 | 5,397 | 4,043 | | 3,844 | |
| HPO + Ozonation | | | | | | |
| AOC (µg C/L) | 139 ± 92 | 100 ± 32 | 134 ± 37 | ND | 103 ± 28 | ND |
| BDOC (µg/L) | 139 ± 91 | 2,042 ± 152 | 1,530 ± 76 | | 139 ± 77 | |
| DOC (µg/L) | 5,020 | 5,843 | 4,834 | | 3,462 | |
| HPO + Chlorination | | | | | | |
| AOC (µg C/L) | 68 ± 44 | 140 ± 46 | 94 ± 17 | ND | 70 ± 23 | ND |
| BDOC (µg/L) | 736 ± 279 | 807 ± 22 | 265 ± 35 | | 143 ± 157 | |
| DOC (µg/L) | 3,665 | 5,097 | 4,365 | | 3,661 | |
| HPI | | | | | | |
| AOC (µg C/L) | NG‡ | NG | 72 ± 17 | ND | 20 ± 8 | 75 ± 38 |
| BDOC (µg/L) | 197 ± 76 | 1626 ± 106 | 205 ± 10 | | 249 ± 77 | |
| DOC (µg/L) | 4,675 | 5,809 | 4,373 | | 4,233 | |
| HPI + Ozonation | | | | | | |
| AOC (µg C/L) | 240 ± 81 | 149 ± 58 | 244 ± 67 | ND | 148 ± 28 | ND |
| BDOC (µg/L) | 521 ± 22 | 2,166 ± 34 | 1,572 ± 61 | | 617 ± 241 | |
| DOC (µg/L) | 4,461 | 5,503 | 4,713 | | 3,822 | |
| HPI + Chlorination | | | | | | |
| AOC (µg C/L) | 137 ± 123 | 122 ± 36 | 102 ± 22 | ND | 83 ± 15 | ND |
| BDOC (µg/L) | 1,002 ± 31 | 1,079 ± 444 | 467 ± 89 | | 202 ± 3 | |
| DOC (µg/L) | 3,613 | 5,523 | 4,372 | | 3,954 | |

*Located in Springfield, Mass.
†ND = not determined.
‡NG = no growth.



©2007 AwwaRF. All Rights Reserved.



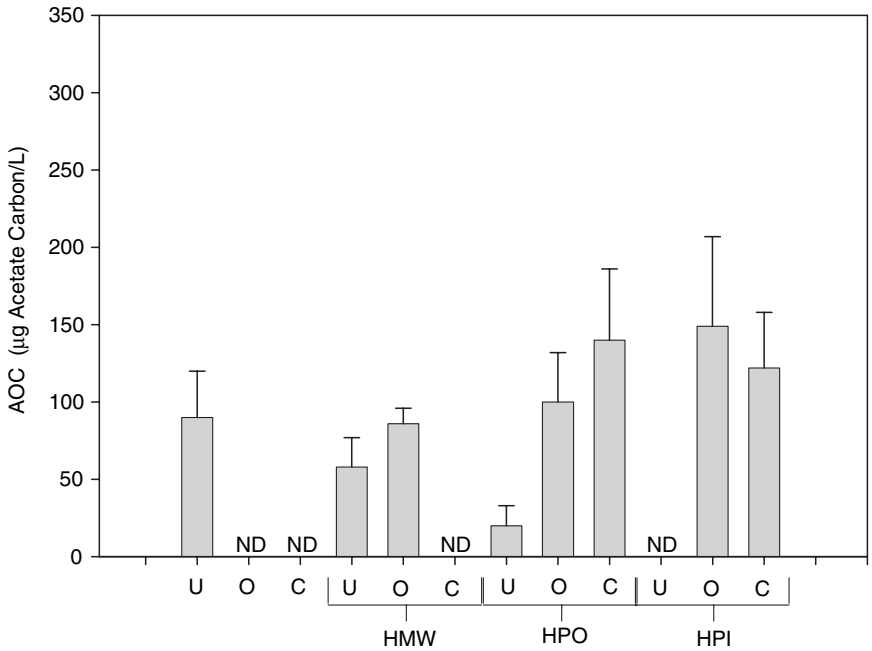

NOTES: NOM fractions before (U = untreated) and after treatment (O = ozonation, C = chlorination). ND = not determined.

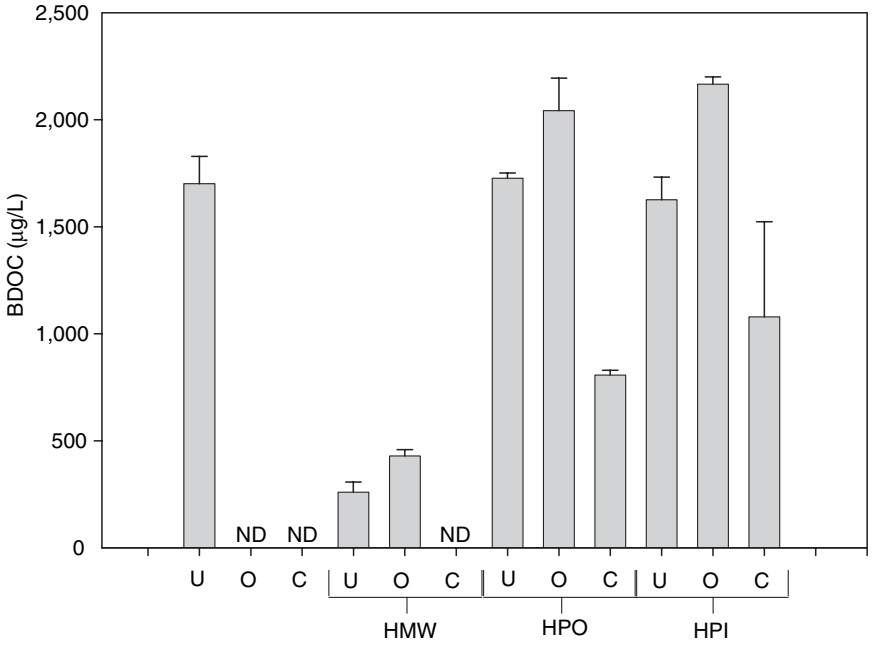
**Figure 5.7 AOC and BDOC of Chickahominy River water (Newport News, Va.) and corresponding NOM fractions**

BDOC levels were below 2 mg/L, with only one raw water sample and two chlorinated water samples reaching approximately 4 mg/L and one raw water sample reaching greater than 8 mg/L BDOC. Apparent negative BDOC results were found in two water samples for the untreated HPO fraction. It is speculated that these samples may not have been adjusted to a neutral pH value at which the indigenous bacteria inoculum could adequately metabolize organic carbon. This negative effect would be amplified with the long incubation time (28 days) required for the BDOC test.



©2007 AwwaRF. All Rights Reserved.



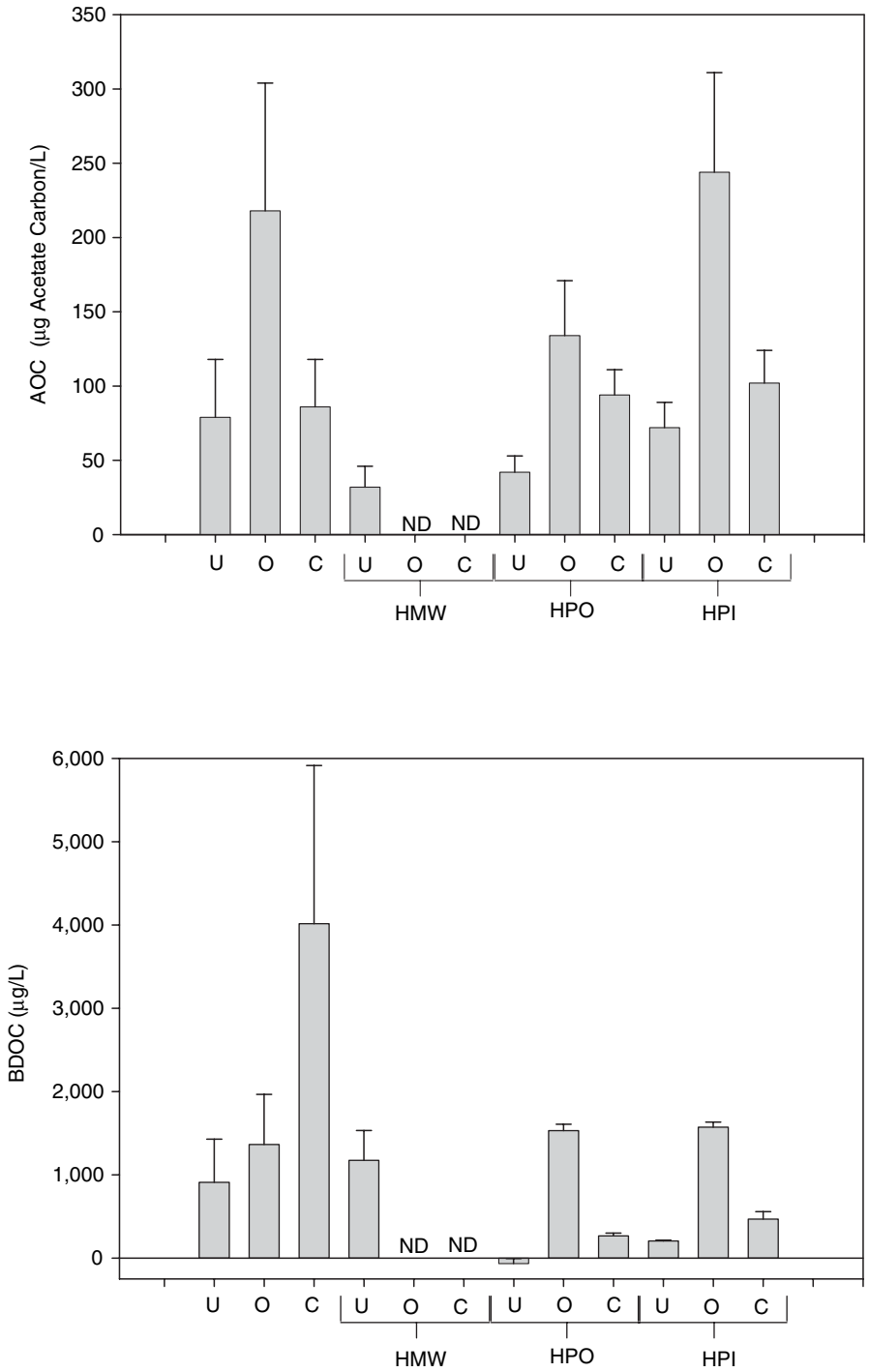




NOTES: NOM fractions before (U = untreated) and after treatment (O = ozonation, C = chlorination). ND = not determined.

**Figure 5.8 AOC and BDOC of Assomption River (Repentigny, Que.) water and corresponding NOM fractions**



©2007 AwwaRF. All Rights Reserved.

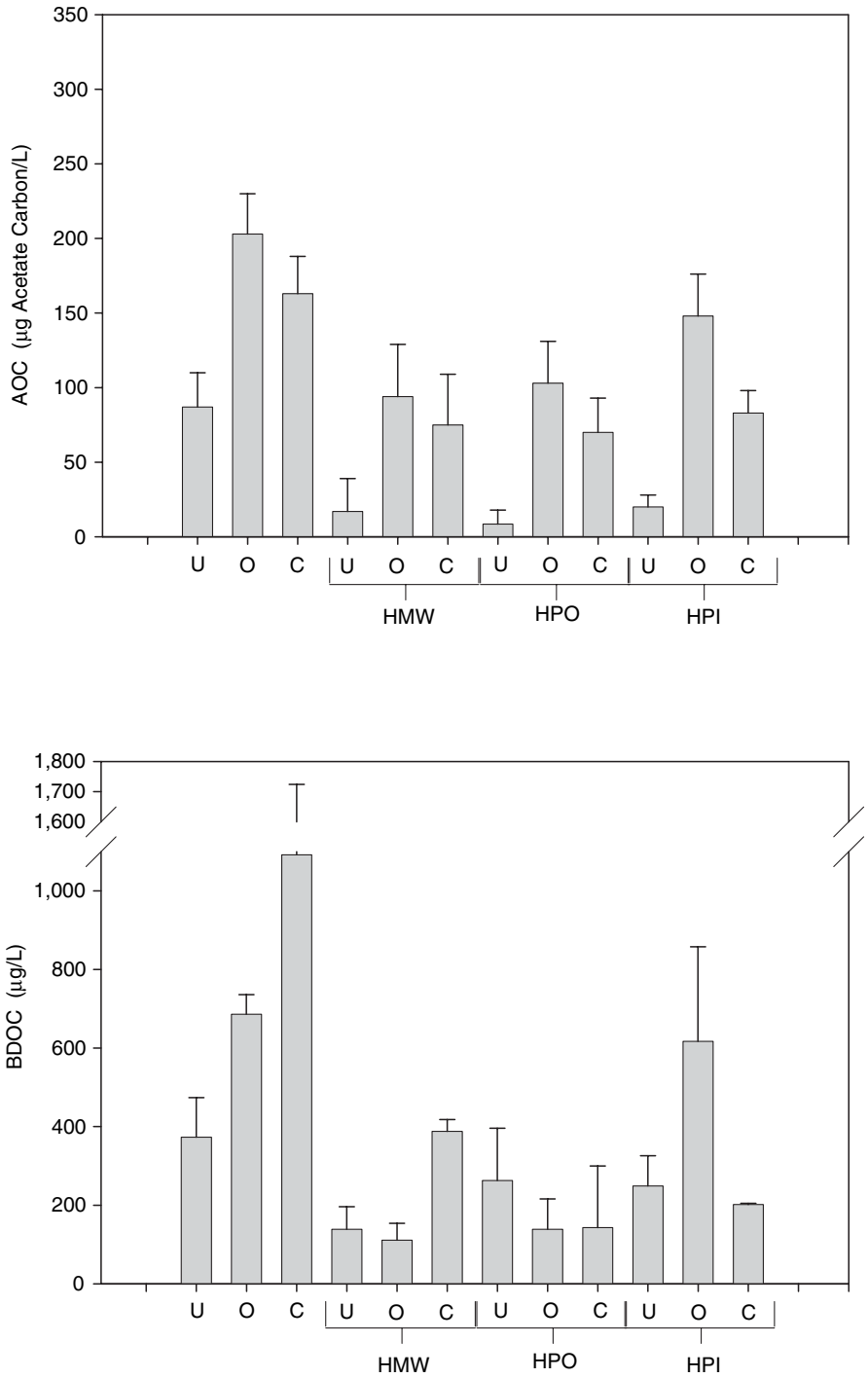



NOTES: NOM fractions before (U = untreated) and after treatment (O = ozonation, C = chlorination). ND = not determined.

**Figure 5.9 AOC and BDOC of Shoal Lake (Winnipeg, Man.) water and corresponding NOM fractions**



©2007 AwwaRF. All Rights Reserved.

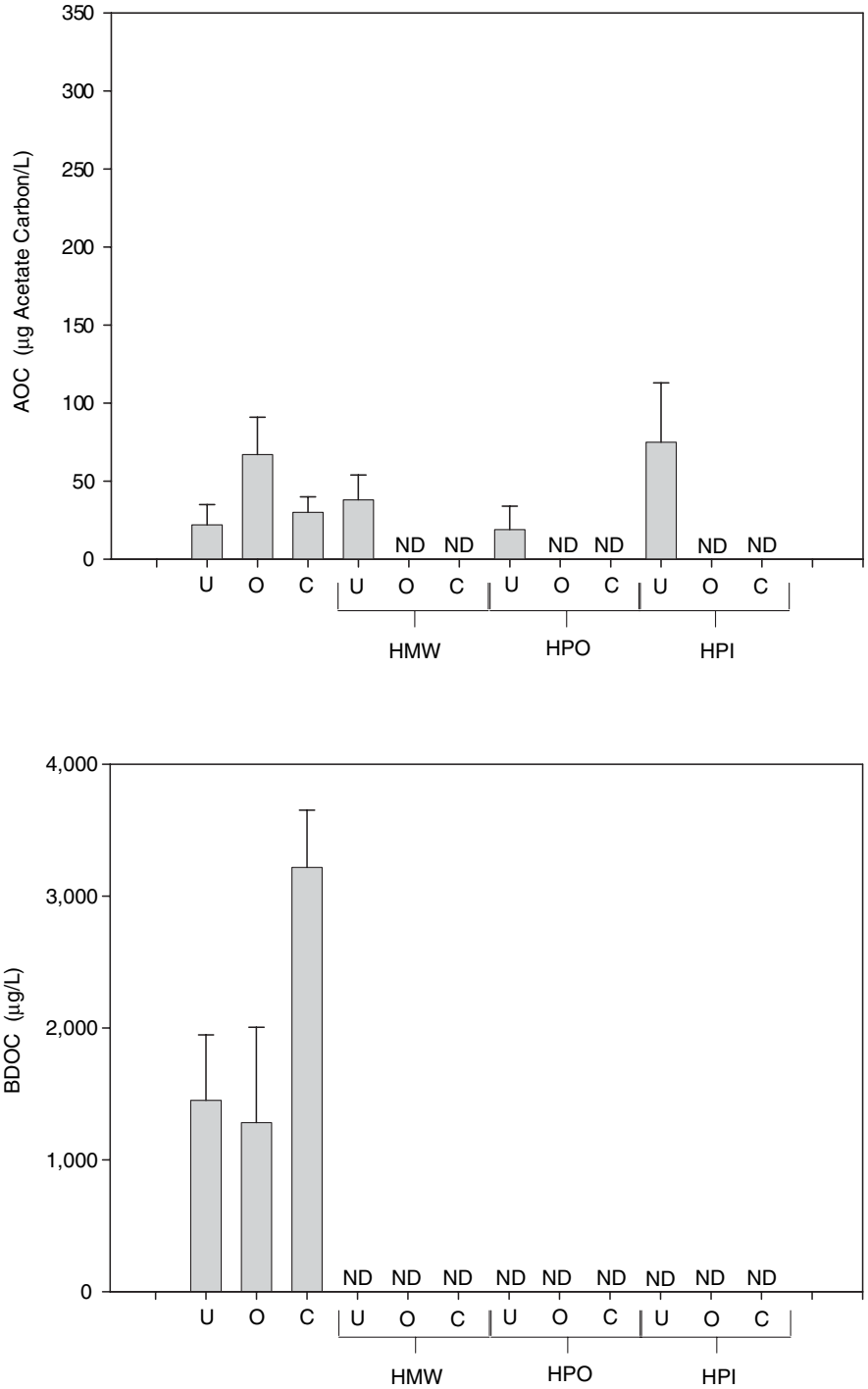



NOTES: NOM fractions before (U = untreated) and after treatment (O = ozonation, C = chlorination). ND = not determined.

**Figure 5.10 AOC and BDOC of Hillsborough River (Tampa, Fla.) water and corresponding NOM fractions**



©2007 AwwaRF. All Rights Reserved.




NOTES: NOM fractions before (U = untreated) and after treatment (O = ozonation, C = chlorination). ND = not determined.

**Figure 5.11 AOC and BDOC of Quabbin Reservoir (Boston, Mass.) water and corresponding NOM fractions**



©2007 AwwaRF. All Rights Reserved.

**Table 5.19**
**The relative contribution of each fraction compared to the raw water sample**

| % Raw | Chickahominy River | Assomption River | Shoal Lake | Cobble Mountain Reservoir | Hillsborough River | Quabbin Reservoir |
|---|---|---|---|---|---|---|
| **AOC** | | | | | | |
| Raw | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| HMW | 33.51 | 64.44 | 40.51 | ND* | 21.52 | 48.10 |
| HPO | 20.62 | 22.22 | 53.16 | ND | 10.76 | 24.05 |
| HPI | NG† | NG | 91.14 | ND | 25.32 | 94.94 |
| **BDOC** | | | | | | |
| Raw | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
| HMW | 0.93 | 6.27 | 129.04 | ND | 15.29 | ND |
| HPO | –7.27 | 41.61 | –7.37 | ND | 28.93 | ND |
| HPI | 4.97 | 39.18 | 22.55 | ND | 27.39 | ND |

*ND = not determined.
†NG = no growth, possibly due to inhibition.

### *BOM Mass Balance*

The relative contribution of each fraction compared to the raw water sample is summarized in Table 5.19. For AOC, HPI fractions contribute the most for all water samples in which this isolate was tested. No measurable AOC was found for HPI fractions for two watersheds. This is thought to be due to some inhibitory substance in these fractions (possibly related to the high salt content). The HMW fractions were the second largest contributors of AOC for four out of five water samples, followed by the HPO fraction. The percentage contribution by each fraction to the BDOC is less consistent, with each fraction contributing different levels in each different water sample.

### Chlorination By-product Yields

Molar concentrations are used in this study to account for the differences in bromide concentrations among the various water samples. Figure 5.12 illustrates the formation of THMs, DHAAs, and THAAs in the different raw waters. Samples from Chickahominy and Assomption river waters were incubated for a period of 48 hours; however, an incubation period of 72 hours was used for the other water samples. DBP data for the former waters were subsequently adjusted to allow comparison with the latter waters (using the power function model described earlier). The molar production of DBPs was normalized by dividing the observed molar DBP concentration (micromoles per liter) by the corresponding DOC concentrations of the water. This is the standard method that was used for comparing DBP precursor data from samples of differing NOM quality.

The Hillsborough water sample produced the highest THM, DHAA, and THAA formation among the five water samples. Regardless of the water sample, THM formation exceeded THAA and DHAA formation. THAA formation exceeded DHAA in all water samples with one exception



©2007 AwwaRF. All Rights Reserved.

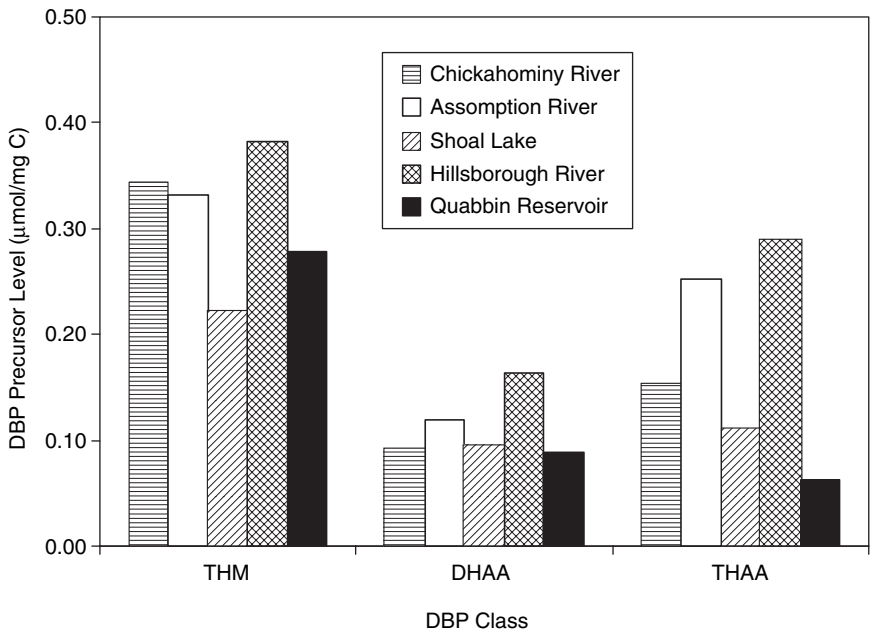



**Figure 5.12 Chlorination by-products precursor content for the raw waters**

for low-SUVA—Quabbin water. The diminished formation of THAA in Quabbin and Hillsborough waters is significant. These two waters are believed to be heavily autochthonous, and the major sources of THAA precursors are thought to be higher terrestrial plants (e.g., lignins).

All NOM fractions isolated in a dried form were chlorinated for a contact period of 72 hours and analyzed for the major DBPs. Although the concentration of DOC was different for each individual NOM solution, DBP data can be normalized on a per-carbon-yield basis (concentrations of DBPs produced per unit of DOC) to facilitate comparisons.

Figure 5.13 shows THM, DHAA, and THAA formation of the three Chickahominy-derived NOM fractions. Figures 5.14–5.17 show the analogous data for Assomption, Shoal, Hillsborough, and Quabbin waters. There is no general trend in precursor content among the three fractions. Results show that HMW fractions isolated from high-SUVA waters (Assomption and Hillsborough) produced the highest DBPs compared to the low-MW fractions (HPI and HPO). The opposite was true for low- and moderate-SUVA waters (Quabbin and Chickahominy). For Shoal Lake water, there is a decreasing trend in DBP precursor content going from HPI to HMW to HPO fractions. DHAA and THM precursor levels decrease when going from HPI to HPO fractions. However, the opposite was true for the fractions extracted from Assomption water. THAA formation from the HPO fraction exceeds its formation from the HPI fraction. The reverse was observed for the fractions extracted from Shoal Lake water.

### Chlorination By-product Precursor Recovery

Figures 5.18–5.22 illustrate the DBP precursor recovery against the organic carbon recovery. The bar for TOC represents the carbon-based distribution of the three fractions and the missing NOM. Results show that the fractionation method used in this study allows recovery of most of the



©2007 AwwaRF. All Rights Reserved.

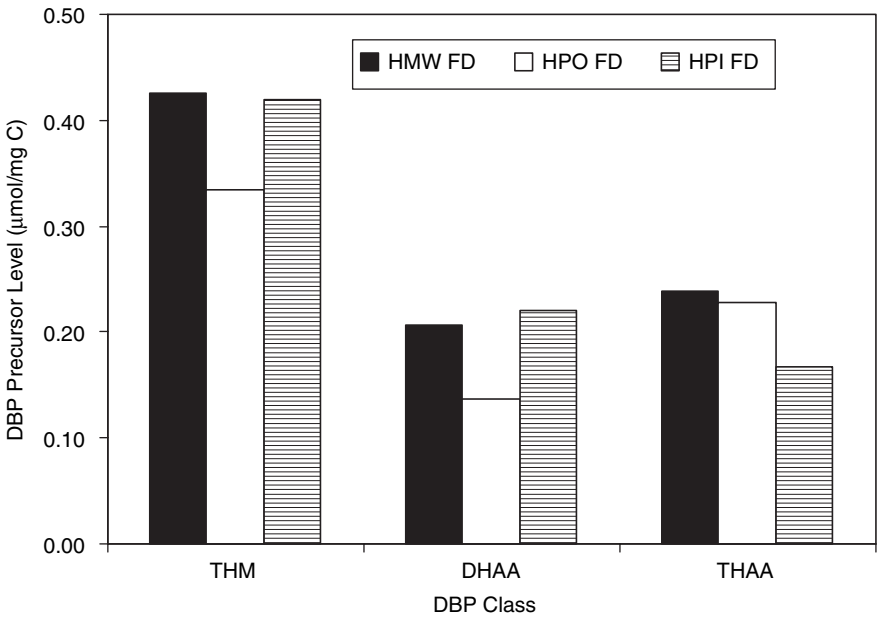



**Figure 5.13 DBP precursor content for NOM fractions extracted from Chickahominy River water**

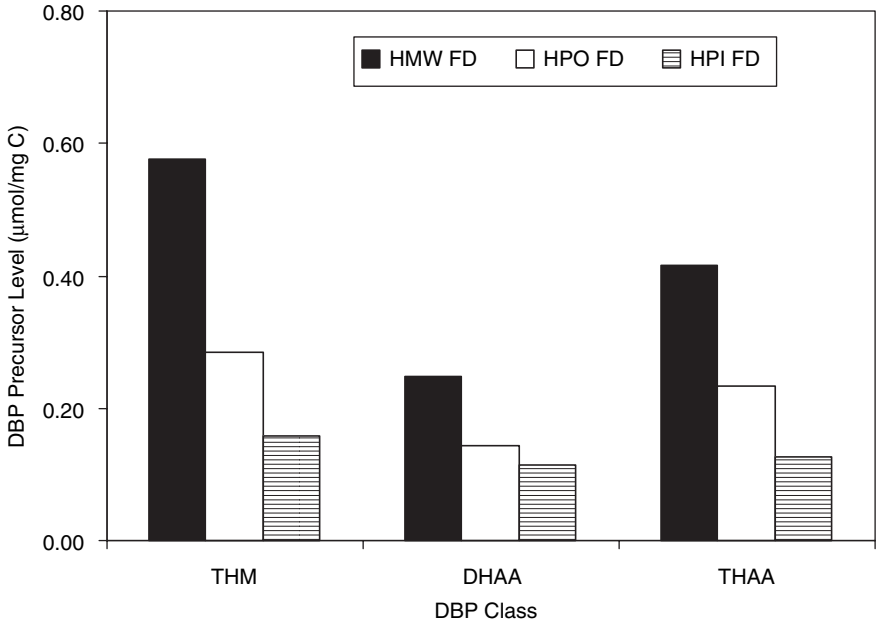



**Figure 5.14 DBP precursor content for NOM fractions extracted from Assomption River water**



©2007 AwwaRF. All Rights Reserved.

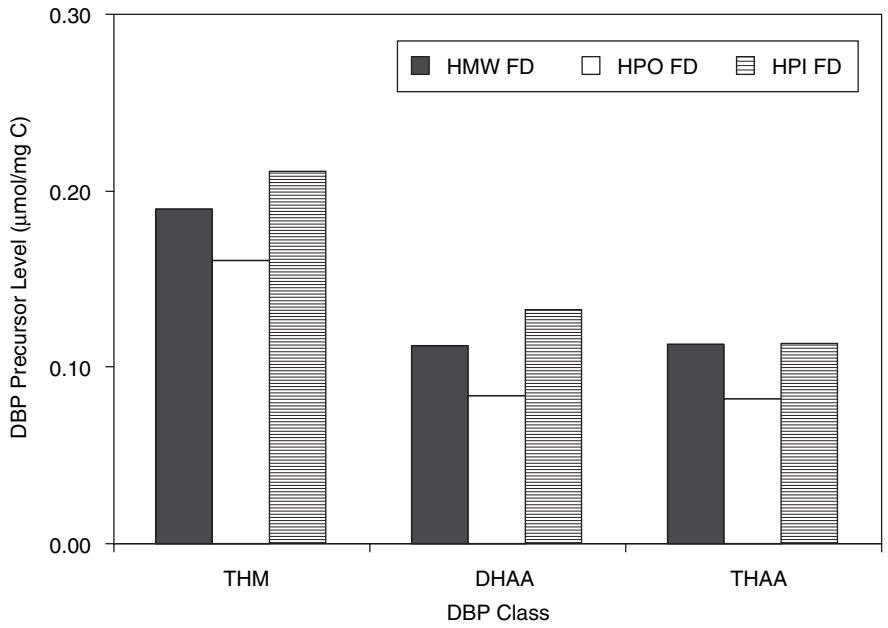



**Figure 5.15 DBP precursor content for NOM fractions extracted from Shoal Lake water**

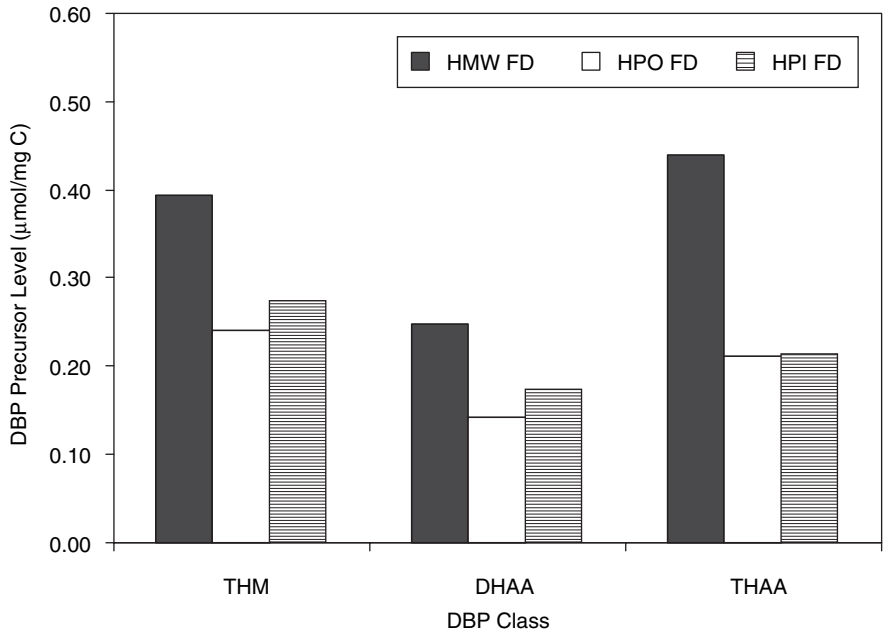



**Figure 5.16 DBP precursor content for NOM fractions extracted from Hillsborough River water**



©2007 AwwaRF. All Rights Reserved.

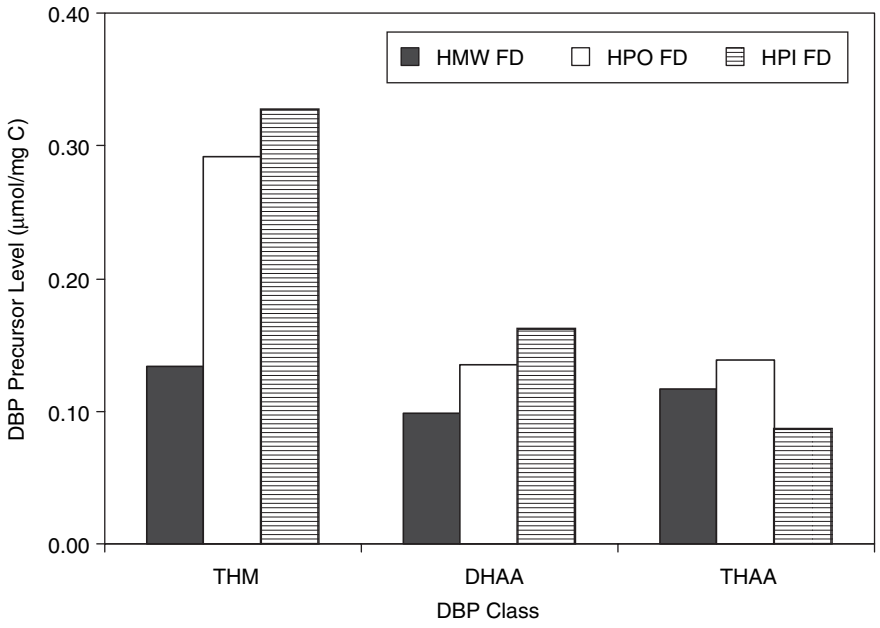



**Figure 5.17 DBP precursor content for NOM fractions extracted from Quabbin Reservoir water**

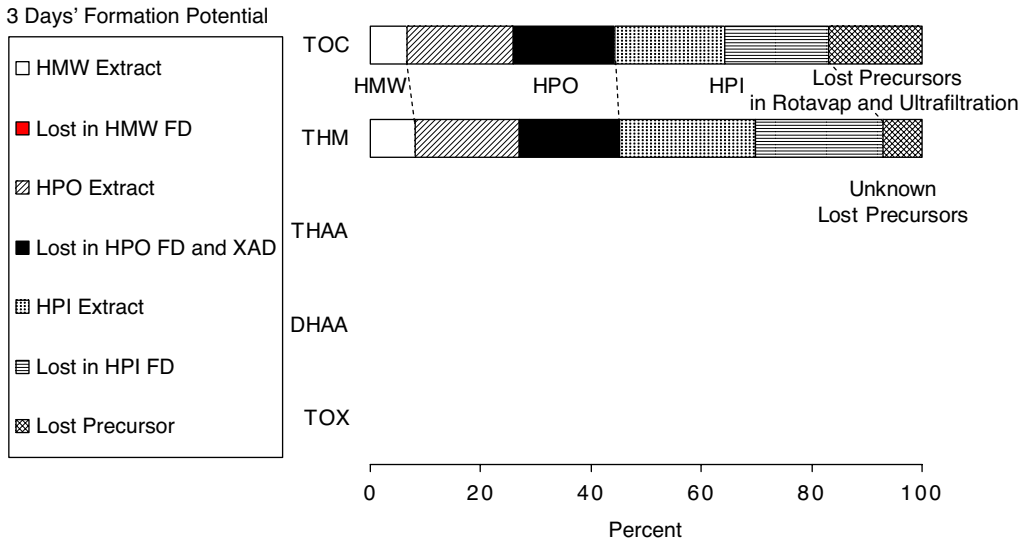



**Figure 5.18 Recovery of organic carbon versus recovery of chlorinated by-product precursor (Chickahominy River water)**



©2007 AwwaRF. All Rights Reserved.

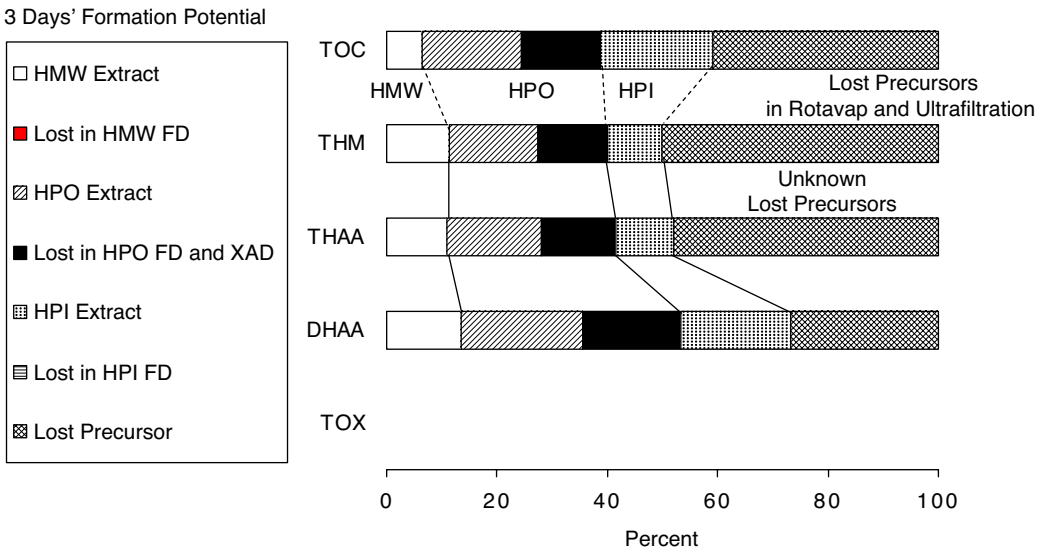



**Figure 5.19 Recovery of organic carbon versus recovery of chlorinated by-product precursor (Assomption River water)**

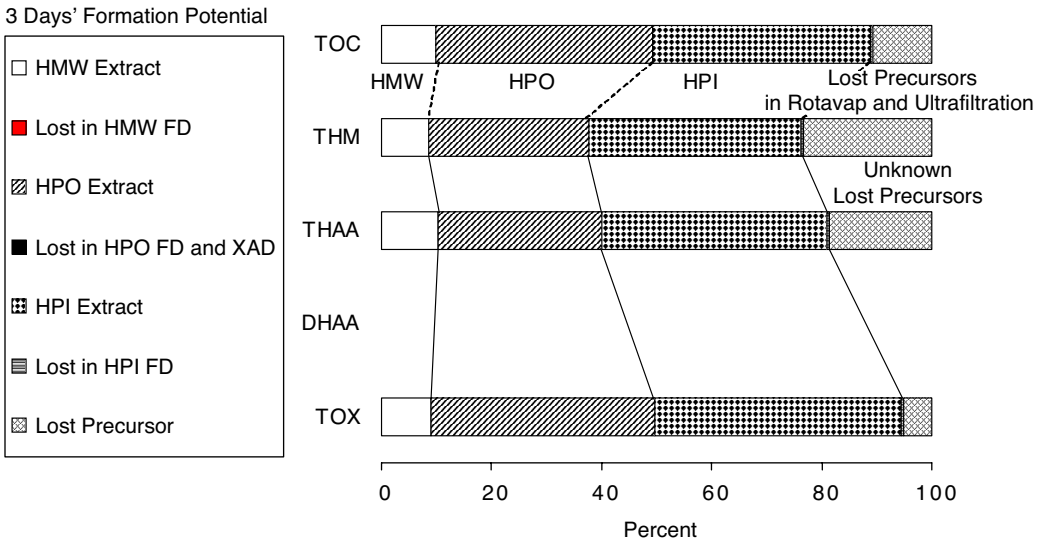



**Figure 5.20 Recovery of organic carbon versus recovery of chlorinated by-product precursor (Shoal Lake water)**

carbon and DBP precursor. The other bars are graphical representations of the cumulative mass of precursor in each fraction. These numbers are simply the product of the normalized DBP formation of each fraction and the mass of that fraction that was actually isolated. If all isolation steps were 100% efficient without altering the reactivity of the NOM, all bars would extend to 100% without any "lost precursor." In reality, the lost precursor ranges from 5% to 50%. Much of this is clearly attributable to losses in carbon during the extraction process. However, the possibility of chemical changes during extraction with accompanying changes in reactivity cannot be ruled out.



©2007 AwwaRF. All Rights Reserved.

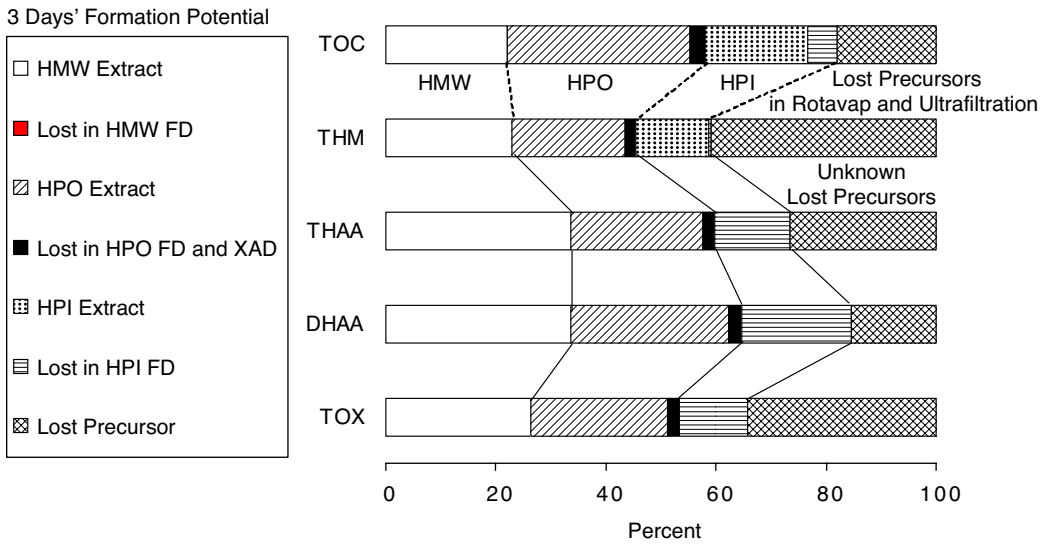



**Figure 5.21 Recovery of organic carbon versus recovery of chlorinated by-product precursor (Hillsborough River water)**

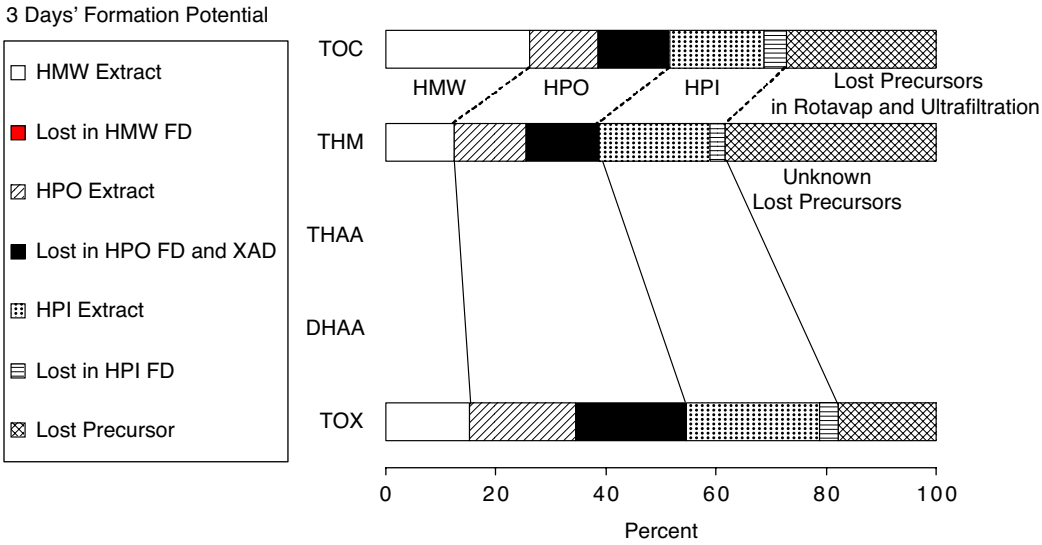



**Figure 5.22 Recovery of organic carbon versus recovery of chlorinated by-product precursor (Quabbin Reservoir water)**

### *Comparison With the Literature Database*

Figure 5.23 shows the THMFP cumulative frequency plot developed from the full literature database in chapter 4. Superimposed over this information are the precursor data for the five waters that are the subject of this chapter. The sample codes are the same as those originally presented in Table 5.9 (1 is Chickahominy River, 2 is Assomption River, 3 is Shoal Lake, 4 is Hillsborough River, and 5 is the Quabbin Reservoir). Figure 5.23 shows that the HPO fractions are low as compared to the overall database (5th–30th percentiles). In contrast, the HPI fractions



©2007 AwwaRF. All Rights Reserved.

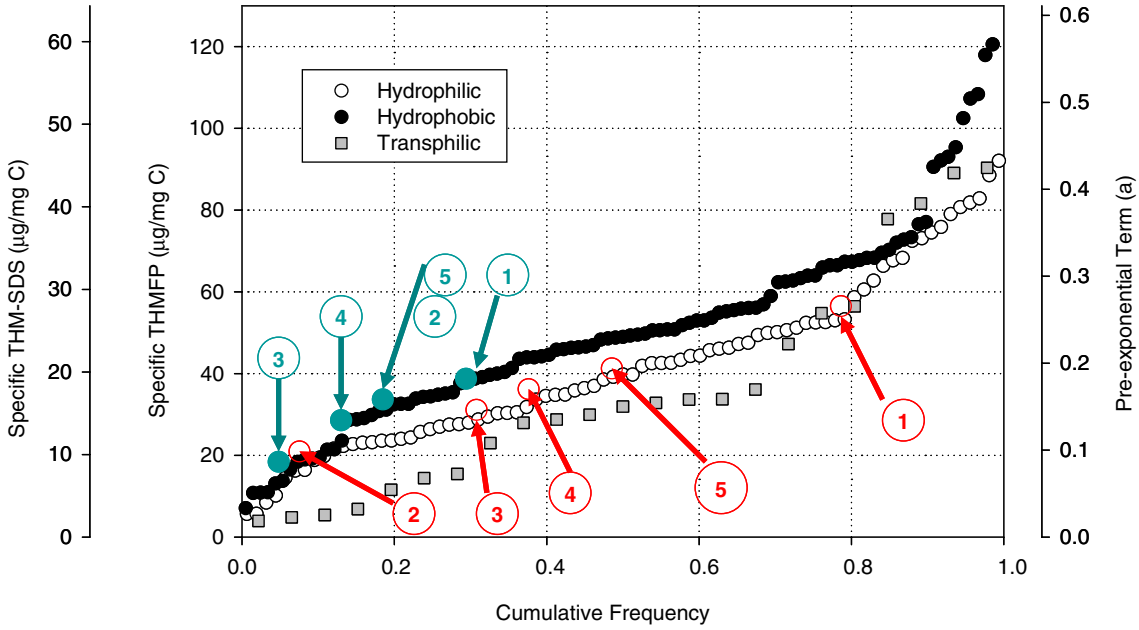



**Figure 5.23 THM precursor content of watershed NOM shown against the backdrop of the full literature database (untreated waters)**

are well distributed, including one very high (HPI fraction of Chickahominy water) and one very low (HPI fraction of Assomption water).

Figures 5.24 and 5.25 show the corresponding cumulative frequency plots for the DHAAs and THAAs, respectively. As with the THM precursor data, the five HPO fractions fall below the database median. In contrast, the HPI fractions are quite divergent, including one that is very high and one very low. The ratio of the DHAAs to the THMs (Figure 5.26) shows a slightly higher value for the hydrophilic fractions. The reverse is true for the THAA/THM ratio (Figure 5.27).

### *Influence of Freeze-Drying*

Figures 5.28 and 5.29 present the DBP precursor content for NOM fractions extracted from Quabbin and Hillsborough waters, respectively, without any further lyophilization. Data for freeze-dried samples are also presented to facilitate comparison. Results showed that freeze-drying enhanced the production of DBPs, suggesting a change in chemical structure of NOM. Decarboxylation (loss of $CO_2$) and dehydration (loss of $H_2O$) reactions probably occur during freeze-drying, and these are likely to increase reactivity with an electrophilic substance such a hypochlorous acid.

## SUMMARY AND CONCLUSIONS

The objective of this study was to investigate some key properties of NOM that pertain to finished water quality (DBP formation and BOM) and examine their relationship to watershed characteristics. Five water samples were selected for this study, all of which are raw water supplies for major metropolitan areas: (1) Chickahominy River, (2) Assomption River, (3) Shoal



©2007 AwwaRF. All Rights Reserved.

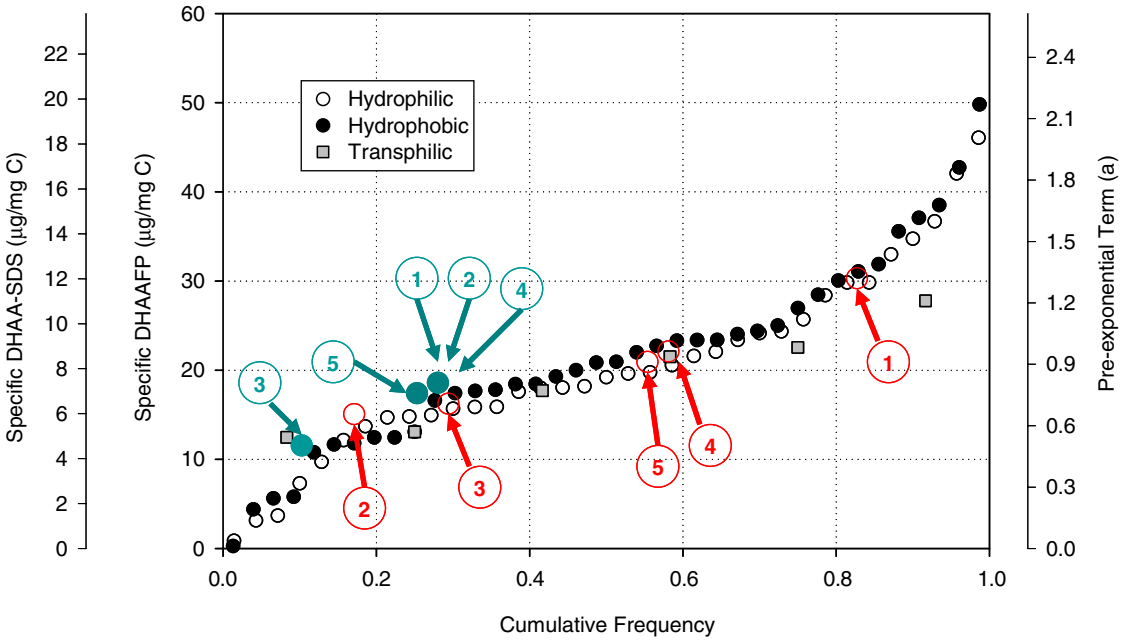



**Figure 5.24 DHAA precursor content of watershed NOM shown against the backdrop of the full literature database (untreated waters)**

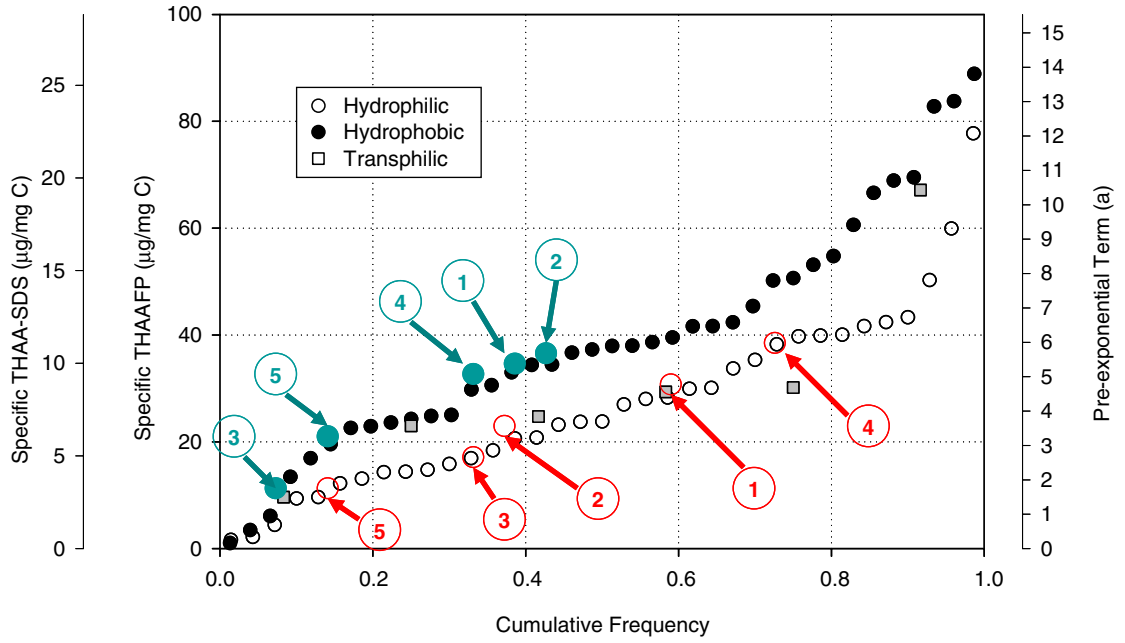



**Figure 5.25 THAA precursor content of watershed NOM shown against the backdrop of the full literature database (untreated waters)**



©2007 AwwaRF. All Rights Reserved.

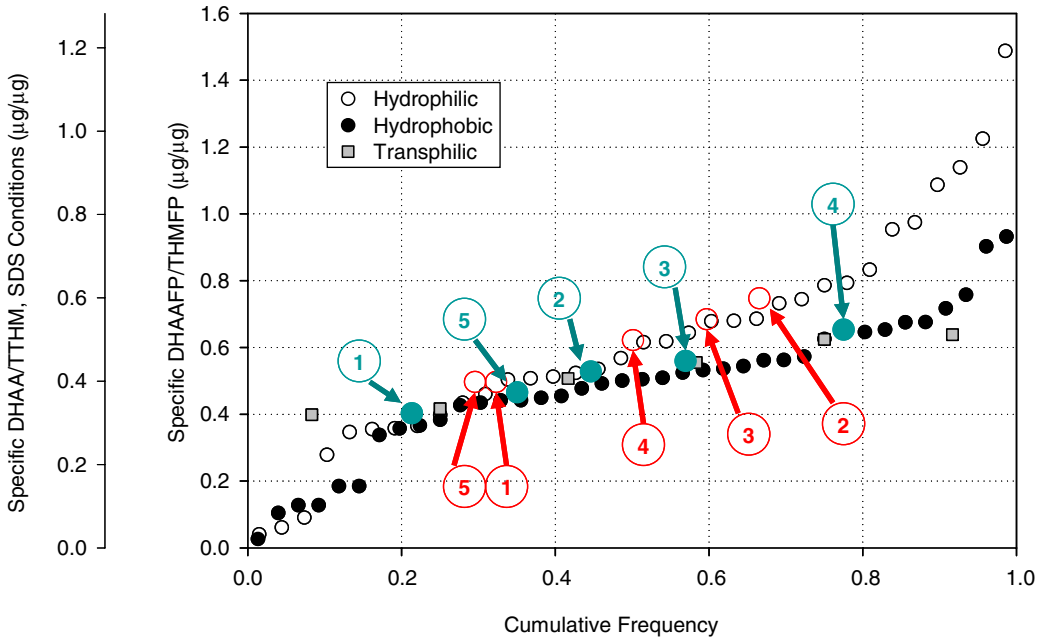



**Figure 5.26 DHAA/THM precursor ratio of watershed NOM shown against the backdrop of the full literature database (untreated waters)**

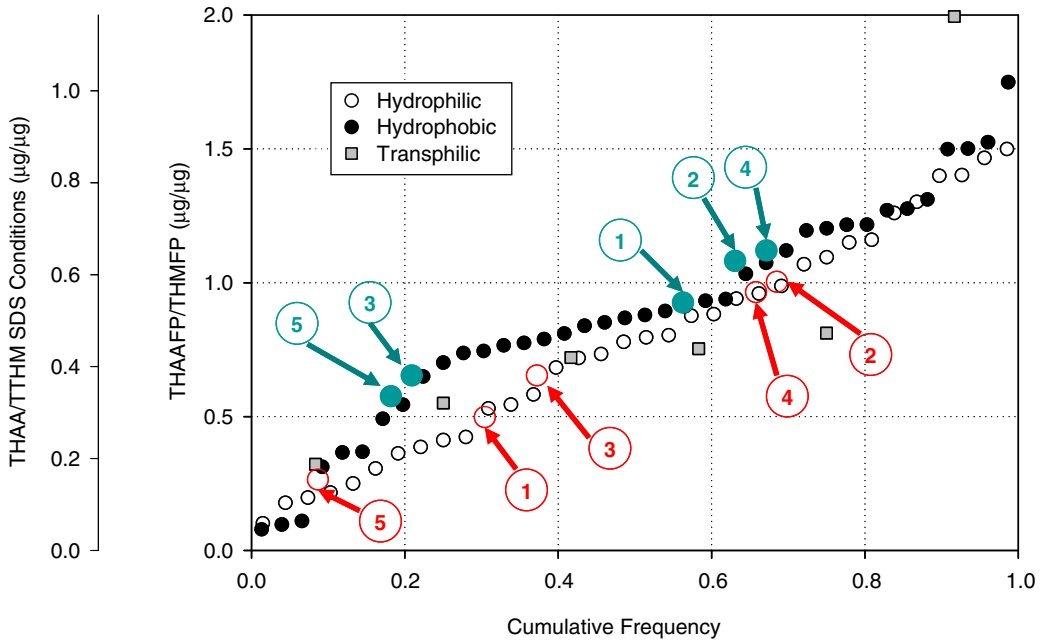



**Figure 5.27 THAA/THM precursor ratio of watershed NOM shown against the backdrop of the full literature database (untreated waters)**



©2007 AwwaRF. All Rights Reserved.

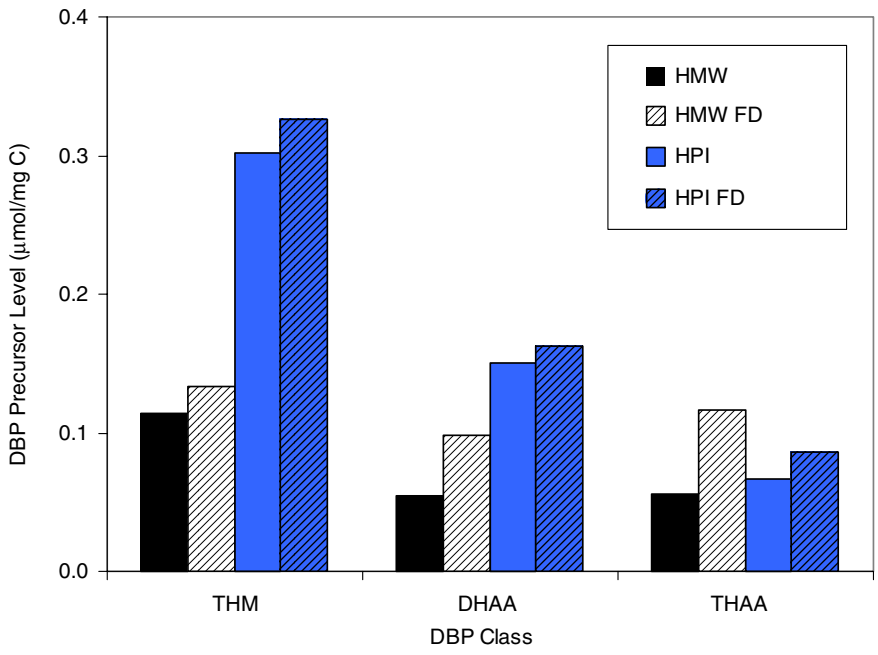



**Figure 5.28 Chlorination by-product precursor content for NOM fractions extracted from Quabbin River water (dry versus aqueous NOM)**

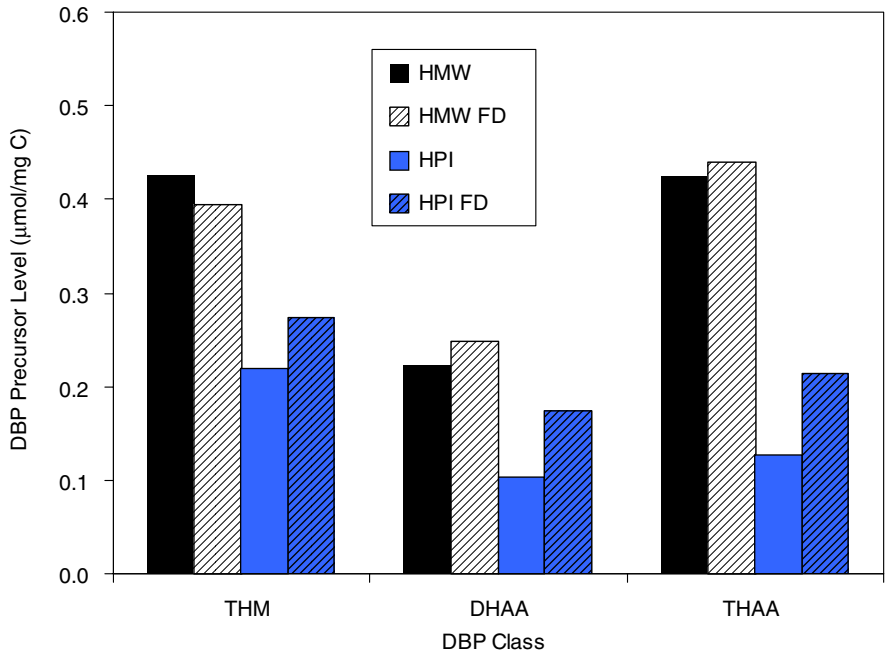



**Figure 5.29 Chlorination by-product precursor content for NOM fractions extracted from Hillsborough River water (dry versus aqueous NOM)**



©2007 AwwaRF. All Rights Reserved.

Lake, (4) Hillsborough River, and (5) Quabbin Reservoir. TOC measurements show that three of these waters have levels that are greater than the national average for surface waters, one is about average, and one (Quabbin) is well below. Based on SUVA measurement, Hillsborough and Assomption waters are classified as high-SUVA water, suggesting high lignin content, which is typical of highly terrestrial and forested sources. Moderate- and low-SUVA water can be attributed to Chickahominy River, Shoal Lake, and Quabbin Reservoir waters.

Laboratory testing of these waters included isolation of NOM fractions based on size (UF) and hydrophobicity (resin adsorption), resulting in three isolates from each water sample: an HMW fraction, an HPO fraction, and an HPI fraction. Each water and the corresponding NOM fractions were tested for their tendency to form chlorination by-products including THMs (THM4: chloroform, bromoform, bromodichloromethane, and chlorodibromomethane); MHAAs (MCAA and MBAA); DHAAs (DCAA, DBAA, and BCAA); and THAAs (TCAA, BDCAA, CDBAA, and TBAA). Samples were also taken for testing of AOC and BDOC.

NOM fractionation based on size (UF) and hydrophobicity (resin adsorption) results in losses at multiple stages of sample treatment. Thus, the overall distribution of NOM in each water is determined as the collection of isolatable fractions (extract) from the freeze-drying vessels and the amount lost either during freeze-drying and sample transfer (lost FD) or freeze-drying and adsorption onto XAD-8 (lost FD and XAD). Distinguished also is a "missing NOM" category, which is organic material that has been lost during preliminary evaporation and UF. This portion of NOM is probably a combination of HPI, HPO, and HMW fractions, but the exact breakdown cannot be determined.

The correlation between DBP precursor content and DOC is somewhat source specific, and this is commonly quantified in the form of a precursor-to-DOC ratio (i.e., the specific DBP precursor level). The Hillsborough River water sample shows the highest specific abundance of THM, DHAA, and THAA precursors among the five water samples. Regardless of the water sample, THM precursor content exceeds THAA and DHAA precursor contents. THAA and DHAA precursor contents decrease going from high-SUVA water to moderate-SUVA water and low-SUVA water. Shoal water showed the lowest THM precursor when compared to other waters.

There is no easily discernable trend in precursor content among the three fractions. The HMW fraction isolated from high-SUVA waters (Assomption and Hillsborough) showed high DBP precursor content as compared to the low-MW fractions (HPI and HPO). The opposite was true for low- and moderate-SUVA waters (Quabbin and Chickahominy). For Shoal Lake water, there is a decreasing trend in DBP precursor content going from HPI to HMW to HPO fractions. DHAA and THM precursor levels decrease when going from HPI to HPO fractions. However, the opposite was true for the fractions extracted from Assomption River water. THAA formation from the HPO fraction exceeds its formation from the HPI fraction. The reverse was observed for the fractions extracted from Shoal Lake water. In general, the HPO fractions showed low abundance of DBP precursor falling below the database median. The HPI fractions are quite divergent, however.

Both ozonation and chlorination resulted in increases in AOC and BDOC of all waters and NOM fractions. Ozonation generally resulted in the highest levels, but in several cases chlorination caused increases that were greater than those observed in the ozonated samples.

Fractionation methods used in this study allowed recovery of most of the carbon and DBP precursor. However, freeze-drying seemed to enhance the production of DBPs, suggesting a change in chemical structure of NOM. Decarboxylation (loss of $CO_2$) and dehydration (loss of $H_2O$) reactions probably occur during freeze-drying, and these are likely to increase reactivity with an electrophilic substance such a hypochlorous acid.



©2007 AwwaRF. All Rights Reserved.

# CHAPTER 6
# LEACHING STUDIES

## INTRODUCTION

The first step in NOM export is DOM generation. The subordinate process of DBP precursor export is imbedded within overall DOM generation. Three principal sources of DOM are atmospheric organic matter (vapor, liquid, and solid), release from nonliving organic matter (e.g., soil), and release from living organic structures (e.g., plants, animals). DOM release from plants is quite likely the most important process, especially in heavily forested catchments.

The connection between land cover and DOM generation is at least partly related to leaching of organic material from plant matter. Initial sources of NOM and organic DBP precursors include atmospheric, as indicted by organic carbon levels in direct rainfall, and leaching of organic material from living and degrading biomass. Many researchers have attempted to quantify and characterize the organic matter that is released by contact with higher plants. Usually these include laboratory leaching studies using foliage and woody tissues from various species.

## BACKGROUND

Sources and dynamics of allochthonous NOM are more complex and diverse than autochthonous NOM, and have yet to be fully understood. Allochthonous NOM originates from terrestrial, wetland, and littoral zone plant sources. It is generally believed that the predominant fraction of allochthonous NOM is humic in nature, stemming from the microbial degradation of fresh and detrital plant material (Wetzel 2001).

Cellulose and lignin materials from plant structural constituents make up the largest source of allochthonous humic compounds. These complex compounds undergo a series of decomposition stages in both soil and water. First, degradation of fresh plant material by fungi occurs under aerobic conditions, degrading polyphenolic lignin into recalcitrant high-MW substances (Wetzel 2001). At this point, the slow process of degradation by extracellular enzymes begins, leading to smaller, more oxidized macromolecules containing carboxylic functional groups. Further degradation then finally results in humic and fulvic acids and other compounds (Wetzel 2001). High concentrations of humics as a result of this microbial activity can be found in upper soil horizons as a result of the breakdown of surface plant material (Wetzel 2001, Thurman 1985).

### Leaching of DOC From Plant Material

Plant litter is known to contain highly soluble components that are transported to soil particles during rain events (Gosz et al. 1973, McDowell and Likens 1988, Meyer et al. 1998—all cited in Cleveland et al. 2004). Litter breakdown is a product of physical, chemical, and biological processes (see Currie and Aber 1997, Neff and Asner 2001). There is a large body of literature on degradation and solubilization of leaf litter. Most of the literature concerns loss of biomass from the forest floor and upper soil horizons. However, some of it concerns export of aquatic NOM. Some significant papers from this literature are summarized in Table 6.1.

Several researchers have examined some form of leaching potential of native plant material. A few of these have examined key properties of these materials such as biodegradability.



©2007 AwwaRF. All Rights Reserved.

**Table 6.1**
**Summary of literature on leaching studies with plant matter**

| Study | Pretreatment | Extraction solvent | Solvent/Solid ratios | Conditions | Separation | Notes |
|---|---|---|---|---|---|---|
| Tao and Lin 2000 | None | Water | 2.2–20 mL/g (100 mL with varying amounts of solid) | 24 hr, with occasional shaking | 0.45-µm filtration of supernatant | Used with soils and sediments; partition model |
| Hongve 1999 | None | Rainfall | 12 months of rain on 5–20 g dry weight | Ambient temperature (0°–18°C); no light | 0.45-µm filtration | Conducted in situ with lysimeters |
| Smolander and Kitunen 2002 | None | Water | 1.5 L per 300 mL soil | 14°C for 3 weeks, no light | 0.45-µm filtration | Used with soils only |
| Dilling and Kaiser 2002 | Sieved (<2 mm) and homogenized | DI water | 2 L with 200 g leaves | 15 min stirring, 18 hr quiescent | 0.45-µm filtration (Supor-450; Gelman, Ann Arbor, Mich.) | |
| Page et al. 2002 | None | Milli-Q water (Millipore) | 1 L with 1 kg vegetation | 2 min in blender at ambient temperature | 0.7-µm GF/F, then 0.45-µm nitrocellulose filter | Used for vegetation and soil; also used for formation potential |
| Strauss and Lamberti 2002 | Dried at room temperature and ground into coarse fragments | Reverse osmosis water | | 24-hr contact at room temperature | Series of sieves, then 0.2-µm membrane filter | Also conducted biodegradation study |
| Magill and Aber 2000 | Air-dried at room temperature; cut to 2 × 4 cm in size | DI water with and without $NO_3$ and $NH_3$ | 3.8 g litter; 200 mL water per week | 1–15 weeks; room temperature | | Tested impacts of different leaching protocols |
| Li et al. 2003 | Soil passed through 0.25-mm sieve | DI water | 500 mL with 1 g soil | Shaken for 4 hr | 0.45-µm membrane filter | Soil only; performed formation potentials |
| Randtke et al. 1987 | Dried and milled soil | 10 mL sodium tripolyphosphate solution | 30 mL with 1 g dried sample | 60 sec vortex mixing, 15 min sonication, 60 sec mix | 10 min centrifugation at 2,500 rpm, decant | |
| Randtke et al. 1988 | Dried | Lake water (3.6 mg C/L) | 4 g dry foliage; water volume not mentioned | 7–28 days; both aerobic and anaerobic conditions | | 7-day DOC was high at 400 mg/L, and decreased due to biodegradation. Performed formation potentials. |



©2007 AwwaRF. All Rights Reserved.

Hongve (1999) studied fresh litter from both coniferous and deciduous trees in Norway by obtaining samples collected from lysimeters during a 12-month period of exposure to natural precipitation. No pretreatment of the collected water was conducted. Deciduous litter produced about six times the amount of leached carbon, based on the same starting dry weight. The difference in production was particularly pronounced during the first few months. Deciduous litter also released large amounts of hydrophilic acids (35%) in the early fall, which seemed to give way to hydrophilic neutrals (up to 60%) later in the year. Hydrophobic acids constituted 40%–50% of organic matter released from the deciduous litter throughout the study. Values were slightly higher for the coniferous litter.

### *Role of Light in Leaching and Biodegradation of Plant Matter*

Wetzel (2003) states that "photochemical modification of organic macromolecules can result in major alterations in biological availability of portions of complex, heterogeneous dissolved organic compounds." This can cause decreases (e.g., "cross-linking of polypeptide chains with polyphenolic humic substances can lead to enzymatic inhibition") or increases ("partial photolysis of humic macromolecules, particularly with the generation of volatile fatty acids and related simple compounds that serve as excellent substrates for bacterial degradation") in biodegradability. Many laboratory studies of plant matter leaching processes have incorporated light exposure.

## Biodegradation of Natural Organic Matter

When NOM reaches surface waters, it is subject to extensive modification by microbial populations in streams, including benthic microflora and planktonic bacteria (Wetzel 2001). The loading of DOC to receiving waters has been shown to cause a dramatic increase in DOC-metabolizing bacteria, resulting in the rapid degradation of DOC (Wetzel 2001). Degradation can take weeks (simple carbohydrates from leaves) to years (woody materials; Wetzel 2001).

The stepwise degradation of DOC in waters has been described by Thurman (1985). Simple low-MW compounds (carbohydrates, amino acids, fatty acids) are rapidly metabolized by heterotrophic bacteria, sometimes in a matter of hours. High-MW compounds are too large to pass through the cell membrane and so must first be broken down by abiotic processes or extracellular enzymes. The entire process can take days to months. Straight-chain aliphatic hydrocarbons are also biodegradable and are generally degraded by microorganisms through beta oxidation. Branched cyclic aliphatic structures (terpenoids, steroids) as well as highly substituted aromatic ring structures (tannins, lignins) are quite recalcitrant and resistant to biodegradation. For this reason, these compounds often account for a large percentage of DOC in waters, with only minor amounts of biodegradable compounds from plants and bacteria (Thurman 1985).

Given enough time, even these recalcitrant organics can become mineralized. Cleveland and co-workers (Cleveland et al. 2004) found that both humic and nonhumic carbon (based on XAD-8 adsorption) from plant leachate was almost completely biodegradable following 100 days of incubation. Losses were generally greater than 90%. There was significant discrimination, however, in the degradative losses. The nonhumic fraction degraded to a greater relative extent, such that the humic fraction became substantially more dominant. For the temperate fir (living foliage), they noted increases from 43% to 80% after 100 days.

An increase in the aromatic carbon content of plant-and-soil-derived NOM has been observed using $^{13}C$ nuclear magnetic resonance as water moves downgradient through a stream



©2007 AwwaRF. All Rights Reserved.

(Wetzel 2001). After stream water has reached the larger recipient water body, the original DOC has been extensively modified (Wetzel 2001). In general, almost 50%–80% of surface waters contain low-MW refractory organic matter having a high degree of aromaticity and high C/N ratios (Krasner et al. 1996).

## EXPERIMENTAL DESIGN

Two types of leaching tests were conducted as part of this research. The first (study 1) was designed to examine differences in tree species. This work involved species of pine, maple, and oak. The second type of test (study 2) was intended for understanding differences in sequential extraction (i.e., successive rainfall events) on a single species (maple) but included fresh leaves as well as aged ones.

### Study 1

Laboratory samples of NOM were prepared using leaves and water in order to mimic the leaching of organic matter from materials present in the environment. The fresh organic matter leachate was then allowed to sit for approximately 80 days in order to simulate the potential biodegradation processes that occur after the NOM has been released into receiving bodies of water.

Approximately 1 week prior to the beginning of the leaching study, red maple leaves (*Acer rubrum*), white oak leaves (*Quercus alba*), and white pine needles (*Pinus strobus*) were collected from three areas of the Wachusett Reservoir watershed (supplying the Boston metropolitan area). Five samples were prepared for each leaf type, including one large 4-L dark bottle (containing 4 g of leaves), two smaller 1-L dark bottles, one 1-L light bottle, and one 1-L dark bottle with biocide (all 1-L bottles containing 1 g of leaves in each). Artificial light generated with fluorescent lamps was used to illuminate some of the samples, incorporating daylight and nighttime cycles.

Throughout the course of the leaching and biodegradation periods, samples were analyzed for $UV_{254}$ and DOC. After 8 days of leaching, leaves were removed and the samples in large bottles were filtered and then subjected to the biodegradation test. The four smaller samples for each leaf type were then tested for their potential to form DBPs. At the end of the biodegradation period, the three larger samples were also tested for DBPFP. The results of these analyses are presented in the following paragraphs.

#### *Leaching Period*

Leaves used for the laboratory leaching of NOM were collected from three areas of the Wachusett Reservoir watershed that contained relatively homogeneous stands of one tree species. The three species collected were red maple (*A. rubrum*), white oak (*Q. alba*), and white pine (*P. strobus*). All samples were collected in June from the top layer of the forest floor, which contained leaves from the previous season's litterfall.

Five bottles were prepared for each leaf type, with two of the bottles prepared in the same manner. Overall, 15 samples were prepared. The bottles and treatments for each leaf type included

- 4-L bottle, dark, 3,200 mL media, 4 g leaves
- Two 1-L bottles, dark, 800 mL media, 1 g leaves



©2007 AwwaRF. All Rights Reserved.

- 1-L bottle, light, 800 mL media, 1 g leaves
- 1-L bottle, dark, biocide, 800 mL media, 1 g leaves

Dark samples were protected from light by surrounding the bottles with aluminum foil. Samples were soaked in Super-Q water seeded with reservoir water in order to simulate field conditions. Sample media was prepared by adding water from the Cosgrove Intake of the Wachusett Reservoir to Super-Q water in a 1:4 ratio. Leaves were dried in an oven at 50°C for less than 24 hours before being weighed. Samples containing biocide were dosed with 1 mg/L $HgCl_2$.

During leaching, samples were kept at 22°C in a constant-temperature incubator that was lighted with seven 30-W fluorescent lamps (GRO-LUX; Sylvania, Danvers, Mass.). Samples were aerated with high-purity synthetic air (21% oxygen, 79% nitrogen) using a coarse glass frit dispersion tube. Samples were leached for 8 days, and $UV_{254}$ and DOC measurements were periodically recorded. $UV_{254}$ measurements were recorded on days 1, 2, 3, 4, 6, 7, and 8; DOC measurements were taken on days 1, 3, 6, and 8.

At the end of the leaching period, the water from all twelve 1-L bottles was decanted from the leaves, filtered through 0.45-µm GF/F, and prepared for bench-scale chlorination and DBPFP analysis. The water in the three 4-L bottles was also decanted and filtered; however, the leaves were removed from the bottle and the water was returned to the bottle for the biodegradation period of the experiment.

### *Biodegradation Period*

The three 4-L samples for biodegradation were returned to the constant-temperature incubator set at 22°C, and then were aerated and held for 81 days. During this time, periodic measurements of $UV_{254}$ and TOC were recorded, including measurements on the following days: days 0, 5, 11, 12, 13, 15, 19, 26, 43, 49, and 81 for $UV_{254}$ and days 0, 5, 19, 49, and 81 for DOC. At the end of the 81-day biodegradation period, all three samples were filtered through a 0.45-µm GF/F, analyzed for $UV_{254}$ and DOC, chlorinated, and then analyzed for DBPFPs.

## Study 2

### *Leaching Period*

Red maple leaves were collected in July 2004 from a relatively homogeneous stand of red maple (*A. rubrum*) located in the Quabbin Reservoir watershed, just south of the Windsor Dam (also serving the Boston metropolitan area). Both "new" and "old" leaves were collected, with new leaves collected directly from red maple trees, and old leaves collected from the top layer of the forest floor, which contained leaves from the last season's litterfall. Leaves were placed in ziplock bags and transported to the UMass laboratory.

Small-scale trial leaching experiments were performed to determine the requisite mass of leaves needed in each bottle. New leaves were also dried at 50°C for less than 24 hours to establish a ratio of ambient leaf mass to dried leaf mass. However, in order to maintain the original structural composition of the biomass, the leaves used for the leaching period of the experiment were not dried (see Figure 6.1).



©2007 AwwaRF. All Rights Reserved.

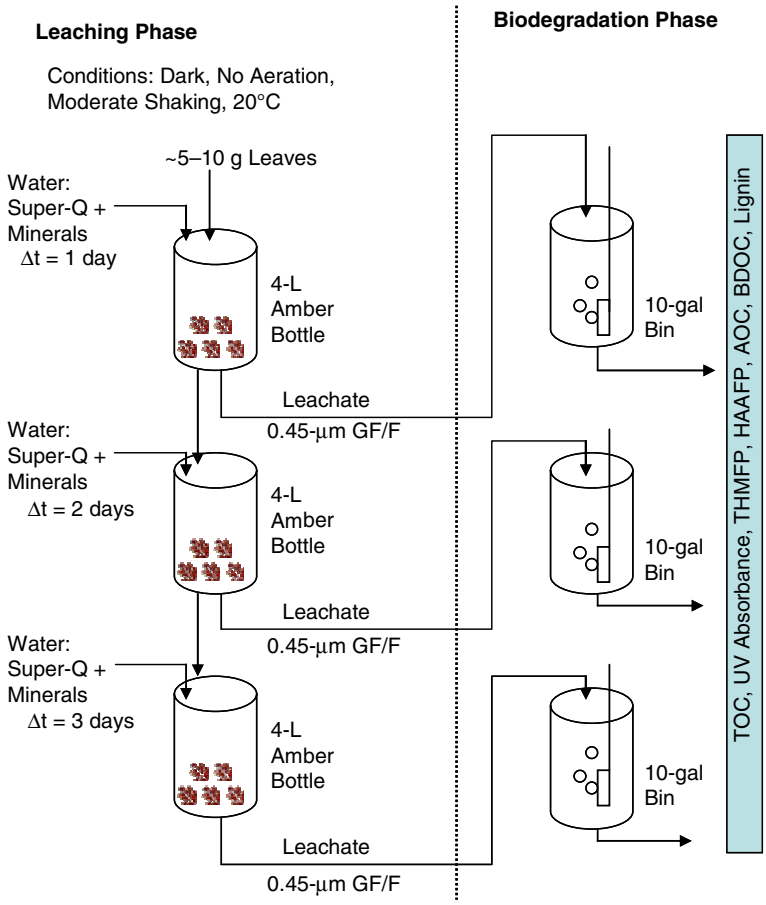



**Figure 6.1 Experimental design for leaching study 2**

Using values determined from the trial leaching experiments, two leaching setups were created, with the following substances added to 4-L amber glass bottles:

- New leaves: 3.2 L mineral water + 10 g new leaves
- Old leaves: 3.2 L mineral water + 20 g old leaves

Mineral water was prepared using standard methods for preparation of moderately hard reconstituted water. The following procedure was used for preparing 1 L of moderately hard reconstituted water:

1. Dissolve 0.060 g $CaSO_4$ (calcium sulfate) in 300 mL of high-purity water.
2. In a separate glass container, dissolve 0.004g KCl (potassium chloride), 0.060 g $MgSO_4$ (magnesium sulfate), 0.096 g $NaHCO_3$ (sodium bicarbonate), and 0.02 mL of selenium stock solution (0.33 g $Na_2SeO_3$, or magnesium selenate) in 1 L of high-purity water.
3. Mix both solutions.
4. Add high-purity water to reach a volume of 1 L.



©2007 AwwaRF. All Rights Reserved.

Leaching bottles were kept in dark conditions in a constant-temperature room set to 20°C, and were moderately agitated at 20 rpm using a laboratory shaker. After 24 hours, carbon levels in each bottle were measured; if the level reached at least 40 mg C/L, the samples were then prepared for the biodegradation period.

### *Biodegradation Period*

Water from the leaching test was decanted from the leaves, filtered through 0.45-µm GF/F, and then prepared for the biodegradation portion of the experiment. Remaining leaves and bottles were prepared for the next phase of leaching by adding 3.2 L of mineral water and starting the leaching process again. New leaves were leached for a total of three phases (cycles), and old leaves were leached for a total of two cycles, with the number of cycles determined based on available carbon leached from the leaves.

After the leachate was filtered, the entire 3.2 L of leachate was placed in a precleaned 10-gallon polyethylene carboy, along with 1.6 L of "seed" water that was added to simulate field conditions and 30.4 L of high-purity water. The seed water was collected from Mountain Street Reservoir, a raw water supply for Northampton, Mass., located in Williamsburg, Mass. Samples were then held in dark conditions at a constant room temperature of 20°C and aerated with filtered, humidified, zero-grade air using a coarse glass frit dispersion tube. Throughout the biodegradation period, samples were collected for the following analytical tests:

- $UV_{254}$
- DOC
- DBPFP
- AOC and BDOC
- Lignin monomers

Samples were not aerated during sample collection. Samples were collected and testing was performed according the schedule shown in Table 6.2.

## RESULTS OF STUDY 1

### Leaching

All three leaf types were leached under three different conditions: dark, dark with biocide (abiotic), and light. Triplicates of the dark conditions were performed in one large-volume (4-L) bottle and two small-volume (1-L) bottles. The results of the leaching of maple leaves for all five samples are presented in Figure 6.2. All five samples responded in the same manner, with slightly higher levels of organic matter leaching from the light and dark-with-biocide bottles. The responses of the three dark samples were almost identical. The results of the leaching of oak leaves are presented in Figure 6.3, with the samples having a similar response. All of the samples leached comparable levels of NOM, with slightly higher levels from the sample containing biocide. Figure 6.4 shows the results of the pine needles, where the pattern is similar. However, the levels of $UV_{254}$ are very different.

The results of TOC leaching for all of the samples over time are presented in Table 6.3. The leaching patterns for the three large samples from each of the leaf types are compared in



©2007 AwwaRF. All Rights Reserved.

**Table 6.2**
**Analytical schedule for study 2 testing**

| | Number of days spent in biodegradation period | | | | | |
|---|---|---|---|---|---|---|
| | 0 | 7 | 14 | 21 | 28 | 35 |
| $UV_{254}$ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| DOC | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| DBPFP | ✓ | ✓ | ✓ | | ✓ | |
| AOC/BDOC | ✓ | ✓ | ✓ | | | |
| Lignins | ✓ | ✓ | ✓ | | | |
| | 42 | 49 | 56 | 63 | 70 | 77 |
| $UV_{254}$ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| DOC | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| DBPFP | | | ✓ | | | ✓ |
| AOC/BDOC | | | | | | ✓ |
| Lignins | | | ✓ | | | ✓ |

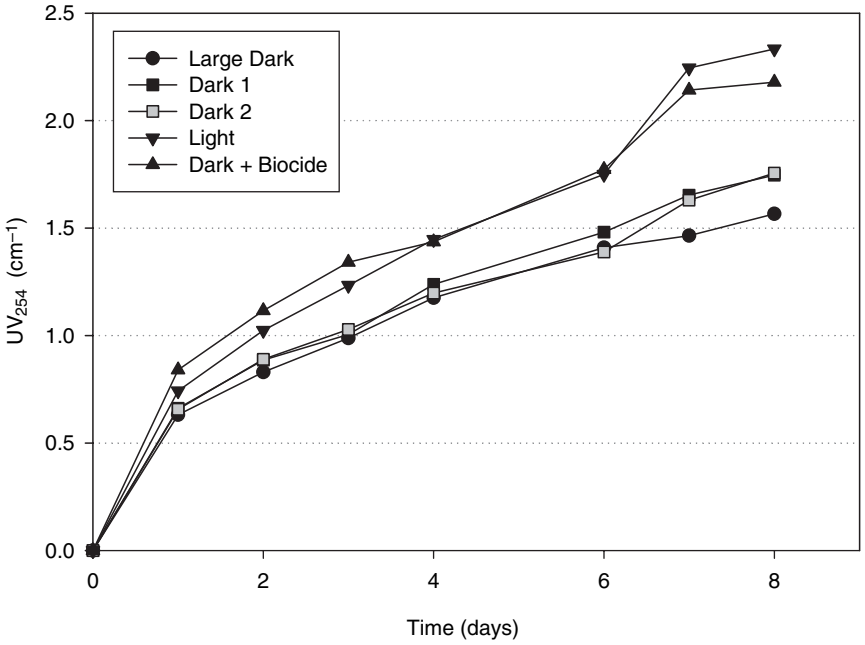



**Figure 6.2 UV absorbance levels over time for leaching of red maple leaves (*A. rubrum*)**



©2007 AwwaRF. All Rights Reserved.

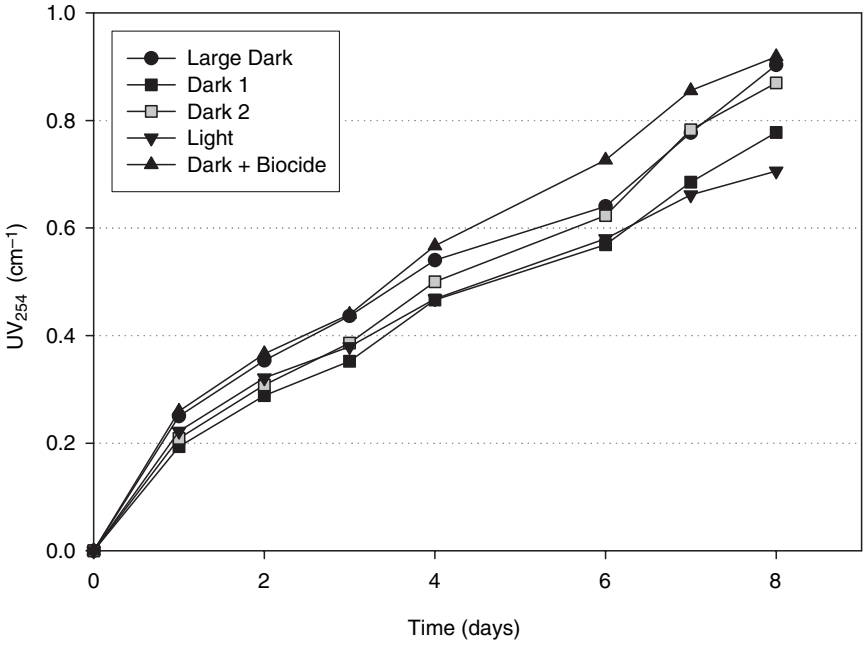



**Figure 6.3 UV absorbance levels over time for leaching of white oak leaves (*Q. alba*)**

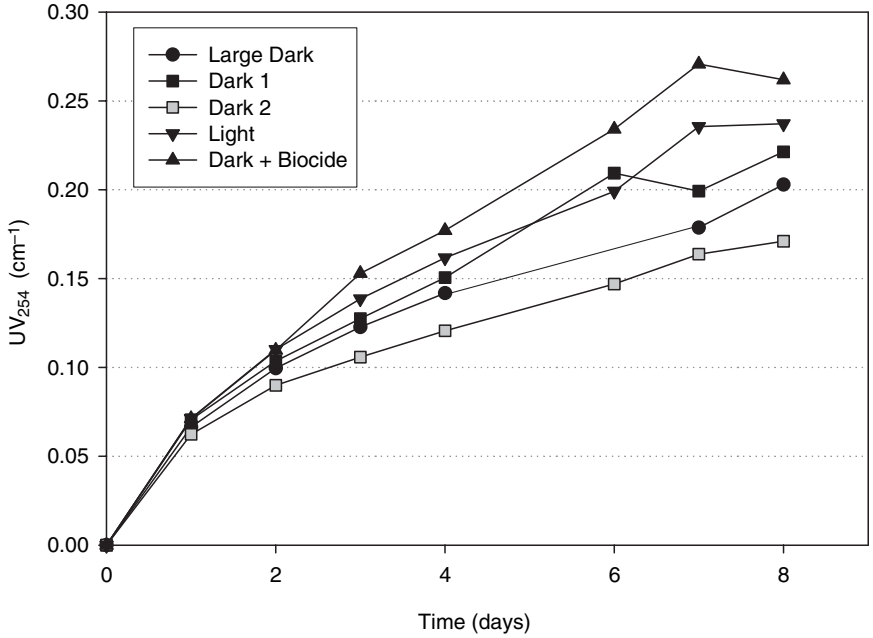



**Figure 6.4 UV absorbance levels over time for leaching of white pine needles (*P. strobus*)**



©2007 AwwaRF. All Rights Reserved.

**Table 6.3**
**Leaching experiment DOC results for all samples**

| Sample number | Description | DOC (mg/L) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Day 1 | SD* | Day 3 | SD | Day 6 | SD | Day 8 | SD |
| 1 | Large maple dark | 15.9 | 0.3 | 20.5 | 0.3 | 27.8 | 0.4 | 31.9 | 0.5 |
| 2 | Large oak dark | 10.33 | 0.16 | 18.1 | 0.9 | 23.4 | 0.3 | 30.3 | 0.4 |
| 3 | Large pine dark | 6.86 | 0.07 | 11.3 | 0.2 | 10.3 | 0.2 | 12.56 | 0.14 |
| 4 | Maple dark 1 | 17.8 | 0.4 | 25.4 | 0.2 | 32.3 | 0.6 | 40.7 | 0.7 |
| 5 | Maple dark 2 | 18.5 | 0.4 | 22.86 | 0.08 | 31.86 | 0.06 | 37.3 | 0.8 |
| 6 | Oak dark 1 | 8.57 | 0.17 | 15.68 | 0.15 | 22.2 | 0.4 | 28.7 | 0.3 |
| 7 | Oak dark 2 | 10.1 | 0.2 | 18.1 | 0.4 | 26.1 | 0.3 | 34.9 | 0.1 |
| 8 | Pine dark 1 | 5.96 | 0.02 | 10.05 | 0.17 | 10.32 | 0.11 | 19.7 | 0.6 |
| 9 | Pine dark 2 | 6.2 | 0.5 | 8.9 | 0.5 | 8.0 | 0.3 | 13.45 | 0.09 |
| 10 | Maple light | 20.64 | 0.12 | 29.5 | 0.04 | 43.8 | 0.8 | 61.2 | 0.6 |
| 11 | Oak light | 9.77 | 0.15 | 15.6 | 1.3 | 22.28 | 0.09 | 27.5 | 0.3 |
| 12 | Pine light | 7.33 | 0.05 | 11.66 | 0.03 | 13.3 | 0.3 | 17.81 | 0.12 |
| 13 | Maple dark + biocide | 23.8 | 0.5 | 32.0 | 0.3 | 43.0 | 0.7 | 48.4 | 0.7 |
| 14 | Oak dark + biocide | 12.1 | 0.2 | 20.2 | 0.4 | 29.9 | 0.5 | 36.5 | 0.6 |
| 15 | Pine dark + biocide | 8.1 | 0.2 | 15.00 | 0.04 | 19.0 | 0.2 | 24.4 | 0.3 |

*SD = standard deviation.

Figure 6.5 using UV absorbance results. The maple leaves leached the highest levels of organic matter, with a final UV absorbance of 1.57 $cm^{-1}$ on day 8, corresponding to a TOC of approximately 32 mg/L. Oak leaves also leached high levels of organic matter, with a final UV absorbance of 0.90 $cm^{-1}$ and a TOC of 30 mg/L. Pine needles leached the lowest levels of organic matter, with a final UV absorbance of 0.20 $cm^{-1}$ and a TOC of 13 mg/L.

As the leaching experiment progressed, small amounts of liquid were lost to evaporation. This caused the organic carbon concentrations to be higher than if no evaporation had occurred. For this reason, the differences in the amounts of NOM leached were also compared in terms of total carbon, or the total mass that leached from each leaf type. This was calculated by estimating the volume of solution present on each day and then multiplying this volume by the DOC in order to calculate total milligrams of carbon.

Figure 6.6 shows the change in total carbon for each leaf type over time for the large samples. The total carbon levels in both maple and oak leachate samples were very high toward the end of the 9-day leaching experiment, with both levels approaching 90 mg C. Levels for the pine needle samples were much lower, with an accumulation of only 36 mg of dissolved carbon. The daily net (incorporating fast degradation with release) release rates for each leaf type and time period are shown in Table 6.4.

$UV_{254}$ and DOC for all of the samples on the final day of leaching are presented in Table 6.5. The DOC for all of the samples on day 9 is also presented in Figure 6.7. UV absorbance



©2007 AwwaRF. All Rights Reserved.

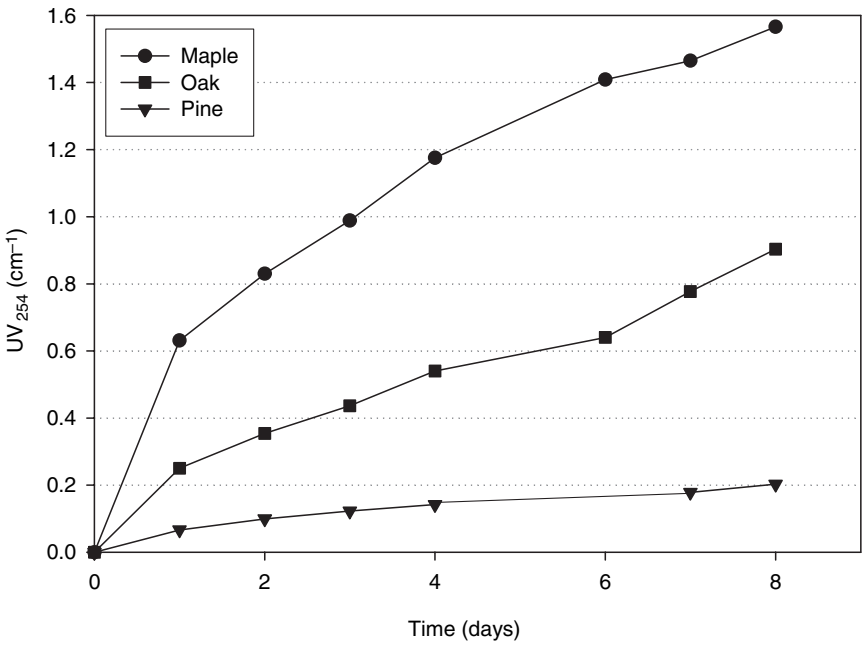



**Figure 6.5 UV absorbance levels over time for leaching of maple, oak, and pine foliage (large bottles, no light)**

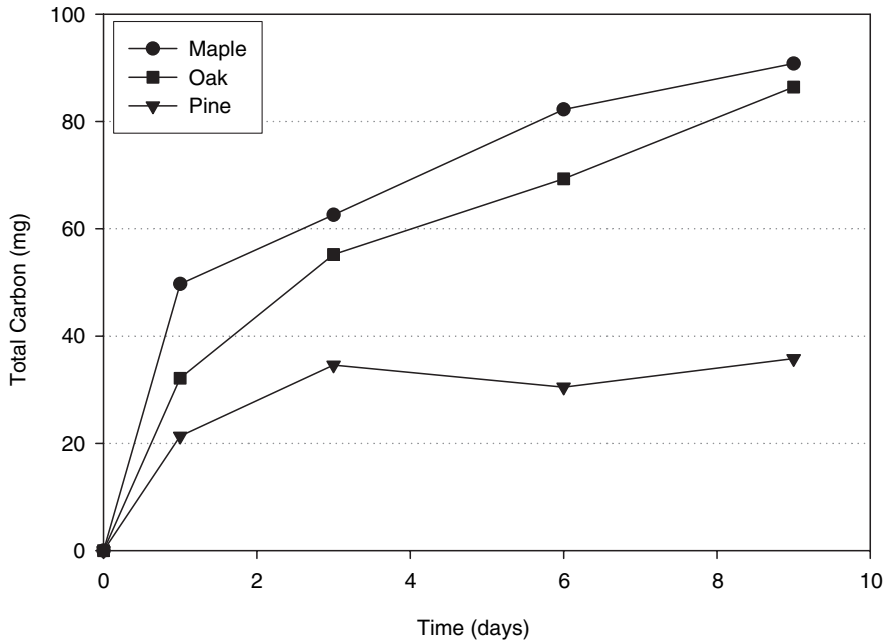



**Figure 6.6 Changes in total carbon over time for leaching of maple, oak, and pine samples (large bottles, no light)**



©2007 AwwaRF. All Rights Reserved.

**Table 6.4**
**Net carbon release rates (mg C/g dry wt/d)**

| Time (days) | Pine | Oak | Maple |
|---|---|---|---|
| 0–1 | 5.4 | 8.2 | 12.7 |
| 1–3 | 1.6 | 2.9 | 1.5 |
| 3–6 | –0.3 | 1.2 | 1.7 |
| 6–9 | 0.4 | 1.4 | 0.7 |

**Table 6.5**
**$UV_{254}$, DOC, and SUVA on final day of leaching for all samples**

| Sample name | $UV_{254}$ ($cm^{-1}$) | DOC (mg/L) | SUVA (L/mg-m) |
|---|---|---|---|
| Large maple dark | 1.570 | 31.90 | 4.92 |
| Maple dark 1 | 1.750 | 40.70 | 4.30 |
| Maple dark 2 | 1.760 | 37.30 | 4.72 |
| Maple light | 2.330 | 61.20 | 3.81 |
| Maple dark + biocide | 2.180 | 48.40 | 4.50 |
| Large oak dark | 0.903 | 30.30 | 2.98 |
| Oak dark 1 | 0.777 | 28.70 | 2.71 |
| Oak dark 2 | 0.870 | 34.90 | 2.50 |
| Oak light | 0.706 | 27.50 | 2.57 |
| Oak dark + biocide | 0.919 | 36.50 | 2.52 |
| Large pine dark | 0.203 | 12.56 | 1.62 |
| Pine dark 1 | 0.221 | 19.70 | 1.12 |
| Pine dark 2 | 0.171 | 13.45 | 1.27 |
| Pine light | 0.237 | 17.81 | 1.33 |
| Pine dark + biocide | 0.262 | 24.40 | 1.07 |

values for the maple samples ranged between 1.57 and 2.33 $cm^{-1}$, and correspondingly, DOC levels ranged between 32 and 61 mg/L, with the highest leached from the light sample and the lowest from the large dark sample. The oak samples had less variability, with UV absorbance varying between $0.7^0$ and $0.9^0$ $cm^{-1}$ and DOC levels between approximately 27 and 37 mg/L. The highest level of organic matter in the oak samples was found in the dark-with-biocide conditions, and the lowest in the light conditions. The pine samples also had less variability, with UV absorbance values between 0.17 and 0.26 $cm^{-1}$ and DOC between 13 and 24 mg/L. The large dark bottle leached the lowest amount and the dark with biocide leached the greatest level of DOC, the latter observation being similar to the oak leaves.

SUVA values were significantly different among leached species. Maple leaves had the highest SUVA values, with an average of about 4.5 L/mg-m. The oak samples had an average



©2007 AwwaRF. All Rights Reserved.

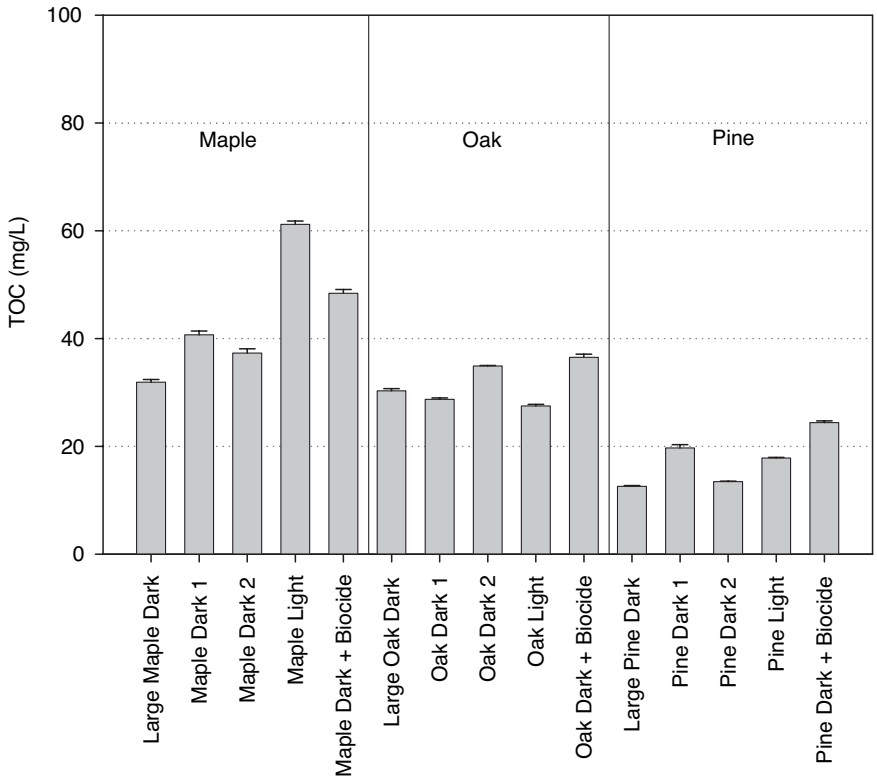



**Figure 6.7 TOC on final day of leaching for all samples**

SUVA that was much smaller, at about 2.7 L/mg-m. The pine samples had the lowest overall average SUVA of about 1.3 L/mg-m. This value was almost four times smaller than the maple samples.

Laboratory leaching studies conducted at UMass clearly show that precursor material can rapidly leach out of leaf litter, even aged litter. The leaves and needles gathered for this study were collected in late winter (March) from the top of the leaf pack. Therefore, they had undergone natural weathering for 4 to 5 months. For each leaf type, five bottles (four amber, with and without biocide, and one clear) were prepared containing a set mass of leaf matter in distilled water, seeded with a small amount of reservoir water (system intake). These bottles were allowed to sit for 7 days in the dark, during which time the $UV_{254}$ and SUVA were monitored daily (Figure 6.8).

DBPFPs were measured for all of the small-volume samples for each leaf type, with the larger-volume samples continuing on to the biodegradation period. Samples were diluted by a factor of 5 before chlorination. The results of this analysis are presented in Table 6.6 and Figure 6.9. The maple leaves had the highest overall formation potentials of the three leaf types, with very high THAAFP levels in particular. All of the DBPFP levels for each of the maple samples were similar, with THMFP levels about 1,800 µg/L, DHAAFP levels about 1,500 µg/L, and THAAFP levels about 5,000 µg/L. Oak DBPFPs were slightly different, with the light samples resulting in lower formation potentials and the dark-with-biocide samples resulting in the highest DBPFPs. In general, the dark oak samples (no biocide) had THMFP levels of about 1,450 µg/L, DHAAFP levels of about 1,350 µg/L, and THAAFP levels of about 3,000 µg/L. Most of the pine samples had comparable, but



©2007 AwwaRF. All Rights Reserved.

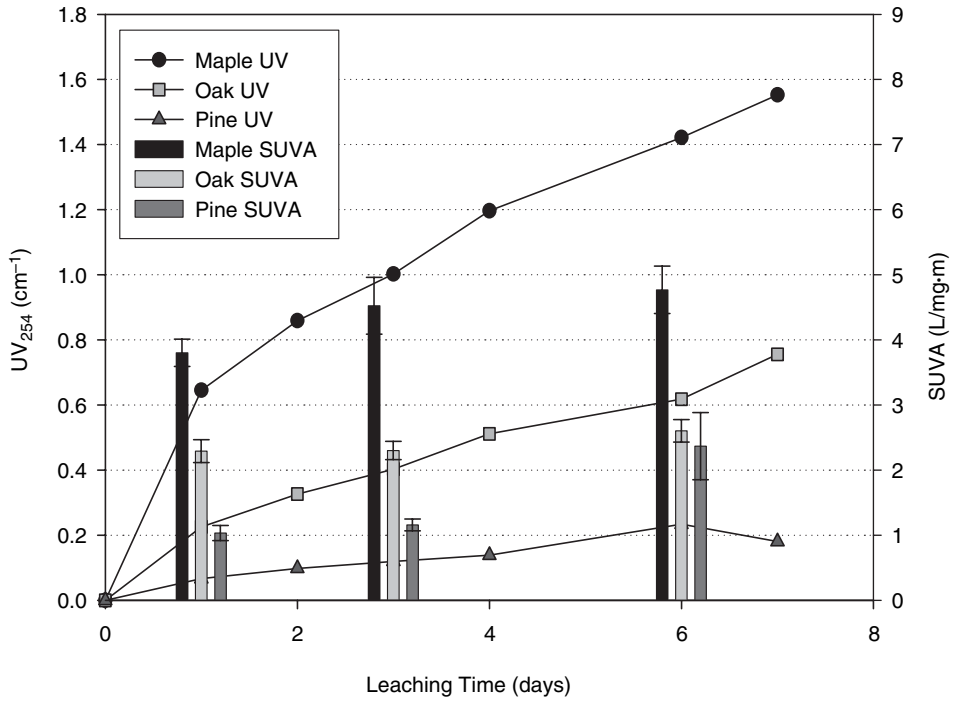



**Figure 6.8 Leaching of organic matter versus time for late fall leaf litter (dark conditions)**

**Table 6.6**
**DBPFP data for last day of leaching**

| Sample name | THM (µg/L) | SD* | DHAA (µg/L) | SD | THAA (µg/L) | SD |
|---|---|---|---|---|---|---|
| Maple dark 1 | 1,850 | 50 | 1,610 | 30 | 5,100 | 100 |
| Maple dark 2 | 1,500 | 400 | 1,500 | 140 | 5,200 | 700 |
| Maple light | 1,806.5 | 0.2 | 1,455 | 9 | 3,500 | 150 |
| Maple dark + biocide | 1,812 | 9 | 1,590 | 40 | 4,435 | 8 |
| Oak dark 1 | 1,376 | 2 | 1,400 | 400 | 3,200 | 700 |
| Oak dark 2 | 1,620 | 90 | 1,300 | 90 | 3,020 | 70 |
| Oak light | 1,313 | 6 | 880 | 160 | 2,100 | 800 |
| Oak dark + biocide | 1,889 | 15 | 1,740 | 50 | 4,400 | 300 |
| Pine dark 1 | 560 | 14 | 300 | 90 | 500 | 200 |
| Pine dark 2 | 453 | 5 | 323.6 | 0.8 | 526 | 2 |
| Pine light | 660 | 30 | 390 | 40 | 750 | 90 |
| Pine dark + biocide | 665 | 13 | 0 | 0 | 0 | 0 |

*SD = standard deviation.



©2007 AwwaRF. All Rights Reserved.

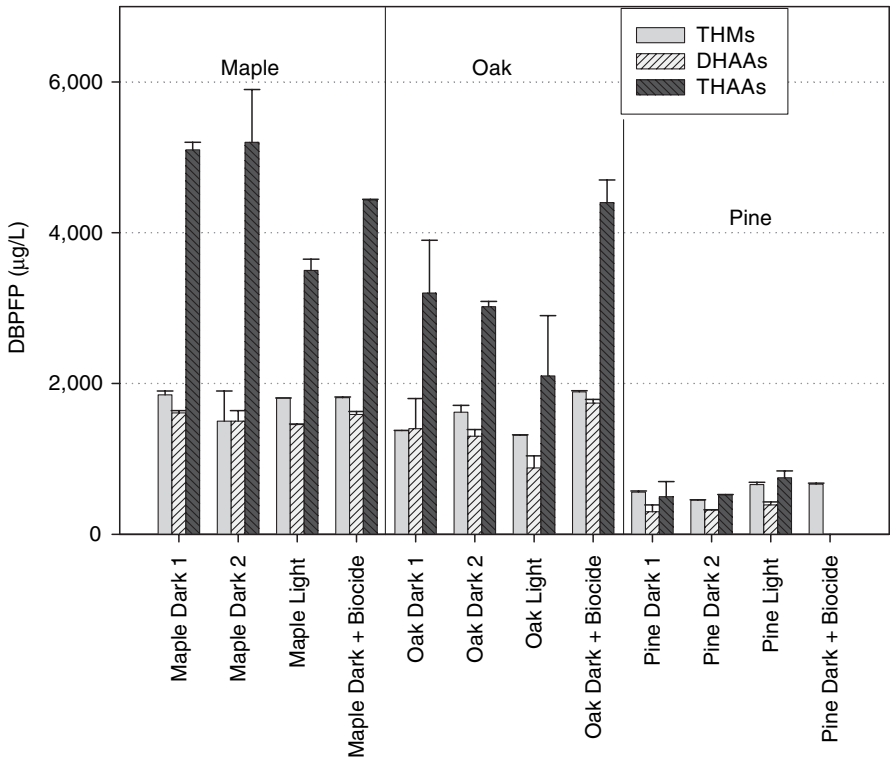



**Figure 6.9 DBPFPs for last day of leaching**

much lower, DBPFP levels, with the exception of the dark-with-biocide samples, where no HAAs were detected. The dark pine samples (no biocide) had an average of about 500 µg/L for THMs, about 300 µg/L for DHAAs, and about 500 µg/L for THAAFPs.

Overall, the THAAFP levels were very high for the maple and oak samples. The THM and DHAA formation potential levels were all below 2,000 µg/L, with DHAAFP in general lower than THMFP. All of the pine sample DBPFPs were lower than the maple and oak samples.

Specific DBPFP levels were calculated for all of the samples, and the results are presented in Table 6.7 and Figure 6.10. Oak samples had the highest specific THMFP, with all samples in the vicinity of 50 µg/mg C. Maple had the second highest specific THMFP ranging between 30 and 45 µg/mg C. Pine had the lowest specific THMFP with about 30 µg/mg C. Specific DHAAFPs followed the same pattern, with oak having the highest specific DHAAFP ranging between 32 and 49 µg/mg C, maple with the second largest formation potentials of 24 and 41 µg/mg C, and pine having the lowest specific DHAAFP of 15 to 24 µg/mg C. The water samples from leached maple had the highest specific THAAFP of about 130 µg/mg C for the dark samples. The oak water samples had a specific THAAFP of about 90 µg/mg C for the dark samples (no biocide), and the pine water had a specific THAAFP of about 33 µg/mg C for the dark samples. Overall, specific THAAFP levels for all three leaf types were the highest of the DBPs, with some levels reaching beyond 100 µg/mg C.

When subjected to chlorination tests, the leached organic matter proved to be a moderately rich source of THM and HAA precursors (Figures 6.11 and 6.12). There was considerably more variability in HAA yields than in THM yields. The oak and maple also showed a much higher ratio



©2007 AwwaRF. All Rights Reserved.

**Table 6.7**
**Specific DBPFP data for last day of leaching**

| Sample name | THM (µg/mg C) | SD* | DHAA (µg/mg C) | SD | THAA (µg/mg C) | SD |
|---|---|---|---|---|---|---|
| Maple dark 1 | 45.4 | 1.5 | 39.5 | 1 | 124 | 3 |
| Maple dark 2 | 40 | 10 | 41 | 4 | 140 | 20 |
| Maple light | 29.5 | 0.3 | 23.8 | 0.3 | 58 | 3 |
| Maple dark + biocide | 37.4 | 0.5 | 32.9 | 0.9 | 91.6 | 1.3 |
| Oak dark 1 | 48 | 0.5 | 49 | 14 | 110 | 20 |
| Oak dark 2 | 46 | 3 | 37 | 3 | 86 | 2 |
| Oak light | 47.8 | 0.6 | 32 | 6 | 75 | 30 |
| Oak dark + biocide | 51.7 | 0.9 | 47.6 | 1.6 | 120 | 8 |
| Pine dark 1 | 28.5 | 1.1 | 15 | 5 | 24 | 10 |
| Pine dark 2 | 33.7 | 0.4 | 24.1 | 0.2 | 39.1 | 0.3 |
| Pine light | 36.8 | 1.8 | 21 | 2 | 42 | 5 |
| Pine dark + biocide | 27.3 | 0.7 | 0 | 0 | 0 | 0 |

*SD = standard deviation.

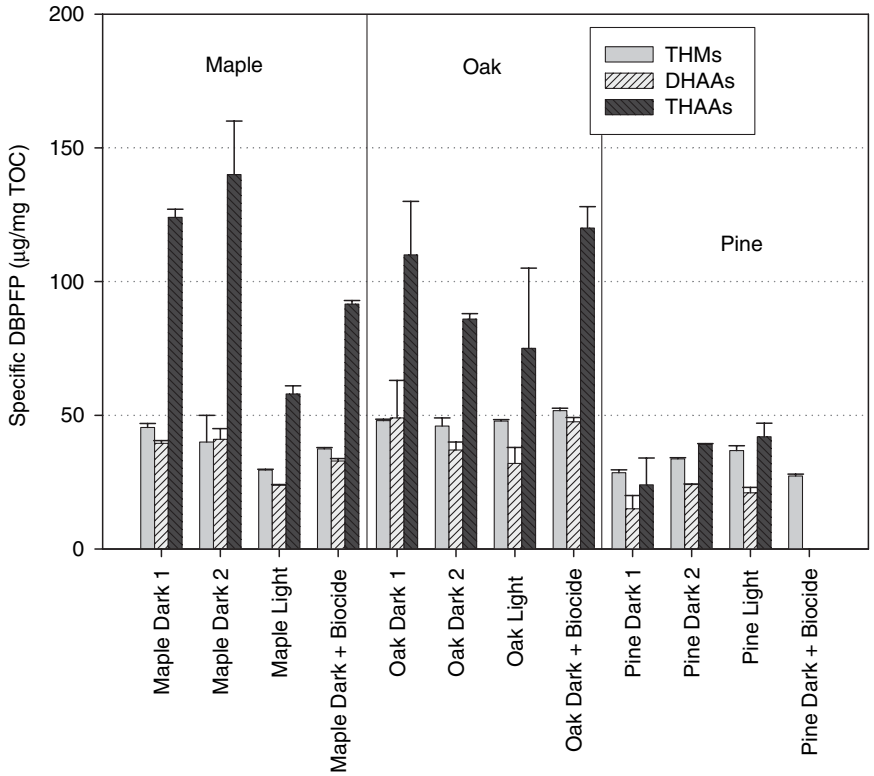



**Figure 6.10 Specific DBPFPs for last day of leaching**



©2007 AwwaRF. All Rights Reserved.

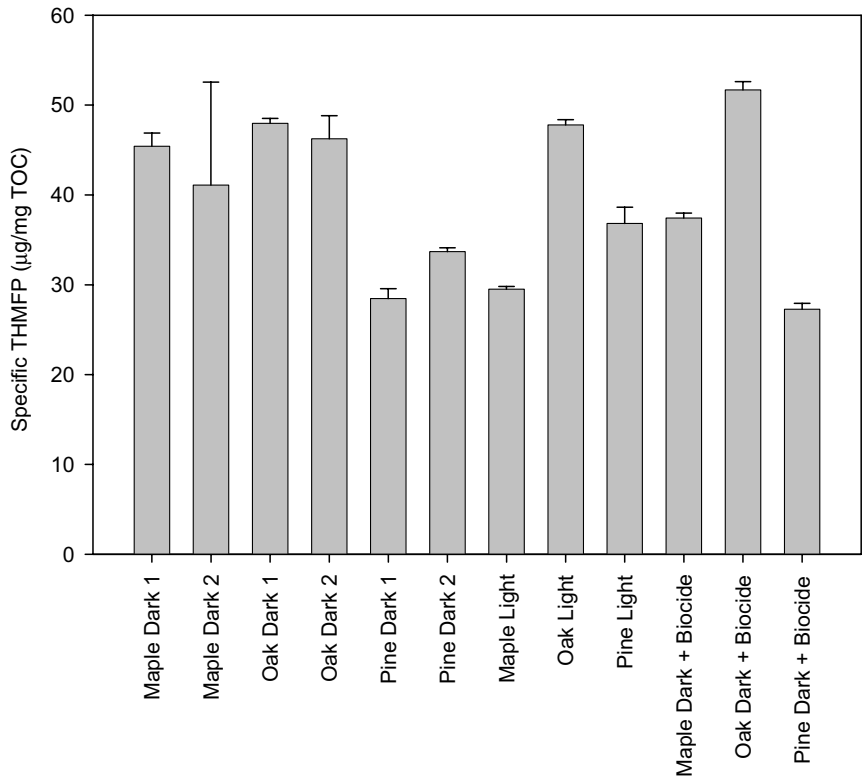



**Figure 6.11 Specific THM formation as a function of leaf type and leaching conditions**

of HAA/THM formation than is usually seen in surface waters. This may be interpreted as reflecting the aromatic nature of these materials (e.g., lignin and tannin structures). This same comparison has been made between humic acids and fulvic acids (Reckhow, Singer, and Malcolm 1990).

## Degradation

After 7 days, the leaves were removed from three of the leaching experiment dark bottles (one of each of the studied species). The bottles were held under aerobic conditions for 80 days (20°C, bubbling, with zero-grade air). During this period, the bottles were monitored for biodegradation. Despite its weathered condition, the material remaining in the bottles was quite biodegradable, as illustrated in Figure 6.13 for one of the monitored samples. Results for the remaining bottles were similar and are thus not presented. Although many weeks may be required for significant carbon loss in a batch suspended culture in the laboratory, indigenous benthic organisms in the stream are likely to accelerate this process. Following this period of biodegradation, samples were retested for TOC levels and DBP precursor content. Figure 6.14 shows the specific DBP precursor levels for the three samples before biodegradation.

The large samples from the leaching period continued on to the biodegradation period, where samples were allowed to sit for more than 81 days in aerated, constant-temperature conditions. During this time, samples were monitored by recording frequent UV absorbance measurements and less-frequent TOC measurements. The results of these measurements are presented in



©2007 AwwaRF. All Rights Reserved.

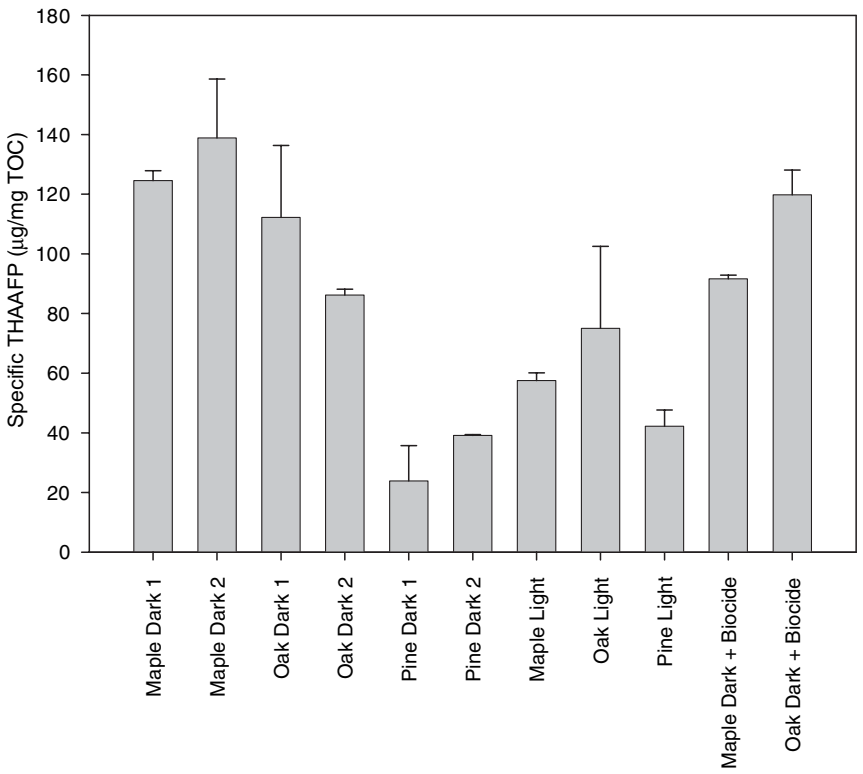



**Figure 6.12 Specific THAA formation as a function of leaf type and leaching conditions**

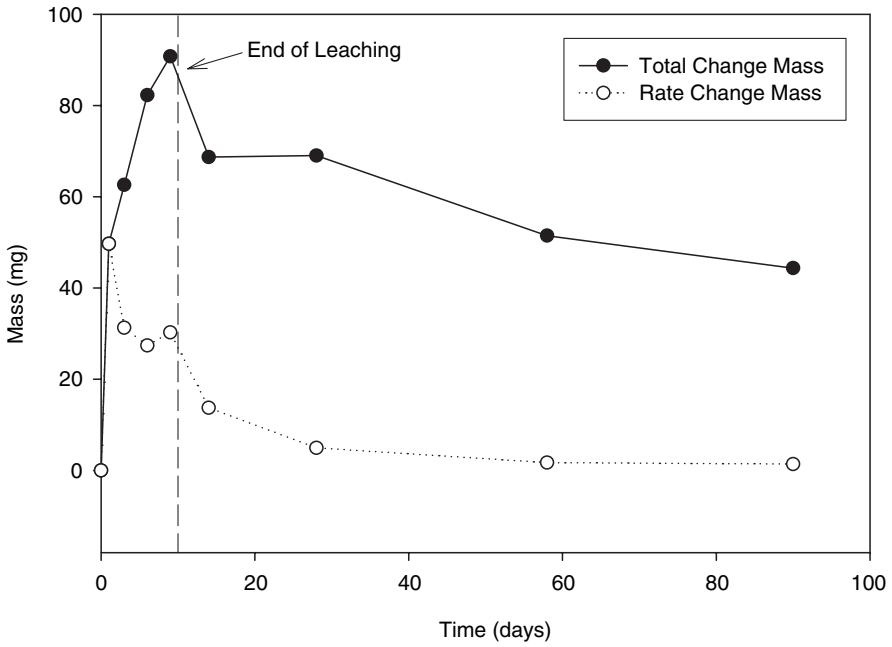



**Figure 6.13 Degradation of leached organic matter versus time (dark conditions)**



©2007 AwwaRF. All Rights Reserved.

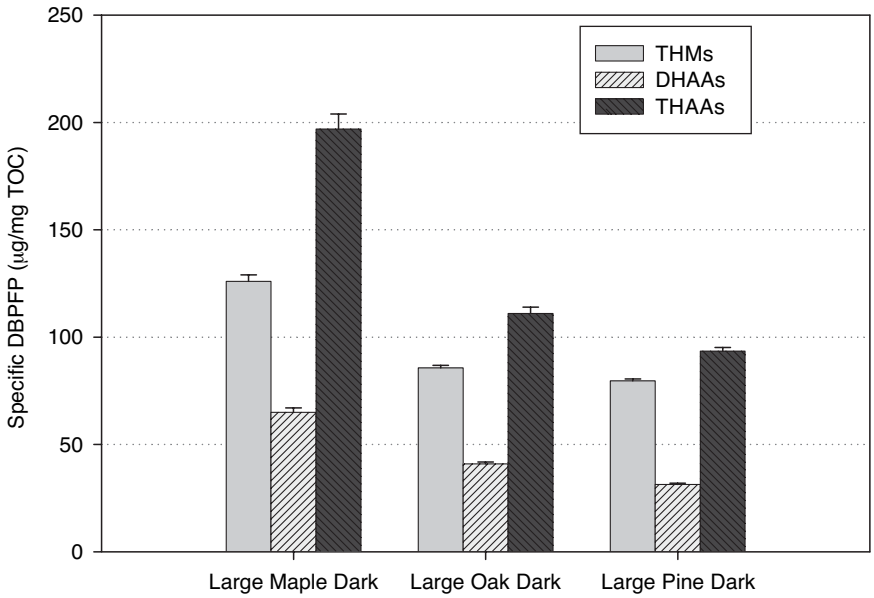



**Figure 6.14 Specific DBPFPs for last day of leaching**

**Table 6.8**
**Biodegradation experiment UV absorbance results for all samples***

| | $UV_{254}$ ($cm^{-1}$) | | |
|---|---|---|---|
| Day | Large maple dark | Large oak dark | Large pine dark |
| 0 | 1.566 | 0.903 | 0.203 |
| 5 | 1.590 | 0.936 | 0.181 |
| 11 | 1.608 | 0.954 | 0.198 |
| 12 | 1.767 | 0.960 | 0.214 |
| 13 | 1.797 | 0.970 | 0.219 |
| 15 | 1.974 | 1.060 | 0.239 |
| 19 | 1.971 | 1.044 | 0.238 |
| 26 | 2.127 | 1.106 | 0.245 |
| 43 | 2.283 | 1.192 | 0.286 |
| 49 | 2.508 | 1.240 | 0.324 |
| 81 | 4.315 | 1.733 | 0.878 |

*Samples were diluted before measurement with the spectrophotometer to ensure that no reading was greater than 1 $cm^{-1}$; values reported are results multiplied by the dilution factor.

Tables 6.8 and 6.9, respectively. At the end of the 81-day biodegradation period, final TOC measurements were taken and DBPFP analyses were performed.

UV absorbance during the 81-day biodegradation period for each sample is presented in Figure 6.15. The figure reveals that UV absorbance increased over time for each sample. The frequent removal of water for analysis, combined with evaporation, led to a large decrease in the



©2007 AwwaRF. All Rights Reserved.

**Table 6.9**
**Biodegradation experiment TOC results for all samples***

| | TOC (mg/L) | | | | | |
|---|---|---|---|---|---|---|
| Day | Large maple dark | SD† | Large oak dark | SD | Large pine dark | SD |
| 0 | 31.86 | 0.5 | 30.33 | 0.43 | 12.564 | 0.14 |
| 5 | 27.5 | 0.3 | 22.91 | 0.12 | 7.21 | 0.06 |
| 19 | 32 | 5 | 32.5 | 0.3 | 9.8 | 0.3 |
| 49 | 33.3 | 0.4 | 24.77 | 0.17 | 8.05 | 0.13 |
| 81 | 56.8 | 0.9 | 38.9 | 0.4 | 22.3 | 0.2 |

*Samples were diluted before measurement on the TOC machine to ensure readings within the prepared standard curve; values reported are results multiplied by the dilution factor.
†SD = standard deviation.

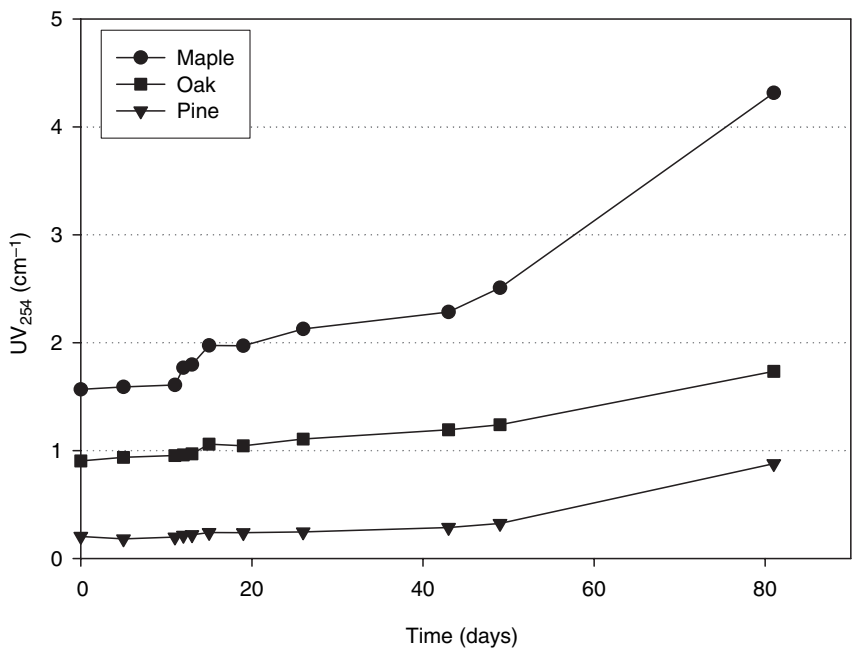



**Figure 6.15 UV absorbance over time for biodegradation of maple, oak, and pine leachate**

overall volume of water in all of the samples. During the biodegradation period, the samples lost approximately 1,700 mL of solution, or almost 70% of the starting volume. For this reason, changes in total carbon were evaluated in order to accurately depict the biodegradation of the organic material in the samples.

Figure 6.16 shows that total carbon levels decreased over time as water levels in the sample containers decreased because of evaporation. Total carbon levels in the maple samples decreased from 90 to 45 mg C, the oak samples fell from 86 to 30 mg C, and the pine samples decreased from 36 to 17 mg C.



©2007 AwwaRF. All Rights Reserved.

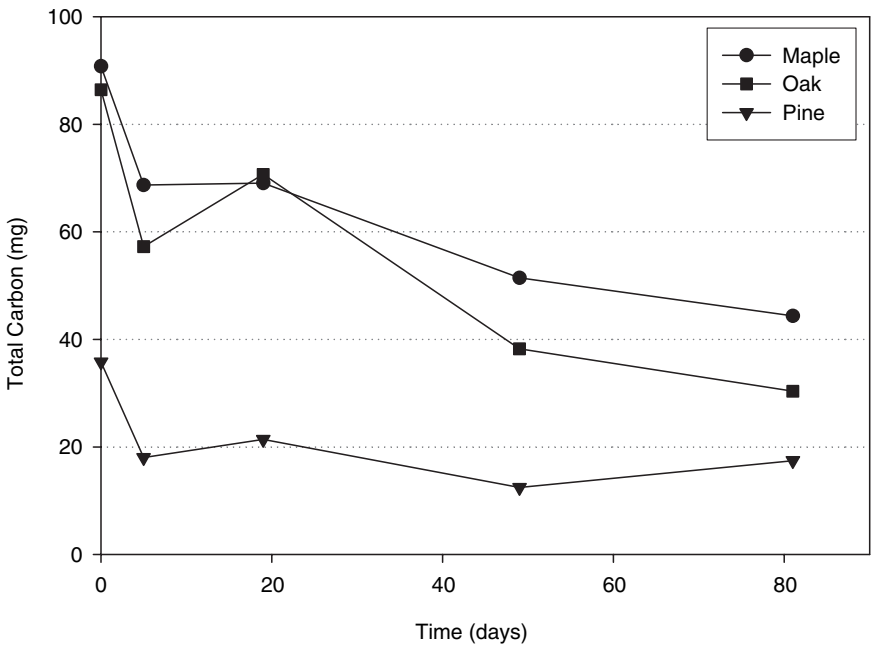



**Figure 6.16 Changes in total carbon over time for biodegradation of maple, oak, and pine leachate**

**Table 6.10**
**$UV_{254}$, TOC, and SUVA on final day of biodegradation for all samples**

| Sample name | $UV_{254}$ ($cm^{-1}$) | TOC (mg/L) | SUVA (L/mg-m) |
|---|---|---|---|
| Large maple dark | 4.31 | 56.78 | 7.59 |
| Lark oak dark | 1.73 | 38.88 | 4.45 |
| Large pine dark | 0.88 | 22.35 | 3.93 |

On the final day of the biodegradation experiment, the final TOC and DBPFP were measured. The results for TOC, UV absorbance, and SUVA on the final day of biodegradation are presented in Table 6.10. The final TOC levels for each of the samples was about 57, 39, and 22 mg/L for maple, oak, and pine, respectively. The final SUVA of the maple sample was very high, at 7.6 L/mg-m. The SUVA of the oak and pine samples on the final day of biodegradation were similar, at 4.45 and 3.93 L/mg-m, respectively.

The results of the DBPFP analysis are presented in Table 6.11 and Figure 6.17. DBPFP levels were extremely high for the maple sample, with THAAFP levels up to 11,200 µg/L. Oak DBPFP levels were almost half those of the maple samples, and the DBPFPs from the pine samples were almost half those of the oak samples. Overall, the THAAFP levels for each of the leaf types were the highest of the measured DBPs for all samples, followed by THMFP and then DHAAFP.



©2007 AwwaRF. All Rights Reserved.

**Table 6.11**
**DBPFP data for last day of biodegradation**

| Sample name | THM (µg/L) | SD* | DHAA (µg/L) | SD | THAA (µg/L) | SD |
|---|---|---|---|---|---|---|
| Large maple dark | 7,130 | 140 | 3,670 | 120 | 11,200 | 360 |
| Large oak dark | 3,330 | 30 | 1,590 | 30 | 4,310 | 120 |
| Large pine dark | 1,780 | 4 | 700 | 11 | 2,090 | 30 |

*SD = standard deviation.

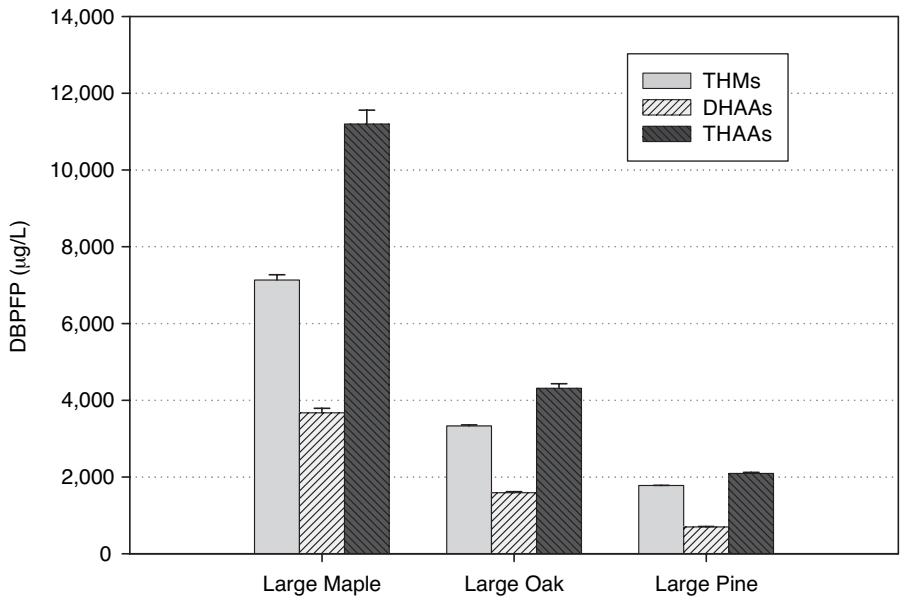



**Figure 6.17 DBPFPs for last day of biodegradation**

Specific DBPFP levels were also high for all biodegradation leaf types, as presented in Table 6.12 and Figure 6.14. As expected, specific THAAFP levels were the highest, with the maple sample forming almost 200 µg/mg C, and both oak and pine samples resulting in close to 100 µg/mg TOC. Specific THMFP levels were the second highest, with the maple sample resulting in about 120 µg/mg C, and both the oak and pine samples close to 80 µg/mg C. Specific DHAAFP levels were the lowest, with maple forming more than 60 µg/mg C, oak up to about 40 µg/mg C, and pine resulting in close to 30 µg/mg C.

Overall, specific DBPFP levels were very large for all of the biodegradation samples for all leaf types. In addition, although the oak sample had DBPFP levels twice those of the pine samples, the specific DBPFP levels were almost identical.

These data clearly show that biodegradation causes the DOC levels to drop (Figure 6.16 and other data not shown) but has little overall impact on the DBP precursor levels. The net result



©2007 AwwaRF. All Rights Reserved.

**Table 6.12**
**Specific DBPFP data for last day of biodegradation**

| Sample name | THM (µg/mg C) | SD* | DHAA (µg/mg C) | SD | THAA (µg/mg C) | SD |
|---|---|---|---|---|---|---|
| Large maple dark | 126 | 3 | 65 | 2 | 197 | 7 |
| Large oak dark | 85.6 | 1.2 | 40.9 | 0.9 | 111 | 3 |
| Large pine dark | 79.6 | 0.9 | 31.3 | 0.6 | 93.5 | 1.7 |

*SD = standard deviation.

**Table 6.13**
**Mass balance of leached material***

| Age | Phase | TOC (mg C/L) | C-Yield (mg C/g dry wt) | Net C rate (mg C/g dry wt/d) | Dissolved nitrogen (mg N/L) | N-Yield (mg N/g dry wt) | Net N rate (mg N/g dry wt/d) | C/N (g/g) | k ($d^{-1}$) |
|---|---|---|---|---|---|---|---|---|---|
| Old | 1 | 4.36 | 7.6 | 7.6 | 0.14 | 0.2 | 0.2 | 31 | 0.015 |
| | 2 | 6.13 | 10.7 | 5.4 | 0.14 | 0.2 | 0.1 | 88 | 0.011 |
| New | 1 | 6.72 | 23.5 | 23.5 | 0.25 | 0.9 | 0.9 | 27 | 0.047 |
| | 2 | 5.37 | 18.9 | 11.8 | 0.10 | 0.4 | 0.4 | 43 | 0.024 |
| | 3 | 7.94 | 27.9 | 7.8 | 0.13 | 0.5 | 0.3 | 95 | 0.016 |

*Total volume is 35 L in each case; k value determined from DOC release.

is an increase in the specific (DOC-normalized) precursor levels. These same phenomena may also occur in nature as fresh plant leachate is aged.

## RESULTS OF STUDY 2

### Leaching

The second set of leaching experiments was conducted so that the rate of decay of leaf mass could be estimated from the standpoint of DOC production (Table 6.13). This may be distinct from the rate based on measured loss of solid leaf mass. The final loss rate is greatest for the first flush from the new leaves, followed by the second flush. The rate for the first flush from the aged leaves was about equal to the third flush from the new ones. As expected, the last (second) flush of the old leaves was least productive.

### Degradation

Degradation occurred at a half-life of as little as 40 days for the first flush of NOM from the new leaves (Figure 6.18). Later flushes not only resulted in lower rates of leaching, but the



©2007 AwwaRF. All Rights Reserved.

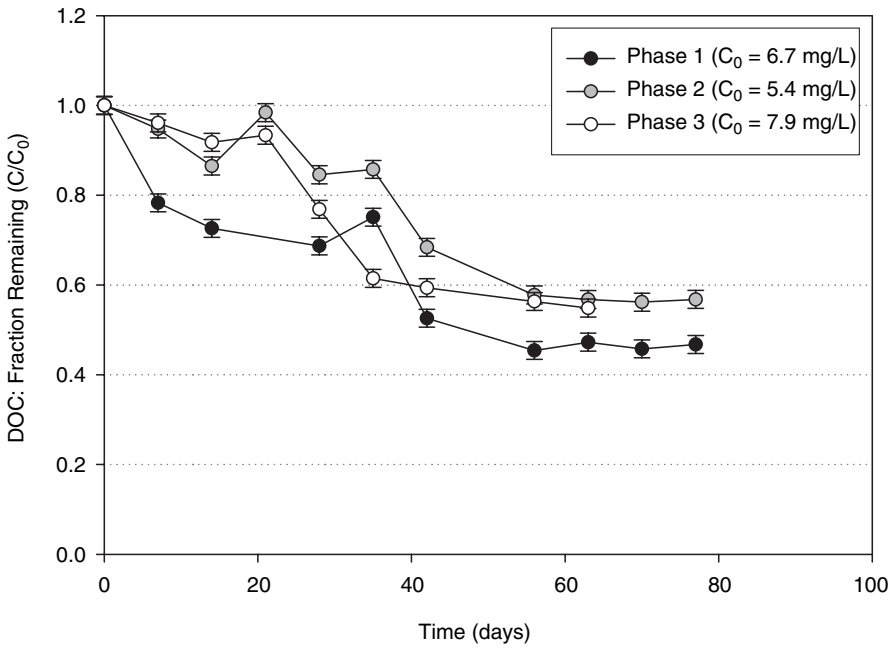



**Figure 6.18 Degradation of DOC for tests with new maple leaves (corrected for evaporative losses)**

leached material was slightly less biodegradable. After almost 80 days of incubation, about 55% of the DOC was degraded from the first flush, and about 45% from the second and third. In contrast, the old leaves only showed a 30% loss due to degradation at the longest incubation times (Figure 6.19).

***Biodegradable Dissolved Organic Carbon***

Figures 6.20 and 6.21 show the BDOC data from samples collected during the biodegradation period of study 2 for the new and old maple leaves, respectively. The x-axis indicates the time in days after sampling the biodegradation tank. All data are arithmetic averages of triplicates, and error bars reflect the standard deviation.

***New Maple Leaves.*** BDOC levels decreased slightly from day 1 to day 7 and then increased at day 14. This may be due to the breakdown of leached compounds into more bioavailable substrates between days 7 and 14. This pattern was observed for samples taken from phase 1 and phase 2. Phase 3, however, shows a threefold increase in BDOC on day 7. This increase may be due to the longer leaching process resulting in the breakdown of compounds at an earlier time point (day 10 after initial leaf immersion). All phases appear to reach similarly low levels by day 77. Overall, the greatest amount of BDOC is released by leaching during phase 1.

***Old Maple Leaves.*** The initial immersion of old maple leaves releases less BDOC into the surrounding water compared to the new leaves (day 1, phases 1 and 2). Although this low biodegradation of leachate remains true for phase 1 (apart from a late increase of BDOC at day 77), a substantial amount of BDOC is released with the phase 2 leachate at day 7 and day 14.



©2007 AwwaRF. All Rights Reserved.

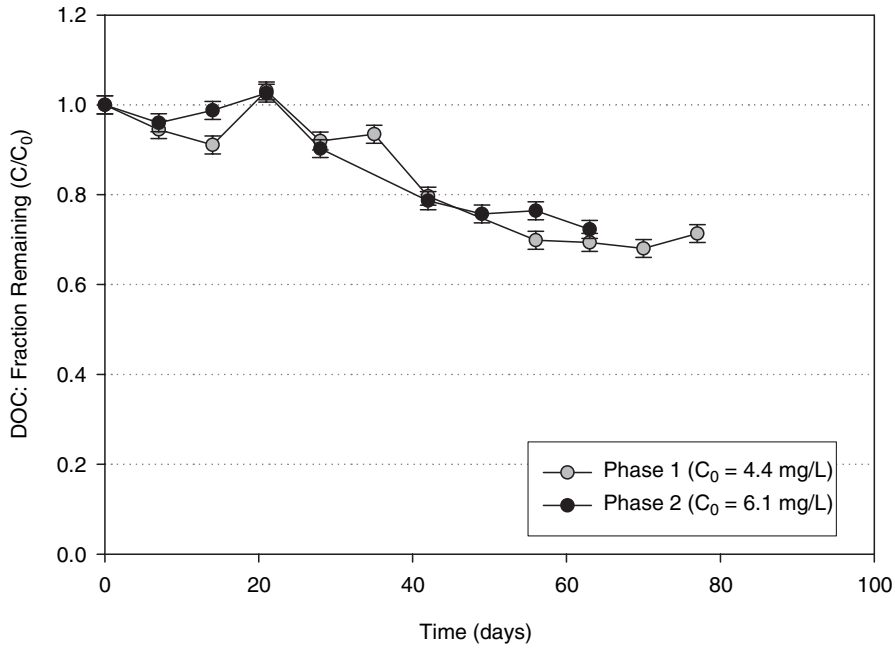



**Figure 6.19 Degradation of DOC for tests with old maple leaves (corrected for evaporative losses)**

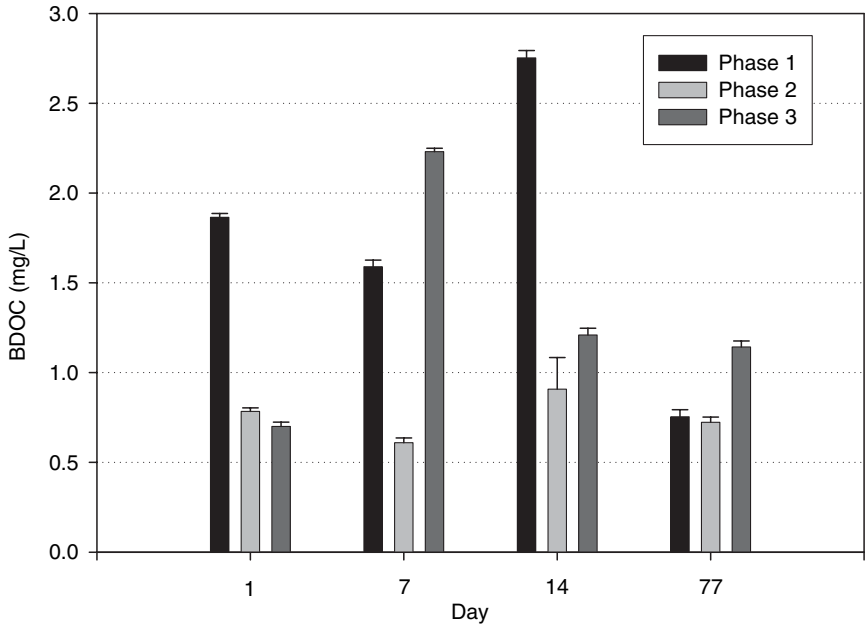



**Figure 6.20 BDOC results for leachate tests with new maple leaves**



©2007 AwwaRF. All Rights Reserved.

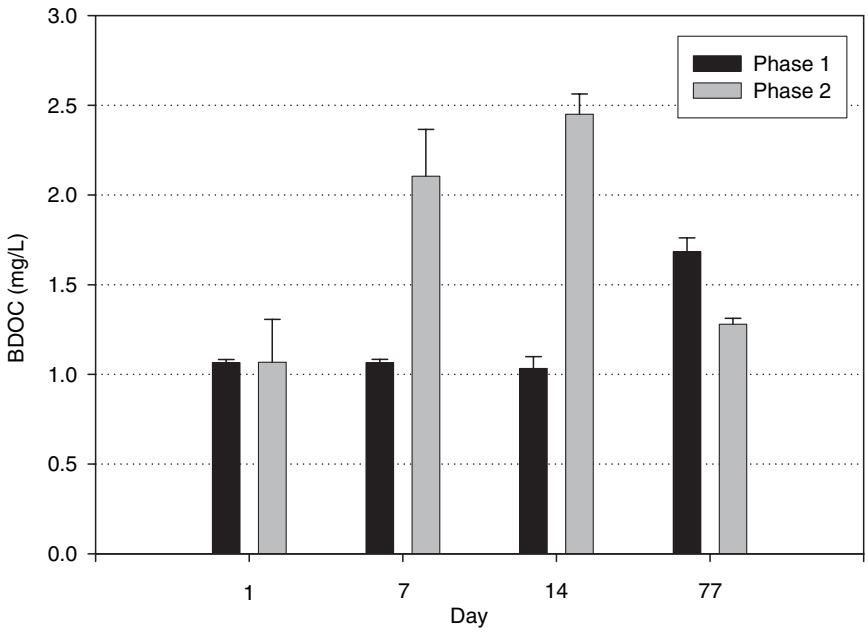



**Figure 6.21 BDOC results for leachate tests with old maple leaves**

This increased release in phase 2 may be due to the longer period of leaching during which the leaves were stirred, which allowed for potentially more leaf breakage followed by the release of plant organic carbon. One interpretation is that this organic carbon may have needed a week's time to break down into easily useable substrates. This pattern is similar to that observed for AOC. Again, because the leaching of NOM into surrounding waters took longer for old maple leaves than for new maple leaves, only two phases were tested for the old maple leaves (phase 1: 24 hours; phase 2: 72 hours).

### *Assimilable Organic Carbon*

Figures 6.22 and 6.23 show the AOC data from samples collected during the biodegradation period of study 2 for the new and old maple leaves, respectively. The x-axis indicates the time in days after sampling the biodegradation tank. All data are arithmetic averages of triplicates, and error bars reflect the standard deviation.

***New Maple Leaves.*** Three phases were tested for the new maple leaves.

**Phase 1:** The AOC for phase 1 leachates decreased more than fourfold from day 1 to day 7. AOC levels appeared to stay constant from day 7 to day 14, and decreased again approximately fourfold between day 14 and day 77. As speculated, the most bioavailable organic carbon is released from the new leaves during the first leaching cycle and is assimilated during the first 7 days of the experiment. A substantial amount of AOC is still available on day 7 and day 14. This bioavailable organic carbon may be present due to the slow breakdown of some larger compounds into more easily assimilable carbon. By day 77 there is still some AOC remaining, although it is substantially diminished.

**Phase 2:** There is a marked decrease of more than 90% in the AOC levels between phases 1 and 2. This indicates that most AOC was leached from the leaves during phase 1. The decrease



©2007 AwwaRF. All Rights Reserved.

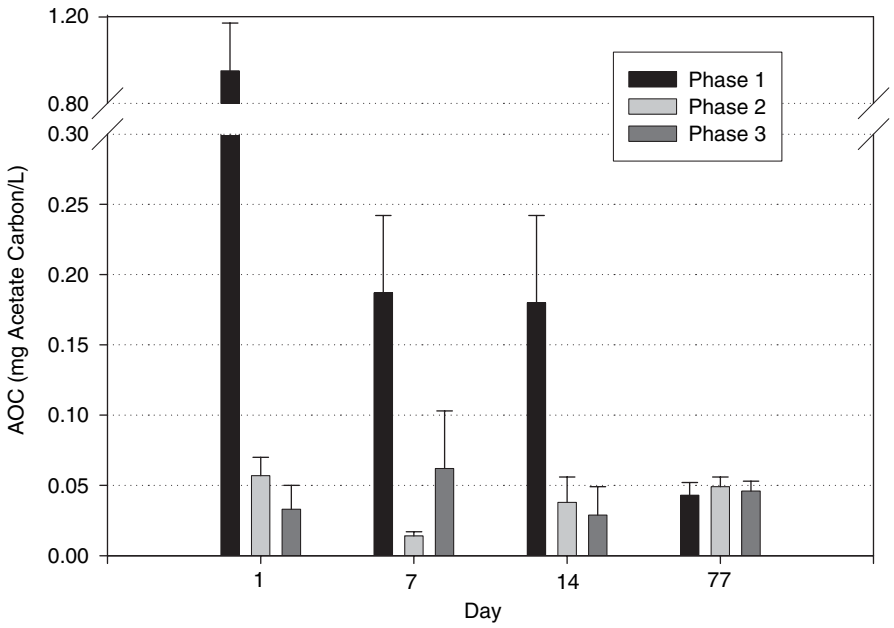



**Figure 6.22 AOC results for leachate tests with new maple leaves**

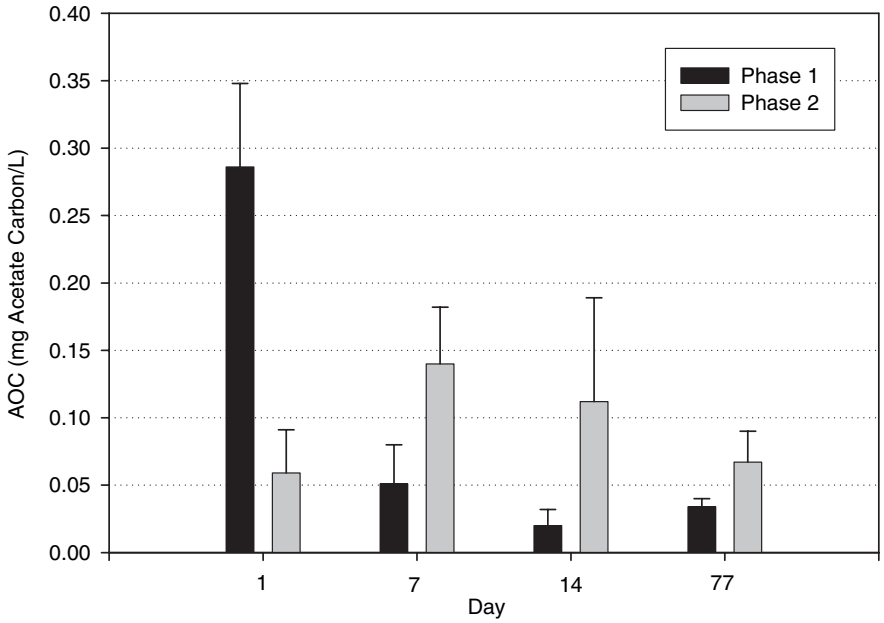



**Figure 6.23 AOC results for leachate tests with old maple leaves**



©2007 AwwaRF. All Rights Reserved.

in AOC from day 1 to day 7 and the subsequent increase at day 14 and day 77 could indicate that although most of the bioavailable organic carbon is available at the start of the biodegradation process, recalcitrant compounds are being broken down to more easily assimilable organic carbon during the course of the experiment.

**Phase 3:** Again, there is a marked decrease in the AOC levels between phases 1 and 3. The fluctuating levels of AOC between time points could indicate the different times in which recalcitrant compounds are being broken down to more easily assimilable organic carbon and/or the adaptation of the seed microorganisms to the available organic compounds.

***Old Maple Leaves.*** Two phases were tested for the old maple leaves.

**Phase 1:** Although there is a smaller amount of AOC leached from the old leaves than from the new leaves, again, the largest amount of AOC is leached at the beginning of the biodegradation process. The AOC levels decrease approximately fivefold from day 1 to day 7 and continue to decrease at day 14. There is a slight increase at day 77, indicating the potential of further breakdown of previously nonassimilable compounds, or the physiological adaptation of the microorganisms' capability for uptake and metabolic use of a particular leachate.

**Phase 2:** The low AOC value for day 1 may be due to experimental error. However, this would not explain the greater amount of AOC available at all other time points in phase 2 compared to phase 1. It is thought that the leaves in phase 2, which were leached for a longer period than those in phase 1, may release a greater amount of nonassimilable compounds that break down, leading to high AOC values or an improved adaptation of the inhabitant community.

## DISCUSSION

### Leaching of NOM

All three tree species used in study 1 released large amounts of organic matter over time, with the highest rate of release in the first day of contact. For the 8-day leaching period, maple leaves released 50% of the long-term (8-day) release in the first day, oak leached 30%, and pine leached 50%. Numerous other studies found similarly rapid carbon release (Thurman 1985, Meyer 1990, Speiran 2000, McArthur and Richardson 2002). Thurman (1985) found that, on average, 20%–40% of the maximum is released within the first 24 hours, much like the results in this study. The yields observed during this study were as high as 24 mg C/g dry weight/d. The highest levels were observed for fresh (green) leaves during the first day of leaching. Older leaves gave lower carbon yields, as did leaves that had previously been leached. Maple leaves showed the highest leaching yields, followed by oak and pine.

Average SUVA values for the maple, oak, and pine samples (4.5, 2.6, 1.3 L/mg-m, respectively) suggest that the maple leaves released mostly hydrophobic aquatic humics, whereas oak leaves leached a mixture of hydrophobic and hydrophilic organic matter and pine needles released primarily nonhumic materials. McArthur and Richardson (2002), working with five different tree species, concluded that primarily water-soluble compounds such as carbohydrates (low-MW nonhumics) and phenolics (aromatic humics) were released on contact with water. Among the tree species, deciduous species leached more than twice the amount of phenolics than coniferous species, resulting in waters with high concentrations of aquatic humics (McArthur and Richardson 2002). Hongve (1999) found that deciduous litter leached both hydrophobic and hydrophilic acids, whereas coniferous litter leached mostly hydrophobic



©2007 AwwaRF. All Rights Reserved.

compounds. These results indicate that the nature of organic matter originating in forested watersheds is highly dependent on the dominant tree species.

### Biodegradation of NOM

All leaf types gave rise to leachate that exhibited extensive biodegradation, and within the first 5 days during study 1, 25%, 34%, and 50% of TOC was lost in the maple, oak, and pine samples, respectively. Overall, 51%, 65%, and 51% of TOC was lost to biodegradation in the maple, oak, and pine samples, respectively. Study 2 also showed about a 50% loss in carbon from maple leaves. From both studies an average half-life for the maple-derived DOC would be about 60 days. Limited data from oak and pine leaf types suggest slightly faster rates. The rates seem to be inversely proportional to the SUVA. This seems to suggest that leachate with a higher aromatic content (lignin, tannin) will degrade more slowly than leachate with a higher aliphatic content. Even as the bulk sample DOC appears to level off (e.g., at 70 days), indications suggest that slow hydrolysis is possible with additional biodegradation (AOC and BDOC values are still substantially larger than zero).

Organic matter from leaves is rapidly digested by microorganisms (Thurman 1985). McArthur and Richardson (2002) observed that microorganism growth rates were the highest during the first hour of biodegradation. Hongve (1999) found that up to 45% of fresh leachate degrades in the first week, and 30% after 7 weeks. Meyer (1990) also recognized this, indicating that the low-MW compounds initially present in fresh leachate are easily digested by microorganisms, utilizing up to 54% of low-MW DOC within the first 3 days. In addition, 21% of the high-MW compounds were also degraded in 3 days. To the contrary, McArthur and Richardson (2002) found that the HMW fraction was more labile.

The organic matter leached from pine needles degraded almost two times more in the first 5 days in comparison to the maple leachate. Pine catchments were studied by Findlay et al. (2001), who found that pine needles leached substantial levels of aromatic compounds and hemicelluloses, both of which are readily (extremely) bioavailable to microorganisms. This may explain the rapid utilization of pine leachate in the first 5 days.

SUVA levels of the leachate after 81 days in the biodegradation phase were noticeably higher, with the maple, oak, and pine samples having SUVAs of about 7.6, 4.5, and 3.9, respectively. These values were two to three times higher than the SUVA levels of the fresh leachate. This indicates that there was not only a loss in TOC but also either an increase in the humic content of the samples or a decrease in the low-SUVA fraction, or both. Overall, at the end of the biodegradation phase, all three samples contained high-SUVA NOM, most likely in the form of recalcitrant aquatic humics.

### Specific DBP Formation Potentials

Specific DBPFP levels were highly variable among specimens for the fresh leachate. The maple samples had the highest specific DBPFP levels, followed by oak and pine. The pine samples had very low levels of all specific DBPFPs in comparison to the maple and oak samples. THAAFPs and specific THAAFPs were the highest of all DBPs, with specific THAAFPs being three and two times higher than specific THMFPs for the maple and oak samples, respectively. Reckhow, Singer, and Malcolm (1990) found that THAA levels often exceeded THM levels for waters with high SUVA values.



©2007 AwwaRF. All Rights Reserved.

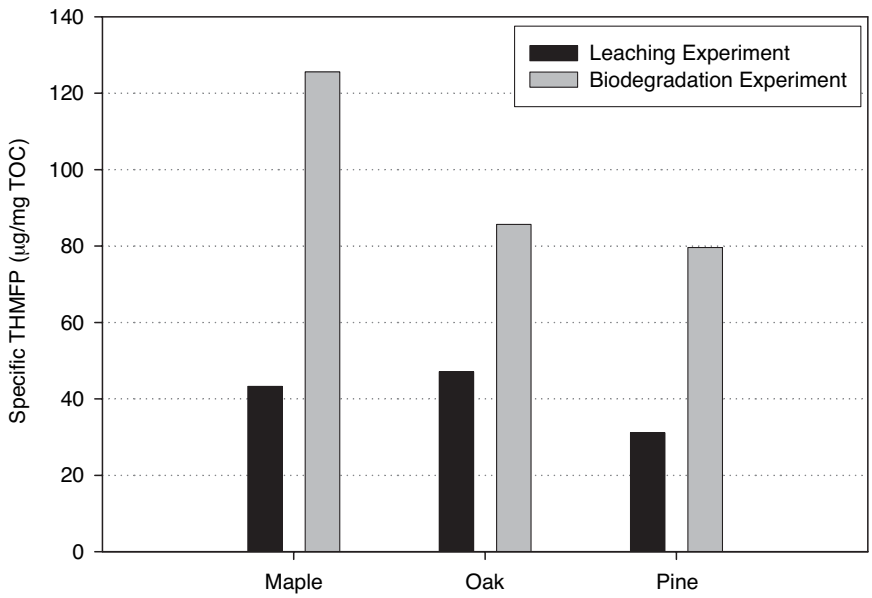



**Figure 6.24 Specific THMFPs for leached and biodegraded organic matter**

Specific DBPFP levels for all biodegradation samples were much higher. Both oak and pine had similar specific DBPFP levels. In Figures 6.24 through 6.26, the specific DBPFP levels are compared to those from the leaching experiment. Maple and pine had the largest increases in specific DBPFP, with levels one-and-a-half to three times greater after biodegradation than in the fresh leachate. The oak specific-THMFP levels increased by 45%; however, the specific DHAAFP and THAAFP levels remained relatively the same. This suggests that the NOM responsible for HAA formation in oak leaves may be highly resistant to biodegradation.

Overall, the 81-day biodegradation phase caused a large increase in SUVA along with the substantial loss of organic carbon mass. This suggests that large quantities of low-UV-absorbing material present in the fresh leachate was degraded to $CO_2$ or perhaps condensed to form highly recalcitrant humic-like material. This caused considerable increases in DBP and specific DBPFPs.

This translates to conditions in the environment, suggesting that if microorganisms in the environment are given enough time, they will remove the low-DBP-forming fraction first, resulting in the accumulation of highly reactive DBP precursor material. Subsurface soils and upstream wetlands are the primary locations where residence times are long enough for this to occur, leading to the accumulation of highly recalcitrant humic material in upstream wetlands and the upper soil horizons of forested areas.

## RELATIONSHIP TO NATURAL PROCESSES

Figure 6.27 shows THM precursor data from the Wachusett Reservoir catchment along with the leachate data for comparison. This figure exhibits a negative correlation between specific precursor content and TOC in the tributary samples. DBPFP for the leachate experiments (indicated with horizontal lines for the oak and maple sample specific THMFP, as their TOC levels were off-scale) are in the medium to low range. In this case the exact TOC level of the laboratory



©2007 AwwaRF. All Rights Reserved.

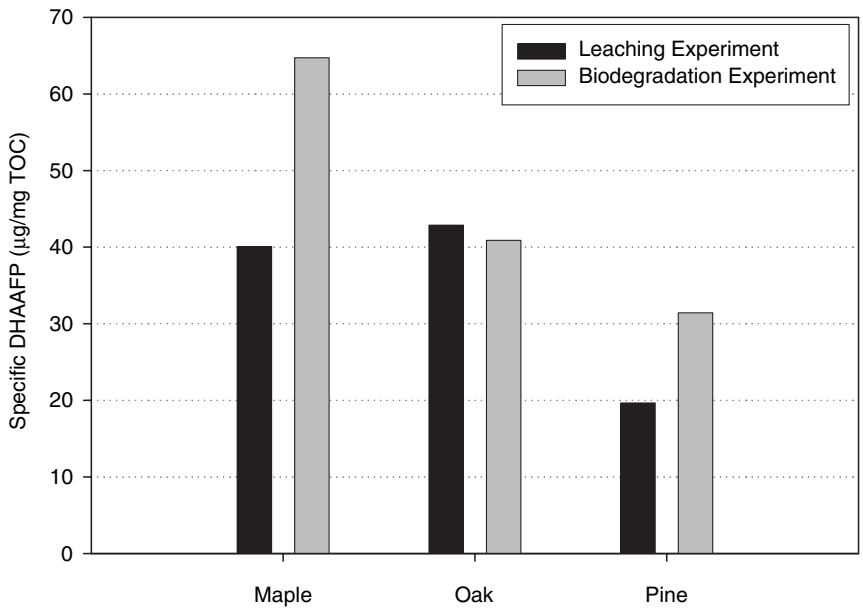



**Figure 6.25 DHAAFPs for leached and biodegraded organic matter**

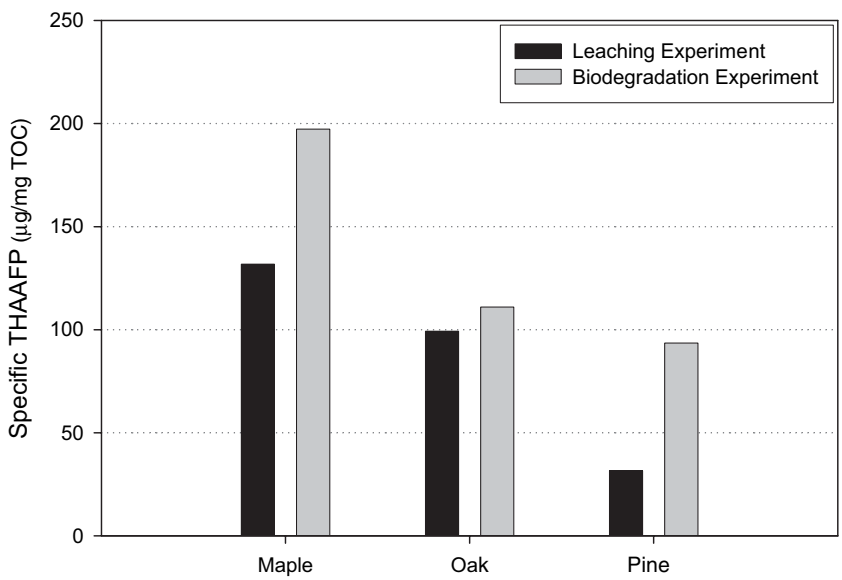



**Figure 6.26 Specific THAAFPs for leached and biodegraded organic matter**

leachate is probably not important, as it is largely determined by the amount of water and leaf matter that was selected for this experiment.

The data in Figures 6.24 to 6.26 suggest that freshly flushed NOM has a lower specific DBP content than aged NOM. This is the material that predominates during and after rain events when the streamflow and DOC levels are elevated. As the amount of runoff and interflow declines and the groundwater level drops, the stream chemistry resembles that of the base flow. This water



©2007 AwwaRF. All Rights Reserved.

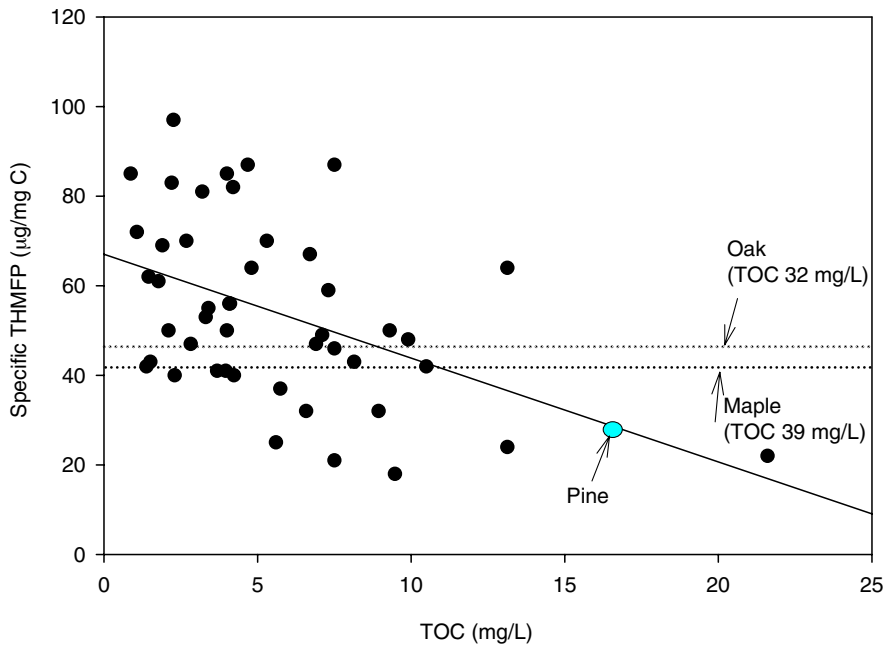



**Figure 6.27 Specific THM formation versus TOC for natural water samples and leaf leachate**

contains less organic matter due to soil loss processes, especially biodegradation. Therefore, as the flow and DOC decline during drier weather, the organic matter is expected to be more aged (i.e., subject to greater degrees of biodegradation). The correlation between flow and NOM biodegradability has been widely reported (e.g., Withers and Drikas 1998). Figure 6.27 suggests that biodegradation preferentially results in loss of nonprecursor material, thereby increasing the specific precursor levels.

There is some support for differential biodegradation of DBP precursors in the soil ecology and water chemistry literature. It is well accepted that polysaccharides are among the most biodegradable and labile of the soil organic matter fractions. Celluloses and hemicelluloses are also biodegradable, although at a slower rate. These compounds are not very reactive with chlorine, and they do not produce large amounts of DBPs on a per-carbon basis (e.g., Hoehn et al. 1980). In contrast, the humic substances, lignin residues, and plant-based phenolics are far less biodegradable and may accumulate during the short term (e.g., Kalbitz et al. 1970). These structures are quite reactive with chlorine because of their activating groups and high electron density in the aromatic rings. As a result, they produce large amounts of DBPs following chlorination. Yields from model compound studies have been as high as 1 mol/mol or 1,600 µg THM/mg C (Larson and Rockwell 1979, Boyce and Hornig 1983, Reckhow and Singer 1985, De Laat 1981).

As NOM continues to age, one would expect that the increase in specific DBP precursor level could reverse itself. There is little doubt that, given enough time, the most reactive compounds (e.g., the activated aromatics) will be lost from the water column. This could be through slow biodegradation, adsorption to settleable particles, photolysis, and so forth. This is probably why oligotrophic reservoirs with long residence times have lower specific THM precursor levels than their contributing tributaries (e.g., Garvey and Tobiason 2003b). This



©2007 AwwaRF. All Rights Reserved.

general phenomenon has also been observed in groundwater recharge systems. Aiken and Leenheer (2004) reported that at very short distances the specific THMFP increased, whereas for longer travel distances the level drops substantially.

## SUMMARY AND CONCLUSIONS

Two sets of long-term experiments were conducted, both intended for quantifying the generation of DOM from leaf litter and following its subsequent behavior, especially pertaining to biodegradation and precursor content. In both cases, leaves were collected and subjected to leaching tests of about 1 week in duration. Then the leachate was seeded and allowed to degrade under aerobic conditions. Samples were periodically removed for analysis of DOC, UV absorbance, biodegradability, and DBP precursor content.

The following results were obtained from the laboratory leaching and biodegradation of NOM from leaf litter:

- Significant concentrations of DOC were leached from maple, oak, and pine samples within the first 24 hours. SUVA levels varied considerably among species for the initial leaching phase, suggesting that the nature of organic matter in forested watersheds is dependent on the dominant tree species.
- Leaching rates ranged from near zero to 24 mg C/g dry weight/d. Maple leaves had the highest yield, and newer leaves produced more DOC than older or previously weathered leaves.
- Large amounts of biodegradation occurred in the first 5 days of postleaching aeration. Organic matter that had leached from pine needles degraded almost two times more quickly than the other species. SUVA levels increased two to three times more than the original level after 81 days of biodegradation for all species.
- High levels of specific DBPFPs were formed for both the fresh and biodegraded organic matter. On average, the biodegraded organic matter formed one-and-a-half to three times higher specific DBP levels than the fresh organic matter. Overall, specific THAAFPs were the highest, and were almost two to three times higher than specific THMFPs.
- AOC and BDOC tests indicate reduced values of biologically useable leachate from old leaves compared with new leaves. Although initial leaching tests offered the highest amount of assimilable carbon to the microorganisms, prolonged immersion of leaf litter produced more biodegradable DOC than short-term immersion.



©2007 AwwaRF. All Rights Reserved.

©2007 AwwaRF. All Rights Reserved.

# CHAPTER 7
# ROLE OF LIGNIN

## INTRODUCTION

A contributor to aquatic NOM, lignin is derived from trees, plants, and agricultural crops. Because of its natural abundance, wide distribution, and resistance to microbial degradation, lignin has been viewed as a useful model for NOM (Lobbes, Fitznar, and Kattner 1999). Lignin is also commonly found in aquatic environments. It is a complex natural polymer with seemingly random couplings, and the exact chemical structure is not known. It is generally accepted that lignin is composed of alkylmethoxy phenols with carbon/carbon and carbon/oxygen linkages between the monomeric units. This irregular structure makes it difficult to biodegrade the polymer (Sakakibara and Sano 2001).

Lignin content has been used as a tracer for terrestrial NOM in aquatic environments because the composition of its phenolic groups is characteristic of the types of vascular plant tissues from which they derive (Fengel and Wegener 1984). Phenolic aldehydes, ketones, and acids of the para-hydroxy, vanillyl, and syringyl groups and two cinnamyl acids are released by the alkaline cupric oxide (CuO) oxidation of lignin-containing NOM samples (Hedges and Ertel 1982, Goni and Hedges 1992, Farella et al. 2001). Different ratios of phenolic monomers are used to estimate sources of plant tissues, soil and sediment, and DOM. The ratio of syringyl to vanillyl phenols (S/V) provides information about the contribution of angiosperm versus gymnosperm plant tissues, whereas the ratio of cinnamyl to vanillyl phenols (C/V) reflects the proportion of woody versus nonwoody materials. Higher acid-to-aldehyde ratios are found in samples with more degraded lignin and have been used to reconstruct the diagenetic history of sample material (Goni, Ruttenberg, and Eglinton 1998).

The alkaline CuO oxidation method is probably the most commonly used for analyzing the composition of lignins in complex sample matrixes. It was first developed by Hedges and Ertel using gas chromatography (GC) for separation and quantification (Hedges and Ertel 1982; Hyotylainen, Knuutinen, and Vilen 1995; Louchouarn, Opsahl, and Benner 2000). More recently, Lobbes used the same degradation method but separated the lignin phenols using HPLC as well as GC (Lobbes, Fitznar, and Kattner 1999). HPLC detection of lignin phenols is generally performed by using only a single absorption wavelength. Goni and Montgomery developed this method further using a microwave digestion system and modified extraction procedure (Goni and Montgomery 2000). The HPLC method was used for identification and quantification of the lignin phenols, in conjunction with multiwavelength detection. This method enables the reliable determination of composition and quantity of small amounts of lignin in water samples (Charriere, Gadel, and Serve 1991; Hyotylainen et al. 1998; Da Cunha et al. 2001).

Lignin structures in NOM have often been associated with the tendency to form halogenated DBPs (e.g., THMs, HAAs) following chlorination. This association has been widely advanced because of the ubiquitous nature of lignin in the terrestrial environment, its recalcitrant nature, and its activated aromatic structures that are known to readily react with chlorine. Despite all of this circumstantial evidence, few researchers, if any, have established a causal link between DBP formation in water supplies and the presence of lignin-based precursors. In this work, baseline information is presented that leads to a direct quantitative assessment of the linkage between lignin content and DBP formation.



©2007 AwwaRF. All Rights Reserved.

**Table 7.1**
**Characterization of key lignin monomers**

| Elution order | Lignin phenols | Abbreviation | Retention time (minutes) | Molecular weight (AMU) | CAS number* |
|---|---|---|---|---|---|
| 1 | 4-Hydroxybenzoic acid | PAD | 17.71 | 138.12 | 99-96-7 |
| 2 | Vanillic acid | VAD | 21.21 | 168.15 | 121-34-6 |
| 3 | 4-Hydroxybenzaldehyde | PAL | 22.40 | 122.12 | 123-08-0 |
| 4 | Syringic acid | SAD | 23.99 | 198.18 | 530-57-4 |
| 5 | Vanillin | VN | 28.22 | 152.15 | 121-33-5 |
| 6 | 4-Hydroxyacetophenone | PON | 29.04 | 136.15 | 99-93-4 |
| 7 | Syringaldehyde | SN | 33.30 | 182.18 | 134-96-3 |
| 8 | Acetovanillone | AcV | 36.03 | 166.18 | 498-02-2 |
| 9 | *p*-Coumaric acid | CAD | 39.34 | 164.16 | 501-98-4 |
| 10 | Acetosyringone | AcS | 40.65 | 196.20 | 2478-38-8 |
| 11 | Ferulic acid | FAD | 44.43 | 194.19 | 537-98-4 |
| 12 | Ethyl vanillin | EVA | 46.43 | 166.18 | 121-32-4 |
| 13 | Cinnamic acid | CIAD | 57.10 | 148.16 | 140-10-3 |

*CAS = Chemical Abstracts Service (a division of the American Chemical Society).

## EXPERIMENTAL METHODS

### Chemicals and Standards

Table 7.1 and Figure 7.1 show the characteristics and structures of the major products of alkaline CuO degradation of lignin. Phenolic compounds, copper(II) oxide powder, and NaOH were all of HPLC grade or reagent grade. DI and carbon-filtered water (Millipore Super-Q system) was used for preparing solutions and for rinsing glassware.

Two commercial lignin products were also acquired in purified form. The organosolve lignin is largely based on hardwoods and is extracted using ethanol, without the assistance of strong acids or bases. It is smaller in average molecular weight (both weight averaged [MW] and number averaged [MN], as shown in Table 7.2) than the alkali product. This second type is produced from a gymnosperm using high-pH digestion.

Stock solutions of each lignin phenol were prepared in water and partitioned in 18-mL polyethylene vials (poly-Q vial; Beckman, Fullerton, Calif.). Internal standard solutions containing 5.49 m*M* of ethyl vanillin and 6.75 m*M* of *trans*-cinnamic acid were prepared in water (20 mg of compound in 20 mL of water, each). External working standards were prepared for quantitative analysis (10, 50, 250, 500, 1,000, 5,000, 10,000, and 25,000 n*M*) and discarded after 24 hours. Twenty microliters of each internal standard was added to the samples without further dilution.



©2007 AwwaRF. All Rights Reserved.

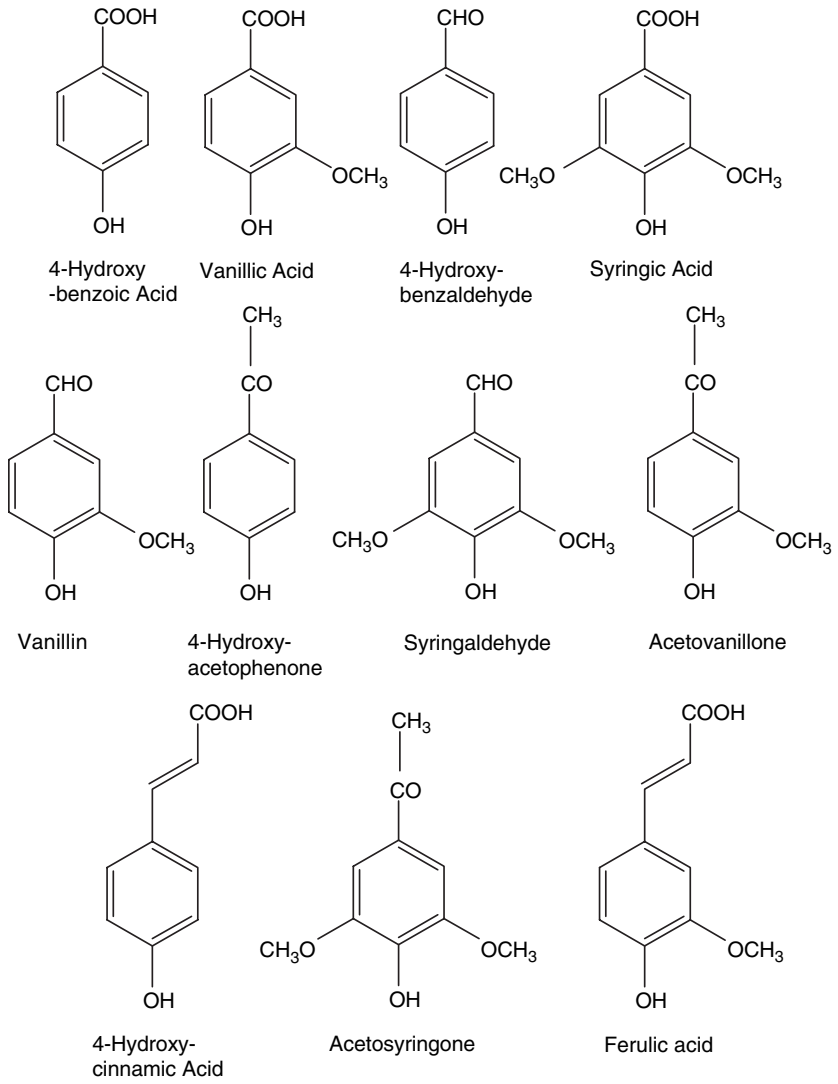



**Figure 7.1 Structures identified from CuO degradation of NOM**

**Table 7.2**
**Characteristics of commercial lignins**

| Type | MW | MN | Source |
|---|---|---|---|
| Organosolve | 3,500 | 800 | 50% maple, 35% birch, 15% poplar |
| Alkali | 14,200 | 1,750 | 100% Norway spruce |



©2007 AwwaRF. All Rights Reserved.

**Table 7.3**
**HPLC gradient program***

| Time (minutes) | % Eluent A | % Eluent B | % Eluent C |
|---|---|---|---|
| 0 | 90 | 0 | 10 |
| 5 | 78 | 0 | 22 |
| 35 | 70 | 0 | 30 |
| 45 | 55 | 0 | 45 |
| 55 | 0 | 0 | 100 |
| 58 | 0 | 100 | 0 |
| 60 | 90 | 0 | 10 |

*Flow rate of the gradient was 0.8 mL/min and the injection (100 µL) was made automatically by an autosampler. Mobile eluent A consisted of 7.4 m*M* phosphoric acid, eluent B was pure acetonitrile, and eluent C was composed of 7.4 m*M* phosphoric acid/methanol/acetonitrile (4:3:3 by volume).

## Oxidation and Extraction

The CuO oxidation was carried out using a slightly modified method of Hedges and Ertel (1982). The sample, which contained 2 to 5 mg C, 2 g CuO, and 0.2 g $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$, was put into a PTFE bomb (45 mL; Parr Co., Moline, Ill.). Under a $N_2$ atmosphere in a glove box, 15 mL of 2*M* NaOH solution (degassed by bubbling $N_2$ overnight) was added to the sample, and the sample was closed with PTFE caps. PTFE bombs were placed in a fitted hole of a stainless-steel bomb. The bombs were oxidized in an oven for 3 hours at 180°C. After cooling, 20 µL of each internal standard solution, ethyl vanillin, and cinnamic acid were added. Then the entire content of the bomb was centrifuged at 3,000 rpm for 10 minutes to separate the solids (sample + CuO) from the hydrolysate. The supernatants were decanted into centrifuge tubes fitted with PTFE-coated caps. This step was repeated after the addition of 5 mL of 1*N* NaOH to each centrifuge tube to ensure efficient transfer of the supernatant. The alkaline solutions were acidified to pH 1 by the addition of ~1 mL of concentrated HCl to each hydrolysate.

## HPLC Analysis

Chromatography and detection of the lignin phenols were conducted using a Waters Alliance 2890 HPLC system (Waters Corp., Millford, Mass.) consisting of an autosampler and a photodiode array detector. A Discovery C18 column (5 µm particle diameter, 250 mm × 4 mm; Supelco Co.) with a guard column was used for separation.

For elution, a gradient program was used (see Table 7.3). This method was modified from Da Cunha et al. (2001) and Charriere, Gadel, and Serve (1991).



©2007 AwwaRF. All Rights Reserved.

### Detection and Quantification

The lignin phenols were identified with the diode array detector using retention times and light absorption characteristics from external standard runs. The spectra of each monomer were measured in the wavelength range from 230 to 360 nm. To simplify the computation of the chromatograms, all peaks were integrated at a wavelength of 275 nm where all phenols showed good absorption. This simplification of using a single wavelength allows the integration of all peak areas from one record. The concentrations of the lignin phenols were calculated by calibrating the peak areas of the external standards from parallel runs. Final concentrations were then calculated considering the recovery of the internal standards, which were routinely added.

### Chlorination

To adjust pH, phosphate buffer solution was added to the samples and then final pH 7 adjustments were made with 1*M* $H_2SO_4$ (sulfuric acid) or NaOH. All samples were dosed with 20 mg $Cl_2$/L and incubated headspace-free for 2 days at 20°C. After the reaction period, chlorine residual concentrations were measured using the DPD ferrous titrimetric method as described in *Standard Methods* (APHA, AWWA, and WEF 1998). Samples showing little or no chlorine residual were retested in a more dilute form. After collecting the samples for DBP analysis, quench reagents were added to the bottles and the samples were stored at 4°C.

### Analysis of DBPs and TOX

THMs, HAAs, HANs, haloketones (HKs), and chloropicrin were analyzed after chlorination for 2 days of incubation. THMs and other neutral extractables (HANs, HKs, chloropicrin, etc.) were measured on all disinfected samples and controls with an HP 5890 series II GC/ECD (gas chromatography–electron capture detector). The full suite of HAAs were extracted by the acidified liquid–liquid extraction method and analyzed on an identical gas chromatograph.

TOX was determined by an adsorption-pyrolysis-titrimetric method based on Standard Method 5320 B (APHA, AWWA, and WEF 1998) with minor modifications. A Euroglas ECS 1200 TOX analyzer was used for pyrolysis and titration following adsorption by the EFU 1700 filtration unit (Euroglas).

## RESULTS AND DISCUSSION

### Lignin Phenolic Yields

The observed lignin parameters based on specific phenolic CuO oxidation products highlight compositional differences in organic matter. This provides a means of discriminating between taxonomic vascular plant groups (gymnosperms vs. angiosperms), tissue types (soft tissues vs. woody tissues), and diagenetic state or alteration of the original lignin material (Hedges et al. 1988; Farella et al. 2001; Goni, Teixeira, and Perkey 2003).

Organosolve lignin and alkali lignin are commercial materials prepared by industrial treatment of various wood products. The lignins were analyzed for lignin monomers following degradation using alkali lignin CuO oxidation. The overall recovery of lignin monomers for this method was 20% for organosolve lignin, based on the carbon content of the identified lignin



©2007 AwwaRF. All Rights Reserved.

monomers and the analytically determined starting carbon content of the commercial product. Table 7.4 shows lignin monomers from the two commercial lignin products. Vanillyl phenols are present in all types of vascular plant tissues, whereas syringyl phenols are only incorporated in significant amounts in angiosperm lignins (i.e., hardwoods). Cinnamyl phenols are thought to be restricted to nonwoody lignin sources. The *p*-hydroxy phenols are more wide-ranging plant products; however, indications suggest that the ketone form (*p*-OH acetophenone) is a unique lignin product (Ertel, Caine, and Thurman 1993). Vanillyl phenols in the degraded alkali lignin and organosolve lignin were measured at 8.5 mg/100 mg C and 10.4 mg/100 mg C, respectively. Syringyl phenols from organosolve lignin were 14.7 mg/100 mg C, whereas syringyl phenols were not detected in alkali lignin. The lack of syringyl phenols in alkali lignin reflects the origin of this material (100% softwood) and is entirely expected.

The S/V ratio is appreciably greater than zero in complex environmental mixtures and is usually indicative of the presence of at least some angiosperm tissue (Goni and Hedges 1992). This was reflected in the S/V ratio of organosolve lignin, which was calculated as 2.18.

To monitor the degradation state of original vascular plant materials, the acid-to-aldehyde ratio of vanillyl phenols ($[Ac/Al]_v$) is commonly used because vanillic acid is known to become more abundant during fungal and microbial degradation of lignin. Thus, the $(Ac/Al)_v$ ratio of aged lignin-based NOM should be substantially elevated above the range for fresh plant tissues. The $(Ac/Al)_v$ ratios of organosolve and alkali lignins was 0.17 and 0.32. The vanillin concentrations for the humic and fulvic acid fractions were too small to allow calculation of this value.

### DBP Yields of Lignin Monomers

DBP yields of lignin monomers are shown in Table 7.5. In this case, the overall vanillyl and syringyl values were calculated averages for the three compounds within each group (i.e., the aldehyde, ketone, and carboxylic acid). The THM yield was always higher for the ketone forms in each group. The HAA6 (sum of six HAAs) yields (0.012–0.076 m*M*/m*M* C) were generally higher than the THM yields (0.001–0.031 m*M*/m*M* C) in lignin monomers. This is especially clear when comparing the averages of the THAAs and THMs (0.036 vs. 0.007), and it reflects the bias toward THAA formation among lignin precursors. Aldehyde forms of the lignin monomers produced higher HAA yields (0.040–0.059 m*M*/m*M* C) than other lignin monomers in each phenolic group. HKs were detected only for the ketones, acetovanillone and acetosyringone. Because none of the lignin monomers contained nitrogen, HANs and chloropicrin were not detected.

The S/V ratio of THM yields was 1.62, which was much higher than the corresponding C/V ratio (0.29). Thus, angiosperm tissue is expected to be richer in THM precursors than other woody and nonwoody materials. The S/V and C/V ratios for THAA yields were similar. HKs also showed the highest S/V ratio (2.8), indicating that angiosperm tissue is an especially rich source for these DBP precursors. From the analysis of TOX and unknown TOX, it was noted that the S/V and C/V ratios were almost identical for the former, whereas the C/V ratio was much higher for the latter. This means that organic matter produced from angiosperm tissue may lead to lower yields of unknown TOX.

In separate experiments, the commercial lignin products (alkali and organosolve) were treated with chlorine and analyzed for DBPs. Table 7.5 shows these data expressed in the same units of millimoles of DBP per millimole of organic carbon in the starting material. Clearly, the commercial lignin products produce, on average, less DBPs on a per-carbon basis than the lignin monomers. However, a more appropriate comparison should take into account the natural abundance of the various monomers in each lignin material.



©2007 AwwaRF. All Rights Reserved.

**Table 7.4**
**Phenolic composition of lignin (mg compound/100 mg C)**

| Monomer | Abbreviation | Organosolve lignin | Alkali lignin | Humic acid | Fulvic acid |
|---|---|---|---|---|---|
| Vanillic acid | VAD | 0.51 | 1.37 | 0 | 0.51 |
| Vanillin | VAL | 5.57 | 6.36 | 0 | 0 |
| Acetovanillone | VON | 2.43 | 2.71 | 0.37 | 0.43 |
| Total V* | | 8.51 | 10.44 | 0.37 | 0.94 |
| Syringic acid | SAD | 0.11 | <0.1 | 3.4 | 4.15 |
| Syringaldehyde | SAL | 8.59 | <0.1 | <0.1 | <0.1 |
| Acetosyringone | SON | 6.01 | <0.1 | 1.56 | <0.1 |
| Total S† | | 14.7 | <0.1 | 4.96 | 4.15 |
| *p*-Coumaric acid | CAD | 1.97 | <0.1 | 0 | <0.1 |
| Ferulic acid | FAD | 6.01 | <0.1 | 1.56 | <0.1 |
| Total C‡ | | 1.97 | <0.1 | 0.4 | <0.1 |
| *p*-Hydroxybenzoic acid | PAD | <0.1 | 0.02 | 0 | 0.49 |
| *p*-Hydroxybenzaldehyde | PAL | 0.68 | 0.05 | 0.04 | <0.1 |
| *p*-Hydroxyacetophenone | PON | 0.65 | 0.71 | 0.89 | <0.1 |
| Total H§ | | 1.33 | 0.78 | 0.93 | 0.49 |
| Total lignin phenols** | | 26.52 | 11.22 | 6.66 | 5.59 |
| C/V | | 0.29 | 0 | 1.1 | 0 |
| S/V | | 2.18 | 0 | 13.49 | 4.4 |
| $(Ac/Al)_v$†† | | 0.17 | 0.32 | NA‡‡ | NA |
| Total lignin carbon§§ | | 20.15 | 7.10 | 4.67 | 3.13 |

*V: Vanillyl phenols (vanillic acid + vanillin + acetovanillone).
†S: Syringyl phenols (syringic acid + syringaldehyde + acetosyringone).
‡C: Cinnamyl phenols (*p*-coumaric acid + ferulic acid).
§H: *p*-Hydroxyl phenols (*p*-hydroxybenzoic acid + *p*-hydroxybenzaldehyde + *p*-hydroxyacetophenone).
**Total lignin phenols: (sum of V + S + C + H).
††$(Ac/Al)_v$: Ratio of vanillic acid to vanillin.
‡‡NA = Not applicable; cannot be calculated because of low vanillin level.
§§Carbon content of total lignin phenols.

## Estimated DBP Yield

The DBP yield of lignin polymers, as an aggregate of the measured monomers, could be estimated from the observed abundances of the monomers and the specific DBP yields for those monomers.

$$\left[\frac{\mathrm{m}M}{\mathrm{m}M \cdot \mathrm{C}}\right]_{\text{polymer}} = \sum^{\text{monomers}} \left[\frac{\mathrm{m}M_{\text{monomers}}}{\mathrm{m}M \cdot \mathrm{C}} \times \frac{\mathrm{m}M_{\text{DBPs}}}{M_{\text{monomers}}}\right]$$



©2007 AwwaRF. All Rights Reserved.

**Table 7.5**
**DBP yields of lignin monomers following chlorination (m*M* DBP/m*M*·C)**

| Monomer | THMs | MCAA | DCAA | TCAA | HAA6 | HKs | TOX (m*M* Cl/ m*M* C) | Unknown TOX (m*M* Cl/ m*M* C) |
|---|---|---|---|---|---|---|---|---|
| Vanillic acid | 0.005 | ND* | 0.002 | 0.037 | 0.039 | ND | 0.158 | 0.029 |
| Vanillin | 0.005 | ND | 0.004 | 0.038 | 0.042 | ND | 0.156 | 0.020 |
| Acetovanillone | 0.011 | 0.002 | 0.004 | 0.006 | 0.012 | 0.009 | 0.106 | 0.019 |
| Average V† | 0.007 | 0.001 | 0.003 | 0.027 | 0.031 | 0.003 | 0.140 | 0.023 |
| Syringic acid | 0.001 | ND | 0.002 | 0.012 | 0.014 | ND | 0.078 | 0.036 |
| Syringaldehyde | 0.002 | ND | 0.006 | 0.053 | 0.059 | ND | 0.230 | 0.052 |
| Acetosyringone | 0.031 | 0.007 | 0.011 | 0.002 | 0.020 | 0.025 | 0.202 | 0.000 |
| Average S‡ | 0.011 | 0.002 | 0.006 | 0.022 | 0.031 | 0.008 | 0.17 | 0.030 |
| *p*-Coumaric acid | 0.001 | 0.005 | 0.015 | 0.004 | 0.024 | ND | 0.143 | 0.094 |
| Ferulic acid | 0.003 | 0.026 | ND | 0.050 | 0.076 | ND | 0.199 | 0.012 |
| Average C§ | 0.002 | 0.015 | 0.007 | 0.027 | 0.050 | ND | 0.171 | 0.053 |
| *p*-Hydroxybenzoic acid | 0.004 | ND | 0.002 | 0.039 | 0.041 | ND | 0.167 | 0.032 |
| *p*-Hydroxybenzaldehyde | 0.004 | ND | 0.002 | 0.038 | 0.040 | ND | 0.186 | 0.056 |
| *p*-Hydroxyacetophenone | 0.013 | ND | 0.003 | 0.032 | 0.035 | ND | 0.195 | 0.053 |
| Average H** | 0.007 | ND | 0.002 | 0.036 | 0.039 | ND | 0.183 | 0.047 |
| Phenolic average†† | 0.007 | 0.005 | 0.005 | 0.028 | 0.038 | 0.003 | 0.166 | 0.038 |
| C/V | 0.29 | 23 | 2.25 | 1.00 | 1.61 | 0 | 1.22 | 2.34 |
| S/V | 1.62 | 3.50 | 1.90 | 0.83 | 1.00 | 2.78 | 1.21 | 1.29 |
| $(Ac/Al)_v$‡‡ | 1.0 | NA§§ | 0.5 | 1.0 | 0.9 | NA | 1.0 | 1.5 |
| Alkali lignin | 0.003 | ND | 0.001 | 0.003 | 0.004 | ND | 0.050 | 0.029 |
| Organosolve lignin | 0.004 | ND | 0.001 | 0.006 | 0.007 | ND | 0.073 | 0.040 |

*ND = not determined.
†V: Vanillyl phenolic average (vanillic acid, vanillin, acetovanillone).
‡S: Syringyl phenolic average (syringic acid, syringaldehyde, acetosyringone).
§C: Cinnamyl phenolic average (*p*-coumaric acid, ferulic acid).
**H: *p*-Hydroxyl phenolic average (*p*-hydroxybenzoic acid, *p*-hydroxybenzaldehyde, *p*-hydroxyacetophenone).
††Phenolic average: Lignin phenolic average (average of V + S + C + H).
‡‡$(Ac/Al)_v$: Ratio of vanillic acid to vanillin.
§§NA = Not applicable; cannot be calculated because of low vanillin level.

Tables 7.6 and 7.7 show the results of this calculation including incremental contributions from each monomer. Because of the small incremental nature of these numbers, the units were changed to from millimoles to micromoles. This calculation indicates that acetosyringone is the major contributor to the THMs in organosolve lignin. In contrast, it is the two aldehydes, vanillin and syringaldehyde, that contribute nearly all of the expected TCAA formation. The alkali lignin are devoid of syringyl phenols, and therefore the THMs come from vanillin and acetovanillone. In this case, nearly all of the TCAA comes from vanillin.



©2007 AwwaRF. All Rights Reserved.

**Table 7.6**
**Estimated DBP yields of monomers within the organosolve lignin polymer (μ*M*/m*M* C)**

| Monomer | THMs | MCAA | DCAA | TCAA | HAA6 | HKs | TOX (m*M* Cl/ m*M* C) | Unknown TOX (m*M* Cl/ m*M* C) |
|---|---|---|---|---|---|---|---|---|
| Vanillic acid | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Vanillin | 0.37 | 0.00 | 0.27 | 2.85 | 3.12 | 0.00 | 11.7 | 1.47 |
| Acetovanillone | 0.35 | 0.05 | 0.13 | 0.20 | 0.39 | 0.29 | 3.46 | 0.63 |
| Average V* | 0.73 | 0.05 | 0.40 | 3.05 | 3.51 | 0.29 | 15.15 | 2.10 |
| Syringic acid | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Syringaldehyde | 0.15 | 0.00 | 0.40 | 3.49 | 3.90 | 0.00 | 15.20 | 3.46 |
| Acetosyringone | 1.34 | 0.31 | 0.46 | 0.10 | 0.87 | 1.10 | 8.8 | 0.02 |
| Average S† | 1.49 | 0.31 | 0.86 | 3.59 | 4.76 | 1.10 | 24.0 | 3.48 |
| *p*-Coumaric acid | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Ferulic acid | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Average C‡ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| *p*-Hydroxybenzoic acid | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| *p*-Hydroxybenzaldehyde | 0.02 | 0.00 | 0.01 | 0.001 | 0.23 | 0.00 | 1.06 | 0.32 |
| *p*-Hydroxyacetophenone | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Total H§ | 0.02 | 0.00 | 0.01 | 0.001 | 0.23 | 0.00 | 1.06 | 0.32 |
| Total lignin phenols** | 2.24 | 0.37 | 1.28 | 6.65 | 8.50 | 1.38 | 40.2 | 5.9 |

*V: Vanillyl phenols (vanillic acid + vanillin + acetovanillone).
†S: Syringyl phenols (syringic acid + syringaldehyde + acetosyringone).
‡C: Cinnamyl phenols (*p*-coumaric acid + ferulic acid).
§H: *p*-Hydroxyl phenols (*p*-hydroxybenzoic acid + *p*-hydroxybenzaldehyde + *p*-hydroxyacetophenone).
**Total lignin phenols: (sum of V + S + C + H).

Comparing the bottom line in both of Tables 7.6 and 7.7 with the bottom two lines in Table 7.5 reveals that the calculated DBP formation is sometimes close to the measured values, although not always (see Figure 7.2). The specific DBP (THMs, MCAA, DCAA, TCAA) predictions for the organosolve lignin based on the monomer contributions (Table 7.6) all match reasonably well to the actual measured values. These numbers differ by as little as 10%, and no more than 40%. Considering the opportunities for compounded errors, these differences are probably not significant. The corresponding values for the alkali lignin showed a negative bias (Table 7.7 and Figure 7.2), and this may reflect the lower levels of formation with this product.

Figure 7.2 also shows a comparison of measured and calculated TOX yields for the two commercial lignin products. For both materials and parameters, the calculated yields were substantially lower than those measured. Careful inspection of the DBP monomer specific yields (Table 7.5) provides some insight here. The vanillyl phenols produce very little "unknown" TOX (i.e., TOX not accounted for in the THMs and HAAs). Under the conditions used in this study, a



©2007 AwwaRF. All Rights Reserved.

**Table 7.7**
**Estimated DBP yields of monomers within the alkali lignin polymer (μ*M*/m*M*·C)**

| Monomer | THMs | MCAA | DCAA | TCAA | HAA6 | HKs | TOX (m*M* Cl/ m*M* C) | Unknown TOX (m*M* Cl/ m*M* C) |
|---|---|---|---|---|---|---|---|---|
| Vanillic acid | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.000 | 0.05 | 0.01 |
| Vanillin | 0.25 | 0.00 | 0.20 | 1.87 | 2.06 | 0.000 | 7.68 | 0.97 |
| Acetovanillone | 0.23 | 0.04 | 0.09 | 0.13 | 0.26 | 0.190 | 2.29 | 0.42 |
| Average V* | 0.48 | 0.04 | 0.28 | 2.01 | 2.34 | 0.190 | 10.01 | 1.40 |
| Syringic acid | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Syringaldehyde | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Acetosyringone | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Average S† | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| *p*-Coumaric acid | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Ferulic acid | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Average C‡ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| *p*-Hydroxybenzoic acid | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.08 | 0.02 |
| *p*-Hydroxybenzaldehyde | 0.00 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 |
| *p*-Hydroxyacetophenone | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Total H§ | 0.00 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.08 | 0.02 |
| Total lignin phenols** | 0.48 | 0.04 | 0.28 | 2.02 | 2.35 | 0.20 | 10.10 | 1.41 |

*V: Vanillyl phenols (vanillic acid + vanillin + acetovanillone).
†S: Syringyl phenols (syringic acid + syringaldehyde + acetosyringone).
‡C: Cinnamyl phenols (*p*-coumaric acid + ferulic acid).
§H: *p*-Hydroxyl phenols (*p*-hydroxybenzoic acid + *p*-hydroxybenzaldehyde + *p*-hydroxyacetophenone).
**Total lignin phenols: (sum of V + S + C + H).

typical water or humic substance will produce by-products where roughly 50%–60% TOX is unknown, whereas these monomers were in the range of 10%–20%. In other words, the vanillyl units are very efficiently fragmented, leaving almost no halogenated products with more than one carbon atom. In contrast, the lignin polymers produced a more typical quantity of unknown TOX (presumably larger halogenated molecules). These larger molecules were either products of chlorination of the aliphatic portions of the polymers or they represented by-products of non-depolymerized fragments.

Incomplete recovery of lignin carbon by the alkaline CuO analysis raises some important questions about mass balance calculations and chlorine reactivity. Only 20% of the organic carbon in organosolve lignin was actually accounted for in the measured lignin monomers, and the corresponding number for alkali lignin was only 7% (Table 7.4). Why is the yield so low and what is the nature of the remaining 80%–93%? Ertel, Hedges, and Perdue (1984) assert that vanillyl phenols are only 33% recovered by the alkaline CuO method. If a correction is made for this



©2007 AwwaRF. All Rights Reserved.

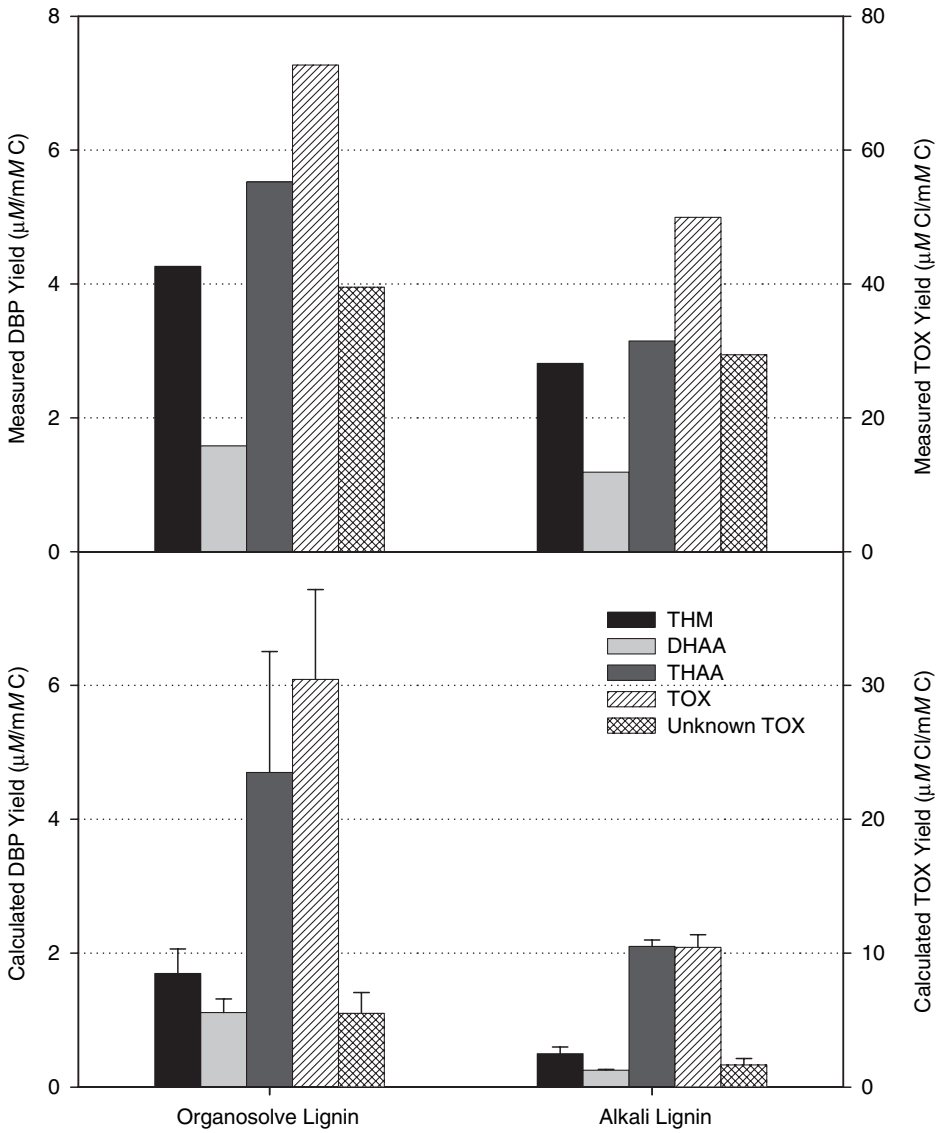



**Figure 7.2 Comparison of measured and calculated DBP yields for the two commercial lignins**

assertion, the recoveries increase to 31% and 20%. However, this may not be justified, as both lignin materials are already fragmented and rendered in a soluble form, and may be more easily depolymerized than native lignin. The literature indicates that most woods contain 19%–33% lignin by weight (Saka 2001). The bulk of the remaining material is cellulose and hemicellulose, neither of which is very reactive with aqueous chlorine.

Several possible explanations exist for discrepancies between the measured DBP formation and that calculated based on the specific monomer yields. First, although relatively nonreactive, the nonlignin carbon (mostly cellulose and hemicellulose) may contribute somewhat to DBP formation. Second, the lignin monomers might be rendered more or less chlorine-reactive as a result of partial degradation by alkaline CuO, leading to differences in DBP species distributions.



©2007 AwwaRF. All Rights Reserved.

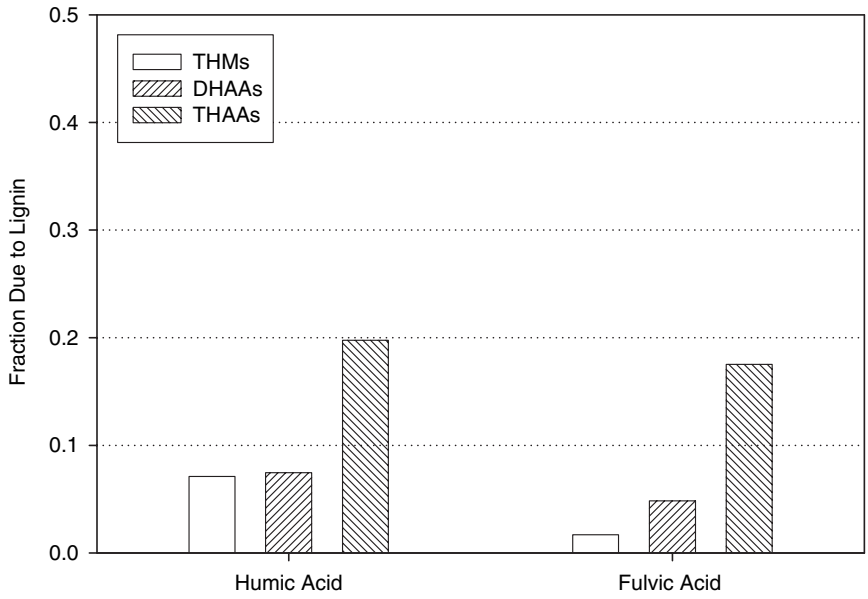



**Figure 7.3 Contribution of lignin to DBP formation in NOM based on specific monomer reactivity**

Data from this study can be used to estimate the contribution of lignin structures to overall DBP formation in natural aquatic organic matter. Calculations based on specific monomer yields and the chlorine reactivity of each can be performed exactly in accordance with the previous calculations on commercial lignin. This "mechanistic model" results in DBP levels that are 8%–20% of the measured values for the humic acid. The aquatic fulvic acid ranges from 2% to 18% (Figure 7.3). If the correction for low recoveries of vanillyl phenols is made, the humic acid yields do not change appreciably, but the fulvic acid yields do increase to a range between 4% and 26% (not shown).

An alternative approach to estimating lignin's role in DBP formation is to use the commercial lignins as a guide to reactivity and ignore individual monomer abundances. This avoids possible errors due to differences in free phenolic reactivity versus reactivity of the phenolics as they naturally exist in the lignin polymer. It does not allow for adjustment in reactivity due to differences in the abundances of the lignin monomers, or differences in the vascular plant types that gave rise to the NOM. The calculation is made by first determining a DBP yield for each DBP compound and express it on a per-lignin-C basis (using the total lignin phenolic determination). This value can then be applied to the humic fractions using the total lignin phenolic measurement made on each of these materials. The results show that about 21% of the THM, 26% of the DHAA, and 37% of the THAA are attributable to lignin sources in both NOM fractions (Figure 7.4). The error bars in Figure 7.4 reflect the two differing estimates based on the two reference lignin materials. These numbers decrease to about 12% of the THM, 15% of the DHAA, and 21% of the THAA when the vanillyl phenolic correction is used (not shown). This is attributable to the relatively low vanillyl content of the NOM fractions as compared to the lignin materials.



©2007 AwwaRF. All Rights Reserved.

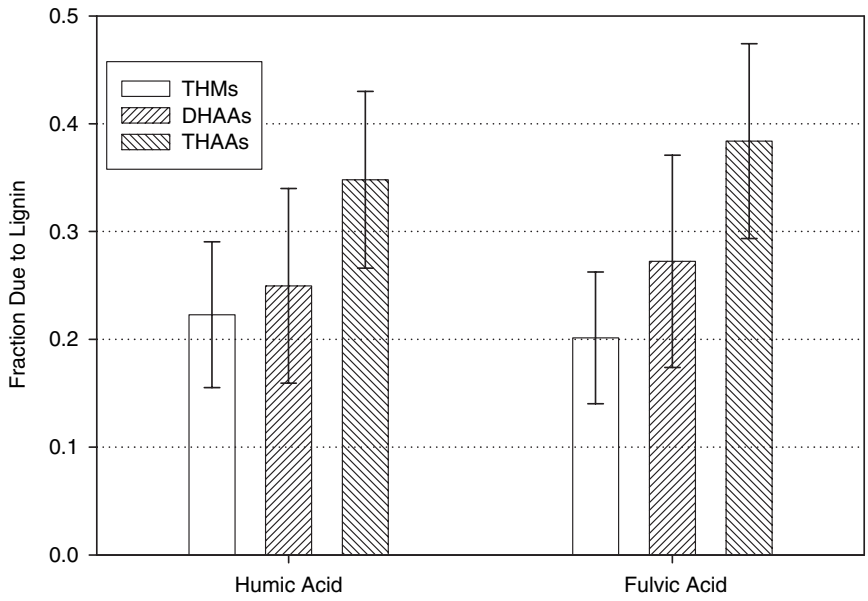



**Figure 7.4 Contribution of lignin to DBP formation in NOM based on total lignin assay**

## SUMMARY AND CONCLUSIONS

An HPLC version of the classic CuO degradation method for lignin analysis was refined and worked well with small quantities of commercial lignin polymers and NOM isolates. Classic lignin parameters (e.g., S/V ratios) were determined, and the values were commensurate with the know origins of the commercial lignins. Total phenolic yields were 11 and 27 mg monomer/ 100 mg C. The same analysis was performed on a commercial terrestrial humic acid and an aquatic fulvic acid (Suwannee River). The total lignin phenolic yields were 7 and 6 mg monomer/100 mg C, respectively. The S/V ratios were quite high for both materials, suggesting large hardwood content with some degradative loss of vanillyl structures.

Lignin monomers were prepared in pure solution and treated in the laboratory with chlorine to determine DBP yields. All phenolic monomers were reactive with aqueous chlorine, producing THMs and HAAs. Most of the syringyl, vanillyl, and *p*-hydroxy phenolics produced large amounts of TCAA as compared to DCAA and the THMs. They also tended to produce only minor amounts of other ("unknown") halogenated by-products. Natural polymeric lignin products also produced THMs and HAAs following chlorination. The amount produced could be roughly estimated from the monomeric composition (as determined by the CuO degradation method) and the specific DBPFP of the monomers. In contrast, the TOX produced by lignin polymers was much greater than predicted from the analysis of monomers.

DBP formation from lignin structures within NOM factions was estimated from two sets of model calculations, one termed "mechanistic" and one "empirical." The mechanistic approach used specific lignin monomer measurements for the NOM fractions and combined these with measured DBP yields from each monomer. The product of these values was summed over all monomers to obtain an overall DBP formation from lignin-based structures. The empirical



©2007 AwwaRF. All Rights Reserved.

approach used two commercial lignins as models of isolated lignin-based precursors. The normalized DBP yields from these materials were adjusted for the lower lignin monomer content of the NOM fraction and compared to the measured (overall) DBP yields. Both calculations were made with and without corrections for incomplete recovery of vanillyl phenols. The median predictions for these four models result in the best estimate at this time. They project a 9% role in THM formation, 12% in DHAA, and 25% for THAA based on the Suwannee River FA sample. The high THAA number suggests that waters influenced by vascular plants will give rise to higher levels of TCAA, such that the ratio of THAA/THM will be greater.



©2007 AwwaRF. All Rights Reserved.

# CHAPTER 8
# CONCEPTUAL MODEL FORMULATION

## INTRODUCTION

This chapter presents a series of arguments about the perceived fate and transport of DBP precursors in the aquatic environment. Data from two Massachusetts sources will be used to infer some key mechanisms. The chapter ends with the depiction of a conceptual model for NOM in general and DBP precursors in particular.

The dominant sources and transport pathways of the organic carbon cycle are well established, as summarized in Figure 8.1. For the purpose of this study, organic carbon transport data have been used to characterize variability, based on study location and vegetative type, in (1) the mass and/or concentration of organic carbon in major continental storage compartments and (2) the characteristic fluxes between these compartments. The focus of the mechanistic organic carbon modeling studies will be summarized in the context of the current understanding of transport and major categories of source and sink terms. For instance, numerous models of vertical transport of organic carbon in soils exist. An initial conceptual model of DBP precursor transport and transformation has been juxtaposed on that of the organic carbon cycle (Figures 8.2 and 8.3). This conceptual model invokes a sequential degradation of NOM types, which in turn leads to observable shifts in the specific DBP precursor content of the bulk NOM. Briefly, the sugars, polysaccharides, cellulose, hemicellulose, and other compounds are rapidly degraded in surface water and in the subsurface. These compounds are poor DBP precursors, and thus the specific precursor content increases as they are degraded. In contrast, the lignin structures are quite resistant to biodegradation. They will persist but will eventually degrade only after a long time (e.g., in large reservoirs).

In the literature, models of DBP precursor transport are strictly export-coefficient based. Such models typically allow the general behavior of NOM within a watershed to be modeled given land use and basic watershed information. However, inter-annual variability, except that anticipated from land use change, is not captured. In addition, a paucity of export coefficients for DBP precursors exists. One hypothesis is that significant spatial variability is present, necessitating development of at least regional export coefficients (although some of the formation potential data presented earlier would lead one to argue that these are not as important). Regional export coefficients for the continental United States should be developed based on actual DBP precursor data collected as part of this and future studies. In this effort, the methods of Waldron and Bent (2001) could be used. This joint USGS and utility study developed export coefficients for THM and HAA precursors based on data collected for the Cambridge public water supply system.

The Cambridge water supply system serves 95,000 permanent residents and 60,000 students each year. The city owns less than 5% of the land surrounding its reservoirs and tributaries and relies heavily on water quality monitoring to ensure safe drinking water. The goals of the joint USGS and Cambridge Water Department study were to characterize current water-quality conditions, to identify tributaries most likely to transport contaminants into the system, and to create baseline loadings for evaluation of future best management practice options. Throughout the 13-month study, several water quality parameters were measured, including fecal coliform, dissolved sodium, DOC, THMFP, and HAAFP, in addition to discharge. Waldron and Bent (2001) developed a series of relationships among the data in order to calculate the monthly



©2007 AwwaRF. All Rights Reserved.

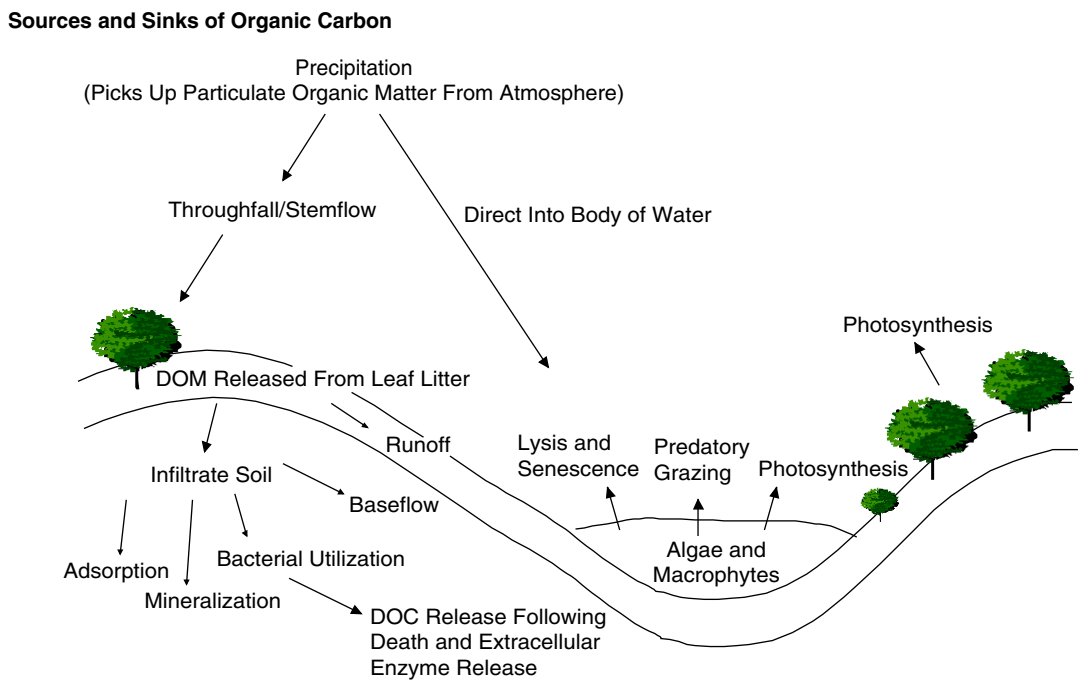



**Figure 8.1 Important watershed processes in the cycling of NOM**

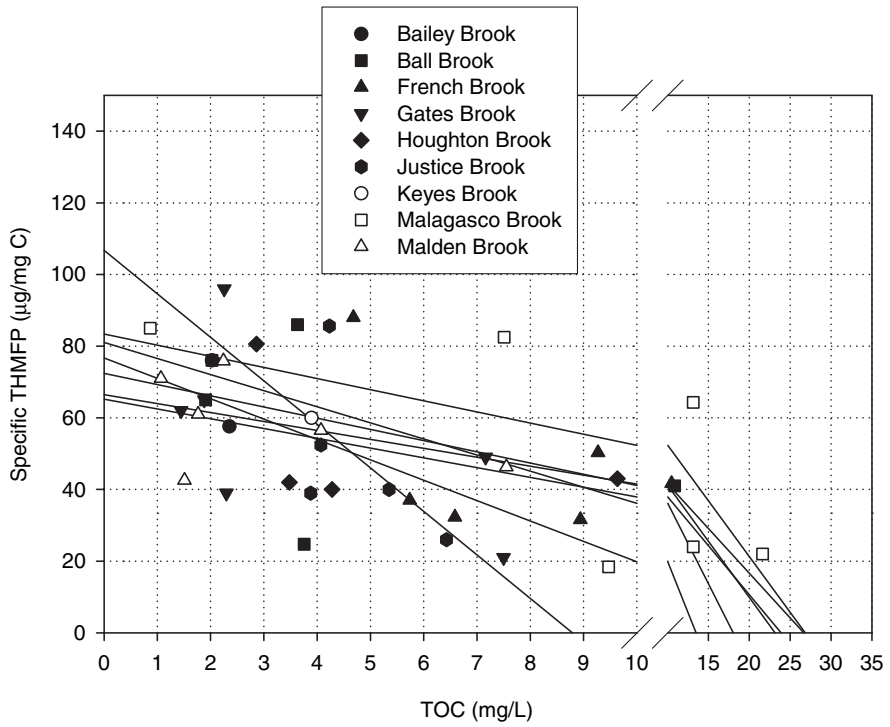



**Figure 8.2 Relationship between specific THM precursors and TOC for individual locations in the Wachusett Reservoir watershed (part 1)**



©2007 AwwaRF. All Rights Reserved.

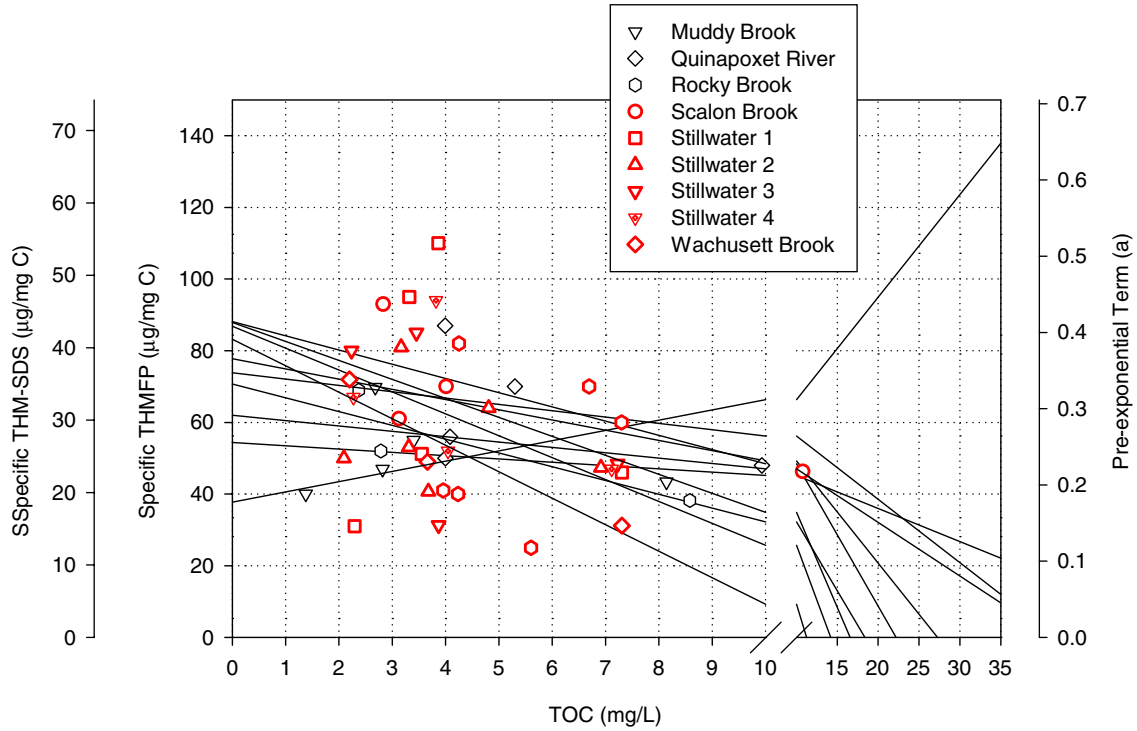



**Figure 8.3 Relationship between specific THM precursors and TOC for individual locations in the Wachusett Reservoir watershed (part 2)**

and annual yields to the water supply system. From the monthly and annual loading data, Waldron and Bent created export coefficients for different land uses within the watershed. Key questions that should be answered include

- How much does export of DOC/TOC and various precursors vary across the nation?
- Are there notable regional differences in export values?
- Are data sufficient to determine the underlying causes of variation, including potential factors such as land use, vegetation type, hydrology, and hydraulics?

Such models would provide utilities with a basic tool for developing watershed management strategies.

## EMPIRICAL OBSERVATIONS FROM FIELD STUDIES

### Long-term Field Studies in the Wachusett Watershed

DBP precursor data from the Wachusett Reservoir watershed were presented in chapter 4 and a key figure from that chapter is referenced here. Figure 4.12 presents the specific THM precursor levels versus TOC concerning a half-dozen sampling campaigns spanning almost 3 years. Note the distinct decreasing trend with increasing TOC in Figure 4.12. This was not expected, based on both fundamental NOM characterizations and empirical observations. Highly



©2007 AwwaRF. All Rights Reserved.

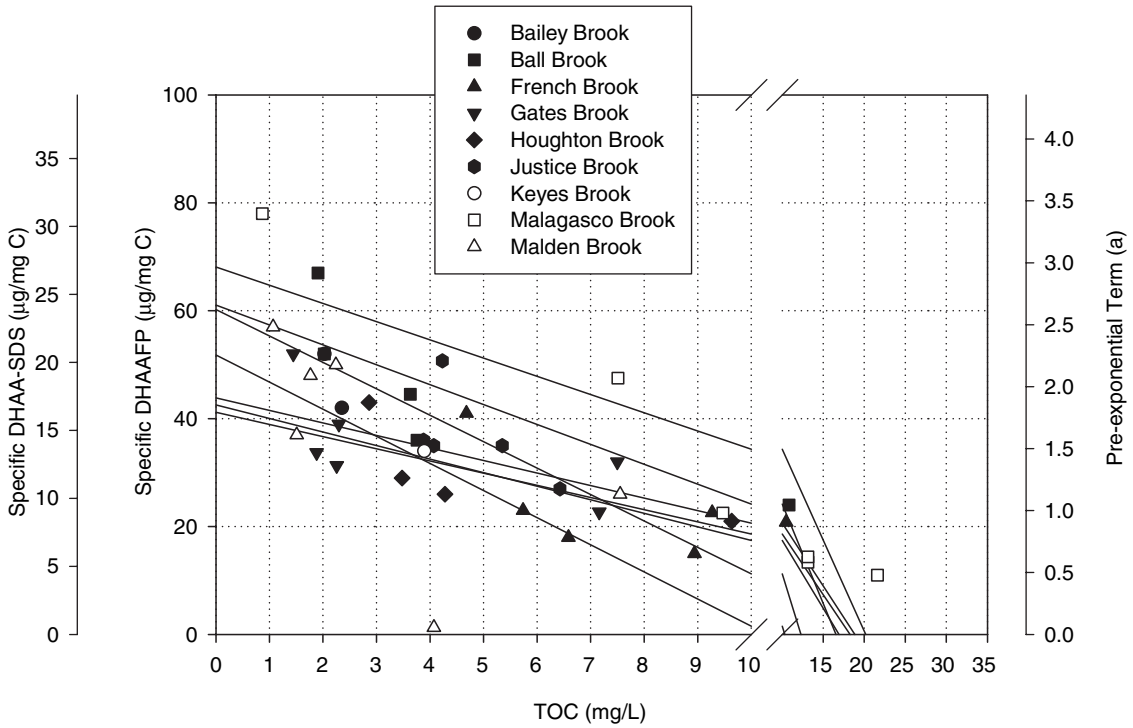



**Figure 8.4 Relationship between specific DHAA precursors and TOC for individual locations in the Wachusett Reservoir watershed (part 1)**

colored waters are dominated by humic substances, and these compounds are typically quite high in specific THMFP (e.g., Reckhow, Singer, and Malcolm 1990). In contrast, groundwaters and many lake waters are higher in the HPI fractions that are relatively low in reactivity.

When these data are examined as separate locations within the watershed, the relationship with TOC becomes clearer. The first nine locations (Figure 8.2) all show decreasing specific THM precursor content with increasing TOC. These are based on linear regressions of as many as nine separate sampling campaigns performed over a period of 28 months. Some were collected during periods of active rainfall, and others during long, dry periods. The remaining nine sites show similar trends (Figure 8.3).

Out of 18 separate sampling sites in this relatively pristine watershed, all but 1 showed a drop in specific THMFP with increasing TOC. This is unlikely to be the result of some systematic impact of TOC on the formation potential test, as all waters were diluted to about 2 mg/L TOC prior to addition of chlorine. The final THMFP was back-calculated to account for the effect of precursor dilution. In this way it is clear that NOM from waters having high TOC levels behaved differently than the NOM in low-TOC waters from the same location.

The inverse relationship observed between specific THM precursor and TOC seems to hold true for specific DHAA precursors (Figures 8.4 and 8.5). The THAA precursors showed more scatter, and the downward trend was not as unequivocal (data not shown).



©2007 AwwaRF. All Rights Reserved.

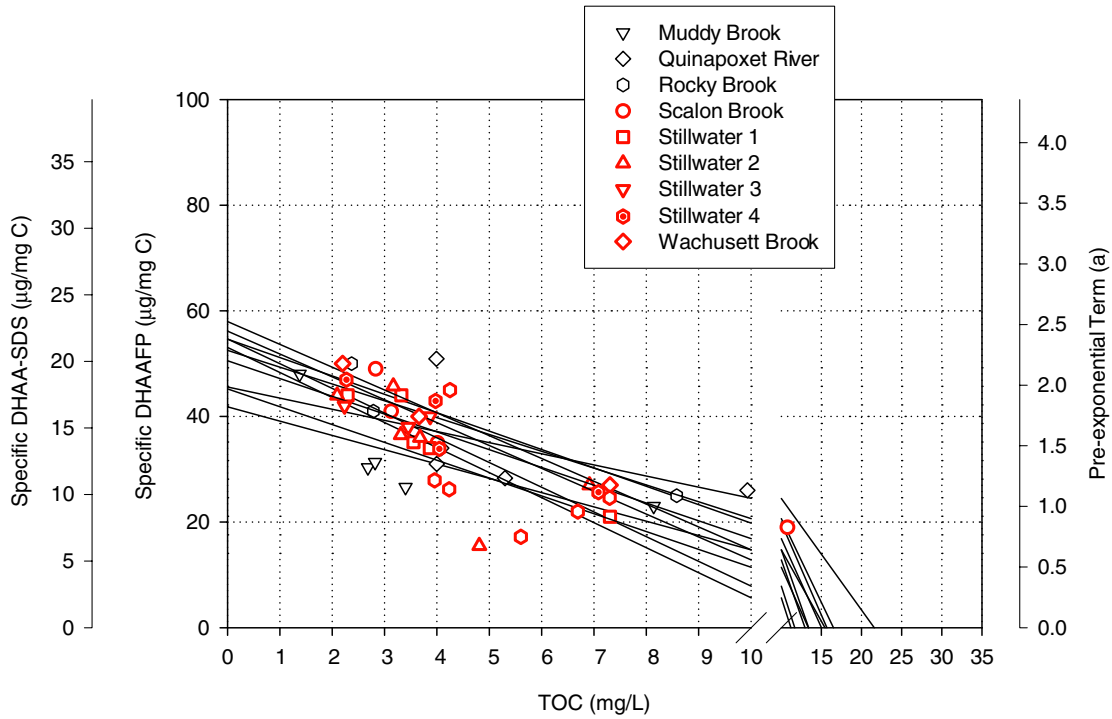



**Figure 8.5 Relationship between specific DHAA precursors and TOC for individual locations in the Wachusett Reservoir watershed (part 2)**

## Wet Weather Studies at Wachusett

It is well known that rain events can lead to flushing of the upper soil horizons and transport of recently deposited organic compounds to surface waters (Aiken and Cotsaris 1995). Precipitation may also lead to increased hydraulic connectivity among isolated wetland and stream channels. For these and other reasons, rain events may dominate the NOM and precursor loading in small watersheds.

Several rain events were captured as part of the long-term Wachusett study. These events showed similar behavior with respect to water quality, so only one will be presented here. On May 2, 2003, samples were collected for TOC and formation potential analysis. Figures 8.6 and 8.7 show the profiles for DOC and specific HAAFP before, during, and after the storm. Eight tenths of an inch of rain fell during a 2-hour period on this day. Figure 8.6 shows data collected for an upstream site (referred to as GBWD) that is in a moderately impacted area consisting of 41% forest, 43.5% urban (mostly residential), 10% agricultural, and 5.5% other land uses. Figure 8.7 shows data collected from a site (referred to as GBBR) that is 3.3 km downstream, isolating a heavily urbanized area (76%, mostly residential). Land use within the entire GBBR watershed (encompassing the GBWD watershed) is 36% forest, 52% urban, 8% agricultural, and 4% other uses. Although this was a relatively small rain event, an initial flush of DOC was evident in both sites, appearing within 30–60 minutes from the onset of rainfall.

In both cases, the DOC and specific HAAFP are nearly mirror images of each other. At the start of the storm, flows are low (not shown), as are DOC levels. The base-flow organic matter is presumably older, depleted of the most biodegradable fractions, and thus higher in specific



©2007 AwwaRF. All Rights Reserved.

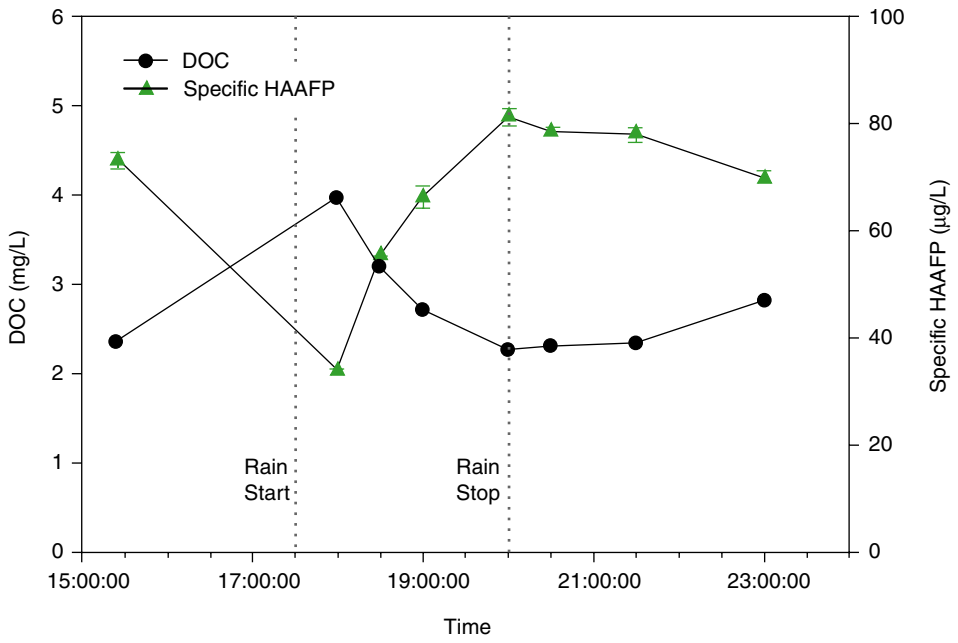



**Figure 8.6 DOC and specific HAAFP as a function of time for a rain event at the upstream, moderately impacted GBWD site (May 2, 2003)**

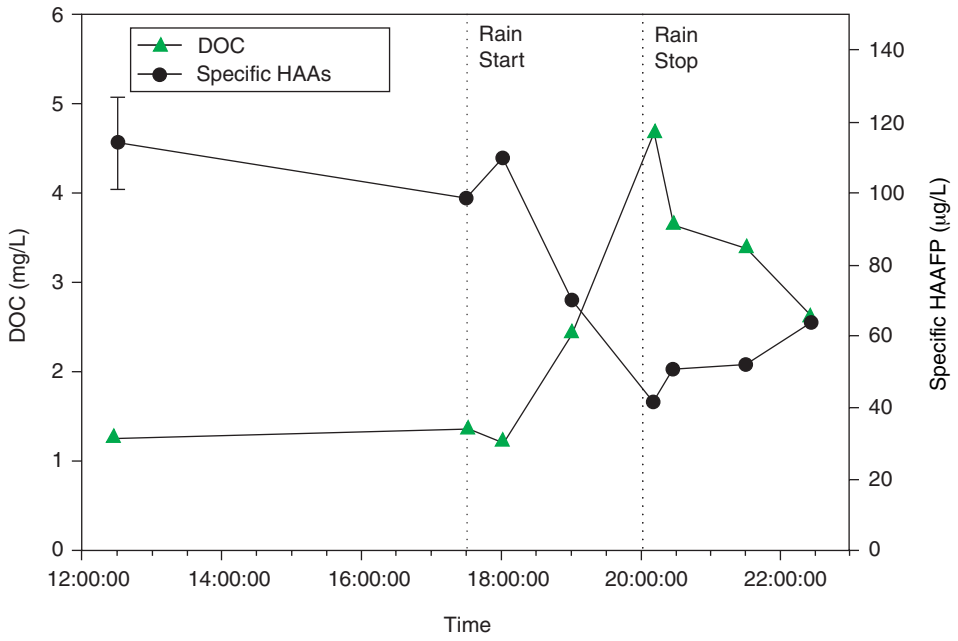



**Figure 8.7 DOC and specific HAAFP as a function of time for a rain event at the downstream, heavily residential GBBR site (May 2, 2003)**



©2007 AwwaRF. All Rights Reserved.

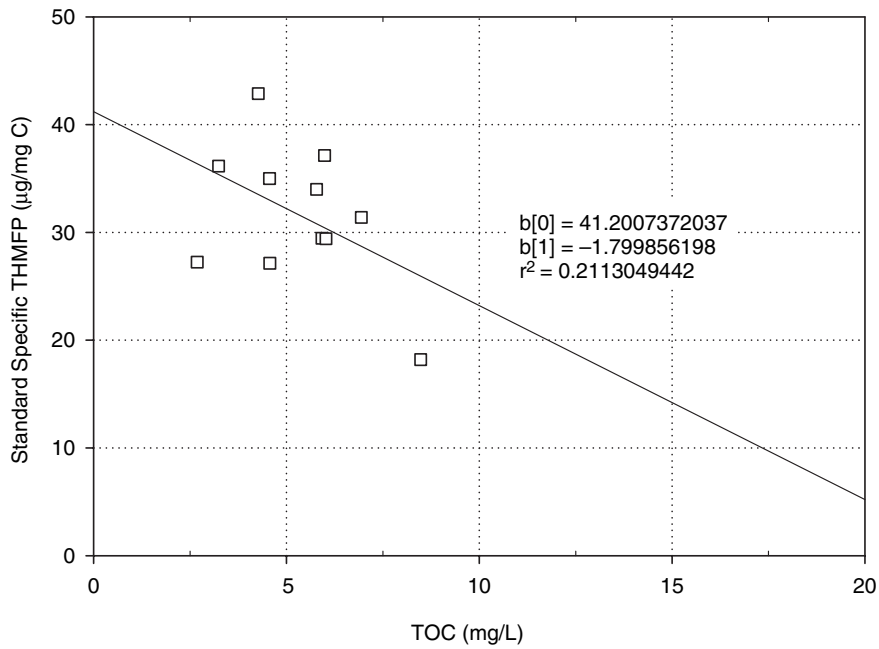



**Figure 8.8 Relationship between specific THM precursor levels and TOC in Cambridge watershed**

HAAFP. As the rain begins, DOC levels rise, and specific HAAFP levels drop. This suggests that although rain events lead to higher organic matter levels, the newly flushed organic matter is lower in HAA precursors than the older, base-flow DOC.

### Cambridge: Another Intensive Regional Study

Work by Waldron and Bent (2001) of USGS supports the previously noted inverse relationship. This group studied the Cambridge, Mass., watershed during a period of 1 year, measuring TOC, THM precursors, and other DBP-related parameters. Figure 8.8 shows the median values from each of 11 locations sampled during the study year. Although the range is not as great as that noted for the Wachusett watershed (perhaps because the median values were used), the negative correlation is nonetheless apparent, albeit rather weak.

## MULTICOMPONENT DEGRADATION OF ALLOCHTHONOUS NOM

The Wachusett and Cambridge data clearly support the leaching study conclusions that freshly flushed (or freshly generated) DOM has a lower specific DBP content than aged DOM. This is the material that predominates during and after rain events when the streamflow and DOC levels are elevated. As the amount of runoff and interflow declines and the groundwater level drops, the stream chemistry resembles that of the base flow. This water contains less organic matter because of soil loss processes, especially biodegradation. Therefore, as the flow and DOC decline during dryer weather, the organic matter is expected to be more aged (i.e., subject to greater degrees of biodegradation). The correlation between flow and NOM biodegradability has



©2007 AwwaRF. All Rights Reserved.

been widely reported (e.g., Withers and Drikas 1998). The data previously presented suggest that biodegradation preferentially results in loss of nonprecursor material, thereby increasing the specific precursor levels.

Sources and dynamics of allochthonous NOM are much more complex and diverse than autochthonous NOM, and have yet to be fully understood. Allochthonous NOM originates from terrestrial, wetland, and littoral zone plant sources. The predominant fraction of allochthonous NOM is humic in nature, stemming from the microbial degradation of fresh and detrital plant material (Wetzel 2001).

Cellulose, hemicellulose, and lignin materials from plant structural constituents are thought to make up the largest source of allochthonous humic compounds. These complex organic polymers undergo a series of decomposition stages in both soil and water. First, degradation of fresh plant material by fungi occurs under aerobic conditions, degrading polyphenolic lignin into recalcitrant high-MW substances (Wetzel 2001). At this point, the slow process of degradation by extracellular enzymes begins, leading to smaller, more oxidized macromolecules containing carboxylic functional groups. Further degradation finally results in humic and fulvic acids and other compounds (Wetzel 2001). High concentrations of humics as a result of this microbial activity can be found in upper soil horizons because of the breakdown of surface plant material (Wetzel 2001, Thurman 1985).

When rainfall rinses these partially degraded organic compounds from the upper horizons, they become mobilized. At this point, degradation occurs in concert with transport. The most degradable compounds disappear first, followed by the next most degradable compounds, and probably ending with humic substances.

A description of how this process would work is presented here. Table 8.1 contains a listing of eight major classes of plant products that are likely to be solubilized and incorporated into aquatic NOM. For each of these eight classes, there exists an estimate of abundance as determined for a particular site-specific NOM (summarized from many sources including Robinson 1991, Thurman 1985, and Saka 2001). One can also propose environmentally important data such as general biodegradation rates. The absolute biodegradation half-lives in Table 8.1 are highly speculative and site specific, but the rank order is probably accurate. For each of these classes, a generalized THMFP was included. These are based on typical structures and known THM yields from model compound studies (e.g., Reckhow and Singer 1985, De Laat 1981, Kim and Reckhow 2006). Combining these data, one deduces that the particular NOM represented by the relative abundance estimates would have a carbon-specific THMFP of about 30 μg/mg C. Recall that this number, arrived at arbitrarily, is on the low end of the observed range in surface waters (refer to chapter 4).

To further this approach, one can postulate the impacts of biodegradation on this material as would occur in nature. Such degradation may occur in a river or reservoir or even in a biological filter (albeit at an accelerated rate). To simulate such a process, the various classes of plant products were allowed to decay based on the arbitrary rates and a first-order rate law. Along with these, the associated THMFP decayed as well (Figures 8.9 and 8.10). Also shown in Figure 8.9 is the modeled ratio of THMFP to residual TOC (i.e., the specific THMFP). This value climbs with time, primarily because of the quick loss of the relatively nonreactive carbohydrates and the persistence of the highly reactive lignins and tannins.

Figure 8.11 shows the modeled THM formation from each of the major plant biochemical types. Initially it is the tannins, lignins, and proteins that are largely responsible for THM formation. However, the proteins are more rapidly degraded, so that the terpenoids become the third



©2007 AwwaRF. All Rights Reserved.

**Table 8.1**
**Arbitrary model parameters used to create multicomponent model**

| Constituent | Initial abundance (% of TOC) | Biodegradation half-life (days) | Specific THMFP (µg/mg C) |
|---|---|---|---|
| Fats and waxes | 2 | 60 | 0 |
| Terpenoids | 12 | 400 | 10 |
| Tannins | 4 | 200 | 400 |
| Lignins | 6 | 150 | 120 |
| Cellulose | 26 | 40 | 3 |
| Hemicellulose | 25 | 25 | 3 |
| Protein | 10 | 8 | 30 |
| Sugars and starches | 15 | 2 | 3 |

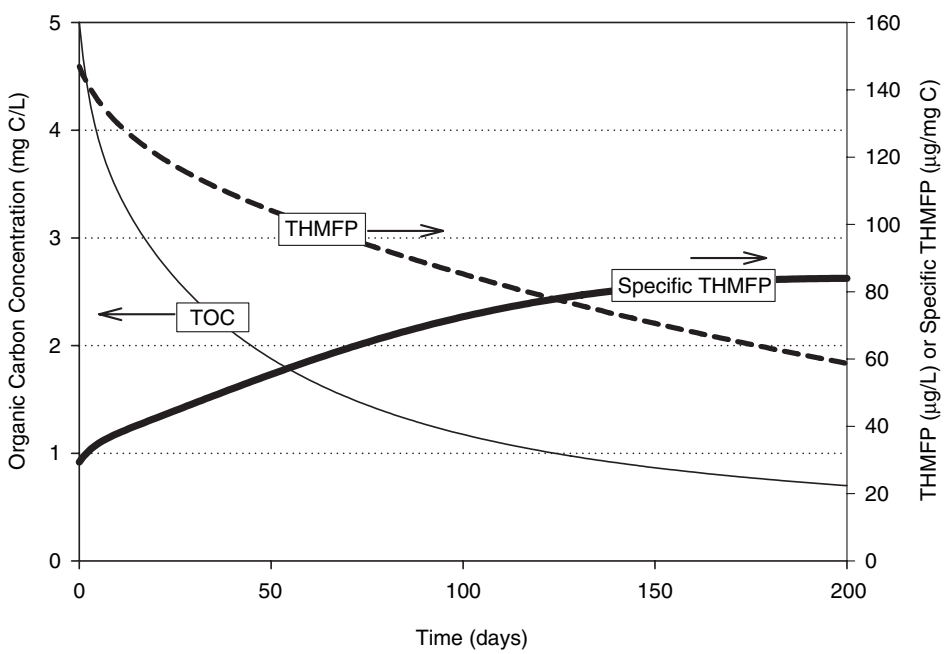



**Figure 8.9 Multicomponent allochthonous NOM degradation: major bulk parameters**

most important group. Based on the parameters chosen, the tannins are dominant and the lignins are also quite substantial. This situation will be different for various DBPs, as lignins contribute a disproportionate amount of THAA (as shown in chapter 7), and proteins are major DHAN precursors (Kim and Reckhow 2006). Nevertheless, the role of tannins in DBP formation should be the focus of future research.

There is some support for differential biodegradation of DBP precursors in the soil ecology and water chemistry literature. It is well accepted that polysaccharides are among the most biodegradable and labile of the soil organic matter fractions. Celluloses and hemicelluloses



©2007 AwwaRF. All Rights Reserved.

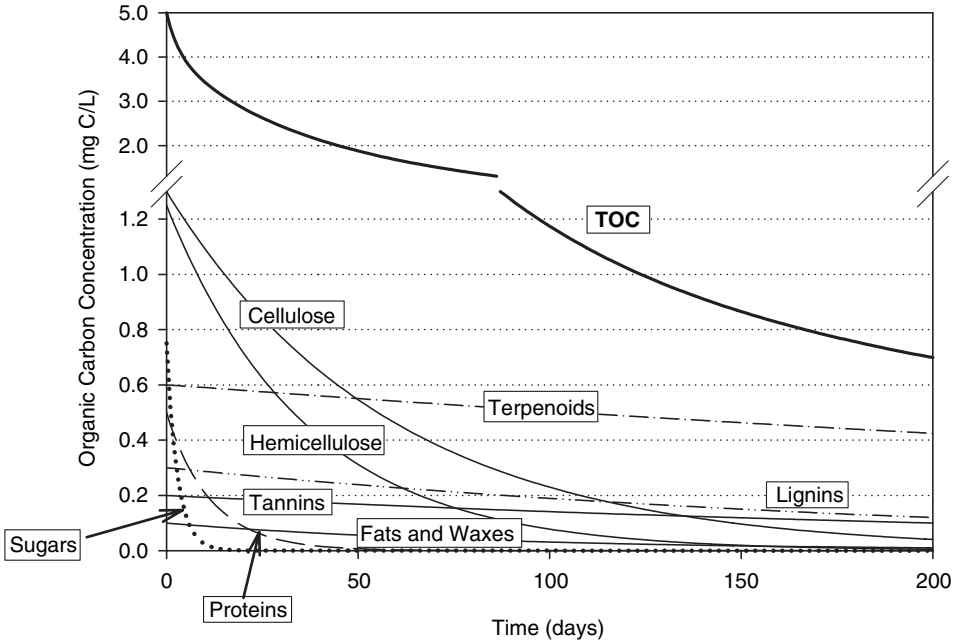



**Figure 8.10 Multicomponent allochthonous NOM degradation: biochemical components**

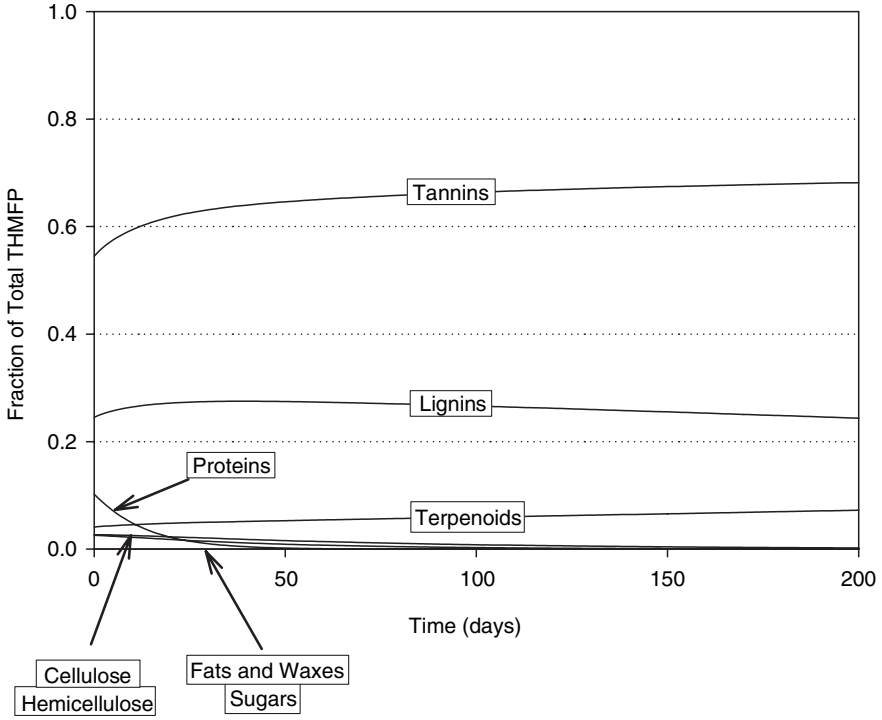



**Figure 8.11 Multicomponent allochthonous NOM degradation: THM precursor composition**



©2007 AwwaRF. All Rights Reserved.

are also biodegradable, although at a slower rate. These compounds are not very reactive with chlorine, and they do not produce large amounts of DBPs on a per-carbon basis (e.g., Hoehn et al. 1980). In contrast, the humic substances, lignin residues, and plant-based phenolics are far less biodegradable and may accumulate during the short term (e.g., Kalbitz et al. 1970). These structures are quite reactive with chlorine because of their activating groups and high electron density in the aromatic rings. As a result, they produce large amounts of DBPs following chlorination. Yields from model compound studies have been as high as 1 M/M or 1,600 μg THM/mg C (Larson and Rockwell 1979, Boyce and Hornig 1983, Reckhow and Singer 1985, De Laat 1981).

This analysis is most applicable to young NOM, particularly the organic matter in flowing surface waters that is freshly leached from terrestrial plants. This is also a gross simplification of the natural processes, as they do not result in direct mineralization of the NOM classes. Biodegradation rates will certainly depend on the immediate environment, the standing biomass, the presence of light and availability of terminal electron acceptors, and so forth. Nevertheless, the authors view this as a useful framework on which uncertainties and additional data can be assembled.

As NOM continues to age, one would expect that the increase in specific DBP precursor level could reverse itself. There is little doubt that, given enough time, the most reactive compounds (e.g., the activated aromatics) will be lost from the water column. This could be through slow biodegradation, adsorption to settleable particles, photolysis, and so forth. This is probably why oligotrophic reservoirs with long residence times have lower specific THM precursor levels than their contributing tributaries (e.g., Garvey and Tobiason 2003a). This general phenomenon has also been observed in groundwater recharge systems. Aiken and Leenheer (2004) reported that at very short distances from the source, specific THMFP increases, whereas for longer travel distances, the level drops substantially.

Figures 8.12 and 8.13 present visual summaries of many of the phenomena already discussed. The steps for NOM degradation may be summarized as follows:

1. Leaf litter releases organic mater with a spectrum of properties. Polysaccharidic-type organics are fast biodegraders but have low specific formation potentials, whereas lignin- or tannin-type organics degrade slowly but have high specific formation potentials.
2. Leachate can lead directly into surface waters or infiltrate into soil.
3. In soil, there is additional opportunity for retention and biodegradation.
4. Groundwater flow can slowly return this material to the surface waters (base flow), but all that remain are the lignin- or tannin-type organics.
5. Under high-flow conditions (Figure 8.13), transport is faster and flow paths change.

## SUMMARY AND CONCLUSIONS

Short-term and long-term temporal trends were used to infer some key fate and transport mechanisms. First, multiseason precursor data from two intensively studied locations were presented and analyzed for selected trends. THM, DHAA, and THAA data from the Wachusett Reservoir watershed and the Cambridge watershed showed consistently decreasing precursor levels with increasing TOC concentrations. Second, rain event data were analyzed for certain chemical parameters during the course of the event. Although the initial TOC after the onset of rainfall is sometimes quite high, the first-flush NOM has a low specific formation potential.

The temporal trends were interpreted within the context of leaching studies and other information in the literature. From this information, a conceptual model was proposed to describe



©2007 AwwaRF. All Rights Reserved.

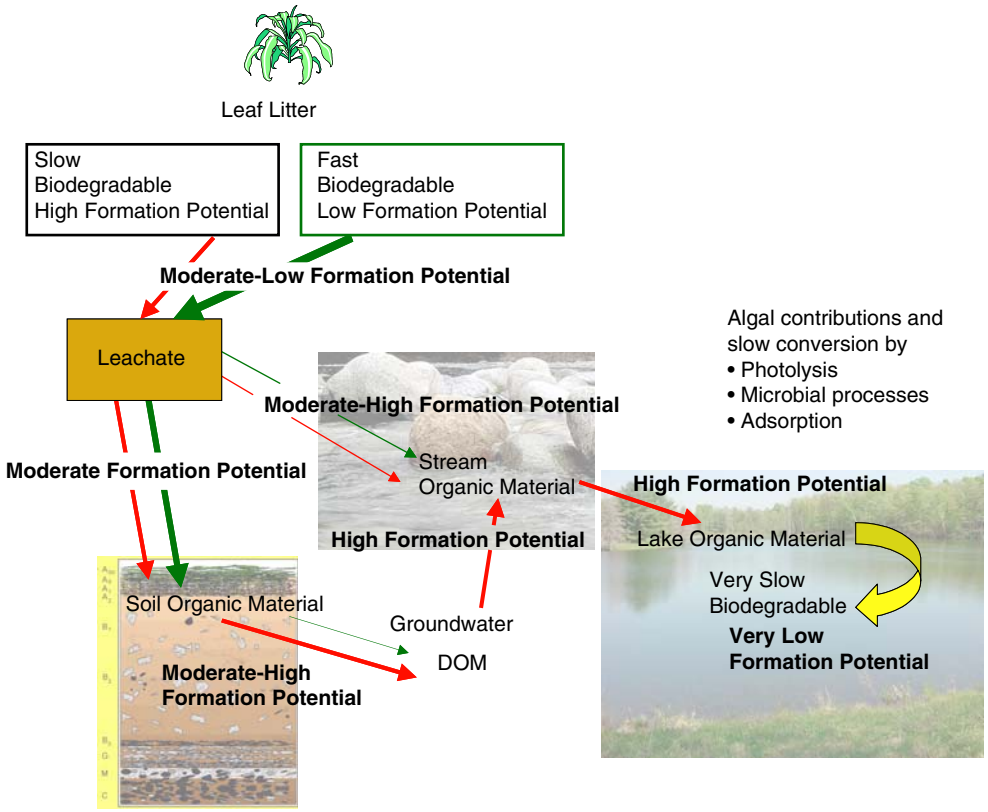



**Figure 8.12 Conceptual model for NOM degradation: low-flow conditions**

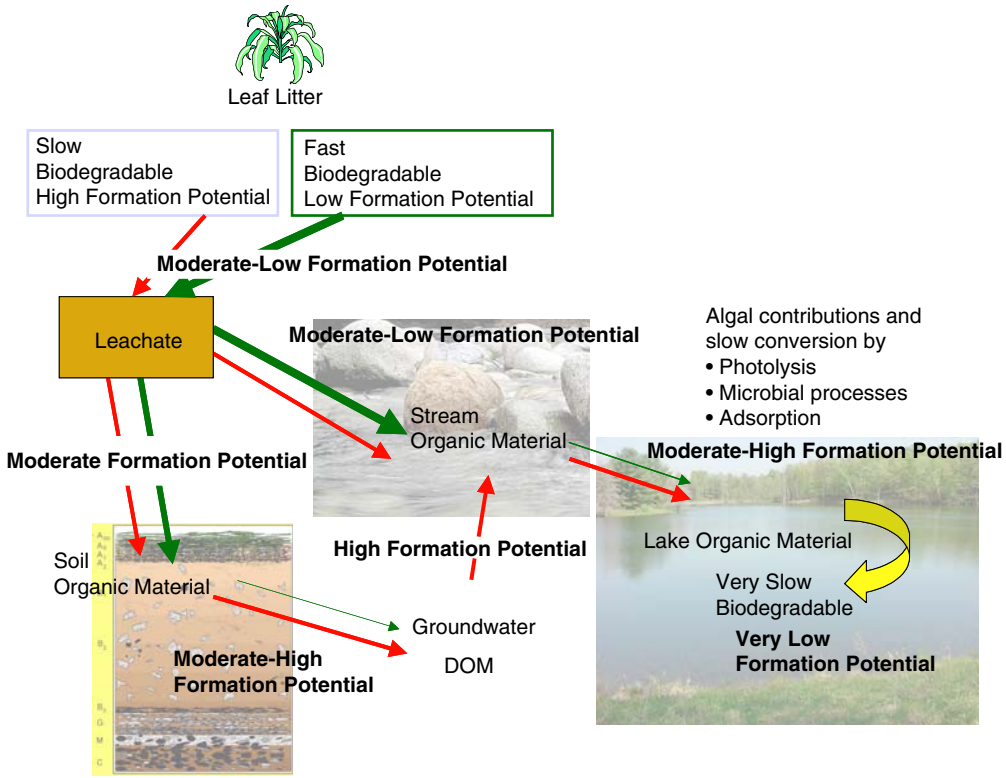



**Figure 8.13 Conceptual model for NOM degradation: high-flow conditions**



©2007 AwwaRF. All Rights Reserved.

the semi-empirical observations. First-flush NOM and NOM that dominates high-flow periods tend to be young or more recently leached from terrestrial plant material. This organic matter contains a suite of plant biochemicals that reflect availability, solubility, and rates of depolymerization or solubilization. As this material ages, it undergoes sequential biodegradation within the timescales of interest. For example, sugars are rapidly degraded, followed by cellulosic material, finally leaving the most recalcitrant biomolecules such as lignins and tannins. These latter compounds tend to be the most productive DBP precursors. As a result, the normalized DBP formation (per unit of DOC) from bulk NOM will increase as the surviving molecules are winnowed down to those that are most reactive with chlorine.

An illustrative, quantitative model was calibrated and predictions made as a means of testing the order-of-magnitude validity of the conceptual model. This included class-specific biodegradation, DBP yield, and natural abundance in fresh DOM. Using plausible values for each, estimates of the concentration of major NOM classes (sugars, proteins, lignins, cellulose, hemicellulose, tannins, terpenoids, fats) versus water age were made. This allowed the calculation of DBP formation from each of the biochemical classes. The final calculations showed that increases in the specific DBP precursor levels with biodegradation time would be expected. This model also showed that lignins and especially tannins should be the major contributors to THM formation, especially after long periods of biodegradation. Proteins may be important contributors during the early stages of DOM formation, but these will be lost and terpenoids may take their place as the third most important source of THM formation. Other DBPs will follow different source patterns, which underscores the importance of considering all reactive biochemical classes.



©2007 AwwaRF. All Rights Reserved.

©2007 AwwaRF. All Rights Reserved.

# CHAPTER 9
# NATIONAL TRENDS FOR ORGANIC CARBON

## OVERVIEW

Organic carbon measured as TOC or DOC is the most widely used measure for quantifying the amount of NOM in water. Impacts of anthropogenic change within a watershed affect NOM levels, which in turn affect DBP precursor levels. As discussed in chapter 8, freshly leached NOM tends to be composed mainly of nonreactive carbohydrates that result in relatively low DBP formation following chlorination. As the water ages, more reactive organic materials such as lignins and phenolics persist, resulting in higher DBP formation following chlorination. In addition, fresh leachate can be highly variable in terms of TOC levels and DBP formation potential as a result of the vegetation composition from which it was derived. The focus of this chapter is identification of spatial variability and temporal trends in TOC concentration across the United States. Understanding how variability in land use, vegetation composition, and hydrology may affect both organic carbon composition and thus precursor levels is important for the development of watershed-based management schemes to limit the formation of harmful DBPs in drinking water. Such management schemes may be a cost-effective means of reducing dependence on conventional treatment systems to provide such controls.

The development of a database tracking organic carbon in raw waters was begun to facilitate study of spatial and long-term temporal variability across the nation. Organic carbon data were collected from existing database sources, such as USGS and USEPA, as well as private utilities participating in the study. Whenever available, specific DBPFP data such as HAAFP or THMFP were included. Because background organic carbon levels and thus DBPFPs are influenced by native vegetation and soils, data were grouped by USEPA ecoregions in order to assess regional trends and variability. USEPA ecoregion level II was used for these assessments (Figure 9.1).

The remainder of this chapter is organized as follows. First, the results of select past studies exploring variability and temporal trends are summarized. Next, data obtained for use in this study are described. Importantly, the study focuses on organic carbon and precursor levels in raw water, not in drinking water distribution systems. The watersheds of participating utilities are reported in terms of the ecoregion in which they are located. The statistical methods are then described and applied to the data available from national data sets to analyze for spatial and temporal trends across the nation by ecoregion. Finally, the presented findings are summarized and discussed. The remaining chapters of the study present empirical and mechanistic models based on both the data presented within this chapter and additional data.

## SELECT RESULTS OF PAST STUDIES

Investigations of nationwide trends have often focused on organic nutrients such as nitrogen and phosphorus as well as other water quality constituents such as suspended solids, lead, and toxic elements. No known U.S. studies have analyzed for temporal trends in organic carbon, though similarities between trends in total phosphorus and organic carbon have been suggested. In Europe, several studies and research investigations have shown an increase in TOC/DOC within water supply lakes and reservoirs as well as within supply tributaries and rivers across various regions. In these studies, trends were determined through correlations, constituent relationships, and seasonal Kendall analyses.



©2007 AwwaRF. All Rights Reserved.

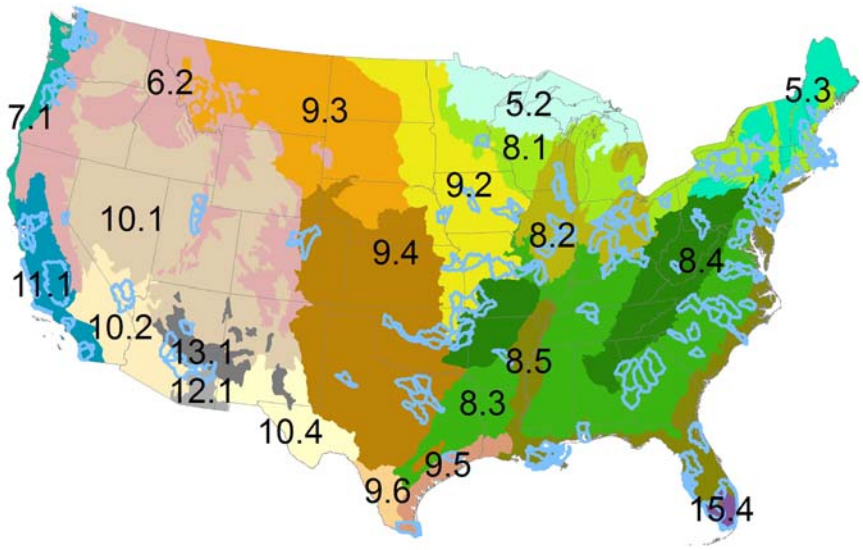



**Figure 9.1 USEPA ecoregion level II designations**

Smith, Alexander, and Wolman (1987) presented trends for various water quality constituents, including phosphorus and nitrogen, based on concentration data collected by USGS at surface water quality monitoring stations across the United States. Data included in the study spanned 8 years (1974–1981). Total nitrogen concentrations exhibited generally increasing trends across the nation, particularly in New England, and were attributed to atmospheric deposition. Trends in total phosphorus concentration were variable across the nation except in the Great Lakes region, where generally decreasing trends were observed. Lettenmaier et al. (1991) found similar trends in phosphorus concentrations, but these were based on USGS data collected over the period from 1978 to 1987. In both studies, it was suggested that increasing trends in total phosphorus were the result of increased population and atmospheric deposition, whereas decreasing trends were attributed to improved wastewater treatment practices. Large swaths of the nation exhibited no discernable trend in either phosphorus or nitrogen. Although neither the Smith Alexander, and Wolman (1987) nor the Lettenmaier et al. (1991) studies explicitly examined trends in TOC, Lettenmaier and colleagues focused on trends in organic nutrients. Overall, many increasing trends were found, although no significant trends were observed at approximately half of the stations analyzed. Increasing trends occurred in New England, along the Arkansas and Ohio rivers, along the Northern Mississippi River, and at several locations along the Gulf Coast and in the Pacific Northwest. Heathwaite, Johnes, and Peters (1996) also suggested that trends of total phosphorus are affected by anthropogenic activity such as point-source discharges.

The increase in TOC/DOC in Europe was first observed nearly 20 years ago in Scandinavia (Steinberg 2003). Global warming environmental impact studies suggest that the increased TOC concentrations result from changes in precipitation. As the temperature has increased in parts of Europe, the predominant form of precipitation has shifted to rain, leading to increased surface runoff and higher TOC loads within water supply sources (Steinberg 2003). Hongve, Riise, and Kristiansen (2004) observed increased acidity and NOM in forested Norwegian lakes between 1989 and 2000. In addition, they noted changes in NOM composition. They hypothesized that increased precipitation formed new transport and infiltration pathways, allowing additional TOC



©2007 AwwaRF. All Rights Reserved.

export and altering its characteristics. Specific absorption values within drinking water reservoirs in Germany also increased during this time period (Steinberg 2003). Freeman et al. (2001) employed a seasonal Kendall analysis across 22 sites within the United Kingdom and found increasing trends at 20 of the sites. The corresponding *p*-values were less than 0.05, indicating a high statistical confidence associated with each prediction. In addition, they observed large increases in carbon export from upland catchments. Hope et al. (1997) conducted an investigation of carbon exports across the United Kingdom from 1989 to 1993. During the 5-year period, variability in carbon fluxes was observed and the trends appeared to be similar to the trends of total phosphorus presented by Heathwaite, Johnes, and Peters (1996).

Researchers have suggested several potential causes for the observed increases in TOC across various regions of Europe. Freeman et al. (2001) hypothesize that increasing temperatures associated with global warming have decreased the average soil moisture, resulting in increased activity of the enzyme phenol oxidase within the soil layers. This increase, coupled with increased rainfall, may account for the observed increases in TOC export. Clair and Ehrman (1996) also suggest that climate change—specifically, increased evapotranspiration due to global warming—promotes microbial activity within the soil layer and thus may increase constituent loadings to adjacent water sources. Hongve, Riise, and Kristiansen (2004) suggest that increases in precipitation and subsequent leaching from forest floors are responsible for increased TOC concentrations within Norwegian lakes. Likewise, Steinberg (2003) hypothesized that increased rainfall accounted for increased TOC loading within the United Kingdom. Increases in TOC across Germany were attributed to forest death and redevelopment; during forest regeneration, large amounts of TOC are available within the soil layer and along the littoral zone (Steinberg 2003).

## AVAILABLE TOC DATA

Water quality data collected by federal, state, and local agencies are stored in USEPA's STOrage and RETrieval (STORET) system. Data collected by Native American tribes, volunteer groups, academics, and others may also be stored in this system at their initiative. Because of the sheer volume of data, the database is operated as two distinct systems. Historical water quality data collected from 1960 up to the end of 1998 are contained in the static Legacy Data Center (LDC). Data collected starting in 1999 are maintained in the STORET system, which is actively populated with new water quality data. Some historical data (pre-1999) have been stored in both the LDC and STORET. However, because this transfer is neither complete nor well documented, both systems must be searched to obtain comprehensive data for a region. The LDC and STORET combined are the largest water quality database for the United States.

In order to build the project-specific database, TOC and DOC data were downloaded from the LDC and STORET. In these systems, data may only be downloaded by county or hydrologic unit. Because of these restrictions and the volume of data contained within the system that needed to be processed, only data available for specific regions of the country were downloaded. These regions were identified by USGS-defined hydrologic unit codes (HUCs) for the watersheds that contained the 52 utility watersheds included in this phase of the study. These 52 utilities were selected in part so that data were collected for at least two utilities, and thus two HUCs, per ecoregion. Data were downloaded from 160 HUCs that contained the 52 utility watersheds. Thus, data acquisition from USEPA for the continental United States is currently incomplete.

As previously noted, surface water quality data collected by USGS have been assessed by other researchers for long-term trends in nutrient and heavy-metals concentrations. Though not as



©2007 AwwaRF. All Rights Reserved.

extensive, the USGS data set also contains a substantial amount of information on TOC and DOC. These data are maintained in a separate database by USGS, called the National Water Information System (NWIS); some of these data are also transferred to STORET. In order to expand the project database, TOC and DOC data collected by USGS were downloaded directly from the NWIS. Data were downloaded by state, resulting in complete data acquisition across the continental United States. Coverage, however, is variable because of state-by-state and year-to-year water quality monitoring schemes.

The data retrieved from the LDC, STORET, and the NWIS were combined and reviewed for extent and coverage. The TOC/DOC database was mined to identify stations consisting of TOC or DOC records collected at least quarterly for a minimum of 4 years. Data were divided into historical (1965–2000) and modern (1995–2004) data sets to capture both long-term and modern trends. These criteria resulted in a total of 307 and 188 stations, respectively, in the historical and modern data sets. These combined data are useful for characterizing regional variability, seasonality, and long-term temporal changes in TOC levels across the nation. The methods used for TOC analysis are generally consistent between the two data sets; however, changes in analysis methods over time (not well documented) and detection limits (observable through data analysis) do exist.

As previously noted, a general survey was sent to 553 utilities. Survey results were used to select a set of 52 utilities for subsequent followup personal communications. Raw water data (TOC, DOC, SUVA, $UV_{254}$, metals, general organics), hydrologic data (reservoir characteristics, streamflow patterns), and GIS boundaries and land use data were requested from these utilities or gathered from other sources. Analyses of these TOC/DOC data are conducted separately because methodologies, sampling frequencies, and detection limits tend to be more variable. Locations of the 52 utilities included in this phase of the study are shown in relation to USEPA ecoregion level II designations in Figure 9.1. The watershed delineation for each utility is outlined in light blue. In Figure 9.2, the location and name of each utility is identified. In addition, HUC locations for each utility are also shown, color coded to match the ecoregions shown in Figure 9.1. Recall that data downloads from the LDC and STORET were limited to these HUC locations, whereas USGS data were obtained for various locations across the United States. The number of USGS stations included, however, was largely constrained by the period of record requirements.

For the majority of the project analyses, TOC and DOC data were lumped together. This was justified because DOC typically comprises the majority of TOC, typically 90% to 95%. In addition, this combination of TOC and DOC data allowed expansion of the spatial and temporal extent of the project database.

## STATISTICAL METHODS

### Background

The Clean Water Act of 1972 presented many organizations with the challenge of assessing the nation's water quality. Assessment techniques were limited and often insufficient; available water quality data afforded investigators a frustrating mosaic of sporadic, heterogeneous measurements across the nation (Lettenmaier et al. 1991, Wolman 1971). Available data hindered initial attempts to accurately describe water quality within the nation's rivers, lakes, and other bodies of water. In order to rectify this situation, significant collection efforts were organized and statistical methods associated with water quality data were developed. Statistical analyses described the trends and behavior of various constituents and related nutrient/pollutant loadings to



©2007 AwwaRF. All Rights Reserved.

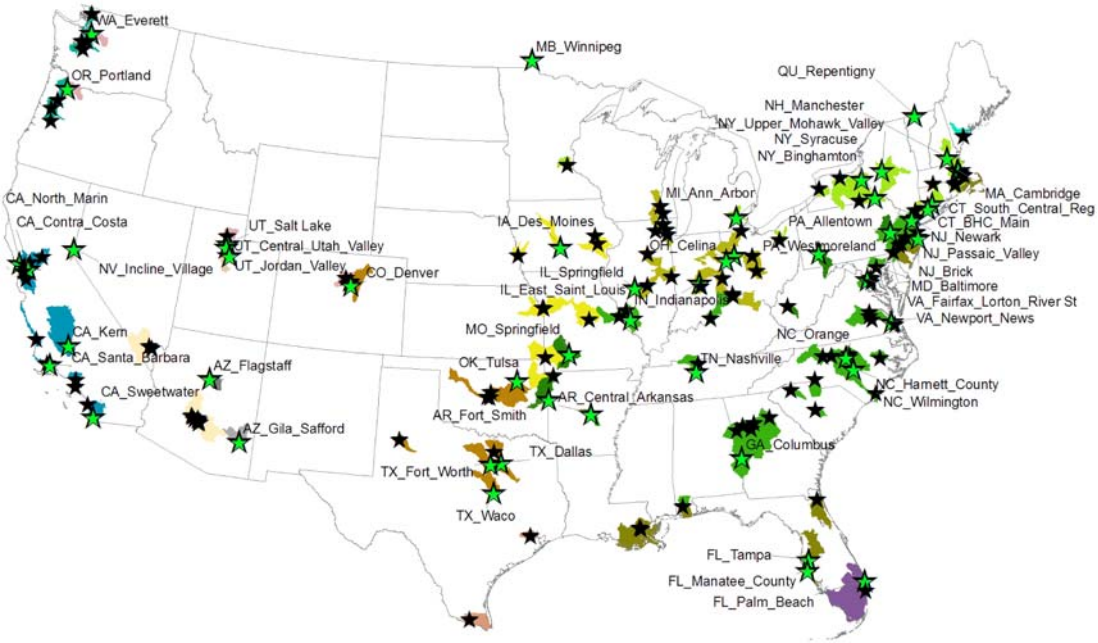



**Figure 9.2 Utility watershed locations and the HUC locations for which USEPA data were downloaded. HUC locations are color coded to correspond to their respective ecoregions in Figure 9.1.**

precipitation and streamflow. The complexity of statistical techniques increased as collected data grew and demands for accuracy increased. Statistical analyses allow investigators to capture relationships between water quality parameters, describe trends across time, and predict future anthropogenic influence on water quality.

Currently, a wide variety of statistical techniques are used to analyze and interpret water quality. The flexibility of statistical models allows for minor adjustments to better describe a unique set of data. Moreover, data transformations present a variety of alterations whereby the altered data better fit an existing statistical model. Model selection is a challenging task and crucial to the accuracy of results. If applied correctly, a statistical model yields accurate and pertinent results.

One objective of this study involved the assessment of long-term temporal variability of NOM across the United States. As noted earlier, previous studies in the United States have focused on long-term variability of phosphorus and nitrogen. These two organic nutrients exhibit some similarities to organic carbon incidence and behavior. Thus, these previous studies provide a framework for the evaluation of temporal trends in TOC. Results from this study provide an initial quantitative description of the long-term variability associated with NOM in the United States.

## Model Selection

Most water quality constituents of interest within the aquatic ecosystem do not behave in a manner defined by conventional statistics. Rather, water quality data stray from normal and log-normal distributions and render most parametric techniques ineffective. Proper model selection eliminates the parametric model (which assumes a normal distribution of the data and accuracy but requires data to stray little from this distribution) in favor of the nonparametric model (which



©2007 AwwaRF. All Rights Reserved.

does not require data to follow a particular distribution and makes no assumptions of the underlying data). A nonparametric model generates consistent results despite the distribution of the data, though the parametric model is more appropriate for normal and log-normal conditions. Nonparametric statistical models, which make no assumption about the underlying distribution, provide excellent flexibility and an array of techniques with which one may successfully analyze water quality data to gain insight into both the terrestrial and aquatic environments. Researchers have made full use of these models to describe the behavior and trends of nutrients, heavy metals, and other water quality parameters. Therefore, in analyzing water quality data it is often necessary to construct or make use of a nonparametric model that is better suited for irregular distributions and fluctuating data behavior.

Effectiveness of either model is also challenged by sampling frequency, erroneous data, censored data, and seasonality. If ignored, each effect significantly alters model predictions and ultimately causes erroneous conclusions regarding the data. The accuracy of predictions of long-term variability is typically improved through data transformations. Most statistical methods generate more accurate descriptions of trends across time with data transformation than without. Log transformations provide more resistant and robust results, especially when the data range across more than one order of magnitude (Helsel and Hirsch 1992). Literature provides ample guidance for each problem and extensive work by Helsel and Hirsch (1992) is summarized in the following discussion.

### *Sampling Frequency*

Constant water quality sampling frequency is extremely rare in large data sets. Without proper funding, samples are collected sporadically, with periods of frequent collection and periods of minimal or no collection. As a result, long-term variability studies encounter varying sampling frequency at a particular site or within a watershed. A few statistical methods ignore or disregard sampling frequency and generate accurate descriptions of trends across time. Most statistical methods, however, require uniform sampling intervals. Thus, most of the available historical data pose challenges to these techniques. To avoid these issues, historical data are typically transformed to a representative data set that is uniform in frequency. Each data set is unique, but the generic transformations create a series of measurements equally spaced in time, where the smallest time increment is often the new sampling frequency (e.g., month, quarter, or year).

### *Inconsistent Data*

Missing data pose a challenge similar to sampling frequency. In some instances, adjustment of the measurement interval surmounts the difficulties associated with gaps in the data. Erroneous data are also detrimental to any statistical analysis and the corresponding description of a data set. Outliers are either a rare response to system stimuli or an erroneous recording and are difficult to identify without intimate knowledge of the system. A transformation (e.g., log normal) reduces the impact of a single outlying or inconsistent measurement but does not entirely remove the data point. Utilization of a local average (an average of the surrounding measurements) may also minimize the influence of outliers (Hirsch, Slack, and Smith 1991). Exclusion of the data point is a viable option, although without confirmation it is a risky procedure because it may result in changes to the data frequency and preclude an unbiased evaluation.



©2007 AwwaRF. All Rights Reserved.

Censored data pose a challenge to statistical analyses but not necessarily to trend analyses. Many measurements that fall below the detection limit heavily skew the distribution. However, most trend analysis methods compare data points in forward progression. Because censored data indicate the minimal presence of the constituent, they do not largely influence the magnitude of change. Thus, censored data, if not the majority of data present, do not adversely affect statistical analysis for trend. Provided that a single censoring threshold exists, the Mann–Kendall test will perform without difficultly and draw comparisons between each paired observation within the data set (Helsel and Hirsch 1992).

### *Seasonality*

The magnitude of change associated with a trend is typically more affected by seasonality than censored data. Seasonal variation of the hydrologic cycle, soil and ground conditions, and anthropogenic activity greatly affect the carbon cycle and fluxes within the terrestrial and aquatic systems. Without proper adjustment for seasonality, a statistical technique may yield erroneous results. Thus, it is necessary to remove the effects of seasonality in order to accurately isolate behavioral changes and properly describe long-term variability. Several statistical methods eliminate variability associated with seasonality.

## SPATIAL TRENDS

### Variability Between Ecoregions

Data downloaded as described from USGS and USEPA databases were used to compare organic carbon levels across ecoregions. Data were not available for all ecoregions (no utility in all). Although the data used for this analysis are not comprehensive for the entire ecoregion, they do provide a representation of between-ecoregion variability in TOC. Box plots of the Legacy (1960–1998) data were generated using USEPA ecoregion level II data (Figure 9.3). The number of data points available per ecoregion is represented by the relative width of each box. Thus, wider boxes are used based on availability of a larger number of data points than narrower boxes. The body of the box represents the 25th, 50th, and 75th quartiles of the data, whereas the whiskers extend to plus or minus two standard deviations of the data. Points outside of this range have been omitted from the graph. To provide a point of comparison, the quartiles for ecoregion 8.3 have been extended across the entire figure. Coastal ecoregions such as 8.5—which also includes the lower Mississippi River corridor—9.5, and 15.4 exhibited generally higher TOC levels. These regions also tend to be heavily urbanized. Descriptive statistics of the TOC data for each ecoregion are summarized in Table 9.1. Table 9.1 and Figure 9.3 reveal a scarcity of data for several ecoregions, such as, for example, ecoregion 5.3. No TOC data were available for several HUCs. The inclusion of additional data would likely result in better representation of ecoregion characteristics.

TOC median data for the modern data were compared using the Mann–Whitney rank sum test to determine the statistical significance of differences in current TOC levels. A subset of resulting $p$-values is provided in Table 9.2. Entries in the table suggesting significant variability between ecoregions at the 10% significance level have been shaded. Legacy data for ecoregions 15.4 and 9.5 are significantly different than TOC data in most other ecoregions. Although the interquartile range of region 8.5 appears higher than most other ecoregions, it is not statistically significant.



©2007 AwwaRF. All Rights Reserved.

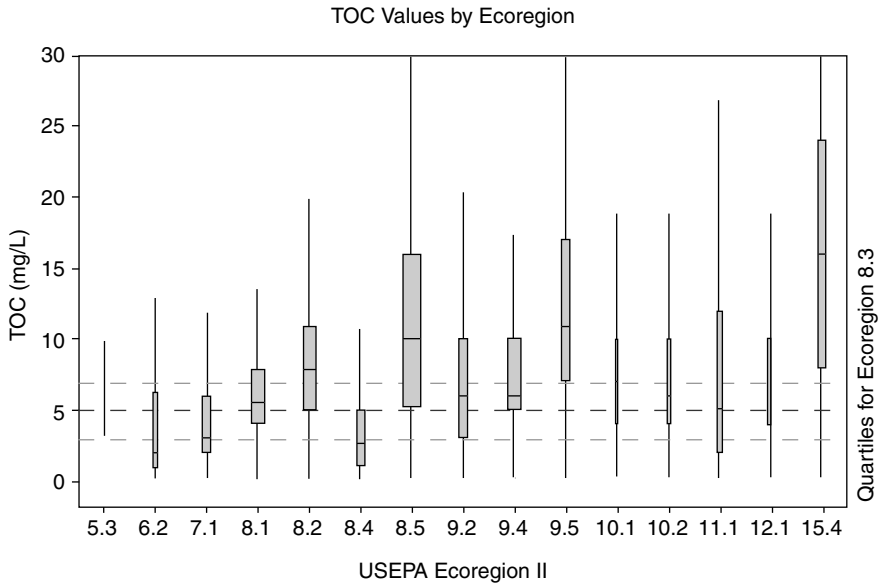



**Figure 9.3 Box-plot summary of data contained in the Legacy (1960–1998) database, organized by ecoregion**

**Table 9.1**
**Descriptive statistics of data contained in the TOC database, by ecoregion**

| Ecoregion | Mean | SD* | Coefficient of variation (%) | 25th quartile | Median | 75th quartile | Maximum | Number of Samples |
|---|---|---|---|---|---|---|---|---|
| 5.3 | 7.3 | 8.0 | 110 | 4 | 6 | 8 | 52 | 35 |
| 6.2 | 5.1 | 6.7 | 132 | 1 | 2 | 6.2 | 60 | 1,222 |
| 7.1 | 5.6 | 9.1 | 162 | 2 | 3 | 6 | 99 | 6,757 |
| 8.1 | 6.5 | 5.1 | 79 | 4 | 5.58 | 7.9 | 99 | 15,099 |
| 8.2 | 9.8 | 9.6 | 97 | 5 | 7.8 | 11 | 98 | 17,603 |
| 8.4 | 3.8 | 4.2 | 111 | 1.1 | 2.7 | 5 | 62 | 7,121 |
| 8.5 | 11.9 | 9.2 | 77 | 5.2 | 10 | 16 | 100 | 26,282 |
| 9.2 | 8.8 | 10.6 | 120 | 3 | 6 | 10 | 98 | 5,795 |
| 9.4 | 8.7 | 8.2 | 94 | 5 | 6 | 10 | 98 | 15,211 |
| 9.5 | 13.5 | 11.3 | 84 | 7 | 11 | 17 | 97 | 7,905 |
| 10.1 | 8.8 | 8.2 | 94 | 4 | 7 | 10 | 80 | 906 |
| 10.2 | 12.1 | 17.1 | 142 | 4 | 6 | 10 | 96 | 805 |
| 11.1 | 9.0 | 11.2 | 125 | 2 | 5.1 | 12 | 96 | 2,486 |
| 12.1 | 10.2 | 14.1 | 138 | 4 | 6 | 10 | 86 | 126 |
| 15.4 | 18.2 | 14.4 | 79 | 8 | 16 | 24 | 99 | 6,150 |

*SD = standard deviation.



©2007 AwwaRF. All Rights Reserved.

**Table 9.2**
**Statistical comparison via the Mann–Whitney test of the median values, modern data (*p*-values shown)**

| Ecoregion | 5.2 | 5.3 | 6.2 | 7.1 | 8.1 | 8.2 | 8.3 | 8.4 | 8.5 | 9.2 | 9.3 | 9.4 | 9.5 | 9.6 | 10.1 | 10.2 | 10.4 | 11.1 | 12.1 | 13.1 | 15.4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5.2 | 0 | | | | | | | | | | | | | | | | | | | | |
| 5.3 | 0.229 | 0 | | | | | | | | | | | | | | | | | | | |
| 6.2 | 0.65 | 0.56 | 0 | | | | | | | | | | | | | | | | | | |
| 7.1 | 0.41 | 0.9 | 0.799 | 0 | | | | | | | | | | | | | | | | | |
| 8.1 | 0.096 | 0.02 | 0.01 | 0.03 | 0 | | | | | | | | | | | | | | | | |
| 8.2 | 0.27 | 0.41 | 0.26 | 0.396 | 0.72 | 0 | | | | | | | | | | | | | | | |
| 8.3 | 0.21 | 0.23 | 0.13 | 0.396 | 0.89 | 0.83 | 0 | | | | | | | | | | | | | | |
| 8.4 | 0.22 | 0.49 | 0.24 | 0.57 | 0.05 | 0.71 | 0.35 | 0 | | | | | | | | | | | | | |
| 8.5 | 0.21 | 0.28 | 0.27 | 0.45 | 0.42 | 0.43 | 0.44 | 0.034 | 0 | | | | | | | | | | | | |
| 9.2 | 0.38 | 0.46 | 0.45 | 0.39 | 0.99 | 0.97 | 0.796 | 0.69 | 0.55 | 0 | | | | | | | | | | | |
| 9.3 | 0.73 | 0.88 | 0.79 | 1 | 0.28 | 0.55 | 0.41 | 0.57 | 0.51 | 0.695 | 0 | | | | | | | | | | |
| 9.4 | 0.83 | 0.18 | 0.49 | 0.3 | 0.006 | 0.12 | 0.1 | 0.07 | 0.26 | 0.14 | 0.71 | 0 | | | | | | | | | |
| 9.5 | 0.19 | 0.01 | 0.005 | 0.01 | 0.002 | 0.033 | 0.11 | 0.007 | 1 | 0.08 | 0.06 | 0.0195 | 0 | | | | | | | | |
| 9.6 | 0.67 | 0.18 | 0.14 | 0.21 | 0.23 | 0.56 | 0.64 | 0.27 | 1 | 0.67 | 0.8 | 0.23 | 0.8 | 0 | | | | | | | |
| 10.1 | 1 | 0.28 | 0.45 | 0.52 | 0.07 | 0.27 | 0.08 | 0.11 | 0.29 | 0.5 | 0.83 | 0.89 | 0.01 | 0.6 | 0 | | | | | | |
| 10.2 | 0.67 | 1 | 0.66 | 1 | 0.32 | 0.84 | 0.81 | 0.87 | 0.8 | 0.88 | 1 | 0.53 | 0.4 | 1 | 0.6 | 0 | | | | | |
| 10.4 | 0.67 | 0.37 | 0.38 | 0.45 | 0.77 | 1 | 1 | 0.52 | 0.93 | 1 | 0.8 | 0.23 | 0.4 | 1 | 0.6 | 1 | 0 | | | | |
| 11.1 | 0.36 | 0.19 | 0.13 | 0.22 | 1 | 1 | 0.68 | 0.496 | 0.49 | 1 | 0.6 | 0.06 | 0.03 | 0.4 | 0.297 | 0.8 | 1 | 0 | | | |
| 12.1 | 0.67 | 0.27 | 0.38 | 0.21 | 1 | 1 | 0.75 | 0.34 | 0.93 | 1 | 0.8 | 0.23 | 0.4 | 1 | 0.6 | 1 | 1 | 1 | 0 | | |
| 13.1 | 0.25 | 0.07 | 0.05 | 0.15 | 0.22 | 0.18 | 0.18 | 0.09 | 0.68 | 0.19 | 0.15 | 0.07 | 0.86 | 1 | 0.036 | 0.64 | 1 | 0.21 | 1 | 0 | |
| 15.4 | 0.15 | 0.02 | 0.01 | 0.04 | 0.006 | 0.06 | 0.15 | 0.02 | 0.68 | 0.1 | 0.11 | 0.03 | 1 | 0.75 | 0.05 | 0.5 | 0.5 | 0.03 | 0.5 | 0.83 | 0 |



©2007 AwwaRF. All Rights Reserved.

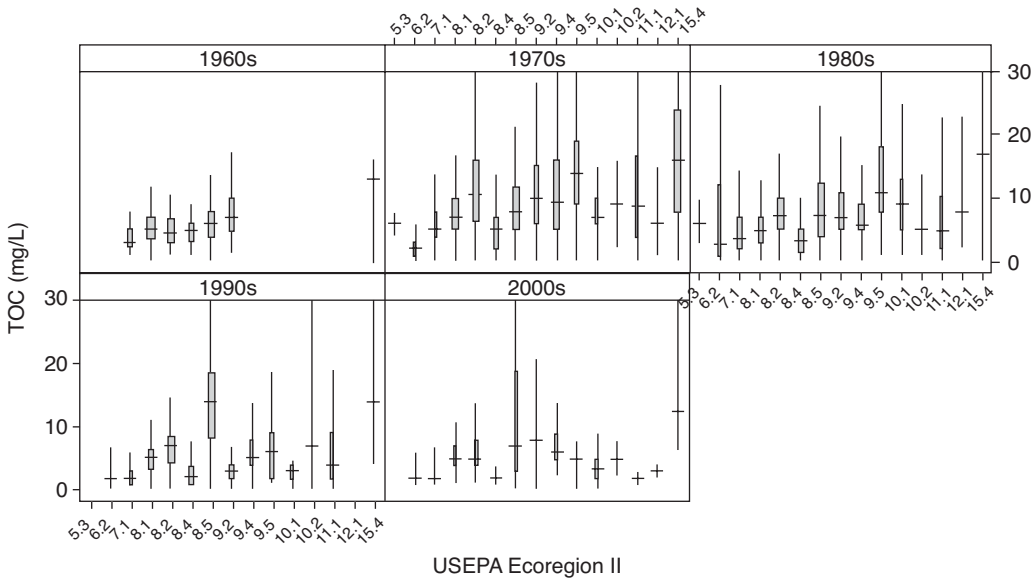



**Figure 9.4 Decadal TOC variability by ecoregion**

Variability among ecoregions was also explored by dividing the ecoregion TOC data by decade (Figure 9.4). TOC data were more prevalent in the 1970s and 1980s (as evidenced by the wider box plots). Although some variation among decades is evident, in general the inter-ecoregion dynamics remain the same among the decades as compared to when the data are lumped (Figure 9.3). Slight decreases in quartile statistics, such as the median TOC value, are suggested by the data for some ecoregions. For example, the median TOC value in ecoregion 15.4 appears to have decreased from about 18 mg/L in the 1970s and 1980s to about 12 mg/L in the period from 2000 to 2004. Decreasing amounts of data, however, make it difficult to lend credence to this apparent trend. Temporal trends will be explored more fully in the following section. The ecoregion-based TOC data were also analyzed on a seasonal basis to assess trends (Figure 9.5). No trends were clearly evident, suggesting that the inter-ecoregion variability is greater than interseasonal variability for an ecoregion. Data were generally more available during the spring through fall seasons than during the winter.

### Variability Within Ecoregions

Ecoregion 8.3 encompasses a large area of the eastern United States, including many major metropolitan areas. Many of the 52 utilities included in this phase of the study lie within ecoregion 8.3. The HUCs within ecoregion 8.3 for which TOC data were obtained through STORET and the NWIS are shown in Figure 9.6. Because of the greater availability of data for this ecoregion within the project-specific database, ecoregion 8.3 was selected for analyses to evaluate inter-ecoregion variability in TOC concentrations.

Ecoregion 8.3 interquartile TOC ranges of each HUC contained within the ecoregion for which data were obtained are presented in Figure 9.7. Both the Legacy and modern data are included. Interquartile ranges for data that have been lumped together are provided as dashed lines on the figure for comparison purposes. As previously noted, the width of each box corresponds



©2007 AwwaRF. All Rights Reserved.

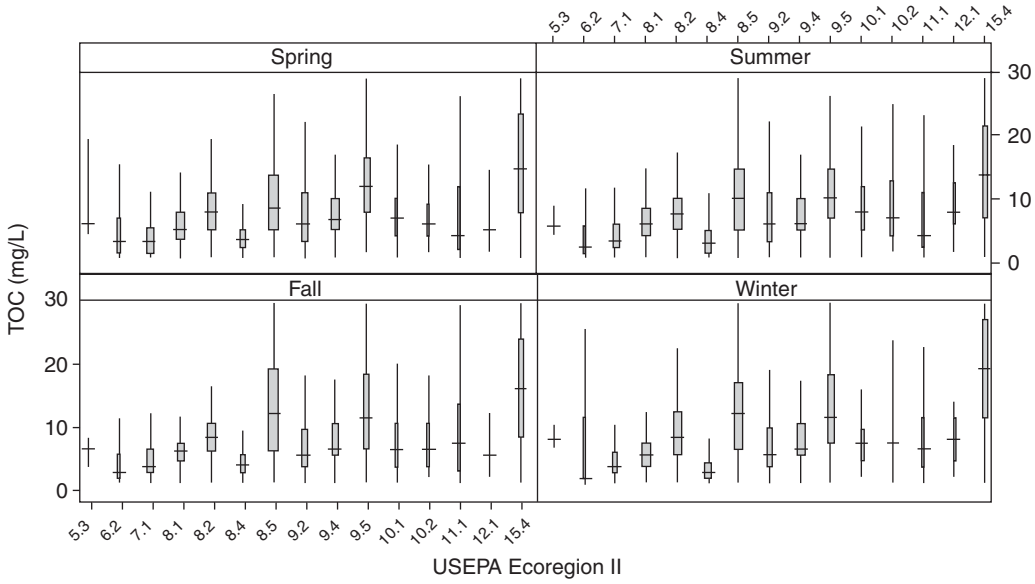



**Figure 9.5 Seasonal TOC variability by ecoregion**

with the amount of data from which the statistics are drawn. In general, data for the HUCs within the ecoregion are well represented by the overall interquartile range for the region. Some exceptions include HUCs 3020103, 3020201, 3030002, and 111100207, whose 50th and 75th quartiles tend to be higher than for most of the remaining HUCs. HUC 2040203 has an interquartile range almost entirely below that for the combined ecoregion data.

Variation in interquartile range by season over time were explored for ecoregion 8.3 by plotting the seasonal interquartile ranges based on data for the period of record—the 1970s, 1980s, and 1990s (Figure 9.8). In the 1970s and 1980s, observed data had similar ranges in the spring, summer, and fall, and a slightly lower range during the winter. In the 1990s, interquartile ranges across all seasons were considerably lower. In particular, the TOC concentrations associated with the 50th and 75th quartiles were much lower than observed in the prior data. It is possible that a true drop in concentrations has occurred. Alternatively, a shift in sampling locations or in methodology may have taken place. The cause of this apparent drop in TOC concentrations cannot be identified without further analysis.

## TEMPORAL TRENDS

### Mann–Kendall Technique

Trend analysis, specifically long-term variability of TOC, is the focal point of this section. Evaluation of long-term variability of NOM requires a comprehensive data set. However, data spanning several years often lack uniform sampling frequency and sample collection methods. In addition, detection limits and laboratory techniques typically change with time. A statistical model, therefore, must be sufficiently robust to provide accurate analysis with moderate to high confidence, despite flaws associated with the data (Lettenmaier 1988). The Mann–Kendall technique, a nonparametric model, describes parameter variability with time. Unlike other temporal



©2007 AwwaRF. All Rights Reserved.

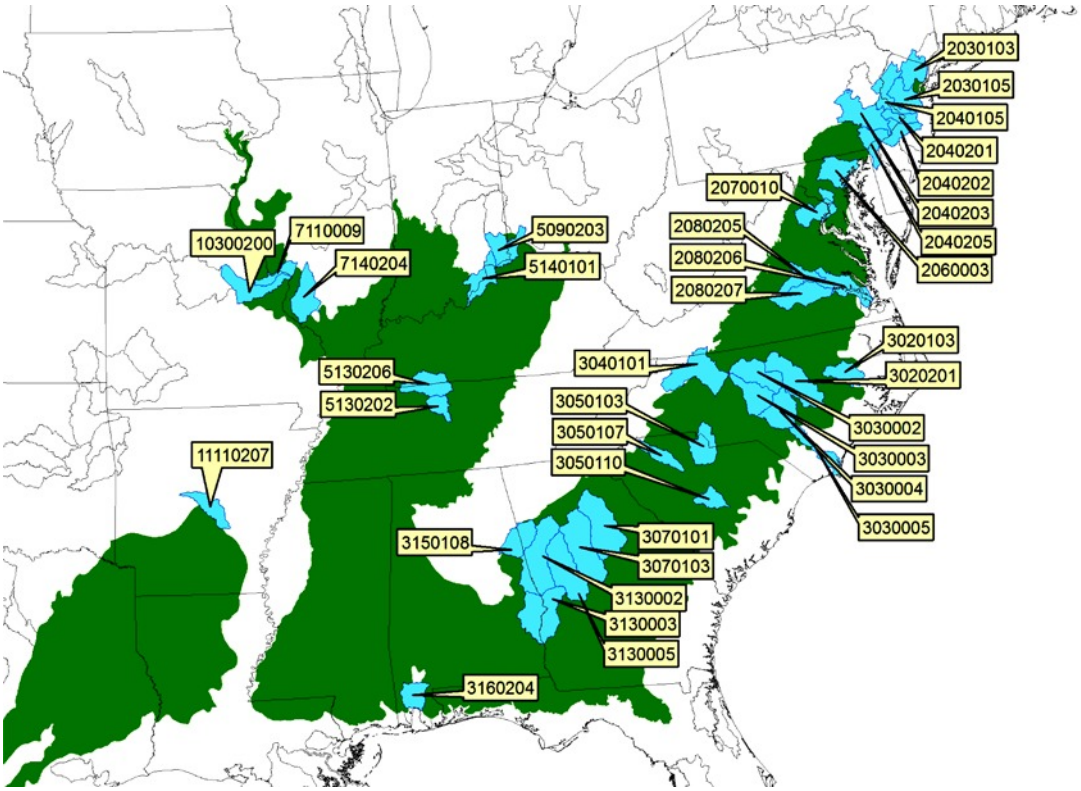



**Figure 9.6 Ecoregion 8.3 HUCs and locations**

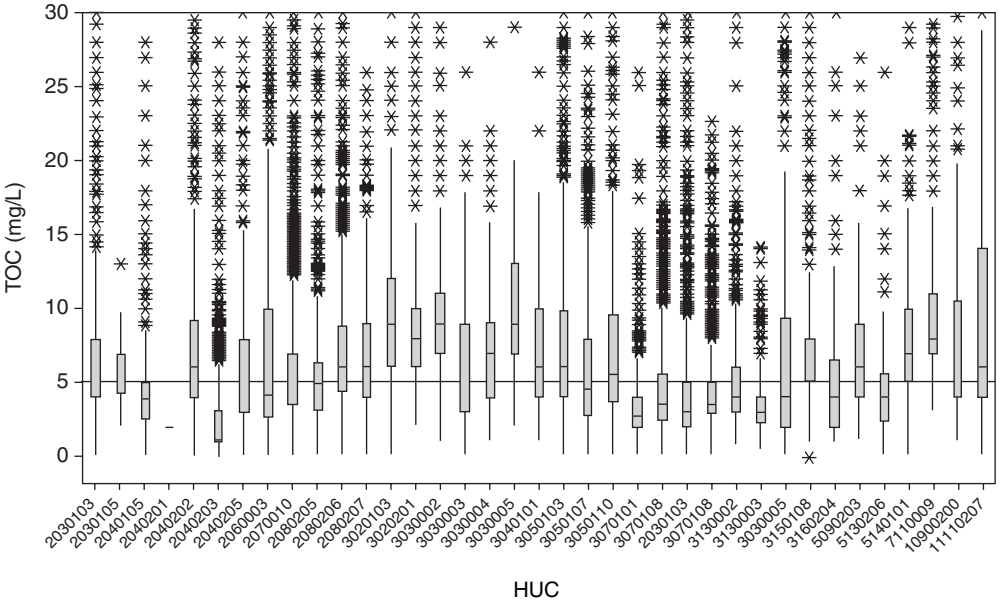



**Figure 9.7 Box-plot representation of interquartile range of data for each HUC within ecoregion 8.3**



©2007 AwwaRF. All Rights Reserved.

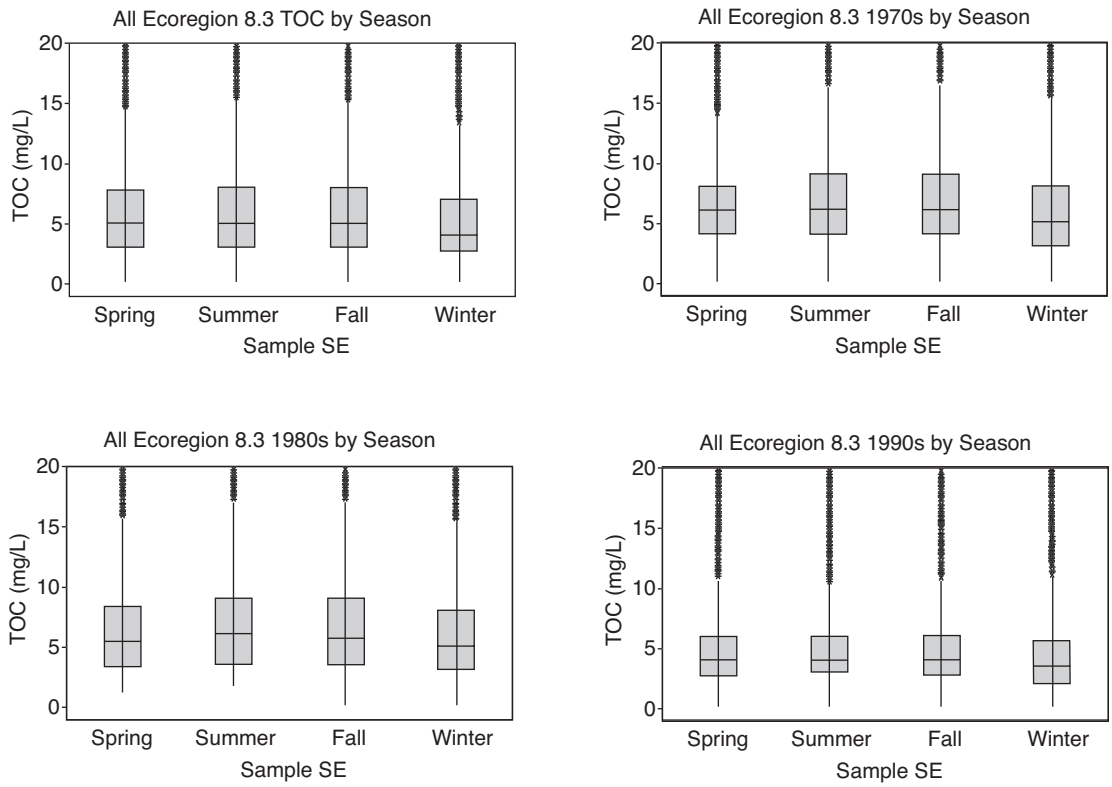



**Figure 9.8 Seasonal trends by decade for ecoregion 8.3**

trend methods and techniques, the Mann–Kendall test performs well among a variety of data, requires minimal data transformations, and is not hampered by seasonality and censored data.

The Mann–Kendall technique operates through the analysis of paired sets of data, in this case a unique time value (the independent variable) and the TOC measurement (the dependent variable). The TOC measurements need not be unique, as the Mann–Kendall technique generates a statistical value representative of the monotonic relationship between the independent (time) and dependent (TOC) variables. The governing equation for the Mann–Kendall technique is defined in Equation 9.1, where the summary-test ($S$-test) statistic is a summation of the "discordant pairs" (decreases) and the "concordant pairs" (increases).

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{sign}[(X_j - X_i) \times (Y_j - Y_i)] \tag{9.1}$$

where $S$ = summary-test statistic
$X_i, X_j$ = independent variables (time)
$Y_i, Y_j$ = dependent variables (organic carbon)

The "sign" function within Equation 9.1 is simply a function that converts discordant-pair results to a value of minus one (–1) and concordant-pair results to a value of positive one (+1).



©2007 AwwaRF. All Rights Reserved.

This function eliminates false predictions via removal of magnitude of change in trend, outliers, and other variability factors (Millard and Neerchal 2001).

$$\text{sign}(g) = \begin{cases} 1, & g > 0 \\ 0, & g = 0 \\ -1, & g < 0 \end{cases} \tag{9.2}$$

where $g$ is the operator within the sign function.

The null hypothesis associated with the Mann–Kendall test states that no trend exists, whereas the alternative hypothesis states that a trend does exist. The null hypothesis is true if the $S$-statistic is near or equal to zero and the alternative hypothesis is rejected. The null hypothesis is rejected, however, if the $S$-statistic is significantly different from zero and the alternative hypothesis is accepted. When the null hypothesis is rejected, a resulting negative $S$-statistic indicates a decreasing trend within the data with time and, conversely, a positive $S$-statistic indicates an increasing trend with time.

Kendall's $\tau$, a correlation coefficient, provides a confidence level associated with the $S$-statistic. Kendall's $\tau$ is calculated as shown in Equation 9.3.

$$\tau = \frac{S}{n(n-1)/2} \tag{9.3}$$

where $n$ = total number of measurements within the data set
$\tau$ = correlation coefficient

An absolute value of $\tau$ less than 0.01 indicates a high confidence in the trend postulated by the Mann–Kendall technique, whereas a $\tau$ value greater than 0.1 indicates little or no confidence in the predicted trend.

The shortcoming of the Mann–Kendall technique is the production of multiple $S$-statistics within a single data set. The statistical technique evaluates the paired data such that an $S$-statistic is reported for each change in trend. An analysis during a 2-year period of record would yield multiple $S$-statistics, and the sum of these values would represent the overall trend associated with the data (Hirsch, Slack, and Smith 1982). Hirsch and colleagues and other investigators implemented a seasonal analysis to combat this problem. The seasonal Kendall technique provides a single S-statistic for the entire period of record and eliminates multiple S-statistics across the period of record provided by the Mann–Kendall technique, as shown in Equation 9.4. The seasonal technique automatically summates the multiple $S$-statistics of the Mann–Kendall test.

$$S_k = \sum_{i=1}^{n} S_i \tag{9.4}$$

where $S_k$ = seasonal Kendall $S$-statistic
$S_i$ = Mann–Kendall $S$-statistic



©2007 AwwaRF. All Rights Reserved.

### *Theil–Sen Estimator of Slope*

Although the suite of Kendall techniques identifies trends that exist with time given an associated level of confidence, these models do not capture the magnitude of change within the data in time. Two techniques provided by Theil (1950) and Sen (1968) estimate the magnitude of change for a given parameter through time. The two techniques are quite similar, and both offer excellent predictions as to the magnitude of increase and decrease with time. Equation 9.5 presents the Theil line, which captures the magnitude of change as predicted by the Mann–Kendall technique.

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 \times x \tag{9.5}$$

where $Y$ = resulting (predicted) parameter value
$\beta_0$ = estimator of intercept
$\beta_1$ = estimator of slope
$x$ = time interval associated with the data

The $\beta_0$ value or intercept of the Theil line is calculated by Equation 9.6:

$$\hat{\beta}_0 = Y_{\text{med}} - \beta_1 \times \hat{X}_{\text{med}} \tag{9.6}$$

where $Y$ = median value of all parameter measurements within the original data set
$X$ = median value of all time measurements within the original data set
$\beta_1$ = estimator of slope

Likewise, the $\beta_1$ value or estimator of slope is calculated by Equation 9.7:

$$\hat{\beta}_1 = \underset{i<j}{\text{median}}\left(\frac{Y_j - Y_i}{X_j - X_i}\right) \tag{9.7}$$

where $Y$ = parameter measurement
$X$ = time measurement

In 1950, Theil proposed the result of the median of all possible pairs of observations that exist within the data set (Equation 9.7) as the estimator of slope. Sen modified the Theil slope estimator by forcing the estimator of slope to ignore any tied pairs within the data set, as shown in Equation 9.8.

$$\hat{\beta}_1 = \underset{\substack{i<j \\ X_i \neq X_j}}{\text{median}}\left(\frac{Y_j - Y_i}{X_j - X_i}\right) \tag{9.8}$$



©2007 AwwaRF. All Rights Reserved.

Essentially, the Sen modification follows the logic of the Mann–Kendall technique in that the independent variable must be unique and progress in chronological order. For this study, the Theil equation paired with the Sen modification was used in conjunction with the seasonal Kendall technique. The slope and intercept values provide a linear model of the existing trend with time. A positive slope corresponds with an increasing trend, and a negative slope corresponds with a decreasing trend.

### *Wilcoxon Rank-Sum Test*

The Mann–Kendall technique and Theil–Sen estimator of trend slope provide insight into the trend and magnitude of change, respectively, within an individual data set but fail to compare trends across various sets of data. To compare trends of organic carbon across many different sample locations, this study made use of a signed-rank test. Developed by Wilcoxon in 1945, the rank-sum test provides a direct comparison between multiple data sets. This statistical method operates via analysis of the difference between paired observations. Equations 9.9 and 9.10 present the Wilcoxon rank-sum test.

$$W = \sum_{i=1}^{n_1} (R_i) = \sum_{i=1}^{n_1} \text{rank}(d_{1i}) \tag{9.9}$$

$$d_i = x_i - y_i \tag{9.10}$$

where $W$ = resulting value of the test
$R$ = signed rank associated with each paired observation
$d_i$ = difference between pairs of observations
$x_i$, $y_i$ = different paired observations

The null hypothesis for this test assumes that the difference between the paired observations is zero. If the difference varies significantly from zero, the null hypothesis is rejected and the alternative hypothesis is accepted; the observations are said to be dissimilar.

## Results

The seasonal Kendall test and Theil estimator of slope were used to assess temporal trends in TOC across the continental United States. These statistical tests were applied to two data sets, a historical (1965–2000) and a modern (1995–2004) data set, to capture both long-term and modern trends, respectively. Intermediate trends—quarterly and annually—were determined through a division of the original data set, such that the seasonal Kendall test operated on a series of smaller data sets. Kendall's $\tau$ provided both direction (increasing and decreasing) and level of significance associated with each trend, whereas the Theil estimator of slope determined the magnitude of change with time. The Mann–Whitney rank-sum test was used to identify similarities in temporal trends across the United States. The use of Microsoft Excel and Visual Basic for Applications (VBA) eliminated the need for proprietary statistical software packages. The following sections present the results of these two analyses, discuss differences between the modern and historical data sets, and discuss the long-term variability of TOC across the country.



©2007 AwwaRF. All Rights Reserved.

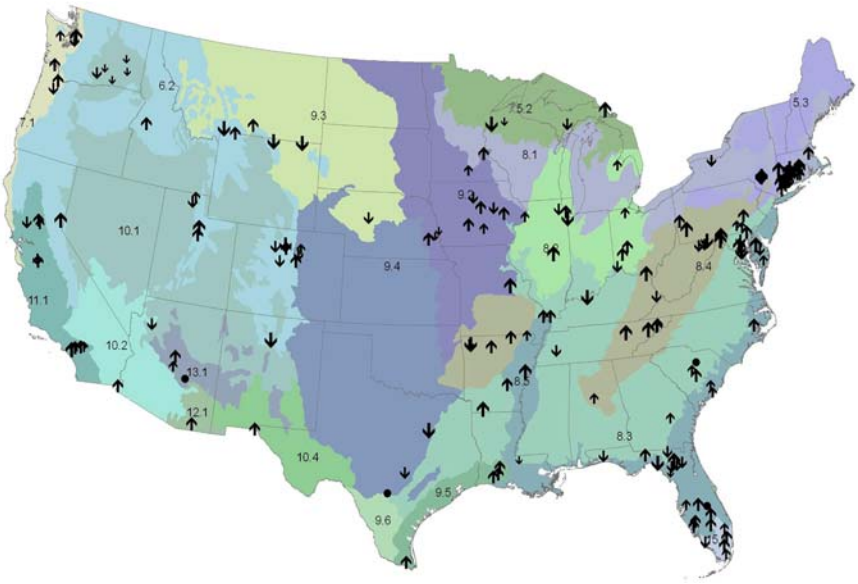



**Figure 9.9 Kendall's τ with ecoregions for modern data set**

### *Modern Trends*

The modern data set encompasses the period 1995–2004 and consists of 116,500 TOC measurements collected at 4,183 sampling locations. The number of sampling locations included in the analysis was reduced by the exclusion of short-term sampling stations (those that had been obtaining samples for less than 4 years) and sporadic sampling locations (those that had been obtaining fewer than four measurements per year). The removal of these stations eliminated bias toward particular seasons or years and reduced the potential for outliers. The seasonal Kendall test was thus applied to 188 stations with 20,840 TOC measurements.

Temporal trends in TOC varied in statistical significance, direction (Figure 9.9), and magnitude (Figure 9.10) across the nation. The significance of the trends and their magnitude varied considerably and no unique national trend was observed.

As shown in Figure 9.9, increasing trends occurred in south-central Florida, southern New England, along the Mississippi River, and along much of the West Coast. Decreasing trends occurred in northern Florida and the inland Pacific Northwest. Mixed results were observed along the Rocky Mountain region. The magnitude of the change of TOC is shown in Figure 9.10. The results are derived from the Theil estimator of slope and correspond to the confidence level of Figure 9.9. The magnitude of change was slightly positive within the Mississippi and Ohio river basins, whereas the magnitude of greatest increase occurred in south-central Florida. Moderate decreases in TOC occurred in the Great Lakes region and the Pacific Northwest. Ecoregion delineations were useful for understanding the spatial variability within the temporal trends. As shown in Figure 9.11, the spatial similarities within ecoregions were typically greater than across the country. The results shown in Figure 9.11 indicate closer spatial agreement (trend similarities) within ecoregions 5.2 and 15.4. The results also present trend dissimilarities within many ecoregions, such as 5.3, 8.5, and 9.2.



©2007 AwwaRF. All Rights Reserved.

**Figure 9.10 Theil estimator of slope for modern data set**

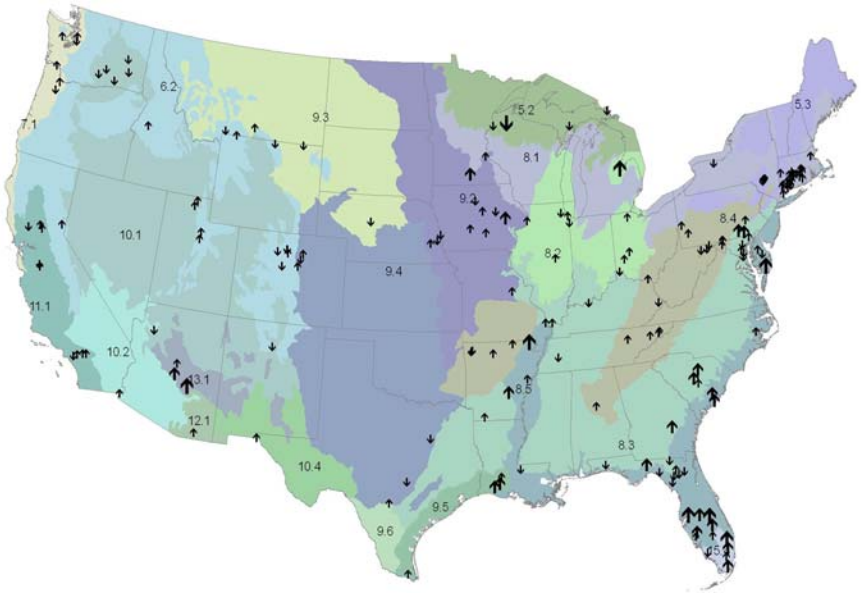



**Figure 9.11 Theil estimator of slope with ecoregions for modern data set**

Table 9.3 presents the modern seasonal Kendall results for ecoregions 5.2 and 5.3 as examples. The $\beta_1$ value, magnitude of change in trend, has the units of milligrams per liter per day. The values of $\beta_1$ indicate relatively small changes with time at each sampling station across the country. The results of the seasonal Kendall's $\tau$ (Figure 9.9), Theil estimates of slope (Figure 9.10), and Table 9.3 do not identify strong similarities in temporal trends between ecoregions (Figure 9.11). The Mann–Whitney rank-sum test was used to examine spatial trends between ecoregions for the



©2007 AwwaRF. All Rights Reserved.

**Table 9.3**
**Sample of seasonal Kendall results and Theil estimator of slope**

| Ecoregion | Station | Count | *S*-Statistic | Kendall's τ | $\beta_1$ (mg/L-day) | $\beta_0$ |
|---|---|---|---|---|---|---|
| 5.2 | 5331833 | 32 | –164 | –0.0827 | –0.0042 | 154.6363 |
| 5.2 | 5333500 | 81 | –6 | –0.0005 | –0.0001 | 8.3536 |
| 5.2 | Bay | 29 | 154 | 0.0948 | 0.0058 | –208.8254 |
| 5.2 | Chippewa | 25 | 0 | 0.0000 | –0.0001 | 4.0266 |
| 5.2 | Delta | 36 | –52 | –0.0206 | –0.0039 | 161.8651 |
| 5.3 | 1198125 | 69 | 38 | 0.0040 | 0.0000 | 4.4000 |
| 5.3 | 1434021 | 158 | 3,835 | 0.0773 | 0.0009 | –29.7996 |
| 5.3 | 1434070 | 114 | 333 | 0.0129 | 0.0000 | 2.5500 |
| 5.3 | 1434072 | 151 | –1,049 | –0.0232 | –0.0005 | 19.4425 |
| 5.3 | 1434073 | 181 | –2,012 | –0.0309 | –0.0004 | 17.5659 |
| 5.3 | 1434076 | 226 | 3,225 | 0.0317 | 0.0004 | –10.8537 |
| 5.3 | 1434080 | 271 | 347 | 0.0024 | 0.0000 | 1.4000 |
| 5.3 | 1434084 | 228 | 2,438 | 0.0236 | 0.0002 | –3.5310 |
| 5.3 | 1434086 | 273 | 242 | 0.0016 | 0.0000 | 1.2000 |
| 5.3 | 1434087 | 157 | –852 | –0.0174 | –0.0002 | 7.3634 |
| 5.3 | 1434092 | 226 | –519 | –0.0051 | –0.0002 | 7.8498 |
| 5.3 | 1434110 | 106 | 1,223 | 0.0549 | 0.0001 | –1.7689 |
| 5.3 | 1434112 | 122 | 593 | 0.0201 | 0.0000 | 1.6000 |
| 5.3 | 1435000 | 159 | –1,239 | –0.0247 | –0.0006 | 21.6077 |
| 5.3 | 143407250 | 121 | 550 | 0.0189 | 0.0000 | 1.7000 |

modern data set. Table 9.4 presents a summary of the Mann–Whitney test. The *p*-values within the table indicate the level of agreement between ecoregions, where values closer to 0 signify close agreement in trend and values close to 1 represent little or no agreement. The positive trend within ecoregion 8.1 was quite similar to ecoregions 9.4, 9.5, and 15.4. Likewise, the positive trend within ecoregion 9.5 was similar to ecoregions 5.3, 6.2, 7.1, 8.1, and 8.4. However, the results within Table 9.4 indicate that few similarities in modern trends exist among ecoregions.

A series of covariance analyses were performed to determine intra-ecoregion similarities. Table 9.5 presents one sample of the many covariance tables. The cross-products of the covariance test differ from the *p*-values of the Mann–Whitney test. The cross-products values represent the level of agreement among stations where values significantly above 1 or below –1 indicate high agreement. The results of the covariance analysis yielded mixed results. Ecoregions 13.1 and 15.4 displayed many similarities among sampling stations, whereas ecoregions 5.3, 8.5, 9.2, and 10.1 displayed few similarities among sampling stations.



©2007 AwwaRF. All Rights Reserved.

## Table 9.4
## Results of the Mann–Whitney rank-sum test for modern data set*

| Ecoregion | 5.2 | 5.3 | 6.2 | 7.1 | 8.1 | 8.2 | 8.3 | 8.4 | 8.5 | 9.2 | 9.3 | 9.4 | 9.5 | 9.6 | 10.1 | 10.2 | 10.4 | 11.1 | 12.1 | 13.1 | 15.4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5.2 | 0 | | | | | | | | | | | | | | | | | | | | |
| 5.3 | 0.229 | 0 | | | | | | | | | | | | | | | | | | | |
| 6.2 | 0.65 | 0.56 | 0 | | | | | | | | | | | | | | | | | | |
| 7.1 | 0.41 | 0.9 | 0.799 | 0 | | | | | | | | | | | | | | | | | |
| 8.1 | 0.096 | 0.02 | 0.01 | 0.03 | 0 | | | | | | | | | | | | | | | | |
| 8.2 | 0.27 | 0.41 | 0.26 | 0.396 | 0.72 | 0 | | | | | | | | | | | | | | | |
| 8.3 | 0.21 | 0.23 | 0.13 | 0.396 | 0.89 | 0.83 | 0 | | | | | | | | | | | | | | |
| 8.4 | 0.22 | 0.49 | 0.24 | 0.57 | 0.05 | 0.71 | 0.35 | 0 | | | | | | | | | | | | | |
| 8.5 | 0.21 | 0.28 | 0.27 | 0.45 | 0.42 | 0.43 | 0.44 | 0.34 | 0 | | | | | | | | | | | | |
| 9.2 | 0.38 | 0.46 | 0.45 | 0.39 | 0.99 | 0.97 | 0.796 | 0.69 | 0.55 | 0 | | | | | | | | | | | |
| 9.3 | 0.73 | 0.88 | 0.79 | 1 | 0.28 | 0.55 | 0.41 | 0.57 | 0.51 | 0.695 | 0 | | | | | | | | | | |
| 9.4 | 0.83 | 0.18 | 0.49 | 0.3 | 0.006 | 0.12 | 0.1 | 0.07 | 0.26 | 0.14 | 0.71 | 0 | | | | | | | | | |
| 9.5 | 0.19 | 0.01 | 0.005 | 0.01 | 0.002 | 0.033 | 0.11 | 0.007 | 1 | 0.08 | 0.06 | 0.0195 | 0 | | | | | | | | |
| 9.6 | 0.67 | 0.18 | 0.14 | 0.21 | 0.23 | 0.56 | 0.64 | 0.27 | 1 | 0.67 | 0.8 | 0.23 | 0.8 | 0 | | | | | | | |
| 10.1 | 1 | 0.28 | 0.45 | 0.52 | 0.07 | 0.27 | 0.08 | 0.11 | 0.29 | 0.5 | 0.83 | 0.89 | 0.01 | 0.6 | 0 | | | | | | |
| 10.2 | 0.67 | 1 | 0.66 | 1 | 0.32 | 0.84 | 0.81 | 0.87 | 0.8 | 0.88 | 1 | 0.53 | 0.4 | 1 | 0.6 | 0 | | | | | |
| 10.4 | 0.67 | 0.37 | 0.38 | 0.45 | 0.77 | 1 | 1 | 0.52 | 0.93 | 1 | 0.8 | 0.23 | 0.4 | 1 | 0.6 | 1 | 0 | | | | |
| 11.1 | 0.36 | 0.19 | 0.13 | 0.22 | 1 | 1 | 0.68 | 0.496 | 0.49 | 1 | 0.6 | 0.06 | 0.03 | 0.4 | 0.297 | 0.8 | 1 | 0 | | | |
| 12.1 | 0.67 | 0.27 | 0.38 | 0.21 | 1 | 1 | 0.75 | 0.34 | 0.93 | 1 | 0.8 | 0.23 | 0.4 | 1 | 0.6 | 1 | 1 | 1 | 0 | | |
| 13.1 | 0.25 | 0.07 | 0.05 | 0.15 | 0.22 | 0.18 | 0.18 | 0.09 | 0.68 | 0.19 | 0.15 | 0.07 | 0.86 | 1 | 0.036 | 0.64 | 1 | 0.21 | 1 | 0 | |
| 15.4 | 0.15 | 0.02 | 0.01 | 0.04 | 0.006 | 0.06 | 0.15 | 0.02 | 0.68 | 0.1 | 0.11 | 0.03 | 1 | 0.75 | 0.05 | 0.5 | 0.5 | 0.03 | 0.5 | 0.83 | 0 |

*Shaded areas represent positive trends within ecoregions.



©2007 AwwaRF. All Rights Reserved.

**Table 9.5**
**Sample of results of covariance analysis within ecoregion 11.1, modern data set**

| | 11051500 | 11060400 | 11073495 | 11075610 | 11273500 | 11274538 | 11421500 | 11424000 | 11451800 |
|---|---|---|---|---|---|---|---|---|---|
| 11051500 | 0.917 | 0.256 | 1.481 | –0.093 | 0.205 | –2.324 | 0.182 | 0.097 | –0.049 |
| 11060400 | 0.256 | 65.416 | 12.060 | 2.414 | –1.273 | –6.478 | –0.789 | 1.185 | 0.529 |
| 11073495 | 1.481 | 12.060 | 24.428 | –0.110 | –1.325 | –7.788 | 0.036 | 1.080 | –0.090 |
| 11075610 | –0.093 | 2.414 | –0.110 | 0.883 | –0.256 | –1.831 | –0.017 | 0.106 | –0.152 |
| 11273500 | 0.205 | –1.273 | –1.325 | –0.256 | 7.628 | –2.691 | –0.050 | –0.251 | 0.064 |
| 11274538 | –2.324 | –6.478 | –7.788 | –1.831 | –2.691 | 50.699 | –0.373 | –1.264 | 0.710 |
| 11421500 | 0.182 | –0.789 | 0.036 | –0.017 | –0.050 | –0.373 | 0.210 | –0.001 | 0.043 |
| 11424000 | 0.097 | 1.185 | 1.080 | 0.106 | –0.251 | –1.264 | –0.001 | 0.313 | –0.045 |
| 11451800 | –0.049 | 0.529 | –0.090 | –0.152 | 0.064 | 0.710 | 0.043 | –0.045 | 0.428 |

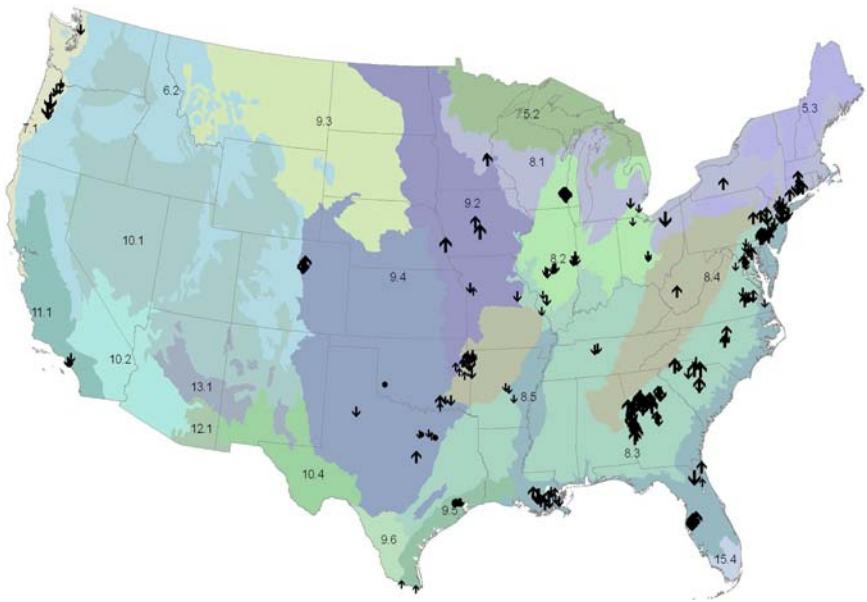



**Figure 9.12 Kendall's τ with ecoregions for historical data set**

### *Historical Trends*

The historical data set encompasses the period 1965–2000 and consists of 218,000 TOC measurements collected at ~6,000 sampling locations. The number of sampling locations included in the analysis was reduced by the exclusion of short-term sampling stations (those that had been obtaining samples for less than 15 years) and sporadic sampling locations (those that had been obtaining fewer than four measurements per year). The removal of more than 5,500 stations eliminated bias toward particular seasons or decades and reduced the potential for outliers. The seasonal Kendall test was thus applied to 307 stations with 35,750 TOC measurements.

Historical temporal trends in TOC varied in direction, statistical significance (Figure 9.12), and magnitude (Figure 9.13). Similar to the modern analysis, the significance of the trends and magnitude varied and no unique trend was observed. As shown in Figure 9.13, decreasing trends occurred in the Pacific Northwest, the Ohio River Valley, and Southern California. Increasing



©2007 AwwaRF. All Rights Reserved.

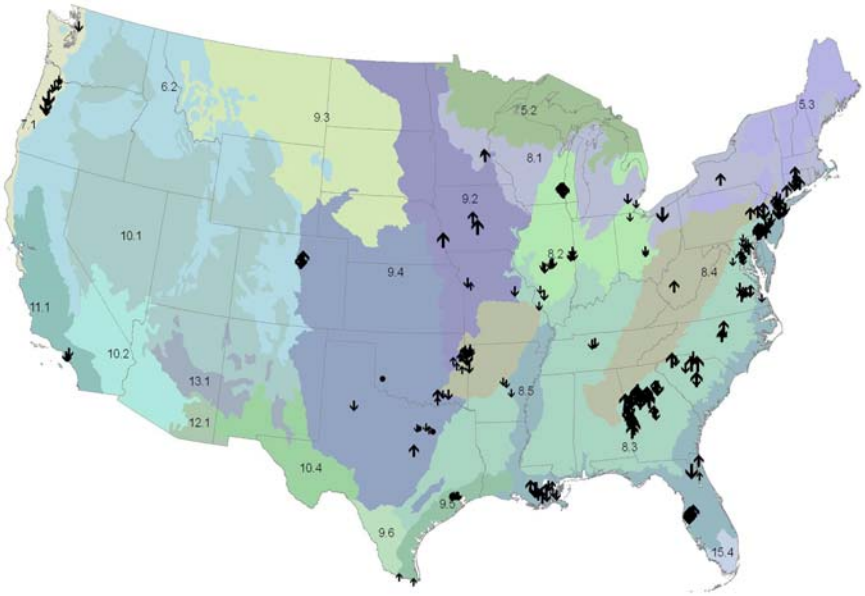



**Figure 9.13 Theil estimator of slope with ecoregions for historical data set**

trends occurred in the Plains states and Appalachia. Mixed results were observed along the East and Gulf coasts as well as inland from Georgia and South Carolina. Decreases in TOC were observed within the Ohio and Tennessee river basins and across Texas. In the Pacific Northwest and southern California, seasonal analysis indicated moderate decreases in TOC. Mixed results were observed within the Mid-Atlantic states. Ecoregion delineations were moderately useful for understanding the spatial variability within the temporal trends. Spatial similarities within ecoregions were typically greater than across the country.

The Mann–Whitney test was employed to determine historical spatial trends and inter-ecoregion similarities. The resulting p-values in Table 9.6 indicate the level of agreement between ecoregions. The increasing trend in ecoregion 9.5 is similar to the trends associated with most ecoregions. Likewise, the decreasing trend in ecoregion 11.1 was similar to ecoregions 8.1, 8.2, and 8.3. With these exceptions, the results within Table 9.6 indicate that few similarities in spatial trends existed and the level of inter-ecoregion agreement was extremely low. Covariance analyses were not conducted for the historical data set.

The historical results of the seasonal Kendall test (Figure 9.12) and the Theil estimates of slope (Figure 9.13) indicate variability across the country. These results were similar to the results of the modern analysis. No unique national trend was observed in either data set, few similarities existed between ecoregions, and the covariance analysis of the historical data was expected to yield results similar to those from the modern data set.

## SUMMARY AND CONCLUSIONS

A database of TOC data for the United States was created as part of the project. The database includes modern and historical information from across the country obtained from USGS and USEPA. Inclusion of all the nationally available USEPA data was beyond the scope of this project. Instead, data for select areas of the country were obtained. These regions were identified by 8-digit USGS-defined HUCs for the watersheds associated with the 52 utilities included in this



©2007 AwwaRF. All Rights Reserved.

**Table 9.6**
**Results (*p*-values) of the Mann–Whitney rank-sum test for historical data set***

| Ecoregion | 5.3 | 6.2 | 7.1 | 8.1 | 8.2 | 8.3 | 8.4 | 8.5 | 9.2 | 9.4 | 9.5 | 11.1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5.3 | 0 | | | | | | | | | | | |
| 6.2 | 0.279 | 0 | | | | | | | | | | |
| 7.1 | 0.013 | 0.046 | 0 | | | | | | | | | |
| 8.1 | 0.088 | 0.672 | 0.005 | 0 | | | | | | | | |
| 8.2 | 0.29 | 0.91 | 0.06 | 0.66 | 0 | | | | | | | |
| 8.3 | 0.02 | 0.19 | 0.05 | 0.16 | 0.404 | 0 | | | | | | |
| 8.4 | 0.026 | 0.263 | 0.105 | 0.182 | 0.585 | 0.823 | 0 | | | | | |
| 8.5 | 0.143 | 0.523 | 0.237 | 0.257 | 0.438 | 0.939 | 0.980 | 0 | | | | |
| 9.2 | 0.280 | 0.885 | 0.005 | 0.284 | 0.406 | 0.022 | 0.062 | 0.211 | 0 | | | |
| 9.4 | 0.149 | 0.685 | 0.077 | 0.961 | 0.916 | 0.331 | 0.382 | 0.598 | 0.338 | 0 | | |
| 9.5 | 0.017 | 0.006 | 0.002 | 0.001 | 0.000 | 0.000 | 0.003 | 0.000 | 0.002 | 0.001 | 0 | |
| 11.1 | 0.100 | 0.049 | 0.011 | 0.008 | 0.012 | 0.006 | 0.018 | 0.016 | 0.031 | 0.042 | 0.613 | 0 |

*Shaded areas represent similarities among trends within ecoregions.

phase of the study. These 52 utilities were selected in part so that data were collected for at least two utilities, and thus two HUCs, per ecoregion. Data were downloaded from 160 HUCs that contained the 52 utility watersheds. Thus, data acquisition from USEPA for the continental United States is currently incomplete. A large body of additional data, scattered among numerous private utilities, local collection agencies, civic organizations, and individual researchers, is not readily available. A subset of this less-readily-available data was obtained through personal communications with individuals from the set of 52 selected utilities. Raw water data (TOC, DOC, SUVA, $UV_{254}$, metals, general organics), hydrologic data (reservoir characteristics, streamflow patterns), and GIS boundaries and land use data were requested from these utilities, although not all were able to provide data. Although the database is incomplete, the data provided the best available data set for evaluating national spatial and temporal trends in TOC/DOC. TOC data were more prevalent in the 1970s and 1980s.

Variability among ecoregions was also explored, both during the period of record and by decade, using quartile statistics. TOC values varied considerably from ecoregion to ecoregion. However, inter-ecoregion variability appeared relatively constant on a decadal basis. One notable exception was ecoregion 15.4, in which median TOC concentrations apparently decreased from about 18 mg/L in the 1970s and 1980s to about 12 mg/L in 2000–2004. Decreasing amounts of data, however, make it difficult to lend credence to this apparent trend. Although comparison of TOC data on an ecoregion basis is convenient, considerable inter-ecoregion variability can exist. Anthropogenic factors are a likely potential cause of inter- and intra-ecoregion variability, although this has not been explored further to date. The ecoregion-based TOC data were also analyzed on a seasonal basis to assess trends. No seasonal trends were evident, suggesting that inter-ecoregion variability is greater than interseasonal variability. Data were generally more available during the spring through fall seasons than during the winter.

The seasonal Kendall test and Theil estimator of slope were used to assess temporal trends in TOC across the continental United States. These statistical tests were applied to historical (1965–2000) and modern (1995–2004) data sets to capture both long-term and modern trends.



©2007 AwwaRF. All Rights Reserved.

The stations included in each analysis varied because of data requirements. An almost equal number of increasing and decreasing trends in TOC were observed in the modern data set. Decreases typically appeared to occur in less-populated regions of the country, whereas increases occurred in the more-populated regions, although no official evaluation of this has been completed to date. Temporal trends were not consistent either within or among ecoregions, suggesting that factors other than variability in native vegetation and soils are important determinants of both TOC levels and trends.

Future work is necessary to focus on individual locations where long-term trends have been observed and to compare the historical and modern results. Trends in ancillary data, such as population- or watershed-specific land use change, may be able to offer additional insight into the underlying causes of observed change. In addition, data for additional stations, particularly long-term data collected by private utilities, may provide additional insight regarding changes in TOC levels and transport across the nation. The next chapter provides a detailed look at land use and other effects on TOC and DBP levels for select utility watersheds.



©2007 AwwaRF. All Rights Reserved.

# CHAPTER 10
# UTILITY DATA ANALYSIS AND MODELING

## OVERVIEW

Ecoregions divide the United States in a manner independent of common geopolitical boundaries (Omernik and Bailey 1997). In chapter 9, it was hypothesized that these divisions are one factor underlying TOC variability across the country. Common land use characteristics often result in similar trends and behavior, whereas differences in trend and behavior result from anthropogenic activity and differing land use. In either situation, land use characteristics significantly influence and affect NOM within a watershed and are of particular interest in this study. This chapter explores whether a more detailed look at land use, and thus anthropogenic activity, is adequate to explain observed variability across the United States.

"Land use" is defined as the various ways in which land may be employed or occupied in terms of human activity, whereas "land classifications" (categories) describe the extent of land use for agriculture, urban development, forests, wetlands, and other land types. The idea of land use classification was created by Francis Marschner and began in the mid-1940s. Marschner's descriptions, via aerial photographs, laid the foundation for classification in the ensuing decades. Through the 1960s, various state, federal, and private agencies classified local and regional land use/land cover (LULC) characteristics, but these varied from region to region. In 1971, a uniform system of LULC classification, called the Anderson classification system, was defined. The broad categories of land use presented by Anderson successfully describe anthropogenic activity and influence within a particular watershed or across a geographical region. The Anderson classification system more accurately describes LULC within a single catchment than does the general description provided by ecoregion descriptions.

Land use directly affects carbon flux from the terrestrial environment to the aquatic ecosystem, both in terms of source and transport pathway. Anthropogenic activity within water supply catchments further influences fate and transport of organic carbon. In addition, because the magnitude of human activity varies across the nation and changes in activity are difficult to discern (Houghton and Hackler 2000), land use change is often used as a surrogate.

The remainder of this chapter is organized as follows. First, the results of select past studies exploring the relationship between export and land use are summarized. Next, empirical modeling methods are briefly described. These methods are combined with data obtained from several utilities to develop national-level period of record and seasonal empirical models. Two site-specific DBP and organic carbon models are then presented, as well as the results from site-specific analyses of microbial communities. Finally, the presented findings are summarized and discussed.

## SELECT RESULTS OF PAST STUDIES

Several studies have used empirical analyses to relate water quality parameters such as total phosphorus, total nitrogen, and DBPs to land use, precipitation, and streamflow. Most of these studies have concluded that catchment-scale land use is a strong predictor of nutrient concentration in streams and rivers (Gergel et al. 2002). The strength of these relationships tend to increase with the quality and representativeness of the underlying data. The relationships may be



©2007 AwwaRF. All Rights Reserved.

used to infer the export/generation of the dependent variable (nutrient) resulting from the presence of the independent variables (land use types), also known as export coefficients.

A national survey of total nitrogen and total phosphorus was conducted in 1993 by Schonter and Novotny (1993). The investigators performed regression analyses using land use, total phosphorus, and total nitrogen data. The resulting relationships suggested that the presence of agriculture had a large impact on both phosphorus and nitrogen export, though wetland areas had little impact. The strong correlation between agriculture and these nutrients is logical given typical annual application of fertilizers. Furthermore, large exports of carbon were closely associated with wetland areas (Schonter and Novotny 1993).

Several studies have shown that wetlands are an important source of NOM. Mulholland and Kuenzler (1979) found that watersheds containing forested wetlands yielded greater export of organic carbon per unit runoff. This increased export was attributed to increased leaching of DOM and increased evapotranspiration (Mulholland and Kuenzler 1979). The resulting empirical models were dissimilar to those for total phosphorus and total nitrogen presented by Schonter and Novotny (1993). Additional work by Gergel, Turner, and Kratz (1999) supports a unique empirical relationship for organic carbon. Gergel and colleagues (1999) also found that wetlands and adjacent soil layers contributed vast amounts of DOC to lakes and streams despite comprising relatively small percentages of the catchment areas. A study by Waldron and Bent (2001) focused on organic carbon export within three urban watersheds. Each catchment, serving one central water treatment plant, was comprised of several land use types including wetlands, agriculture, varying urban and residential development, forested land, and light industry. The authors presented a compilation of export coefficients from sub-basins within the different water supply catchments. Similar to the findings of Mulholland (1981), Mulholland and Kuenzler (1979), and other studies, organic carbon export was most closely correlated with wetland areas. High correlation was also found between organic carbon and forested land, pastures, and medium-density residential areas. The results of this study indicate that both natural and anthropogenic activity impact organic carbon export within urban catchments and provide guidance for implementing a water quality monitoring program.

Regression analysis has also been used to identify relationships between DBPs and water treatment processes. Widespread application of chlorine as a disinfectant and the ubiquity of NOM result in the potential formation of DBPs within the water treatment and distribution system. Although TOC may not be the best surrogate of NOM, it is the most widely used. Edzwald, Becker, and Wattier (1985) found that UV absorbance was a good surrogate for TOC and THMFP when they examined TOC removal across a direct filtration process for two water sources. Through linear regression, a strong correlation among UV, TOC, and THMFP measurements in two different water sources was found. McClellan et al. (1996) also used regression analysis to relate drinking water treatment processes and organic DBP precursor removal. In contrast to the earlier study by Edzwald, Becker, and Wattier (1985), McClellan et al. (1996) examined three treatment processes: direct filtration, ozone with GAC, and conventional anthracite over sand with prefiltration ozone. The raw water characteristics for the two source waters were similar, but the resulting (finished) levels of DBPs and DOC varied significantly. Empirical relationships were able to accurately describe the effect of each treatment process on constituent removal and DBP formation. Although these study results were site specific, similar relationships have been observed in other studies. For example, Amy and co-workers (Amy, Chadik, and



©2007 AwwaRF. All Rights Reserved.

Chowdhury 1987; Amy, Minear, and Cooper 1987) developed a series of linear and nonlinear relationships to best describe and predict TTHM formation from raw water. The original models were adjusted following initial experimentation to more accurately describe TTHM production and remove site-specific constraints. The final regression analyses included raw water data from various water sources across the country.

## EMPIRICAL METHODS

The diversity of land use within many watersheds presents ambiguity as to the exact origins of NOM within the ecosystem as well as to the influence on organic nutrient fluxes and export. Statistical modeling provides one means of describing organic carbon sources and transport within a watershed. Although regression analysis can be used to develop relationships between land use characteristics and NOM export, they provide little insight into the underlying transport and reaction mechanisms. However, empirical relationships, though simple, can depict the occurrence and movement of NOM through a watershed and ultimately into a treatment facility, and thus provide a means of evaluating the net result of land management strategies. The appropriate regression analyses and proper model construction vary based on the available data and their distribution.

For this study, data provided by 10 surveyed water utilities were used to develop national period-of-record and seasonal empirical relationships between median TOC concentration and land use. These 10 utilities provided the most extensive data sets, including LULC characteristics within the source water catchment, TOC measurements, and limited DBP precursor information. For this analysis, either the TOC or DBP precursor data were assumed to be the dependent variables and various LULC characteristics were assumed to be the independent variables. Single regression techniques seem inappropriate given multiple independent variables; therefore, multiple regression techniques were implemented to analyze the relationships between organic matter or DBP precursors and LULC characteristics.

Multiple linear regression (MLR) methods assume linearity between $x$ and $y$ variables, use of representative data within the model, constant variance of the residuals, independence of residuals, and residuals with a normal distribution. MLR models resemble the general form presented in Equation 10.1. Each model differs given the number of independent variables. The general MLR model shown in Equation 10.1 provides the basis for all MLR model construction in this study.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \tag{10.1}$$

where $y$ = response variable
$\beta_0$ = intercept
$\beta_1$ = slope coefficient of the first explanatory variable, $x_1$
$\beta_2$ = slope coefficient of the second explanatory variable, $x_2$
$\beta_k$ = slope coefficient of the $k$th explanatory variable, $x_k$
$\varepsilon$ = error associated with the model

The regression procedure that was employed influences the performance of the regression model and, ultimately, regression results. Multiple regression methods may be applied through



©2007 AwwaRF. All Rights Reserved.

use of three different stepwise* procedures. Each procedure evaluates the significance of an explanatory variable via partial F-tests or *t*-tests.

The first method, "forward selection," begins with an intercept value. Explanatory variables are added in a stepwise fashion to the regression model and evaluated for significance. This method excludes variables with low or no significance whereas those variables with high significance remain in the model. Helsel and Hirsch (1992) have identified two flaws associated with this method: (1) the model includes coefficients not significantly different from zero; and (2) the model removes variables that independently do not influence the dependent variable, though a combination of these variables significantly influences the dependent value.

The second method, "backward elimination," begins with every explanatory variable present within the MLR. The model is adjusted via elimination of the least significant variable (lowest *t*-value) with each iteration. The procedure concludes when all remaining variables within the model are determined to be significant. Similar to the forward selection, Helsel and Hirsch (1992) indicate that this model does not guarantee significant variables within the final model nor do groups of significant variables remain.

Finally, a stepwise regression combines the processes of forward and elimination regression. This method operates via addition and subtraction of explanatory variables from the MLR model while monitoring the significance of all explanatory variables—those within and excluded from the model. The stepwise regression successfully eliminates insignificant variables from the final model but does not evaluate all possible combinations of explanatory variables; thus, it is susceptible to eliminating groups of significance (Helsel and Hirsch 1992). Despite this drawback, the stepwise regression is the most appropriate technique. The regression capitalizes on the strengths of the two preceding methods and yields a satisfactory model.

In this chapter, the stepwise procedure was implemented to produce multiple linear regression models relating LULC characteristics (explanatory variables) to TOC concentrations (dependent variables). These regression models offered insight to the empirical relationships between TOC and land use characteristics. Multiple regression modeling became an additional tool with which to describe the behavior and variability of TOC across the country.

Multicollinearity may influence the outcome of the regression model yet is commonly found to some extent in most multiple regression models (Helsel and Hirsch 1992). "Multicollinearity" is defined as the condition where one or more explanatory variables are closely related to one or more other explanatory variables (Helsel and Hirsch 1992). The effects of multicollinearity include large slope coefficients, unrealistic slope sign, instability of coefficients, and nonagreement among automatic regression procedures. To combat this issue, data may be centered, variables eliminated, additional data collected, or ridge regression techniques may be used. "Ridge regression" is the reduced variance associated with slope estimates through the introduction of bias. Cross-correlation analyses were used to reduce the presence of multicollinearity through elimination of variables.

---

* Stepwise procedures are automated model selection methods in which the computer algorithm determines which model is preferred.



©2007 AwwaRF. All Rights Reserved.

**Table 10.1**
**TOC data for each utility within the regression analyses**

| Utility location | Period of record Median (mg/L) | Period of record SD* | Winter | Spring | Summer | Fall |
|---|---|---|---|---|---|---|
| | | | Median (mg/L) | | | |
| Tampa, Fla. | 13.52 | 9.69 | 5.73 | 4.69 | 22.41 | 17.53 |
| Des Moines River, Des Moines, Iowa | 4.08 | 1.67 | 3.84 | 4.3 | 4.183 | 4.145 |
| Racoon River, Des Moines, Iowa | 3.84 | 3.73 | 3.16 | 4.315 | 4.31 | 3.537 |
| Elgin, Ill. | 7.80 | 3.16 | 6.5 | 7.9 | 8.8 | 7.8 |
| Fort Worth, Texas | 5.08 | 0.34 | 5.075 | 5.2 | 5.02 | 1.965 |
| Jordan Valley, Utah | 2.15 | 1.82 | 2.06 | 2.11 | 2.28 | 6.03 |
| Newport News, Va. | 6.06 | 1.41 | 5.41 | 6.16 | 7.2 | 3 |
| Portland, Ore. | 1.89 | 0.30 | 1.84 | 1.76 | 1.955 | 4.91 |
| University Lake, Orange, N.C. | 6.05 | 1.33 | 4.84 | 5.55 | 6.685 | 2.14 |
| Baltimore, Md. | 2.60 | 1.38 | 2.2 | 2.5 | 3.1 | 6.33 |

*SD = standard deviation.

## NATIONWIDE MODEL FOR ORGANIC CARBON

### Available Land Use and TOC Data

A detailed data set is required to develop empirical relationships between TOC and LULC characteristics. TOC data collected within 10 watersheds across the country were used for this analysis. The data collected from each water supply catchment incorporated more than 5 years of TOC data, typically collected approximately once a month. Sampling locations were most often located at the intake to the water treatment facility and thus captured the combined effects of multiple tributary inputs and the influence of mixed surrounding LULC characteristics. The variability associated with the TOC data for each watershed is presented in Table 10.1. These 10 utilities provided a robust data set for which land use data was readily available.

MLR analyses were conducted to develop empirical relationships between TOC and LULC characteristics. The regression analyses were divided into two phases, a national analysis based on data collected from the 10 water supply catchments in Table 10.1 and an individual watershed analysis based on data from two additional watersheds. Each phase incorporated two analyses: the seasonal analyses developed relationships between seasonal changes in TOC and LULC characteristics, whereas the period-of-record analysis ignored seasonality and identified significant LULC influences on TOC. Variability in sampling frequencies among watersheds prevented analyses that are more detailed (e.g., annually or monthly). The results of the two analyses were then compared to identify the influence/effect of seasonality on the empirical models and behavior of TOC.

The national regression analyses developed empirical relationships among five broad LULC categories and TOC within the 10 water supply catchments. LULC characteristics differed across water supply catchments, and the numerous categories, presented by Anderson, prohibited



©2007 AwwaRF. All Rights Reserved.

exact comparisons among watersheds. General LULC characteristics such as forest, agriculture, wetlands, urban development, and water encompassed these more refined categories and created broad similarities among catchments.

### National Period-of-Record Analysis

The representative TOC values of this analysis were the median values of the data collected within each watershed. The median TOC values were related to the five general land use categories: agriculture, forested land, wetlands, urban development, and water. The national regression thus developed empirical relationships among these land use categories and TOC. Equation 10.2 presents the resulting empirical model for the national watershed survey.

$$TOC_{overall}(mg/L) = 2.71 + 0.0001\ \text{agriculture} - 0.0035\ \text{urban} + 0.0225\ \text{wetland} \qquad (10.2)$$

The resulting empirical model indicated that wetlands contributed the most TOC within each watershed. The relationship also revealed slight influence by agriculture and a negative, or decreasing, influence due to urban land use.

The empirical model was then used to generate predictions of TOC based on the land use composition (in square kilometers) of each utility. The accuracy of the empirical model was determined through a comparison between actual (measured) TOC concentration at the water treatment facility and predicted TOC values, as shown in Equation 10.3.

$$\%\ \text{error} = \frac{\text{predicted TOC} - \text{measured TOC}}{\text{measured TOC}} \times 100 \qquad (10.3)$$

where predicted TOC = value that is generated by the model
measured TOC = median TOC concentration collected at the utility intake

Figure 10.1 presents the results of this comparison.

The results shown in Figure 10.1 represent an overall empirical model error of 29%. The predictions of the empirical model were variable across the country. Large errors (>30%) were found at utilities in Newport News, University Lake, Fort Worth, and Elgin, whereas similarities between predicted TOC values and measured concentration were observed in Des Moines, Portland, Tampa, and Jordan Valley. The error associated with the empirical model was attributed to the representative value of TOC. A single average value of TOC across the period of record within each watershed did not adequately represent the underlying data. The variability of TOC across the period of record within each water supply catchment is shown in Figure 10.2. These box plots highlight the inability of a single value to represent a watershed across time. The erroneous predictions at Newport News, Fort Worth, and Elgin utilities, observed in Figure 10.1, correspond to large variability within the underlying TOC data (Figure 10.2). Period-of-record results thus suggest the need to incorporate temporally varying independent variables in the MLR to enable inclusion of the raw, rather than just summary, data in the analysis.

### Seasonal Analysis

The second phase of the national regression analysis determined whether accounting for seasonality could reduce the previous model errors. This regression analysis divided the available



©2007 AwwaRF. All Rights Reserved.

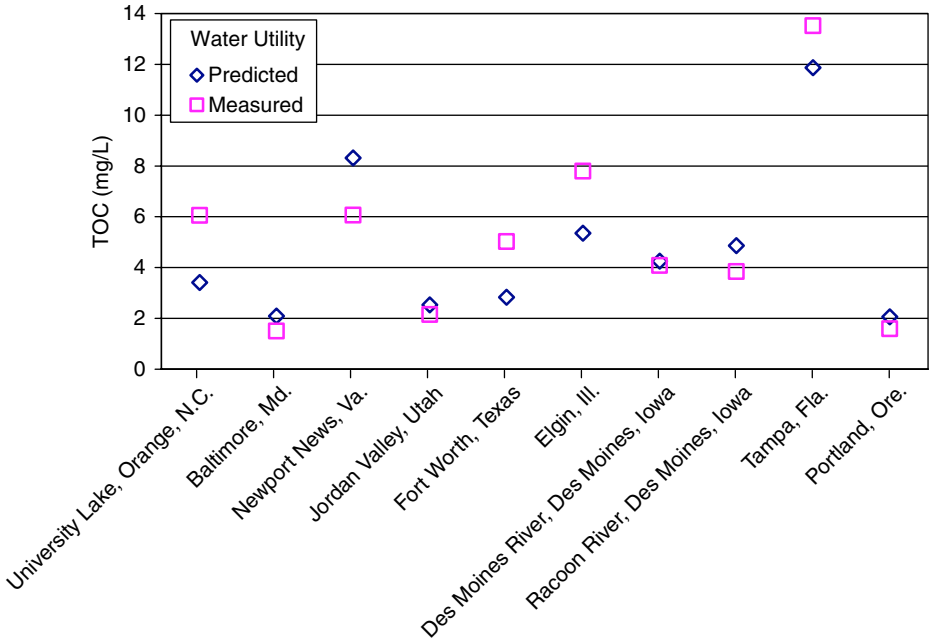



**Figure 10.1 National regression comparison between empirical TOC predictions and actual TOC concentrations**

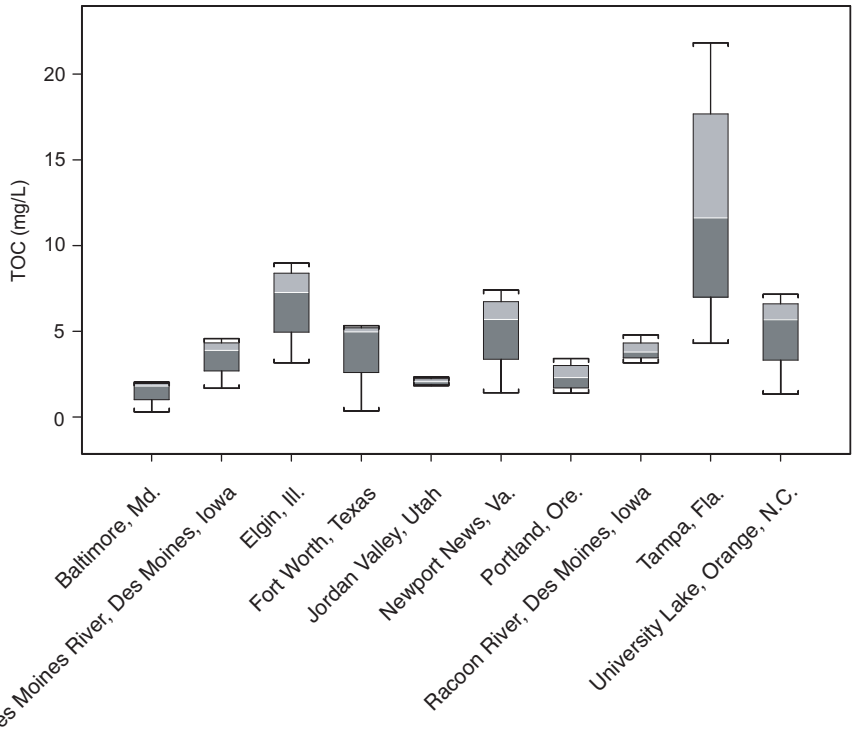



**Figure 10.2 Variability of available TOC data for each utility**



©2007 AwwaRF. All Rights Reserved.

data (Table 10.1) into seasonal (quarterly) averages across the period of record and generated a series of empirical models. The seasonal division provided more accurate descriptions of the relationships between TOC and LULC characteristics. Equations 10.4 to 10.7 present the resulting empirical models for each season.

$$TOC_{winter} = 0.0004 \text{ agriculture} + 0.0007 \text{ forest} + 0.011 \text{ wetland} \quad (10.4)$$

$$TOC_{spring} = 0.0006 \text{ agriculture} + 0.0008 \text{ forest} + 0.0101 \text{ wetland} \quad (10.5)$$

$$TOC_{summer} = 3.04 + 0.0419 \text{ wetland} - 0.009 \text{ water} - 0.005 \text{ urban} \quad (10.6)$$

$$TOC_{fall} = 3.03 + 0.0312 \text{ wetland} - 0.0059 \text{ water} - 0.0037 \text{ urban} \quad (10.7)$$

The seasonal empirical models indicated strong seasonal effects on relationships between land use and TOC. LULC characteristics influenced TOC differently each season. Agriculture, forested lands, and wetlands influenced TOC during the winter and spring seasons, though urban development, water sources, and wetlands influenced TOC during the summer and fall seasons. The average seasonal error was 28% and ranged between 25% and 32%. The results of the empirical models indicated that wetlands were the most influential LULC type nationally. This land use characteristic was present in each empirical model, and was most influential during the spring and winter seasons. Urban development, representative of most anthropogenic activity across the national analysis, also influenced TOC concentration within water supply catchments, but to a lesser extent than wetlands. The negative sign associated within the urban development coefficient indicated TOC was inversely proportional to this land use type; as urban development increased, TOC decreased.

The seasonal analysis did not yield results substantially different from the first analysis; accounting for seasonality alone was insufficient to noticeably improve model results. The errors associated with the empirical models most likely resulted from additional geographic differences among watersheds such as climate, soil type, vegetation, and impact of anthropogenic activity. Figures 10.3 to 10.6 present the accuracy of the seasonal relationships between land use and TOC. These figures present winter, spring, summer, and fall relationships, respectively.

The winter results presented in Figure 10.3 indicate that differences between measured TOC concentration and predicted TOC values were less then 0.5 mg/L in nearly 50% of the watersheds, and differences were less than 1.0 mg/L in 70% of the watersheds. Despite geographic differences among water supply catchments, the accuracy of the empirical relationships was relatively high during the winter. The accuracy decreased for empirical relationships during the spring months. Predictions of TOC concentration within the Des Moines, Newport News, and Jordan Valley catchments slightly affected the seasonal error, though the TOC predictions for the Elgin and Baltimore watersheds largely affected the seasonal error, where errors exceeded 40% and 85%, respectively. The seasonal TOC average within the watersheds did not capture the variability of the underlying data (Table 10.1) during the spring season. This fluctuation accounted for the large errors at the Baltimore and Elgin locations. Empirical relationships were the most accurate during the summer months, when the seasonal error was 25%. Figure 10.5 displays the close agreement between predictions of TOC and measured TOC concentrations. The accuracy of the summer empirical model is further improved when the prediction for Fort Worth is excluded. This



©2007 AwwaRF. All Rights Reserved.

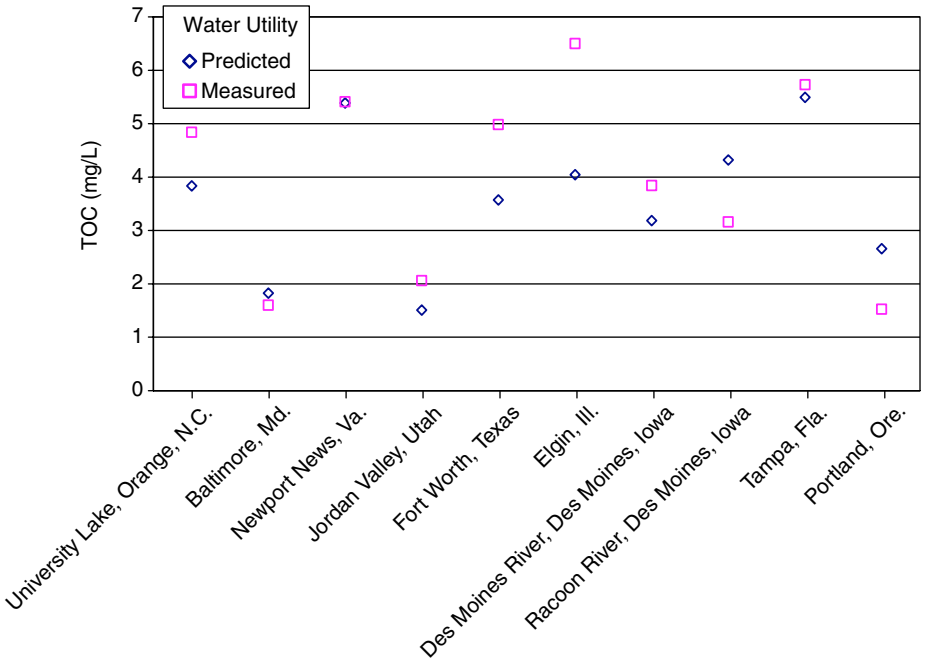



**Figure 10.3 National winter regression comparison between empirical TOC predictions and actual TOC concentration**

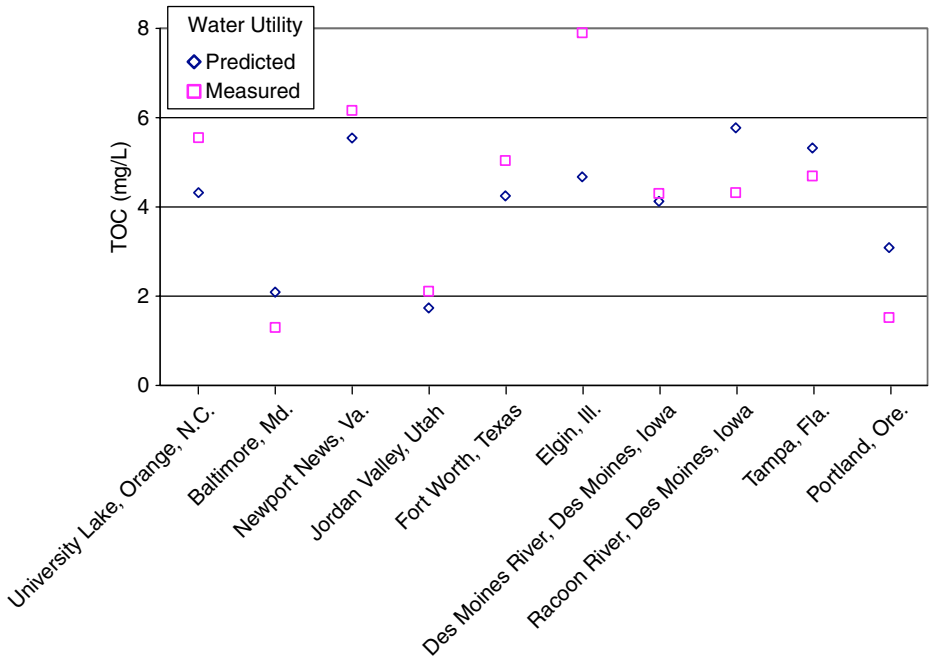



**Figure 10.4 National spring regression comparison between empirical TOC predictions and actual TOC concentration**



©2007 AwwaRF. All Rights Reserved.

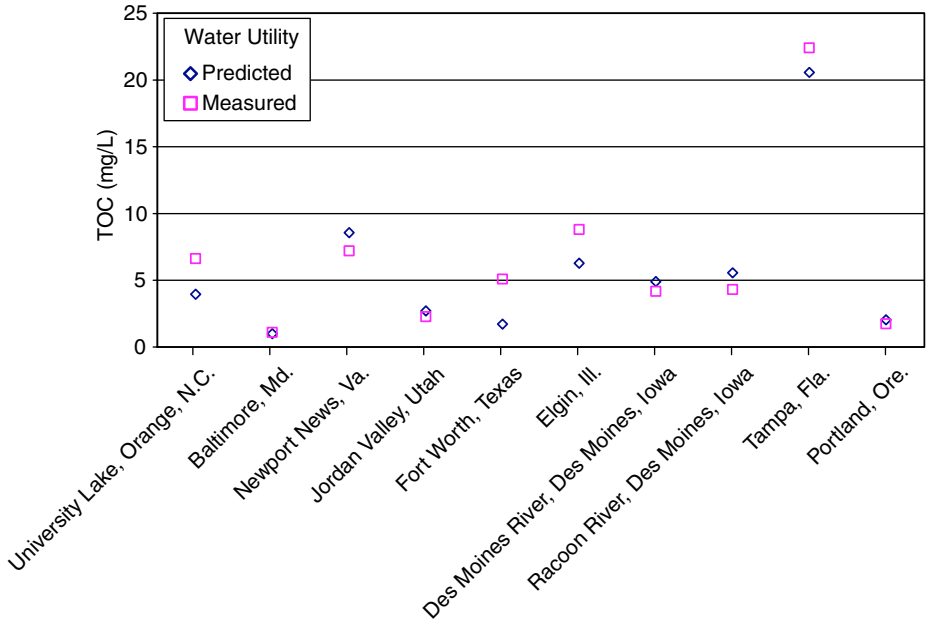



**Figure 10.5 National summer regression comparison between empirical TOC predictions and actual TOC concentration**

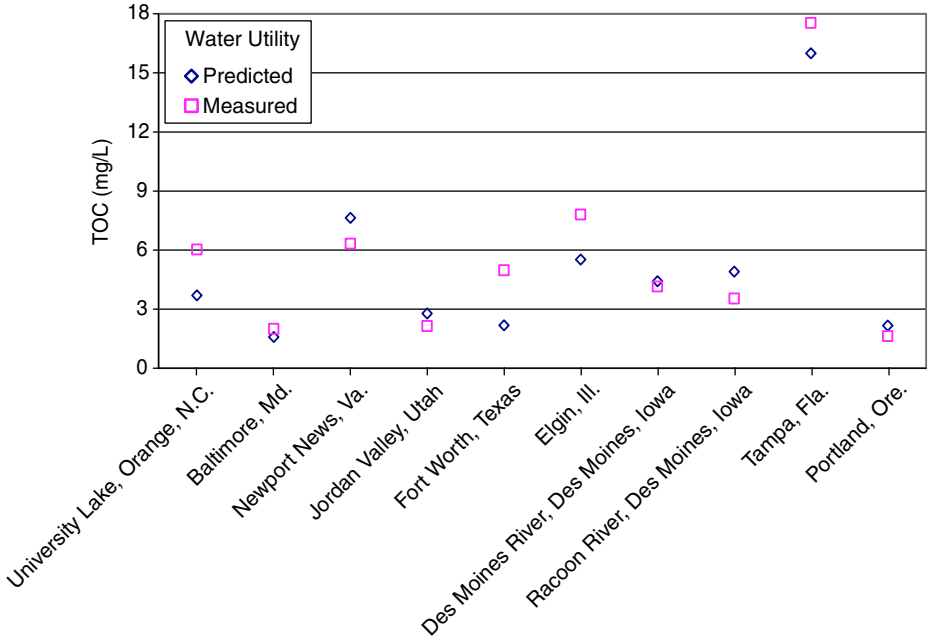



**Figure 10.6 National fall regression comparison between empirical TOC predictions and actual TOC concentration**



©2007 AwwaRF. All Rights Reserved.

prediction accounted for 20% of the model error. The variability of the underlying data during the summer was low and contributed to the relatively small errors between predicted TOC values and summer average TOC concentrations.

Variability of the underlying TOC data during the fall season (Figure 10.6) introduced error into fall prediction values. Most noticeable were disagreements between measured and predicted TOC values in Texas, Illinois, and North Carolina catchments. Without intimate knowledge of these water supply catchments, it was difficult to assess potential causes of error. The geographic differences among the three watersheds were obvious, and the conclusion was made that separate phenomena caused the observed variability within each basin.

The seasonal empirical models were similar to the original empirical model, but the resulting conclusions were different. Errors associated with predicted TOC concentrations within the Newport News watershed were often less than 20% (average 10%). The results of the seasonal analysis were more accurate than the original analyses that included a single representative TOC value. A comparison between the two analyses (Table 10.2) indicates the influence of seasonality on TOC behavior. The variability of the underlying data was better captured during the seasonal analysis, and better predictions across each watershed resulted.

## SITE-SPECIFIC STUDIES FOR DBP PRECURSORS AND ORGANIC CARBON

The land use–based regression techniques implemented in the previous analyses were also applied to the individual watershed studies. Two urban water supply catchments were selected for presentation in this report based on the large, high-quality data sets available. In each case, export modeling included a wide variety of LULC types (~15 of the 21 Anderson categories and many more subcategories) in addition to significant anthropogenic activity. Each watershed was divided into several sub-basins for which TOC data were available. LULC characteristics within each sub-basin were then determined, and the MLR analysis included multiple sub-basins as opposed to the multiple watersheds that were used in the national analysis. The regression analysis provided more detailed empirical models as the number of LULC categories increased for the individual watershed analyses.

The emphasis of this site-specific presentation was to expand on the generalizations of the national survey and determine whether similar empirical relationships existed given a more refined series of LULC categories and TOC data that is more detailed. Similar to the national analysis, the site-specific presentation incorporated the two different sets of analysis, period-of-record analysis, and seasonal analysis.

### Brick, New Jersey: Metedeconk River

#### *Watershed Description*

As with all of southern New Jersey, the Metedeconk River watershed is in ecoregion 8.5.1, the Middle Atlantic coastal plain. It is considered part of the southeastern United States coastal plain (ecoregion 8.5), and more generally the eastern temperate forest ecoregion (8), characterized by moderate temperatures, humid summers, and moderate levels of rainfall distributed throughout the year. The Metedeconk, along with Long Island and Cape Cod, is at the northern extreme of



©2007 AwwaRF. All Rights Reserved.

## Table 10.2
## Errors between measured TOC concentration and predicted TOC values for national regression

| | Overall | | | Winter | | |
|---|---|---|---|---|---|---|
| | Predicted | Measured | % Error | Predicted | Measured | % Error |
| University Lake, Orange, N.C. | 3.41 | 6.05 | 44 | 3.83 | 4.84 | 21 |
| Baltimore, Md. | 2.09 | 1.50 | 40 | 1.83 | 1.60 | 14 |
| Newport News, Va. | 8.32 | 6.06 | 37 | 5.39 | 5.41 | 0 |
| Jordan Valley, Utah | 2.53 | 2.15 | 18 | 1.51 | 2.06 | 27 |
| Fort Worth, Texas | 2.83 | 5.02 | 44 | 3.57 | 4.98 | 28 |
| Elgin, Ill. | 5.35 | 7.80 | 31 | 4.04 | 6.50 | 38 |
| Des Moines River, Des Moines, Iowa | 4.25 | 4.08 | 4 | 3.19 | 3.84 | 17 |
| Racoon River, Des Moines, Iowa | 4.86 | 3.84 | 27 | 4.32 | 3.16 | 37 |
| Tampa, Fla. | 11.87 | 13.52 | 12 | 5.49 | 5.73 | 4 |
| Portland, Ore. | 2.06 | 1.59 | 30 | 2.66 | 1.53 | 74 |
| | | Average | 29 | | Average | 26 |

| | Spring | | | Summer | | |
|---|---|---|---|---|---|---|
| | Predicted | Measured | % Error | Predicted | Measured | % Error |
| University Lake, Orange, N.C. | 4.32 | 5.55 | 22 | 3.96 | 6.63 | 40 |
| Baltimore, Md. | 2.09 | 1.30 | 61 | 1.00 | 1.10 | 9 |
| Newport News, Va. | 5.55 | 6.16 | 10 | 8.57 | 7.20 | 19 |
| Jordan Valley, Utah | 1.74 | 2.11 | 18 | 2.71 | 2.28 | 19 |
| Fort Worth, Texas | 4.25 | 5.04 | 16 | 1.72 | 5.09 | 66 |
| Elgin, Ill. | 4.67 | 7.90 | 41 | 6.28 | 8.80 | 29 |
| Des Moines River, Des Moines, Iowa | 4.12 | 4.30 | 4 | 4.91 | 4.18 | 17 |
| Racoon River, Des Moines, Iowa | 5.77 | 4.32 | 34 | 5.56 | 4.31 | 29 |
| Tampa, Fla. | 5.32 | 4.69 | 13 | 20.57 | 22.41 | 8 |
| Portland, Ore. | 3.09 | 1.52 | 103 | 2.03 | 1.74 | 17 |
| | | Average | 32 | | Average | 25 |

| | Fall | | |
|---|---|---|---|
| | Predicted | Measured | % Error |
| University Lake, Orange, N.C. | 3.70 | 6.03 | 39 |
| Baltimore, Md. | 1.58 | 2.00 | 21 |
| Newport News, Va. | 7.64 | 6.33 | 21 |
| Jordan Valley, Utah | 2.77 | 2.14 | 30 |
| Fort Worth, Texas | 2.18 | 4.97 | 56 |
| Elgin, Ill. | 5.52 | 7.80 | 29 |
| Des Moines River, Des Moines, Iowa | 4.42 | 4.15 | 7 |
| Racoon River, Des Moines, Iowa | 4.91 | 3.54 | 39 |
| Tampa, Fla. | 16.00 | 17.53 | 9 |
| Portland, Ore. | 2.17 | 1.63 | 33 |
| | | Average | 28 |



©2007 AwwaRF. All Rights Reserved.

**Table 10.3**
**Key characteristics of the Metedeconk River watershed**

| Characteristic | Value |
|---|---|
| Watershed area | 231 $km^2$ |
| Annual precipitation | 114 cm |
| Mean January temperature | –1 °C |
| Mean July temperature | 24 °C |

this tier III ecoregion. Table 10.3 presents some of the important characteristics of this watershed. USEPA (2006) describes 8.5.1 as follows:

> This ecological region is a flat coastal plain, with many swampy or marshy areas. Forest cover is predominantly loblolly shortleaf pine with patches of oak, gum, and cypress near major streams, as compared to the mainly longleaf-slash pine forests of the warmer coastal plain to the south (8.5.3). The central and southwestern parts of this region have poorly drained soils and only about 15% of the land is in cropland, whereas in the northeastern parts soils are relatively well drained and 20% to 40% of the land is in cropland.

The watershed is characterized by large amounts of urban and developed land, low elevations, and large amounts of wetlands. This is particularly evident from satellite photos near the confluence of the two branches. The water supply intake is off Forge Pond. Figure 10.7 shows the full set of LULC categories for the Metedeconk River watershed. Eight of the 30 watershed sampling locations are also marked. Upper regions of both branches are heavily forested, whereas lower regions are much more urban. The approach to modeling NOM in individual watersheds in this study was to use the simplest model and move on to more complex ones if supported by available data.

***Long-term Analysis of TOC***

The Metedeconk River system consists of two main branches that converge nearly 1 mile upstream of the water treatment plant intake. Both branches are gauged by USGS. Three sets of analyses were performed within this watershed. The first set used data combined across the whole catchment, whereas the second and third sets were used to analyze data for the North and South branches separately. The results of these latter analyses more closely followed the results of the national survey. Equation 10.8 presents the empirical model of the catchment analysis.

$$TOC_{overall} = 6.99 - 0.009\ urban - 0.0021\ forest - 0.031\ water + 0.0033\ wetland \qquad (10.8)$$

Figure 10.8 presents the comparison between predicted TOC values and measured TOC concentration for the Metedeconk system empirical model. The resulting error of the model was 18%. Significant variability was observed at sampling locations referred to as SN, SE, NP, and NC, whereas strong agreement occurred at stations SA, NJ, NI, and NH.



©2007 AwwaRF. All Rights Reserved.

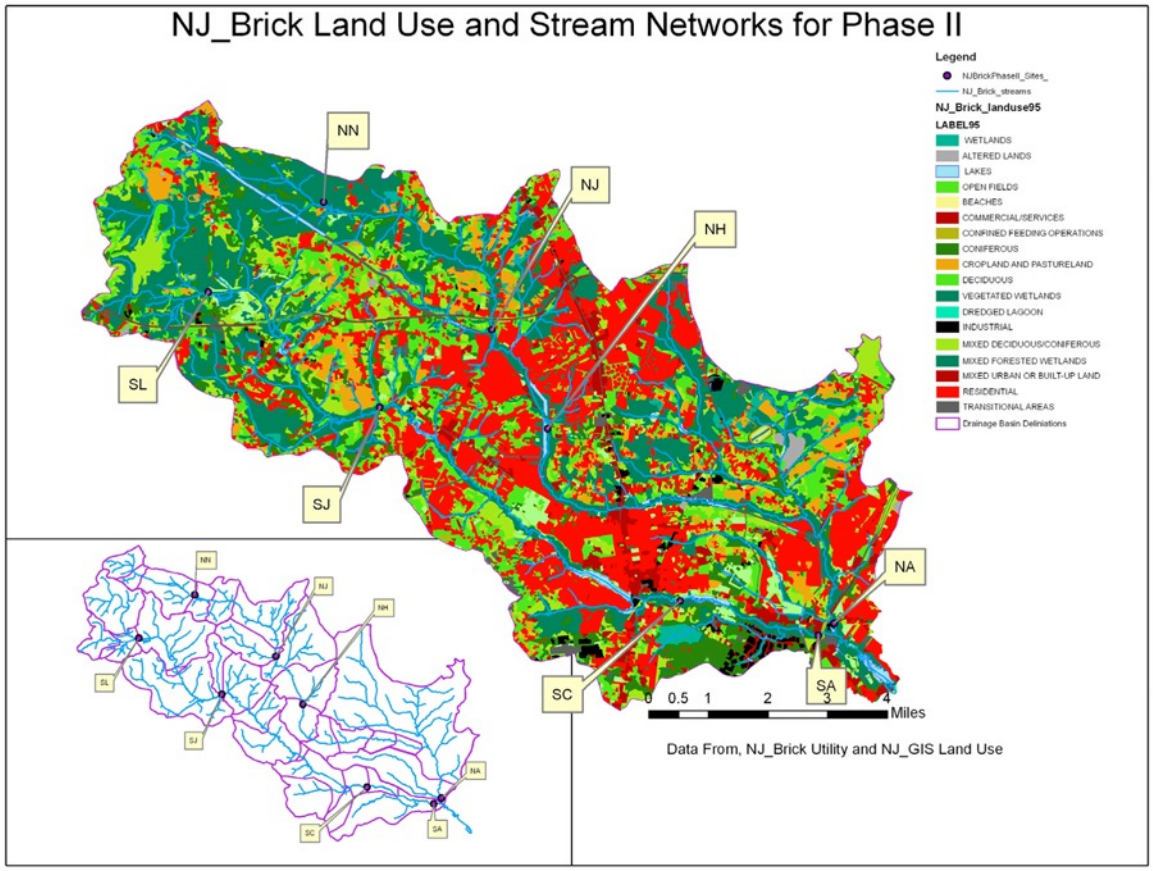



**Figure 10.7 Full set of LULC categories across the Metedeconk watershed**

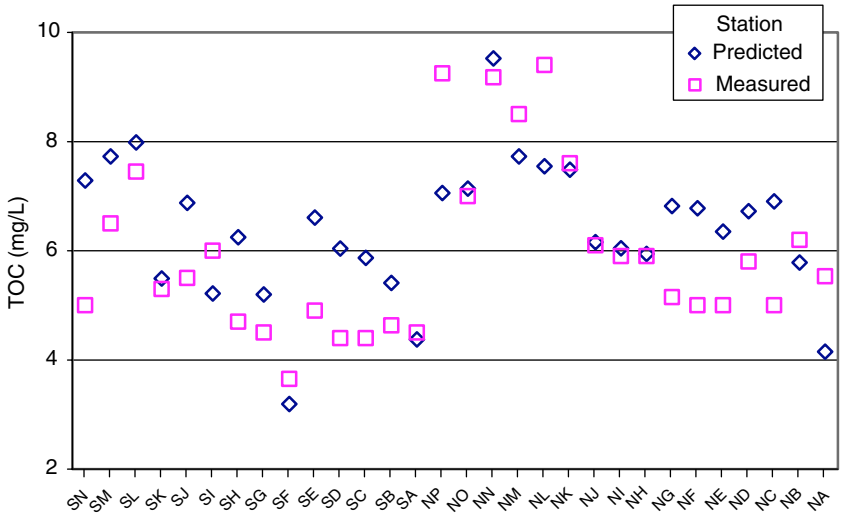



**Figure 10.8 Metedeconk system regression comparison between empirical TOC predictions and actual TOC concentration**



©2007 AwwaRF. All Rights Reserved.

The empirical model included several land use types similar to the national empirical models such as forested land, water, and wetlands. However, the urban development land use characteristic differed between analyses. This characteristic included all commercial, industrial, and residential areas for the national analysis and included only commercial development and light industrial areas during the watershed analysis. This difference yielded dissimilar coefficients between analyses. A comparison between the overall national model and the overall Metedeconk model indicated that the urban land use coefficient in the national empirical model was –0.0035 whereas the Metedeconk urban coefficient was –0.0094. The Metedeconk coefficient was approximately three times larger and also negatively influenced the TOC concentration. Likewise, the wetlands coefficient differed—0.0255 for the national analysis compared with 0.0033 for the Metedeconk watershed. The national coefficient far exceeded the coefficient within the Metedeconk empirical model.

### *Metedeconk System Seasonal Analysis*

Equations 10.9 to 10.12 present the empirical models developed through seasonal analysis. As with the national analysis, seasonal variability influenced each empirical model and the corresponding accuracy at each sampling location.

$$TOC_{winter} = 6.18 - 0.007 \text{ urban} - 0.0023 \text{ forest} - 0.0188 \text{ water} + 0.0037 \text{ wetland} \quad (10.9)$$

$$TOC_{spring} = 8.08 - 0.0044 \text{ forest} + 0.013 \text{ rural} - 0.036 \text{ water} + 0.0041 \text{ wetland} \quad (10.10)$$

$$TOC_{summer} = 7.28 - 0.024 \text{ industry} - 0.051 \text{ water} + 0.0022 \text{ wetlands} \quad (10.11)$$

$$TOC_{fall} = 5.69 - 0.026 \text{ water} - 0.006 \text{ urban} \quad (10.12)$$

The seasonal empirical models differed from one another. Water was present within each model and was the only such similarity, though the wetlands land use type was present in all but one model. The error associated with each model varied and ranged between 14% and 33%. The average error of 23% was slightly higher than in the overall analysis.

### *Metedeconk Branch Analyses*

The empirical models for the North and South branches of the Metedeconk River were somewhat erroneous—30% and 31%, respectively. The system empirical model yielded an error of only 18%, much less than the error of either branch model. The empirical models for the North and South branches (NBTOC and SBTOC, respectively) are shown in Equations 10.13 and 10.14.

$$NBTOC_{overall} = 0.149 \text{ fields} + 0.025 \text{ forest} + 0.017 \text{ suburban} + 0.038 \text{ wetlands} \quad (10.13)$$

$$SBTOC_{overall} = 0.031 \text{ fields} + 0.013 \text{ suburban} - 0.064 \text{ rural} + 0.0138 \text{ wetlands} \quad (10.14)$$

The results of the two empirical models were extremely inconsistent; hence, the high error values. The error between the predicted TOC value and the measured TOC concentration varied between branches. Figure 10.9 presents the comparison for the North Branch.



©2007 AwwaRF. All Rights Reserved.

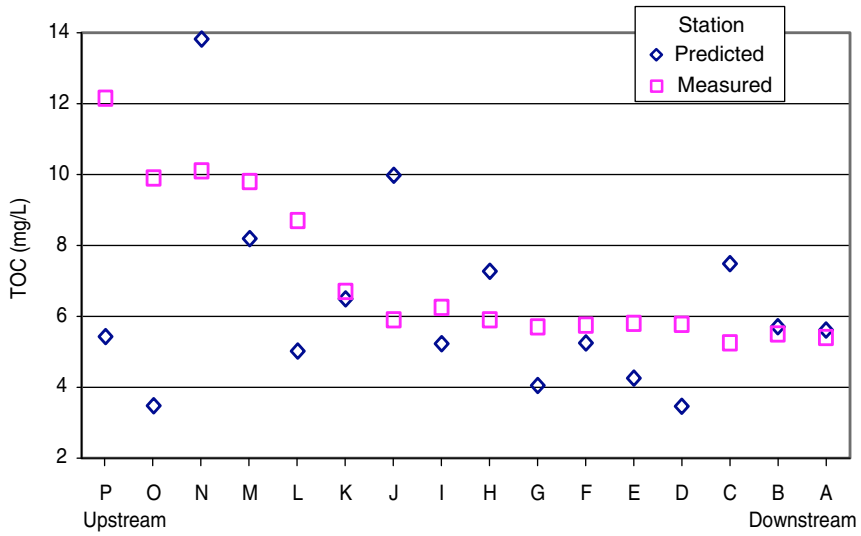



**Figure 10.9 North Branch regression comparison between empirical TOC predictions and actual TOC concentration**

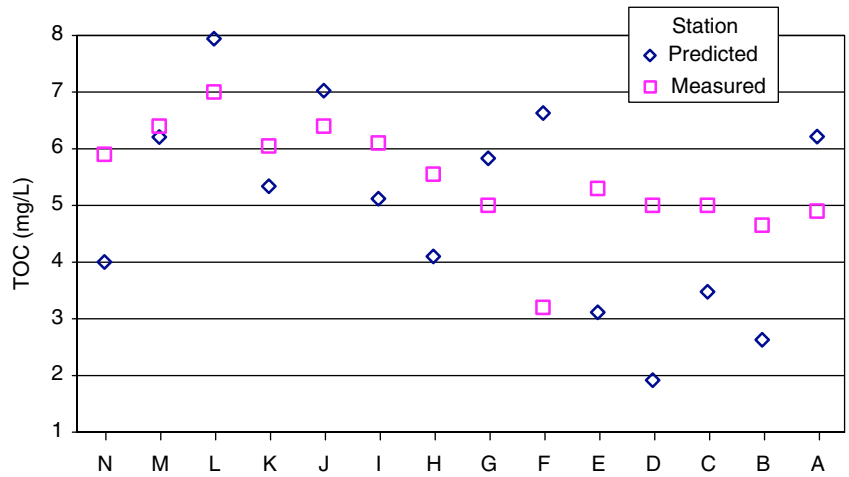



**Figure 10.10 South Branch regression comparison between empirical TOC predictions and actual TOC concentration**

Predictions were quite accurate at sampling stations A, B, K, and F. However, the predictions at stations P, O, N, and J significantly influenced the average error value. The empirical model did not perform well in the headwaters of the North Branch. There was large variability in the underlying data at these sampling stations, and the single average TOC value did not capture this variability, which resulted in poor predictions of TOC. Similar results were observed within the South Branch (Figure 10.10). Significant variability existed along the length of the channel.

Predictions were accurate at sampling stations M, K, J, and I, but predictions at stations F, E, D, and B were not. The empirical model did not perform well for the downstream-most lake (stations F, E, D) and confluence (stations B and A). There was large variability in the underlying



©2007 AwwaRF. All Rights Reserved.

data at these locations, and the average TOC value did not capture this variability; therefore, the model yielded poor predictions of TOC.

### *Metedeconk Branch Seasonal Analysis*

The results of the North Branch and South Branch seasonal regression analyses were significantly different than the seasonal empirical relationships developed with the entire catchment. Equations 10.15 to 10.22 present side-by-side comparisons of the seasonal empirical models for each branch of the Metedeconk River.

$$NBTOC_{winter} = 0.14 \text{ field} + 0.24 \text{ forest} + 0.15 \text{ suburb} - 3.61 \text{ water} + 0.034 \text{ wetlands} \quad (10.15)$$

$$SBTOC_{winter} = 0.0291 \text{ fields} + 0.013 \text{ suburban} - 0.066 \text{ rural} - 0.014 \text{ wetlands} \quad (10.16)$$

$$NBTOC_{spring} = 0.19 \text{ field} + 0.031 \text{ forest} + 0.021 \text{ suburb} - 4.93 \text{ water} + 0.046 \text{ wetlands} \quad (10.17)$$

$$SBTOC_{spring} = 0.037 \text{ fields} + 0.014 \text{ suburban} - 0.069 \text{ rural} + 0.015 \text{ wetlands} \quad (10.18)$$

$$NBTOC_{summer} = 0.084 \text{ fields} - 1.52 \text{ water} + 0.021 \text{ wetlands} \quad (10.19)$$

$$SBTOC_{summer} = 0.015 \text{ suburban} - 0.056 \text{ rural} + 0.015 \text{ wetlands} \quad (10.20)$$

$$NBTOC_{fall} = 0.12 \text{ field} + 0.02 \text{ forest} + 0.014 \text{ suburb} - 3.14 \text{ water} + 0.028 \text{ wetlands} \quad (10.21)$$

$$SBTOC_{fall} = 0.042 \text{ urban} - 0.037 \text{ industry} + 0.005 \text{ wetlands} + 0.068 \text{ misc.} \quad (10.22)$$

The suburban, open fields, and wetlands land use types largely influenced the empirical models of the South Branch, whereas open fields, forest, and wetlands land use types strongly influenced the North Branch models. Predictions of TOC within each branch varied greatly among seasons. The summer months yielded the most erroneous results, whereas winter proved the most accurate of the four seasons. The variability associated with the underlying winter data showed the least variations among the seasons and accounted for the increased accuracy of the empirical model during this season.

Figures 10.11 and 10.12 present the comparison between the predicted and measured TOC concentrations for the winter and summer seasons, respectively. Figures are not presented for spring and fall. The prediction errors for both branches were 28% during the winter. The behavior of each branch resembled that of the overall analyses presented in Figures 10.9 and 10.10. Upstream errors were quite high in the North Branch model and downstream errors were higher in the South Branch model. The summer season was clearly the most difficult to model. Significant discrepancies between predictions and observations within the North Branch extend from the headwaters (typical of other analyses) through nearly 40% of the basin. Predictions were slightly higher toward the confluence, and the overall impact increased the seasonal error. In the South Branch, erroneous predictions resulted around the downstream lake and the confluence as well as just below the headwater tributaries; the average error within that region was 45%.



©2007 AwwaRF. All Rights Reserved.

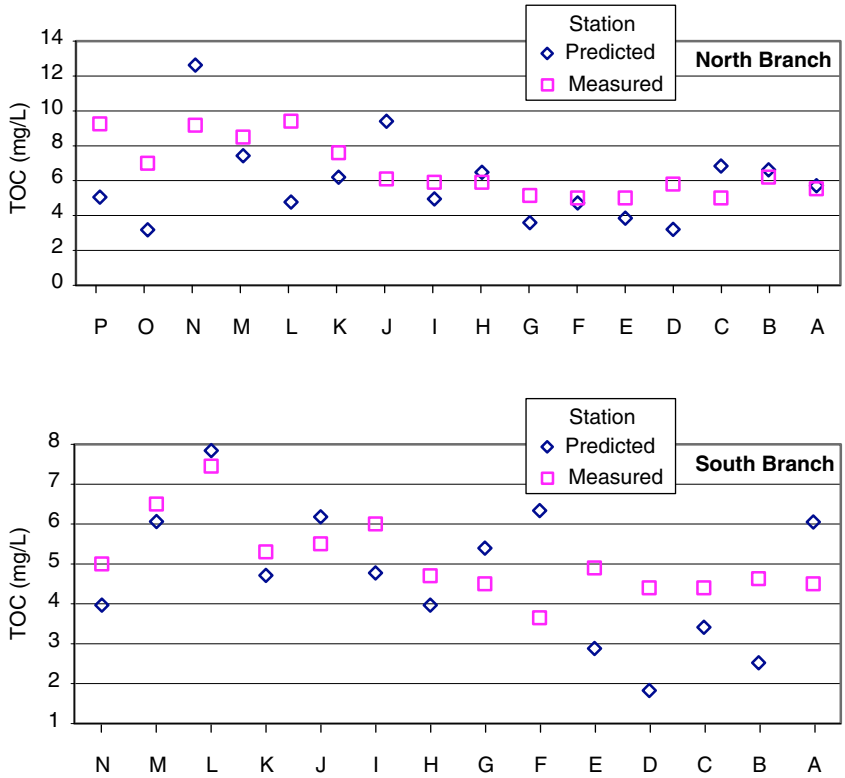



**Figure 10.11 Winter MLR for North and South branch comparison between empirical TOC predictions and actual TOC concentration**

Figures 10.9 and 10.10 highlight the variability associated with the seasonal analyses of the North and South branches of the Metedeconk River. The poor predictions were attributed to several factors within the watershed. The data indicated rapid changes in TOC across the system, due in part to surface runoff and various tributary contributions. Furthermore, the branches pass through several miles of wooded and wetland riparian zones. Without knowledge of the fate and transport mechanisms within the river, it was difficult to isolate the specific factors that influenced TOC behavior within the river and watershed. Clearly, small watersheds are more subject to transient behaviors, whereas larger ones benefit from temporal and spatial "buffering."

Table 10.4 presents a summary of the accuracy of the various empirical models within the Metedeconk River system. This table highlights the differences and similarities among the various analyses not easily gained from the many respective figures.

The results of the seasonal analyses of the North and South branches of the Metedeconk more closely resembled the results of the national seasonal analyses than the overall branch analyses. The seasonal analyses (national and Metedeconk) were more accurate than the overall analyses. The overall analysis for the North and South branches was less accurate; these errors were attributed to the variability of the underlying data across the period of record. Moreover, the analyses of the Metedeconk system, overall and seasonal, were less accurate than the river branch analyses. The resulting empirical models did not accurately capture measured changes of TOC at the intake nor the behavior of TOC within the watershed.



©2007 AwwaRF. All Rights Reserved.

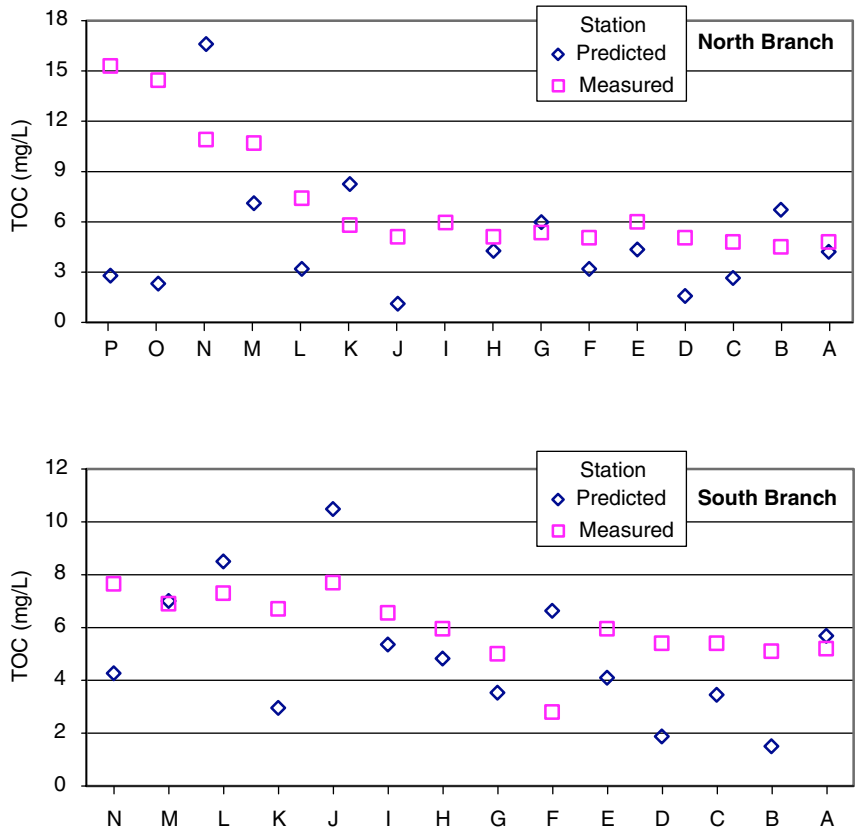



**Figure 10.12 Summer MLR for North and South branch comparison between empirical TOC predictions and actual TOC concentration**

### *Short-term Analysis of DBP Precursors*

Surface water samples were collected at eight locations within the drinking water drainage basin as well as at the plant influent. They are identified as follows: plant effluent (sample 1), varying locations along the North Branch of the Metedeconk River (samples 2–5), and varying locations along the South Branch (samples 6–9). Site codes and descriptions of the sampling locations are summarized in Table 10.5; date and time of samplings are also included.

All samples were analyzed for basic organic parameters (TOC, DOC, and UV absorbance) as well as DBP precursor content. Briefly, the latter is determined by a laboratory test where all samples are chlorinated (dose: 20 mg $Cl_2$/L) at pH 7 and allowed to react for an incubation period of 3 days at 20°C. All samples except sample 5 were diluted to 20% of their initial concentration before chlorination, and subsequent concentrations were mathematically adjusted back to the original undiluted values. Sample 5 was diluted to 7.5%.

***Laboratory Results.*** Table 10.6 summarizes the basic organic characteristics of the water samples. All the water samples are characterized by low DOC ranging from 5.5 to 11.9 mg C/L. The three SUVA values are all near 5, which suggests a very high lignin content, typical of sources dominated by terrestrial and forested environments. Sample NN has an exceptionally high TOC and a UV absorbance commensurate with that organic level.



©2007 AwwaRF. All Rights Reserved.

## Table 10.4
## Errors between measured TOC concentration and predicted TOC values for North and South branches

| North Branch | Winter | | | Spring | | | Summer | | | Fall | | | Overall | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Predicted | Measured | % Error | Predicted | Measured | % Error | Predicted | Measured | % Error | Predicted | Measured | % Error | Predicted | Measured | % Error |
| A | 5.71 | 5.53 | 3 | 6.70 | 6.8 | 2 | 4.21 | 4.8 | 12 | 4.36 | 4.2 | 4 | 5.62 | 5.4 | 4 |
| B | 6.63 | 6.2 | 7 | 4.78 | 6.4 | 25 | 6.72 | 4.5 | 49 | 4.02 | 3.5 | 15 | 5.71 | 5.5 | 4 |
| C | 6.84 | 5 | 37 | 9.00 | 6.8 | 32 | 2.65 | 4.8 | 45 | 5.77 | 3.75 | 54 | 7.48 | 5.25 | 43 |
| D | 3.21 | 5.8 | 45 | 4.09 | 6.65 | 38 | 1.58 | 5.05 | 69 | 2.73 | 4.65 | 41 | 3.46 | 5.77 | 40 |
| E | 3.85 | 5 | 23 | 5.28 | 7.4 | 29 | 4.35 | 6 | 27 | 3.35 | 5.2 | 36 | 4.26 | 5.8 | 27 |
| F | 4.72 | 5 | 6 | 6.47 | 8.2 | 21 | 3.19 | 5.05 | 37 | 4.21 | 4.5 | 7 | 5.25 | 5.75 | 9 |
| G | 3.59 | 5.15 | 30 | 5.20 | 7.3 | 29 | 5.98 | 5.35 | 12 | 3.43 | 5 | 31 | 4.05 | 5.7 | 29 |
| H | 6.48 | 5.9 | 10 | 9.11 | 7.2 | 27 | 4.27 | 5.1 | 16 | 6.25 | 4.5 | 39 | 7.27 | 5.9 | 23 |
| I | 4.96 | 5.9 | 16 | 5.92 | 7.55 | 22 | –0.27 | 5.95 | 105 | 4.09 | 5.55 | 26 | 5.23 | 6.25 | 16 |
| J | 9.40 | 6.1 | 54 | 11.44 | 7.1 | 61 | 1.11 | 5.1 | 78 | 7.76 | 4.65 | 67 | 9.98 | 5.9 | 69 |
| K | 6.20 | 7.6 | 18 | 7.46 | 8.8 | 15 | 8.25 | 5.8 | 42 | 4.50 | 4.95 | 9 | 6.50 | 6.7 | 3 |
| L | 4.78 | 9.4 | 49 | 5.67 | 9.9 | 43 | 3.20 | 7.4 | 57 | 3.48 | 5.3 | 34 | 5.02 | 8.7 | 42 |
| M | 7.43 | 8.5 | 13 | 9.98 | 12.7 | 21 | 7.11 | 10.7 | 34 | 6.11 | 6.4 | 5 | 8.19 | 9.8 | 16 |
| N | 12.62 | 9.175 | 38 | 16.62 | 12 | 39 | 16.61 | 10.9 | 52 | 9.79 | 7 | 40 | 13.82 | 10.1 | 37 |
| O | 3.18 | 7 | 55 | 4.16 | 10.9 | 62 | 2.32 | 14.45 | 84 | 2.63 | 9.3 | 72 | 3.48 | 9.9 | 65 |
| P | 5.06 | 9.25 | 45 | 6.35 | 13.8 | 54 | 2.78 | 15.3 | 82 | 4.06 | 9.55 | 57 | 5.43 | 12.15 | 55 |
| | | Average | 28 | | Average | 32 | | Average | 50 | | Average | 34 | | Average | 30 |

| South Branch | Winter | | | Spring | | | Summer | | | Fall | | | Overall | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Predicted | Measured | % Error | Predicted | Measured | % Error | Predicted | Measured | % Error | Predicted | Measured | % Error | Predicted | Measured | % Error |
| A | 6.1 | 4.5 | 34 | 7.02 | 5.4 | 30 | 5.7 | 5.2 | 9 | 4.4 | 4.0 | 9 | 6.2 | 4.9 | 27 |
| B | 2.5 | 4.6 | 45 | 3.03 | 4.8 | 37 | 1.5 | 5.1 | 70 | 0.5 | 3.4 | 84 | 2.6 | 4.7 | 43 |
| C | 3.4 | 4.4 | 22 | 3.87 | 5.9 | 34 | 3.5 | 5.4 | 36 | 4.1 | 4.2 | 1 | 3.5 | 5.0 | 30 |
| D | 1.8 | 4.4 | 58 | 2.18 | 5.8 | 62 | 1.9 | 5.4 | 65 | 3.5 | 4.5 | 20 | 1.9 | 5.0 | 62 |
| E | 2.9 | 4.9 | 41 | 3.57 | 6.3 | 43 | 4.1 | 6.0 | 31 | 4.5 | 4.5 | 0 | 3.1 | 5.3 | 41 |
| F | 6.3 | 3.7 | 74 | 7.51 | 3.6 | 109 | 6.6 | 2.8 | 137 | 5.1 | 3.1 | 66 | 6.6 | 3.2 | 107 |
| G | 5.4 | 4.5 | 20 | 6.87 | 6.2 | 11 | 3.5 | 5.0 | 29 | 3.0 | 4.5 | 33 | 5.8 | 5.0 | 17 |
| H | 4.0 | 4.7 | 16 | 4.57 | 6.9 | 34 | 4.8 | 6.0 | 19 | 1.9 | 4.7 | 59 | 4.1 | 5.6 | 26 |
| I | 4.8 | 6.0 | 20 | 5.89 | 6.3 | 7 | 5.4 | 6.6 | 18 | 4.4 | 4.7 | 6 | 5.1 | 6.1 | 16 |
| J | 6.2 | 5.5 | 12 | 8.31 | 7.5 | 11 | 10.5 | 7.7 | 36 | 8.0 | 5.7 | 41 | 7.0 | 6.4 | 10 |
| K | 4.7 | 5.3 | 11 | 6.53 | 7.5 | 13 | 3.0 | 6.7 | 56 | 5.1 | 4.8 | 7 | 5.3 | 6.1 | 12 |
| L | 7.8 | 7.5 | 5 | 8.84 | 7.5 | 19 | 8.5 | 7.3 | 16 | 2.7 | 5.3 | 49 | 7.9 | 7.0 | 13 |
| M | 6.1 | 6.5 | 7 | 6.95 | 7.8 | 11 | 7.0 | 6.9 | 2 | 3.6 | 4.9 | 27 | 6.2 | 6.4 | 3 |
| N | 4.0 | 5.0 | 21 | 4.44 | 6.2 | 28 | 4.3 | 7.7 | 44 | 2.8 | 4.9 | 43 | 4.0 | 5.9 | 32 |
| | | Average | 28 | | Average | 32 | | Average | 41 | | Average | 32 | | Average | 31 |



©2007 AwwaRF. All Rights Reserved.

**Table 10.5**
**Sampling locations and related information***

| Number | Sampling location† | Site code | Description | Sample Time (a.m.) |
|---|---|---|---|---|
| 1 | Plant intake | Intake | Metedeconk River at Brick utilities intake | 7:23 |
| 2 | North Branch, site A | NA | North Branch Metedeconk at Highway 88, Brick Township | 8:05 |
| 3 | North Branch, site H | NH | North Branch Metedeconk at Bethel Church Road, Jackson Township | 10:37 |
| 4 | North Branch, site J | NJ | North Branch Metedeconk at Fort Plains Road, Howell Township | 10:28 |
| 5 | North Branch, site N | NN | North Branch Metedeconk at Nomoco Road, Freehold Township | 9:47 |
| 6 | South Branch, site A | SA | South Branch Metedeconk at Chambers Bridge Road, Lakewood Township | 8:15 |
| 7 | South Branch site C | SC | South Branch Metedeconk at Clover Street, Lakewood Township | 8:33 |
| 8 | South Branch site J | SJ | South Branch Metedeconk at Cooks Bridge Road, Jackson Township | 8:58 |
| 9 | South Branch site L | SL | South Branch Metedeconk at Cedar Swamp Road, Jackson Township | 9:13 |

*All samples were taken on May 13, 2005.
†Locations refer to the North Branch and South Branch of the Metedeconk River.

**Table 10.6**
**General characteristics of water samples prior to chlorination**

| Site code | TOC (mg C/L) | DOC (mg C/L) | $UV_{254}$ ($cm^1$) | SUVA (L/m·mg) |
|---|---|---|---|---|
| Intake | 5.5 | ND* | 0.268 | 4.9 |
| NA | 5.5 | ND | 0.261 | 4.7 |
| NH | 6.4 | ND | 0.312 | 4.9 |
| NJ | 7.6 | ND | 0.384 | 5.1 |
| NN | 11.9 | ND | 0.597 | 5.0 |
| SA | 6.0 | ND | 0.266 | 4.4 |
| SC | 5.5 | ND | 0.265 | 4.8 |
| SJ | 5.8 | ND | 0.313 | 5.4 |
| SL | 7.3 | ND | 0.434 | 5.9 |

*ND = not determined.



©2007 AwwaRF. All Rights Reserved.

**Table 10.7**
**Chlorine demand data**

| Site code | Chlorine demand (mg $Cl_2$/L) | Specific chlorine demand (mg $Cl_2$/mg C) |
|---|---|---|
| Intake | 13.87 | 2.52 |
| NA | 13.48 | 2.45 |
| NH | 15.5 | 2.42 |
| NJ | 19.21 | 2.53 |
| NN | 22.79 | 1.92 |
| SA | 13.48 | 2.25 |
| SC | 12.44 | 2.26 |
| SJ | 15.3 | 2.64 |
| SL | 16.82 | 2.30 |

Table 10.7 shows the chlorine residuals and the chlorine demands for the 3-day incubation period. Chlorine residual is the chlorine remaining in the water after the chlorine demand has been satisfied. Chlorine demand is determined as the difference between the amount of chlorine added (20 mg $Cl_2$/L, in this case) and the amount of residual chlorine remaining after the given contact period. Specific chlorine demand (chlorine demand/DOC) ranging from 1.9 to 2.6 was measured for all water samples. These are relatively high values that are commensurate with the high SUVAs of these samples. These samples were all diluted before treatment and the chlorine demands were readjusted to the undiluted sample.

**Major By-products.** The THM, THAA, and DHAA concentrations are presented in Tables 10.8, 10.9, and 10.10, respectively. These concentrations are all high for formation potential data, commensurate with the high sample TOC values. Figure 10.13 illustrates the formation of specific THM (THM/DOC) in all water samples. The production of THMs was normalized by dividing the observed DBP concentration (in micrograms per liter) by the corresponding TOC concentrations of the water. This is a common method for comparing DBP precursor data from samples of differing NOM concentration and quality. In this case, sample SJ produced the highest specific THM formation at greater than 80 µg/mg C. The five samples averaged about 72 µg THM/mg C. Figure 10.14 shows how these data compare with those collected from a diverse set of sources across North America. Brick's samples tend to be just slightly higher in precursor content than average even when normalized to the high TOC.

Figure 10.15 illustrates the concentration of total DHAAs and THAAs formed in all water samples. Once again, the production of DHAA and THAA were normalized by dividing the DBP concentration (in micrograms per liter) by the corresponding DOC concentrations of the water. THAA formation for all samples was substantially higher than DHAA formation. The nine samples exhibited an average DHAA and THAA of about 34 and 81 µg/mg C, respectively. The THAA levels are about at the expected level considering the high THM precursor levels. However, the DHAA formation is quite small given the THM and THAA concentrations. This is evident in the moderate THAA/THM and low DHAA/THM ratios (Figures 10.16 and 10.17, respectively). The general trend across North America is to have lower ratios for waters of low TOC, and Brick's waters seem to fit that trend.



©2007 AwwaRF. All Rights Reserved.

**Table 10.8**
**THM concentrations (μg/L)***

| Site code | Chloroform ($CHCl_3$) | Bromodichloromethane ($CHCl_2Br$) | Chlorodibromomethane ($CHClBr_2$) | Bromoform ($CHBr_3$) | THM4 |
|---|---|---|---|---|---|
| Intake | 390.8 | 19.0 | 0.3 | 0.0 | 410.1 |
| NA | 375.5 | 15.3 | 0.2 | 0.0 | 391.0 |
| NH | 412.8 | 16.7 | 0.2 | 0.0 | 429.7 |
| NJ | 581.2 | 14.8 | 0.1 | 0.0 | 596.1 |
| NN | 771.8 | 11.2 | 0.1 | 0.0 | 783.1 |
| SA | 393.7 | 14.0 | 0.2 | 0.0 | 407.9 |
| SC | 364.7 | 13.3 | 0.1 | 0.0 | 378.2 |
| SJ | 474.8 | 13.3 | 0.1 | 0.0 | 488.2 |
| SL | 531.2 | 9.2 | 0.0 | 0.0 | 540.4 |

*Chloroform and bromodichloromethane concentrations are as measured; chlorodibromomethane and bromoform are estimated based on the standard geometric speciation model.

**Table 10.9**
**THAA concentrations (μg/L)***

| Site code | TCAA | BDCAA | CDBAA | TBAA | THAA |
|---|---|---|---|---|---|
| Intake | 411.4 | 18.5 | 0.3 | 0.0 | 430.2 |
| NA | 389.0 | 14.7 | 0.2 | 0.0 | 403.8 |
| NH | 466.0 | 17.4 | 0.2 | 0.0 | 483.7 |
| NJ | 656.6 | 15.5 | 0.1 | 0.0 | 672.2 |
| NN | 908.7 | 12.2 | 0.1 | 0.0 | 920.9 |
| SA | 438.6 | 14.5 | 0.2 | 0.0 | 453.3 |
| SC | 426.9 | 14.4 | 0.2 | 0.0 | 441.5 |
| SJ | 548.2 | 14.2 | 0.1 | 0.0 | 562.6 |
| SL | 616.9 | 9.9 | 0.1 | 0.0 | 626.8 |

*TCAA was measured; all others were estimated from the THM speciation and the measured TCAA.

**Bromine Incorporation/Substitution Factors.** The bromine incorporation factor (BIF) is a measure of the extent of bromine addition to NOM as compared to chlorine. Quite literally, it is the average number of bromine atoms in a given class of DBP molecules. For a sample having only chloroform, the BIF would be zero. On the other extreme, a sample with only bromoform would have a BIF of 3. A similar measure of bromine level is the bromine substitution factor (BSF). This is the percentage of bromine observed among the species as compared to the total halogen in a group of DBPs, that is, Br/(Cl+Br). Like the BIF, the BSF can be calculated from the speciation of THMs or any class of mixed chloro/bromo compounds. For trihalogenated species, the BSF is



©2007 AwwaRF. All Rights Reserved.

**Table 10.10**
**DHAA concentrations (µg/L)**

| Site code | DCAA | BCAA | DBAA | DHAA |
|---|---|---|---|---|
| Intake | 187.2 | 7.6 | 2.9 | 197.7 |
| NA | 171.1 | 5.7 | 1.0 | 177.8 |
| NH | 203.1 | 5.0 | 1.6 | 209.7 |
| NJ | 270.8 | 3.9 | 1.6 | 276.3 |
| NN | 347.7 | 3.9 | 2.7 | 354.3 |
| SA | 197.0 | 4.6 | 1.6 | 203.2 |
| SC | 169.2 | 4.2 | 1.1 | 174.5 |
| SJ | 226.3 | 4.1 | 1.6 | 232.0 |
| SL | 238.7 | 3.2 | 1.6 | 243.5 |

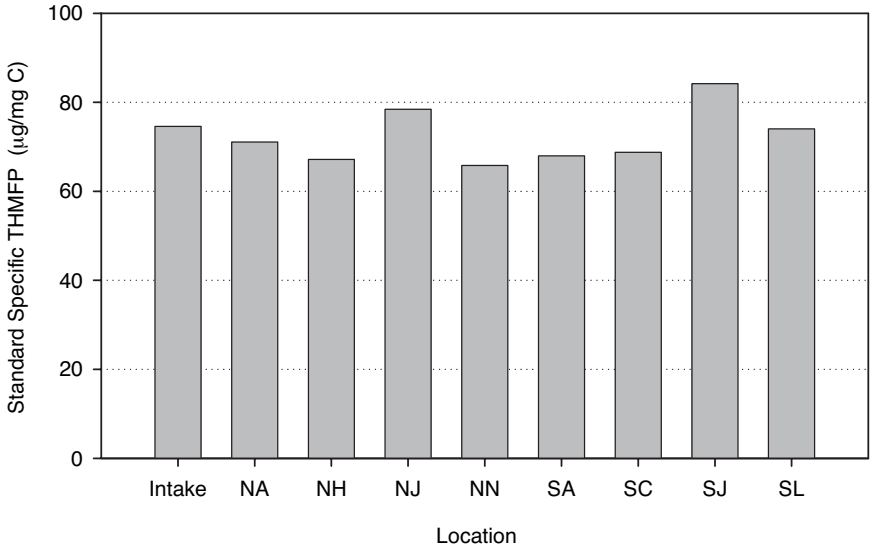



**Figure 10.13 Specific THM4 yield following chlorination of water samples**

equal to exactly one-third the BIF. Table 10.11 shows the BIF and BSF values as calculated from the THM speciation for the full suite of samples. The samples showed an average BIF of about 0.023, which is extremely low. The intake sample stood out as having a higher BIF than the others. This might be due to a higher ambient bromide level in this sample, or some specific affinity for bromine incorporation.

Figure 10.18 shows rainfall data for the Manasquan River at Point Pleasant, just a few miles north of the Brick intake, which is situated at the southeast corner of the current study watershed. Streamflow for the Metedeconk's North Branch is slightly below the long-term average for May 13 (Figure 10.19). There were a series of small rainfall events between April 23 and May 1 (the tail end of these is visible in Figure 10.19) that undoubtedly resulted in elevated streamflow on May 2 and 3 as seen in Figure 10.19. Then, for the 5 days prior to the sampling



©2007 AwwaRF. All Rights Reserved.

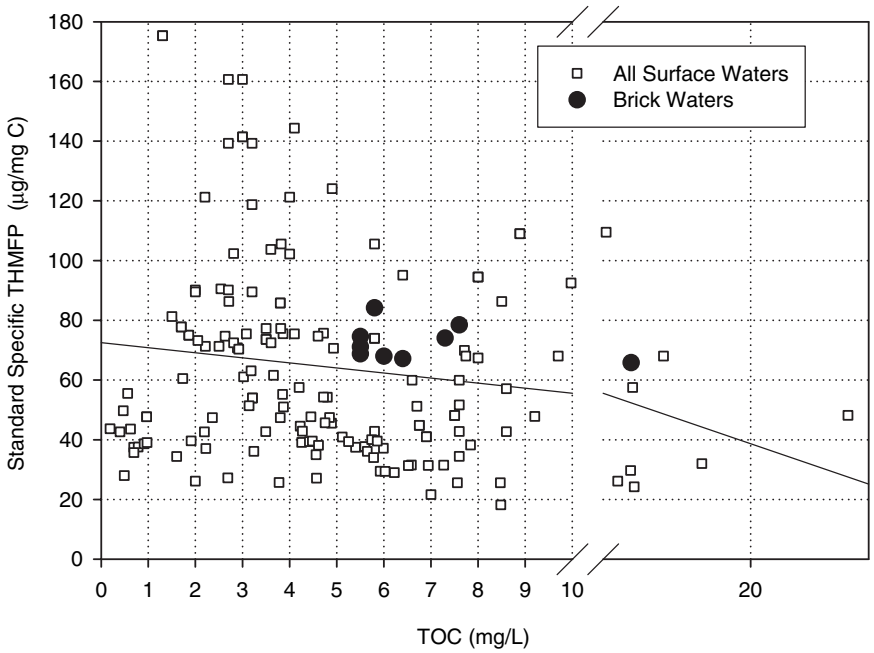



**Figure 10.14 Standard specific THMFP versus TOC in surface waters**

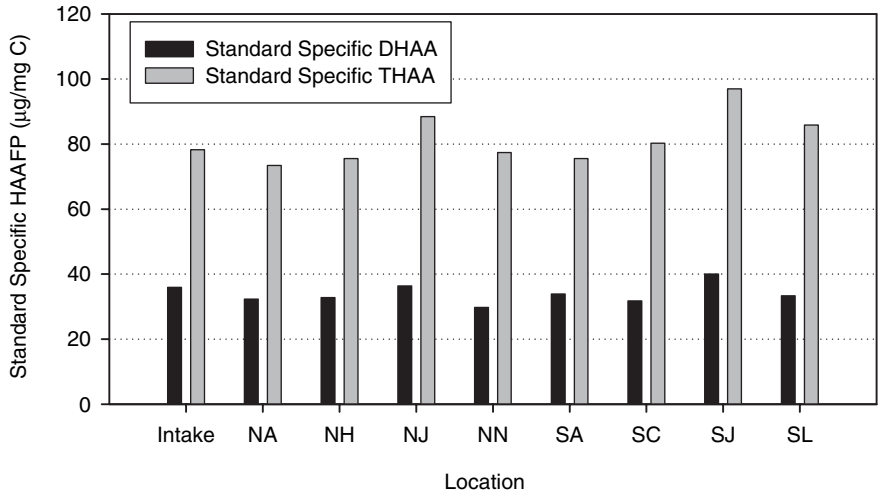



**Figure 10.15 DHAA and THAA yields following chlorination of water samples**

date, there was no measurable rainfall. Streamflows dropped during this period and did not increase again until the storm on May 20.

Land use characteristics for the Metedeconk River watershed are available for 53 categories (Table 10.12). This level of detail is inappropriate for most modeling purposes and cannot be supported by the limited number of dependent variables in the synoptic database used for this study. As a result, it had to be collapsed to a set of LULC categories that were less numerous and of lower resolution. This was also done for the long-term data analysis, but it becomes much more



©2007 AwwaRF. All Rights Reserved.

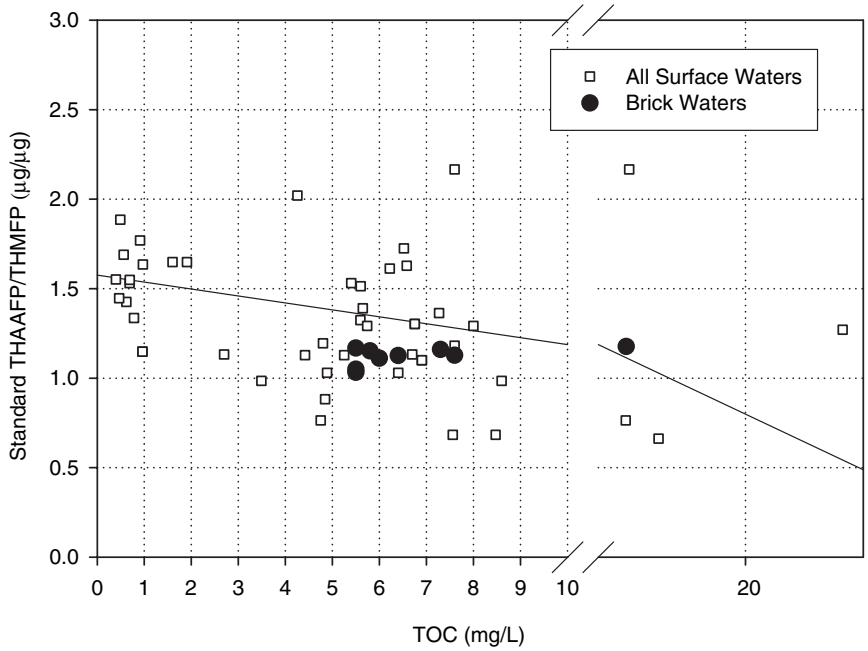



**Figure 10.16 Ratio of THAA precursors to THM precursors in surface waters**

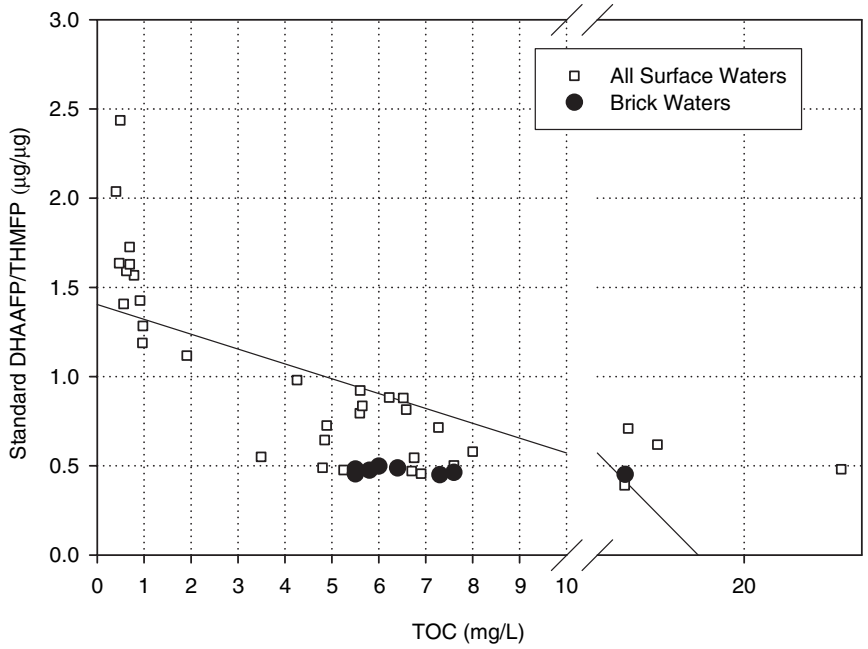



**Figure 10.17 Ratio of DHAA precursors to THM precursors in surface waters**

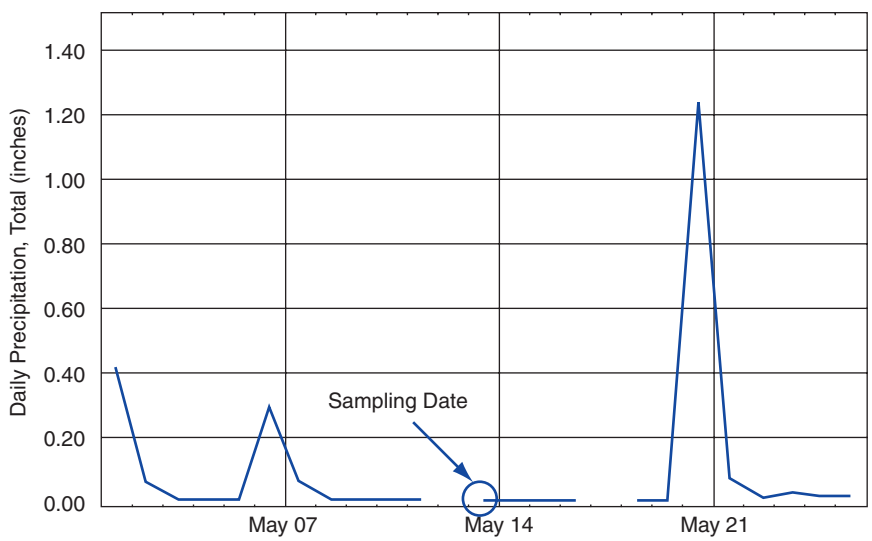


©2007 AwwaRF. All Rights Reserved.

**Table 10.11**
**Bromine incorporation factor and substitution factor for THMs***

| Sample identification | BIF | BSF (%) |
|---|---|---|
| Intake | 0.035 | 1.2 |
| NA | 0.029 | 1.0 |
| NH | 0.029 | 1.0 |
| NJ | 0.018 | 0.6 |
| NN | 0.011 | 0.4 |
| SA | 0.026 | 0.9 |
| SC | 0.026 | 0.9 |
| SJ | 0.020 | 0.7 |
| SL | 0.013 | 0.4 |

*Based on measured chloroform and bromodichloromethane concentrations and the standard geometric speciation model.




*Source:* USGS 2006.

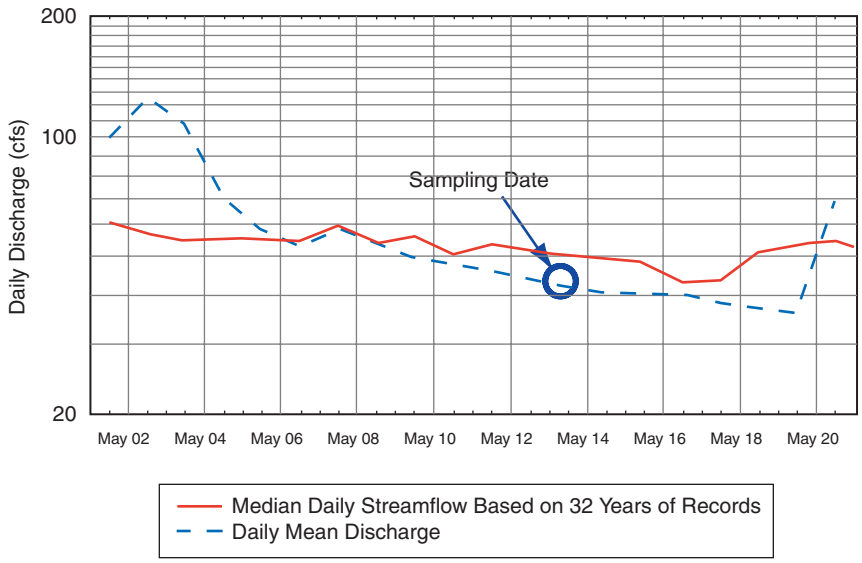
**Figure 10.18 Rainfall for the Brick region during the sampling period in May 2005**

acute with the smaller data set that exists for the DBP precursors. Consideration of the tertiary level, however, did allow examination of some of the more critical land cover types that might have been lost in a collapsed set. In particular, the Atlantic white cedar swamp was proposed as a possible major source of NOM. For this reason, it was retained for later analysis. Table 10.12 shows that deciduous and coniferous wetlands are dominant land cover types.

The first set of collapsed land use data was limited to numerical-based categories at the secondary (100) level. This brought the total number of categories used down to 24 (Table 10.13). The coniferous and deciduous wetlands were merged with several other related categories, giving an even more dominant category of "interior wetlands." The second stage in collapsing the original 53 categories brought the number down to 7 categories (Table 10.14). These are all at the



©2007 AwwaRF. All Rights Reserved.




*Source:* USGS 2006.

**Figure 10.19 Streamflow for the North Branch of the Metedeconk River during the sampling period in May 2005**

primary (1000) level, except for the Atlantic white cedar swamp category, which was retained in a separate group because of its perceived link to NOM generation.

The fundamental export model used here was a simple additive model founded on area-based export coefficients (Equation 10.23). Models based on a single synoptic survey were developed with TOC concentration as the dependent variable, rather than TOC loading. This is because runoff and streamflow are essentially fixed for these studies, and therefore loading is strongly autocorrelated with the product of concentration and catchment area.

$$\text{Concentration} = a + b(\text{frac}_{\text{LULC1}}) + c(\text{frac}_{\text{LULC2}}) + d(\text{frac}_{\text{LULC3}}) + \ldots \tag{10.23}$$

where: $a$ = constant term

$b$, $c$, $d$ = parameters for LULC category 1, 2, 3, etc.

$\text{frac}_{\text{LULC1}}$, $\text{frac}_{\text{LULC1}}$, $\text{frac}_{\text{LULC1}}$ = fraction of total drainage basin area that is covered by the particular LULC category (1, 2, 3, etc.)

The first model development and calibration was of the static export type. As previously stated, precursor concentrations were treated as a form of TOC and linked to TOC through the specific precursor content (i.e., precursor concentration per milligrams of carbon). Therefore the modeling effort first focused on TOC or DOC and then was amended with models for specific precursor content. The approach taken was to use forward stepwise MLR, solved using the Marquardt–Levenberg algorithm. Independent variables were introduced one at a time, and the model coefficients were tested for significance. Those that were significantly different from zero were retained, and the next independent variable was tested. This iterative process was done using a Windows-based statistical software package (SigmaStat 3.0; SPSS Inc., Chicago, Ill.). TOC was



©2007 AwwaRF. All Rights Reserved.

## Table 10.12
## Tertiary LULC categories (53) for eight sub-basins in the Brick watershed

| Category number* | Land use category | Sub-basin | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | NA | NH | NJ | NN | SA | SC | SJ | SL |
| 2140 | Agricultural wetlands (modified) | 0.94 | 0.99 | 1.13 | 0.06 | 0.09 | 0.10 | 0.15 | 0.16 |
| 7400 | Altered lands | 0.65 | 0.29 | 0.46 | 0.38 | 0.06 | 0.06 | 0.11 | 0.29 |
| 5300 | Artificial lakes | 0.30 | 0.23 | 0.21 | 0.16 | 0.78 | 0.59 | 0.48 | 0.41 |
| | Athletic fields (schools) | 0.31 | 0.13 | 0.00 | 0.00 | 0.17 | 0.08 | 0.00 | 0.00 |
| 6221 | Atlantic white cedar swamp | 0.00 | 0.00 | 0.00 | 0.00 | 0.85 | 0.19 | 0.19 | 0.03 |
| 7100 | Beaches | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1200 | Commercial/services | 2.31 | 2.22 | 0.79 | 0.26 | 2.32 | 1.97 | 0.23 | 0.19 |
| 2300 | Confined feeding operations | 0.24 | 0.27 | 0.34 | 0.19 | 0.09 | 0.10 | 0.04 | 0.00 |
| 4430 | Coniferous brush/shrubland | 0.55 | 0.40 | 0.42 | 8.53 | 1.19 | 1.03 | 1.02 | 1.38 |
| 4220 | Coniferous forest (>50% crown closure) | 3.53 | 4.49 | 6.68 | 0.12 | 7.48 | 6.76 | 9.47 | 6.78 |
| 4210 | Coniferous forest (10%–50% crown closure) | 0.49 | 0.03 | 0.04 | 0.00 | 1.67 | 0.19 | 0.04 | 0.05 |
| 6232 | Coniferous scrub/shrub wetlands | 0.12 | 0.17 | 0.27 | 0.78 | 0.32 | 0.25 | 0.44 | 0.93 |
| 6220 | Coniferous wooded wetlands | 3.25 | 6.71 | 10.65 | 24.55 | 6.53 | 6.99 | 11.01 | 15.13 |
| 2100 | Cropland and pastureland | 3.94 | 3.85 | 4.70 | 1.82 | 3.36 | 3.75 | 5.85 | 5.02 |
| 4420 | Deciduous brush/shrubland | 0.37 | 0.30 | 0.30 | 0.17 | 0.36 | 0.35 | 0.41 | 0.13 |
| 4120 | Deciduous forest (>50% crown closure) | 11.31 | 12.32 | 13.58 | 10.85 | 12.79 | 14.06 | 12.62 | 12.11 |
| 4110 | Deciduous forest (10%–50% crown closure) | 0.67 | 0.34 | 0.36 | 0.00 | 0.33 | 0.31 | 0.00 | 0.01 |
| 6231 | Deciduous scrub/shrub wetlands | 0.54 | 0.79 | 0.60 | 0.76 | 0.34 | 0.37 | 0.51 | 0.40 |
| 6210 | Deciduous wooded wetlands | 14.59 | 17.36 | 18.36 | 23.96 | 7.50 | 7.76 | 8.31 | 12.84 |
| 7430 | Disturbed wetlands (modified) | 0.53 | 0.22 | 0.24 | 0.04 | 0.37 | 0.39 | 0.61 | 0.38 |
| 7300 | Extractive mining | 0.06 | 0.05 | 0.02 | 0.00 | 0.24 | 0.26 | 0.07 | 0.23 |
| 2150 | Former agricultural wetland (becoming shrubby, not built up) | 0.10 | 0.20 | 0.32 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | Herbaceous wetlands | 0.81 | 0.76 | 0.36 | 0.13 | 0.67 | 0.75 | 1.07 | 1.10 |
| 1300 | Industrial | 0.57 | 0.16 | 0.20 | 0.12 | 1.60 | 0.87 | 0.37 | 0.54 |
| 1850 | Managed wetland in built-up, maintained recreational area | 0.05 | 0.07 | 0.00 | 0.00 | 0.26 | 0.27 | 0.36 | 0.00 |
| 1750 | Managed wetland in maintained lawn green space | 0.30 | 0.43 | 0.53 | 0.19 | 0.14 | 0.13 | 0.10 | 0.12 |
| | Military reservations | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 4440 | Mixed deciduous/coniferous brush/shrubland | 1.03 | 0.86 | 1.07 | 0.12 | 1.87 | 1.95 | 1.77 | 0.52 |
| 4312 | Mixed forest (>50% coniferous with >50% crown closure) | 2.30 | 2.18 | 2.91 | 2.70 | 5.33 | 5.59 | 8.11 | 10.70 |
| 4311 | Mixed forest (>50% coniferous with 10%–50% crown closure) | 0.20 | 0.05 | 0.06 | 0.00 | 0.11 | 0.11 | 0.07 | 0.06 |
| 4322 | Mixed forest (>50% deciduous with >50% crown closure) | 3.82 | 2.87 | 3.79 | 2.99 | 5.01 | 5.07 | 5.31 | 3.37 |
| 4321 | Mixed forest (>50% deciduous with 10%–50% crown closure) | 0.29 | 0.00 | 0.00 | 0.00 | 0.12 | 0.12 | 0.03 | 0.08 |
| 6252 | Mixed forested wetlands (coniferous dominant) | 2.15 | 2.41 | 3.83 | 2.97 | 3.99 | 4.38 | 7.07 | 14.15 |
| 6251 | Mixed forested wetlands (deciduous dominant) | 5.58 | 6.15 | 8.28 | 4.09 | 4.71 | 5.13 | 7.28 | 5.78 |
| 6234 | Mixed scrub/shrub wetlands (coniferous dominant) | 0.05 | 0.11 | 0.18 | 0.12 | 0.28 | 0.32 | 0.55 | 0.94 |
| 6233 | Mixed scrub/shrub wetlands (deciduous dominant) | 0.46 | 0.44 | 0.70 | 1.33 | 0.43 | 0.34 | 0.57 | 0.34 |
| 1600 | Mixed urban or builtup land | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 5200 | Natural lakes | 0.00 | 0.00 | 0.00 | 0.00 | 0.38 | 0.42 | 0.00 | 0.00 |
| 4410 | Old field (<25% brush covered) | 0.88 | 0.44 | 0.44 | 0.32 | 0.54 | 0.50 | 0.35 | 0.07 |
| 2200 | Orchards/vineyards/nurseries/horticultural area | 0.15 | 0.25 | 0.35 | 0.89 | 0.05 | 0.06 | 0.08 | 0.21 |
| 2400 | Other agriculture | 1.32 | 0.76 | 0.88 | 0.23 | 0.15 | 0.17 | 0.28 | 0.42 |
| 1700 | Other urban or builtup land | 2.00 | 1.77 | 0.84 | 0.59 | 1.87 | 1.51 | 1.15 | 0.81 |
| 4230 | Plantation | 0.09 | 0.02 | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1800 | Recreational land | 1.42 | 0.46 | 0.47 | 0.00 | 1.90 | 1.85 | 1.23 | 0.00 |
| 1110 | Residential, high density, multiple dwelling | 2.42 | 0.37 | 0.19 | 0.00 | 2.13 | 2.30 | 0.51 | 0.00 |
| 1140 | Residential, rural, single unit | 6.14 | 6.87 | 8.18 | 4.54 | 4.09 | 4.44 | 4.77 | 1.86 |
| 1130 | Residential, single unit, low density | 3.48 | 3.78 | 2.70 | 1.71 | 6.57 | 7.19 | 5.37 | 1.06 |
| 1120 | Residential, single unit, medium density | 16.34 | 13.99 | 0.63 | 0.05 | 8.96 | 8.99 | 0.38 | 0.03 |
| 5100 | Streams and canals | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.03 | 0.02 | 0.00 |
| 7500 | Transitional areas | 0.46 | 0.34 | 0.32 | 0.00 | 0.71 | 0.72 | 0.32 | 0.18 |
| 1400 | Transportation/communications/utilities | 2.01 | 1.88 | 2.05 | 0.64 | 1.12 | 1.11 | 1.14 | 1.04 |
| 7600 | Undifferentiated barren lands | 0.03 | 0.00 | 0.00 | 0.00 | 0.10 | 0.11 | 0.19 | 0.15 |
| 1461 | Wetland rights-of-way (modified) | 0.86 | 1.23 | 1.54 | 3.63 | 0.03 | 0.02 | 0.00 | 0.00 |
| | Total | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |

*The category numbers may be described in terms of primary, secondary, tertiary, and quaternary levels. For instance, in category 2140, the 2 represents the primary or 1000 level, 1 represents the secondary or 100 level, 4 represents the tertiary level, and 0 represents the quaternary level. Minor categories do not have LULC numbers.



©2007 AwwaRF. All Rights Reserved.

**Table 10.13**
**Secondary LULC categories (24) for eight sub-basins in the Brick watershed**

| Category number | Land use category | Sub-basin | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | NA | NH | NJ | NN | SA | SC | SJ | SL |
| 1100 | Residential | 28.37 | 25.01 | 11.70 | 6.30 | 21.75 | 22.91 | 11.03 | 2.96 |
| 1200 | Commercial and services | 2.31 | 2.22 | 0.79 | 0.26 | 2.32 | 1.97 | 0.23 | 0.19 |
| 1300 | Industrial | 0.58 | 0.16 | 0.20 | 0.12 | 1.60 | 0.87 | 0.37 | 0.54 |
| 1400 | Transportation, communication, and utilities | 2.86 | 3.11 | 3.60 | 4.27 | 1.15 | 1.13 | 1.14 | 1.04 |
| 1600 | Mixed urban or builtup land | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 1700 | Other urban or builtup | 2.30 | 2.20 | 1.37 | 0.78 | 2.01 | 1.64 | 1.26 | 0.93 |
| 1800 | Recreational land | 1.78 | 0.65 | 0.47 | 0.00 | 2.32 | 2.20 | 1.60 | 0.00 |
| 2000a | Cropland (not wetland) | 4.98 | 5.04 | 6.16 | 1.88 | 3.45 | 3.85 | 5.99 | 5.17 |
| 2000b | Agricultural wetlands (current and former) | | | | | | | | |
| 2200 | Orchard, vineyards, nurseries, and horticultural areas | 0.15 | 0.25 | 0.35 | 0.89 | 0.05 | 0.06 | 0.08 | 0.21 |
| 2300 | Confined feeding operations | 0.24 | 0.27 | 0.34 | 0.19 | 0.09 | 0.10 | 0.04 | 0.00 |
| 2400 | Other agriculture | 1.32 | 0.76 | 0.88 | 0.23 | 0.15 | 0.17 | 0.28 | 0.42 |
| 4100 | Deciduous | 11.98 | 12.66 | 13.94 | 10.85 | 13.12 | 14.38 | 12.62 | 12.12 |
| 4200 | Coniferous | 4.11 | 4.53 | 6.74 | 0.12 | 9.15 | 6.95 | 9.50 | 6.83 |
| 4300 | Mixed deciduous/coniferous | 6.60 | 5.10 | 6.76 | 5.69 | 10.57 | 10.89 | 13.51 | 14.21 |
| 4400 | Brushland/scrubland | 2.82 | 2.00 | 2.23 | 9.14 | 3.96 | 3.83 | 3.56 | 2.09 |
| 5100 | Streams and canals | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.03 | 0.02 | 0.00 |
| 5200 | Natural lakes | 0.00 | 0.00 | 0.00 | 0.00 | 0.38 | 0.42 | 0.00 | 0.00 |
| 5300 | Artificial lakes and reservoirs | 0.30 | 0.23 | 0.21 | 0.16 | 0.78 | 0.59 | 0.48 | 0.41 |
| 6200 | Interior wetlands | 27.55 | 34.91 | 43.21 | 58.69 | 25.64 | 26.48 | 37.00 | 51.64 |
| 7100 | Beaches | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 7300 | Extractive mining | 0.06 | 0.05 | 0.02 | 0.00 | 0.24 | 0.26 | 0.07 | 0.23 |
| 7400 | Altered lands | 1.18 | 0.51 | 0.70 | 0.42 | 0.42 | 0.45 | 0.72 | 0.67 |
| 7500 | Transitional areas | 0.46 | 0.34 | 0.32 | 0.00 | 0.71 | 0.72 | 0.32 | 0.18 |
| 7600 | Undifferentiated barren lands | 0.03 | 0.00 | 0.00 | 0.00 | 0.10 | 0.11 | 0.19 | 0.15 |
| | Total | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |

dependent on the fraction of wetlands (Equation 10.24), and no other variables could be introduced to significantly improve this.

$$\mathrm{TOC} = 1.28 + 15 f_{\mathrm{wetlands}} \quad (r^2 = 0.74) \tag{10.24}$$

The strength of this relationship was similar when either the fraction of wetlands in the entire watershed or only in the riparian zone was used (Figure 10.20).

The specific THMFP was also dependent on wetland area (Equation 10.25), and no other LULC variables were significant.



©2007 AwwaRF. All Rights Reserved.

**Table 10.14**
**Primary LULC categories* for eight sub-basins in the Brick watershed**

| Category number | Land use category | Sub-basin | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | NA | NH | NJ | NN | SA | SC | SJ | SL |
| 1000 | Urban or builtup land | 38.21 | 33.35 | 18.14 | 11.74 | 31.15 | 30.71 | 15.62 | 5.66 |
| 2000 | Agricultural land | 6.69 | 6.32 | 7.74 | 3.19 | 3.74 | 4.18 | 6.38 | 5.80 |
| 4000 | Forestland | 25.51 | 24.29 | 29.67 | 25.80 | 36.81 | 36.04 | 39.19 | 35.26 |
| 5000 | Water | 0.30 | 0.23 | 0.21 | 0.16 | 1.18 | 1.04 | 0.50 | 0.41 |
| 6000 | Wetlands | 27.55 | 34.91 | 43.21 | 58.69 | 24.79 | 26.28 | 36.81 | 51.61 |
| 6221 | Atlantic white cedar swamp | 0.00 | 0.00 | 0.00 | 0.00 | 0.85 | 0.19 | 0.19 | 0.03 |
| 7000 | Barren, disturbed, and transitional land | 1.74 | 0.90 | 1.04 | 0.42 | 1.47 | 1.55 | 1.31 | 1.23 |
| | Total | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |

*Six primary categories plus the Atlantic white cedar swamp category.

$$\frac{\text{THMFP}}{\text{TOC}} = 58.5 + 36 f_{\text{wetlands}} \quad (r^2 = 0.47) \tag{10.25}$$

Interestingly, the predictions call for about 58 μg/mg C in the absence of wetlands, which is close to the national average for this parameter. At the highest wetland levels (100% coverage), the specific THM content would be about 95 μg/mg C, near the upper limit for North American waters. The implications of this are that wetlands generate organic carbon that is especially high in THM precursors (~95 μg/mg C, in this case).

The THAA precursor content was quite consistent at about 80 μg/mg C (see Figure 10.15). As a result, the export model was expected to show little dependency on any particular LULC type. Surprisingly, only the urban category was entered into the model with the stepwise routine, as shown in Equation 10.26.

$$\frac{\text{THAAFP}}{\text{TOC}} = 93.1 - 54 f_{\text{urban}} \quad (r^2 = 0.63) \tag{10.26}$$

When this variable was excluded from consideration, the routine found three significant land use types, with wetlands once again being the most important (Equation 10.27).

$$\frac{\text{THAAFP}}{\text{TOC}} = 27.9 + 100 f_{\text{wetlands}} - 163 f_{\text{ag}} + 1{,}900 f_{\text{barren}} \quad (r^2 = 0.917) \tag{10.27}$$

This gives wetlands a positive coefficient, suggesting that NOM coming from wetlands is especially high in THAA precursors, just as it is high in THM precursors. This is all in addition to the fact that wetlands generate more NOM to begin with.

The DHAA precursors were also quite consistent at about 35 μg/mg C (see Figure 10.15). Again, this led to the presumption that the export model would show little dependency on any particular LULC type. In fact, this was the case, and there was no statistical justification for anything more than a constant DHAA/TOC ratio. These model coefficients are summarized in Table 10.15.



©2007 AwwaRF. All Rights Reserved.

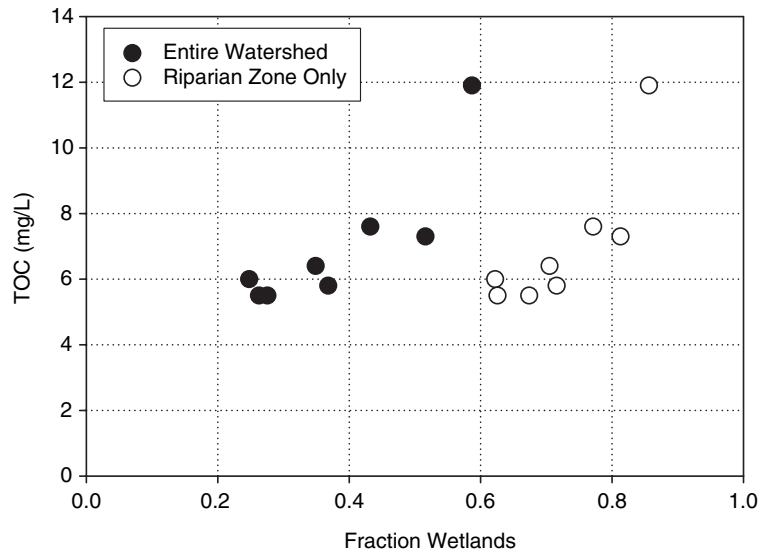



**Figure 10.20 Relationship between wetland area and TOC concentration**

**Table 10.15**
**Model summary for full watershed LULC**

| Dependent variable | Constant a | Independent variable 1 | | Independent variable 2 | | Independent variable 3 | |
|---|---|---|---|---|---|---|---|
| | | Name | Coefficient 1 | Name | Coefficient 2 | Name | Coefficient 3 |
| TOC | 1.28 | Wetlands | 15 | | | | |
| THMFP/TOC | 58.5 | Wetlands | 36 | | | | |
| THAAFP/TOC | 93.1 | Urban | –54 | | | | |
| THAAFP/TOC* | 27.9 | Wetlands | 100 | Agricultural | –163 | Barren | 1,900 |
| DHAAFP/TOC | 34.0 | | | | | | |

*Excluding urban land from consideration.

Land use impacts are generally believed to become attenuated by travel time and distance to the receiving body of water. For this reason, the choice was made to run a parallel analysis considering only LULC in a predetermined riparian zone. This required that the land use analysis be rerun, restricting the attention to land area within 200 feet of a major stream channel (Figure 10.21). The criteria used to identify major streams were based on the stream class. All streams of class "x-3" and above were included, where "x" is the highest class stream in the basin.

This analysis resulted in similar, though not identical, export models. In general, the goodness of fit was no better, yet no worse, than with the full catchment LULC. The TOC export model showed a higher sensitivity to wetlands area when considering only riparian zones of wetlands:buffer. This is evident in the higher coefficient shown in Equation 10.28.

$$\text{TOC} = -8 + 20.7 f_{\text{wetlands:buffer}} \quad (r^2 = 0.68) \tag{10.28}$$



©2007 AwwaRF. All Rights Reserved.

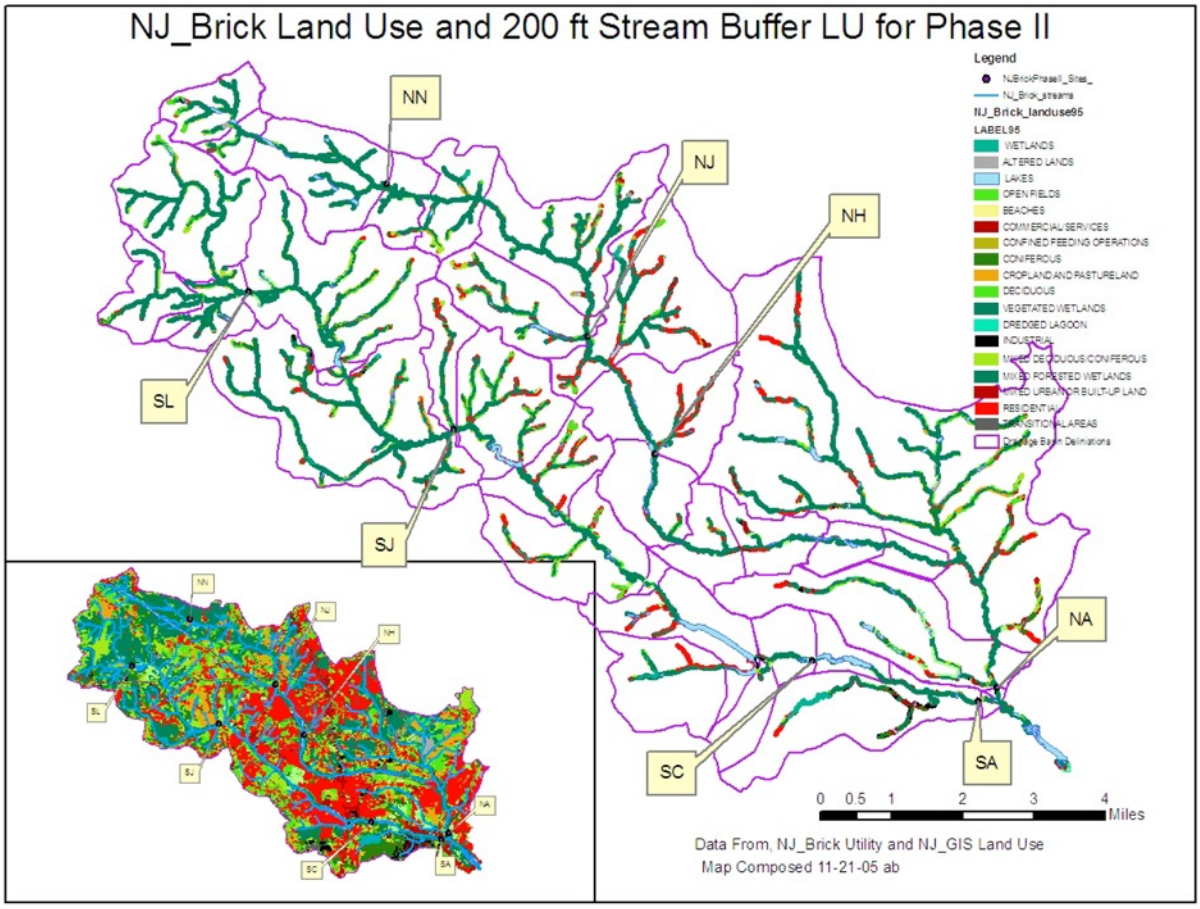



**Figure 10.21 Full set of LULC categories in riparian zones across the Metedeconk watershed**

**Table 10.16**
**Model summary for riparian zone LULC**

| | | Independent variable 1 | | Independent variable 2 | |
|---|---|---|---|---|---|
| Dependent variable | Constant | Name | Coefficient 1 | Name | Coefficient 2 |
| TOC | –8 | Wetlands | 20.7 | | |
| THMFP/TOC | 34.3 | Wetlands | 52.4 | | |
| THAAFP/TOC | 91 | Urban | –113 | | |
| THAAFP/TOC* | 27.9 | Wetlands | 100 | Barren | –2,600 |
| DHAAFP/TOC | 34.0 | | | | |

*Excluding urban land from consideration.



©2007 AwwaRF. All Rights Reserved.

The THM model (Equation 10.29) also showed a higher sensitivity for wetlands area. Here, the predicted range was from 34 µg/mg C to 87 µg/mg C for locations with 0% and 100% wetlands coverage, respectively. This is slightly lower than the range calculated based on the full catchment LULC, yet it still falls within the observed range of values found across North America.

$$\frac{\text{THMFP}}{\text{TOC}} = 34.3 + 52.4 f_{\text{wetlands:buffer}} \quad (r^2 = 0.49) \tag{10.29}$$

The THAA model is almost identical to the one based on all areas in the watershed (compare Equations 10.30 and 10.26). Excluding urban land from the analysis results in a model (Equation 10.31) that is very similar to Equation 10.27, except that agricultural land isn't retained. Table 10.16 contains a summary of these riparian zone export models.

$$\frac{\text{THAAFP}}{\text{TOC}} = 91 - 113 f_{\text{urban:buffer}} \quad (r^2 = 0.65) \tag{10.30}$$

$$\frac{\text{THAAFP}}{\text{TOC}} = 29.1 + 98 f_{\text{wetlands:buffer}} - 2{,}600 f_{\text{barren:buffer}} \quad (r^2 = 0.835) \tag{10.31}$$

***Export/Attenuation Model.*** The models presented so far for TOC, THM/TOC, and THAA/TOC provide explanations for variable export from different parts of the watershed. However, this set of static export models does not account for several other factors that are sometimes considered with lumped or nondistributed models:

1. Degradation that increases with travel time
2. Adsorption/degradation that occurs because of contact with different soils
3. Effects of temperature and season
4. Effects of antecedent rainfall and soil moisture or groundwater table (flow paths)

For this work, it was proposed to incorporate various elements of items 1 and 2 to expand the simple export model to a combined export/attenuation model. Factor 1 can be roughly correlated to drainage area and slope. A simple form that requires a minimum of site-specific information is the exponential model based only on drainage area, shown in Equation 10.32.

$$t = r(\text{DA})^z \tag{10.32}$$

where $t$ = travel time
DA = drainage area
$r$, $z$ = coefficients

Then, taking the TOC degradation rate as a simple first-order process with rate constant, $k$:

$$\text{TOC} = \text{TOC}_i e^{-kt} \tag{10.33}$$



©2007 AwwaRF. All Rights Reserved.

and combining, Equation 10.34 is obtained:

$$\mathrm{TOC} = \mathrm{TOC}_i e^{-q\mathrm{DA}^z} \tag{10.34}$$

where TOC = concentration of TOC
$\mathrm{TOC}_i$ = theoretical initial concentration of TOC exported prior to degradation
$q$ = coefficient

It should be recognized that some of the TOC will degrade quickly, whereas other fractions will be relatively refractory. A simple way of handling this issue is to assume a fixed biodegradable fraction that is subject to loss in accordance with Equation 10.34. This results in the following model for TOC loss, Equations 10.35 to 10.37.

$$\mathrm{TOC} = \mathrm{TOC}_r + \mathrm{TOC}_l \tag{10.35}$$

$$\mathrm{TOC} = \mathrm{TOC}_r + \mathrm{TOC}_{li} e^{-q\mathrm{DA}^z} \tag{10.36}$$

$$\mathrm{TOC} = [(1 - f_L) + f_L e^{-q\mathrm{DA}^z}]\mathrm{TOC}_i \tag{10.37}$$

where $\mathrm{TOC}_r$ = refractory TOC
$\mathrm{TOC}_l$ = labile TOC
$f_L$ = fraction of TOC that is labile

In this approach, TOC can be estimated from a simple static export model based on the breakdown of LULC characteristics for the watershed. However, this larger export/attenuation model did not result in improved TOC prediction from the Brick watershed, nor was the quantity of data sufficient to support the additional parameters.

***Summary.*** A 1-day sampling event was conducted in May 2005, encompassing nine locations in the Metedeconk River watershed. These samples were analyzed for TOC and DBP precursors. Relevant geographical and land use data were assembled in parallel with the water quality analyses. In general, the organic carbon contents were high, and DBP precursor levels were similarly high. The specific DBP precursor levels (normalized to 1 mg/L TOC) were well within the typical range for North American waters.

Export coefficient modeling showed a strong dependence on wetlands area for TOC concentrations in the Metedeconk River watershed. In addition, there was a similar relationship with specific THM and HAA content. This implies that wetlands are important sources of organic carbon, but especially of THM and HAA precursors. The proximity of these wetlands with the Metedeconk River channel, however, did not seem to have a measurable affect on DBP levels.

### Cambridge System Analysis

The second catchment analyzed was the Cambridge, Mass., water supply basin. This catchment is composed of three reservoirs and associated tributaries that are divided among 11 sub-basins. Each sub-basin and corresponding TOC concentration contributed to the system analysis. Similar to the Metedeconk system, the Cambridge watershed is urbanized and susceptible to anthropogenic activity. Equation 10.38 presents the resulting empirical relationships.



©2007 AwwaRF. All Rights Reserved.

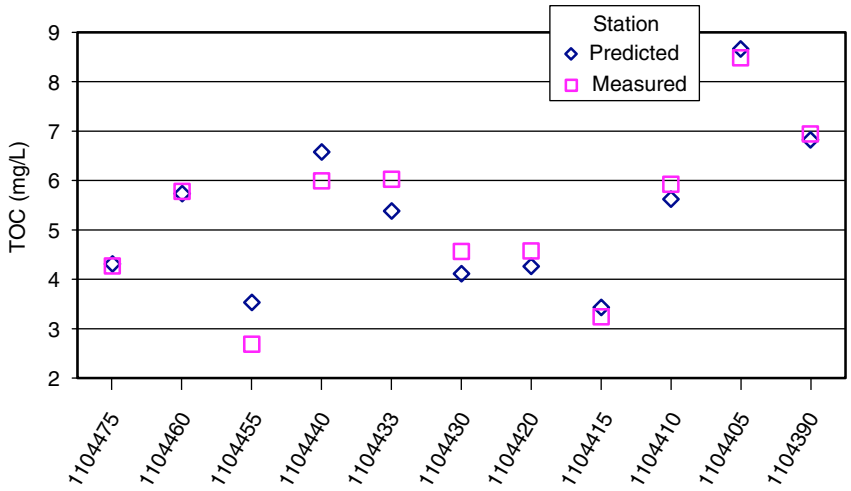



**Figure 10.22 Cambridge regression comparison between empirical TOC predictions and actual TOC concentration**

$$\begin{aligned} TOC_{overall} = {} & 6.89 + 51.54 \text{ agriculture} - 27.31 \text{ urban} + 13.87 \text{ field} \\ & - 3.22 \text{ forest} + 22.68 \text{ high residential} - 10.95 \text{ suburban} \\ & + 8.42 \text{ water} + 99.14 \text{ wetland} \end{aligned} \tag{10.38}$$

Nonforested wetlands most significantly influenced TOC within the water supply catchment. This land use type comprised only 1.5% of the total drainage basin, but its location in stream headwaters directly contributed to the amount of NOM present in the reservoirs and tributaries. The urban development land use type also strongly influenced organic matter and, similar to the wetland land use type, was only a small fraction of the entire watershed (3%). Forested land represented nearly 50% of the catchment and also influenced TOC. The resulting empirical model was quite accurate, with an overall average error of 8%. Figure 10.22 presents the comparison between predicted TOC values and measured TOC concentration.

The largest error (31%) was associated with sample station 1104455 (Figure 10.22). The difference was attributed to the sub-basin LULC composition. The sub-basin was composed of impervious surfaces, significant commercial and industrial development, and a large amount of high-density residential space, and these LULC characteristics significantly increased the predicted value of TOC. High-residential areas also positively contributed to the TOC prediction, and the large high-density residential areas within the sub-basin contributed to an overprediction in TOC.

The erroneous results at station 1104433 derived from headwater wetlands. The empirical model underpredicted TOC within this sub-basin. The empirical model did not capture the significant influence of headwater wetlands on TOC. This land use coefficient and small corresponding wetlands area within the sub-basin did not properly estimate the TOC concentration.

The average error between predicted TOC values and measured TOC concentration within the Cambridge watershed (8%) was quite low, and the difference between the error of this analysis and the national analysis (26%) was significant. The small error was attributed to the low variability within the underlying data. Significant fluctuations in TOC data were highlighted by the erroneous predictions at sampling stations 1104455 and 1104433.



©2007 AwwaRF. All Rights Reserved.

### *Cambridge Seasonal Analysis*

The initial seasonal empirical models generated large errors. The average seasonal error was 42% and exceeded the overall error (8%) by 34%. The nonforest wetland coefficient was negative and was attributed to the erroneous results. This behavior contradicted the findings of a USGS study within the same watershed (Waldron and Bent 2001). Therefore, the regression analysis was performed again, and each resulting model demanded that the wetland land use coefficient be positive. The resulting empirical modes are shown in Equations 10.39 to 10.42.

$$TOC_{winter} = -48.88 \text{ agriculture} + 5.63 \text{ forest} - 9.39 \text{ water} \tag{10.39}$$

$$TOC_{spring} = 9.89 - 25.02 \text{ urban} - 4.24 \text{ industry} + 17.74 \text{ high residential} - 7.64 \text{ suburban} + 6.47 \text{ water} + 125.14 \text{ wetlands} \tag{10.40}$$

$$TOC_{summer} = 9.28 + 108.84 \text{ agriculture} - 38.92 \text{ urban} + 17.04 \text{ fields} - 8.10 \text{ forest} + 42.21 \text{ high residential} - 15.05 \text{ suburban} + 16.24 \text{ water} + 125.09 \text{ wetlands} \tag{10.41}$$

$$TOC_{fall} = 7.01 - 65.21 \text{ agriculture} - 15.49 \text{ urban} + 4.66 \text{ forest} + 44.12 \text{ fields} - 5.4 \text{ industry} - 2.9 \text{ suburban} + 99.27 \text{ wetlands} \tag{10.42}$$

The accuracy of the empirical models improved following the correction to the wetlands coefficient. The error associated with the modified seasonal empirical models was reduced to 18% but was significantly larger than the period of record analysis. Unlike the seasonal results within the Metedeconk catchment, the winter and summer seasons within the Cambridge water supply catchment were the most and least erroneous seasons, respectively.

Figure 10.23 presents the comparison between predicted TOC values and measured TOC concentration during the winter season. The winter seasonal error, 27%, was attributed to the erroneous predictions at several sampling locations. The empirical model significantly underpredicted TOC concentration at station 1104440 and overpredicted TOC concentration at station 1104430. The erroneous results (high prediction) were attributed to the large forested area within the sub-basin. Curiously, station 1104440 had similar LULC composition, and the erroneous prediction was below the measured TOC concentration. However, an extremely high TOC measurement in January greatly affected the seasonal average and, without this outlier, the TOC prediction falls above the actual TOC concentration.

The summer seasonal error, 10%, was significantly lower than the winter seasonal error (27%), but the empirical model yielded an extremely erroneous prediction for station 1104460, Figure 10.24. The agriculture land use significantly influenced the summer empirical model. The sub-basin above 1104460 was dominated by agriculture, and this LULC composition was attributed to the erroneous prediction.

The study within the Cambridge basin was conducted during a 13-month period of time and included a “wet” hydrologic year, whereas the national survey employed at least 5 years of data. Without additional data, it was difficult to completely assess the accuracy of the empirical models within the Cambridge catchment and draw a comparison between it and the national



©2007 AwwaRF. All Rights Reserved.

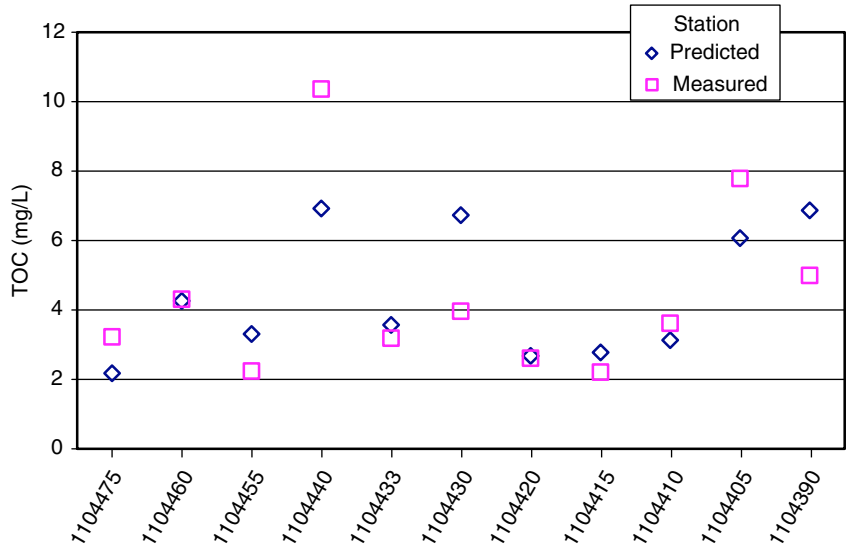



**Figure 10.23 Cambridge winter regression comparison between empirical TOC predictions and actual TOC concentration**

models. As a result, the empirical models from the national survey or the Metedeconk River system were believed to provide more consistent and accurate results.

***Trihalomethane Formation Potential Analysis***

As a part of the USGS study, THMFP data were also collected. Thus, a regression analysis was performed to identify the relationships between THMFP and LULC. The THM data collected during the study are shown in Table 10.17. Specific yield (THMFP/DOC) was used as the dependent value within the regression analysis. The resulting empirical model is presented in Equation 10.43:

$$\begin{aligned} THMFP_{overall} = {} & 235.112 + 3{,}896.63 \text{ agriculture} - 2{,}466.73 \text{ urban} \\ & + 1{,}225.62 \text{ fields} - 301.84 \text{ forest} + 2{,}086.34 \text{ high residential} \\ & - 726.67 \text{ suburb} + 749.99 \text{ water} + 6{,}420.4 \text{ wetland} \end{aligned} \tag{10.43}$$

The empirical model indicated strong influence by the agriculture and nonforested wetland land use characteristics. The open field and high-density residential land use types also influenced the empirical relationships between THMFP and LULC. The average error associated with the empirical model was 5%, and the comparison among predicted THMFP values and measured TMHFP concentrations is shown in Figure 10.25. The data obtained for this analysis precluded accurate seasonal analyses. However, the results for the period of record were extremely accurate and indicated the possibility of further such model development.



©2007 AwwaRF. All Rights Reserved.

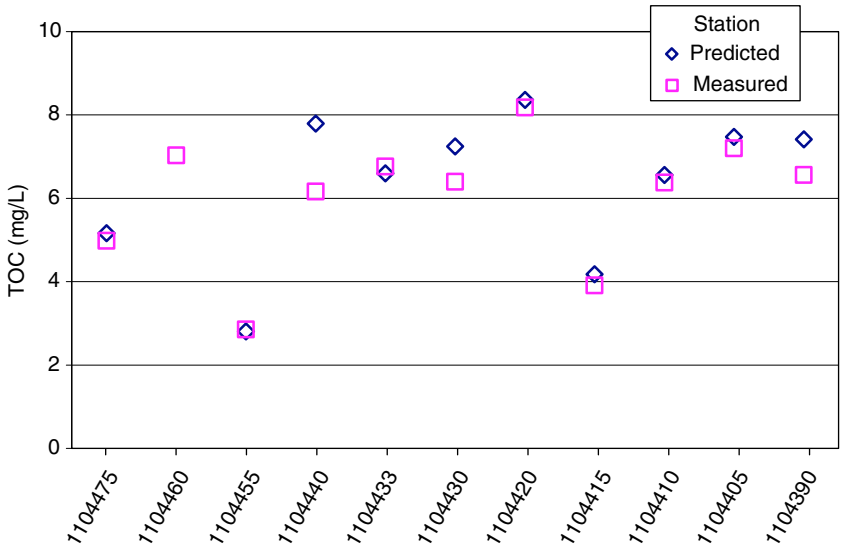



**Figure 10.24 Cambridge summer regression comparison between empirical TOC predictions and actual TOC concentration**

**Table 10.17**
**THMFP data for Cambridge water supply catchment**

| Station | Median THMFP (µg/L) | DOC (mg/L) |
|---|---|---|
| 1104390 | 240.0 | 6.94 |
| 1104405 | 224.5 | 8.48 |
| 1104410 | 177.5 | 5.92 |
| 1104415 | 163.5 | 3.24 |
| 1104420 | 200.0 | 4.57 |
| 1104430 | 194.0 | 4.56 |
| 1104433 | 178.0 | 6.03 |
| 1104440 | 273.5 | 5.99 |
| 1104455 | 85.0 | 2.69 |
| 1104460 | 225.0 | 5.78 |

## SITE-SPECIFIC ANALYSIS OF MICROBIAL COMMUNITIES

### Introduction

Natural water communities represent microbial assemblages with high diversity, which might be correlated to the diversity, availability, and degradability of NOM. As most naturally occurring microorganisms cannot be grown in culture, the focus here was based on a cultivation-independent technique to examine community structure and diversity. The species composition



©2007 AwwaRF. All Rights Reserved.

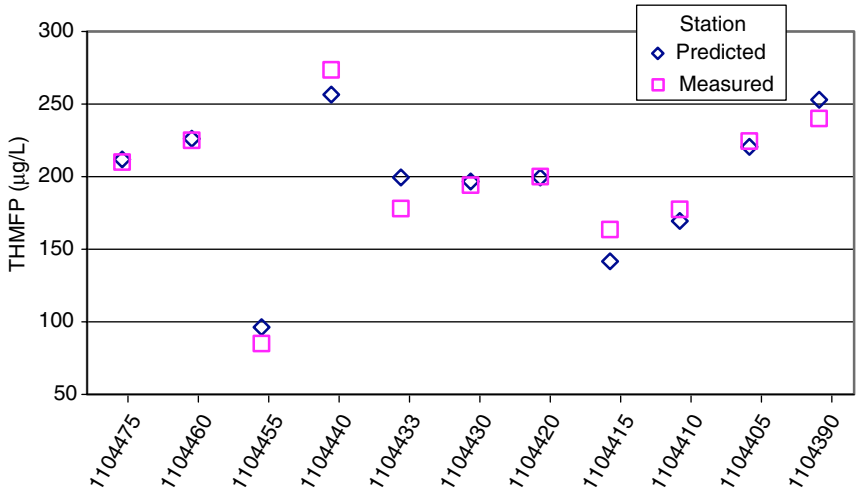



**Figure 10.25 Cambridge THMFP regression comparison between empirical TOC predictions and actual TOC concentration**

and diversity of the microbial assemblage in watershed water was analyzed using molecular microbial analysis based on marker genes of the previously extracted total community nucleic acid. This is the state-of-the-art classification of microbes based on their marker gene 16S rRNA. In order to understand concepts of functionality and selective carbon pool turnover within NOM as the source for DBPs, and how they link to DBP stability or regrowth potential, the establishment of the dominant diversity within a sample must be a key requirement.

Past studies of drinking water microbial quality most often incorporated culture-based techniques (Besner et al. 2002). However, certain organisms, including pathogens such as *Legionella pneumophila* and *Mycobacterium* spp., are difficult to isolate and quantify from environmental samples. Direct sequence analysis of 16S rRNA clone libraries has been shown to complement the information gained from culture-based work, and studies of waters repeatedly show that molecular identification of isolates by sequencing 16S rRNA is more complete than biochemical tests (Singh et al. 2003). In this study, molecular analysis of the bacterial community and the identity of bacteria present in raw water samples were provided. Water from the Newport News watershed was examined by a direct DNA analysis of 16S rRNA-cloned sequences.

The method of analysis was as recently reported by Stout and Nüsslein (2005), with only minor exceptions. For the phylogenetic analysis, all sequences created in the clone library were aligned with ClustalW (<http://www.ebi.ac.uk/clustalw/>), and the phylogenetic dendrogram was created with Jalview v1.8 (<http://www.ebi.ac.uk/jalview/>).

## Results

DNA sequencing and phylogenetic analysis were performed on the extensive clone library of more than 100 sequences, of which 54 clones represented dominant and rare phylotypes (Figures 10.26–10.28). The phylogenetic tree indicates the relationship of the sequences detected (indicated by "fHM…") with those already present in public sequence databases for the marker gene 16S rRNA. Molecular community analysis based on the molecular phylogenetic fingerprints of 16S rRNA clone libraries showed a diverse species composition within this raw water sample



©2007 AwwaRF. All Rights Reserved.

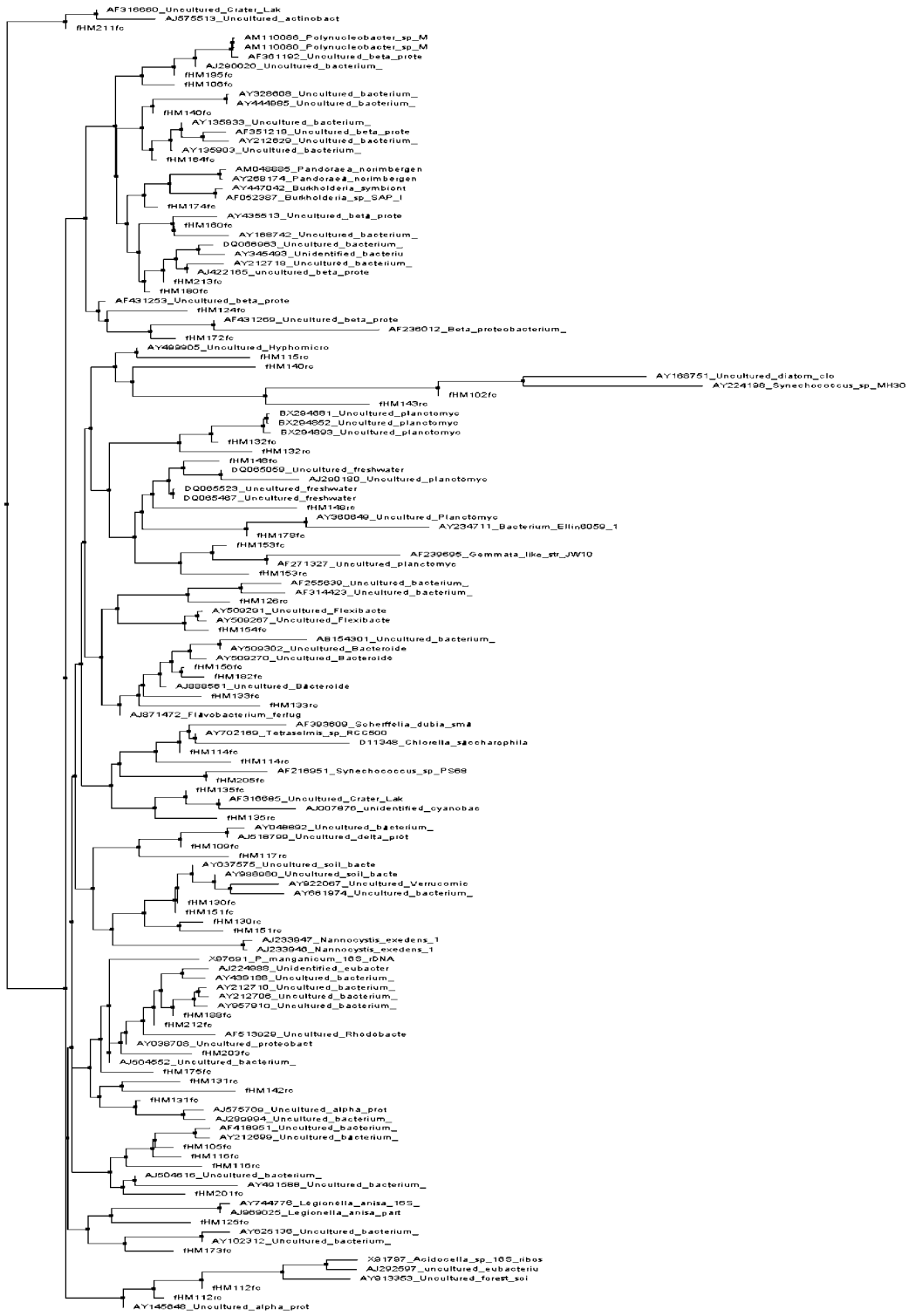



**Figure 10.26 Phylogenetic analysis of bacterial 16S rRNA in raw water from the Newport News watershed (full chart)**



©2007 AwwaRF. All Rights Reserved.

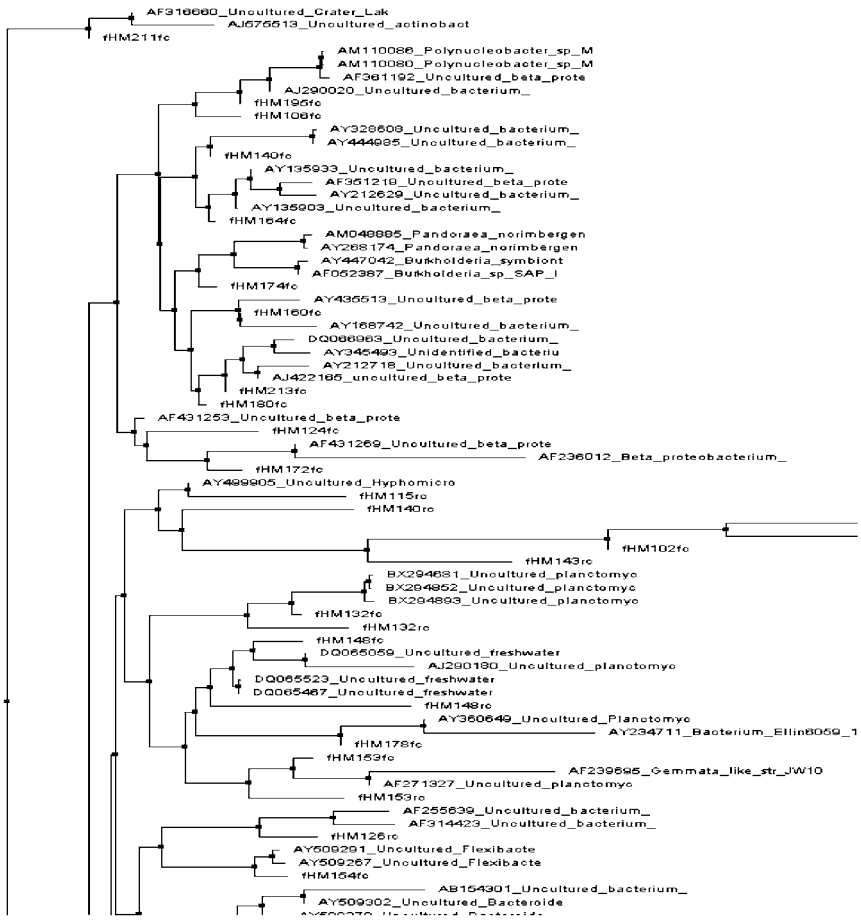



**Figure 10.27 Phylogenetic analysis of bacterial 16S rRNA in raw water from the Newport News watershed (expansion of top half of Figure 10.26, for clarity)**

(more water samples are currently under investigation). Most of the clones did not exhibit a high match with any of the sequences found in the databases. Therefore, some sequences with low similarity were affiliated based on their positions in the current phylogenetic trees. All clones were members of the domain Bacteria or chloroplasts of algae. Almost all sequences were related to those marker genes previously found in freshwater systems, mostly lakes but also some rivers, hyporheic zones, and groundwaters.

The presence of cyanobacteria is not unexpected, and so is the finding of bacteria commonly detected in soils. However, the presence of close relatives to *Legionella anisa*, as indicated by the sequence fHM125fc, shows contamination of an organism closely related to *L. pneumophila*, a known waterborne-opportunistic pathogen that causes pneumonia (Legionnaires' disease) in immunocompromised individuals. In fact, recent literature (van der Mee-Marquet et al. 2006) suggests that *L. anisa* is one of the most frequent species of *Legionella*—other than *L. pneumophila*—in the environment and may mask water contamination by *L. pneumophila*. This suggests that there is a risk of *L. pneumophila* infection in immunocompromised patients if water is contaminated with *Legionella* species other than *L. pneumophila*.



©2007 AwwaRF. All Rights Reserved.

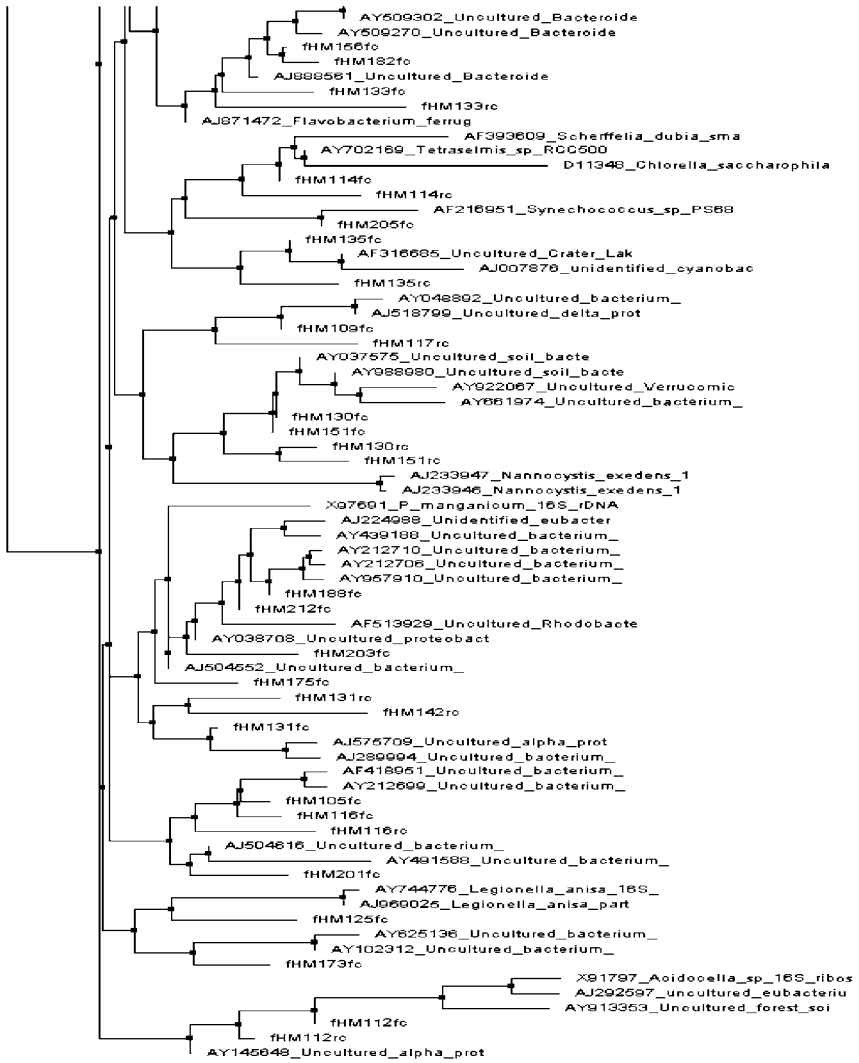



**Figure 10.28 Phylogenetic analysis of bacterial 16S rRNA in raw water from the Newport News watershed (expansion of bottom half of Figure 10.26, for clarity)**

*Legionella* species are common inhabitants of natural and artificial freshwater systems, though they are difficult to isolate from drinking water distribution systems (Fields 1996). Although molecular techniques allow for accurate identification of *Legionella* species without depending on culturing methods, more work is needed to examine the occurrence and pathogenicity of these bacteria in water distribution systems so that public health risk can be determined.

One of the sequences was closely related to the common genus, *Hyphomicrobium*, which exhibited a sequence identical to a strain recently described in fish-farm sediments (Bissett, Bowman, and Burke 2006). In a drinking water distribution system analysis, Williams et al. (2004) found *Hyphomicrobium* to dominate chloraminated water system biofilms, suggesting that these organisms survive well following exposure to disinfectant. These results are not surprising considering that this bacterial group has been isolated in many oligotrophic aquatic environments including mineral drinking water (Gonzalez, Gutierrez, and Grande 1987). No pathogenic



©2007 AwwaRF. All Rights Reserved.

*Hyphomicrobium* species has been isolated thus far, so this bacterial group does not represent a public health risk. However, it is reasonable to speculate that they may possibly play an important role in carbon utilization in distribution systems, as this group is known to grow under carbon-limited conditions (Corpe and Jensen 1996), and thus this detail is important to this study.

Several of the sequences were related to not-yet-classified organisms that are also found in oligotrophic freshwaters, groundwater, or freshwater reservoirs. Some were associated with metal-rich particles, which is common in waters with seasonal metal precipitation, and can be biogenic and thus temperature and DOC dependent. Other sequences were most closely related to not-yet-cultivated microbes found with an increased level of humics within lake beds. Some of the sequences also indicated chloroplasts from the common genus of algae, *Chlorella*.

Only a few of the clone sequences within this raw waters sample could be phylogenetically identified to the species level by sequence analysis. For most other sequences, no close relatives could be found in all databases investigated. Many sequences were affiliated with candidate divisions and subdivisions, but it was impossible to speculate about their nearest relatives. This analysis shows a clear dominance of multiple unknown species, and also genera, which underscores the need for additional diversity studies of raw waters to establish the ability to better correlate microbial species' presence with NOM fractions as substrate groups, and a more comprehensive understanding of their influence in subsequent DBP formation.

## SUMMARY AND CONCLUSIONS

### Empirical Analyses

This investigation employed several analyses to construct empirical models across the nation and within individual watersheds. Development of the empirical relationships was limited by the available data, and variability within these data introduced errors within empirical models. The results of regression analysis were thus only as accurate as the collected data and their statistical representativeness. In general, errors decreased as the empirical analysis focused on more regional areas. In other words, catchment-specific results based on multiple sampling points within the watershed had a greater predictive capability than the national models developed from data obtained across several watersheds. In all cases, results suggest that incorporation of temporally varying independent variables in the relationships may be beneficial. Inclusion of such parameters would enable greater inclusion of the raw, rather than just summary, data in the analysis. All models have exhibited an inability to fully explain underlying variability in TOC and precursor data based on land use alone. The importance of various aspects of the regression analyses are discussed further in the following paragraphs.

#### *Wetlands*

Regardless of analysis method, wetlands persisted across 13 watersheds (11 national and 2 individual) as a strong indicator of TOC contributions within water supply catchments. Several studies prior to this investigation demonstrated the large impact of headwater wetlands on riparian ecosystems. Though highest concentrations of NOM in wetlands occurred during winter and summer, significant contributions of NOM occurred during periods of increased flow (spring and fall; Mulholland and Kuenzler 1979, Hongve 1999, Meyer and Tate 1983). Several sampling stations within the Cambridge supply system contained wetlands in the tributary headwaters.



©2007 AwwaRF. All Rights Reserved.

Most notable were stations 1104433 and 1104390, where the measured TOC concentrations often exceeded that of the predicted concentrations. This behavior indicated the significance and impact of wetlands within a watershed (Kuenzler 1979).

Wetland area comprised a large percentage of land and yielded the higher TOC values within the Metedeconk River catchment. Both branches of the river pass through miles of forested land with wetland areas immediately adjacent to each branch. The wetlands land use characteristic was retained within each empirical model from this watershed. The empirical models from the national survey also described the strong influence of wetlands on TOC concentration. This land use type was dominant in each series of relationships between LULC and TOC. Furthermore, as with the individual analyses, the percentage of area associated with wetlands did not exceed 30% of the total catchment and typically ranged between 2% and 8%.

### *Anthropogenic Influence*

Anthropogenic activity was retained within the national empirical models and was represented by such LULC characteristics as urban development, industry, and residential areas. The national seasonal empirical models included anthropogenic activity during the summer and fall seasons as well as within the overall model. Similarly, the Cambridge empirical models were all influenced by urban development except the winter model. The overall model also indicated the influence of two residential land use types. The seasonal models for each branch of the Metedeconk River did not include urban development but were influenced by a variety of residential land use types as well as by the field land use characteristic. The overall system empirical relationships included urban development, industry, and agriculture. The variability among these models highlighted the dependence of a regression analysis on available data.

### *Seasonality*

Seasonality also influenced the concentration of TOC within watersheds. The empirical models from the 13 watersheds were indicative of vast seasonal differences. Average seasonal error varied across each data set. The empirical models for the Metedeconk River system exhibited the most significant seasonal variation between predicted TOC values and measured TOC concentration, whereas the national empirical models generated the least seasonal difference between predicted TOC values and measured TOC concentration. Without hydrologic and hydraulic data that corresponded to the TOC data, it was difficult to identify probable causes for variation, such as rainfall, snowmelt, and discharge. Furthermore, without concurrent streamflow data, the TOC was less understood, especially the seasonal variability of TOC predictions observed within the Metedeconk River system. The fluctuation in prediction values yielded high seasonal error. Discharge data are available for some of the study watersheds included in the analyses, such as the Metedeconk River system. It is anticipated that acquisition and inclusion of these data in the analyses will be done in the future. The current focus was on development of a regional export coefficient for the continental United States.

### *Similarities Between TOC and THMFP*

The Cambridge empirical models for TOC and the model for THMFP were quite similar. Each constituent was strongly influenced by the nonforest wetlands LULC characteristic and was



©2007 AwwaRF. All Rights Reserved.

affected by open fields and agriculture. The impact of anthropogenic activity varied slightly between the two constituents. The high-density residential land use type largely influenced THMFP concentration, whereas the suburban land use type was more influential for TOC. The forested LULC characteristic that was retained within the THMFP model was not present in the overall Cambridge system model but was found within the seasonal models. The similarities among empirical models suggested that the behavior of TOC and THMFP were similarly influenced by LULC characteristics and anthropogenic activity. The Brick models also show a strong parallel between TOC and THM or HAA precursors. This is partially attributed to the relative homogeneity of NOM characteristics across this catchment. DBP models for this case study were based on a single synoptic survey.

### *Land Use Region*

For the Metedeconk watershed, regression analyses were conducted based on both land use of the full catchment and land use within 61 m of the major stream channels. These analyses resulted in similar, though not identical, export models. In general, a higher sensitivity to wetlands area was found when only riparian zones were considered. However, the goodness of fit remained essentially the same. Similar analyses for other watersheds are necessary to determine whether limiting land use to the riparian zone can improve empirical model performance.

## Short-Term DBP Precursor and Microbial Analyses

DBP precursor content was analyzed for surface water samples collected at several locations within the Metedeconk watershed. Similar analyses were conducted for several additional utilities, but thorough review of these data with respect to empirical analyses has not been completed. The Metedeconk watershed contains elevated levels of TOC and specific THMFP, in comparison to the national average, but moderate THAA/THM and low DHAA/THM ratios. Based on the data collected from a single survey, reasonable predictive equations for THAAFP/TOC and TOC could be developed, although relationships for THMFP and THAAFP ratios were weaker. Additional data collection efforts and incorporation of temporal dependence would likely improve model performance. A watershed-based focus appears to be more appropriate than a national level–based focus.

Water samples from the Newport News watershed were examined by direct DNA analysis of 16S rRNA-cloned sequences to establish the dominant microbial diversity in the samples. Only a few of the clone sequences within this raw water could be phylogenetically identified to the species level. A clear dominance of multiple unknown species, and also genera, underscores the need for additional diversity studies of raw waters. Such studies will facilitate the ability to correlate microbial species' presence with NOM fractions as substrate groups, and thus provide a more comprehensive understanding of their influence on subsequent DBP formation.

## FUTURE WORK

Future work is necessary to expand on the analyses presented in this chapter. In particular, the DBP precursor empirical analyses need to be extended across the watersheds for which DBP formation data were obtained as part of this study. In addition, the incorporation of temporally variable independent variables, such as precipitation or streamflow, may provide a more robust means of predicting loading than export from land use alone.



©2007 AwwaRF. All Rights Reserved.

# CHAPTER 11
# MECHANISTIC MODELING

## OVERVIEW

"Mechanistic modeling" is defined as a deterministic method that generates a predictable system response to stimuli given an exact set of data related to initial system conditions. Mechanistic modeling differs greatly from the statistical modeling described in chapter 10. Mechanistic models provide more detailed analyses and results than do empirical models. Theoretical relationships and equations that intimately describe the underlying physical, chemical, and biological processes within a system compose a mechanistic model. The governing equations rely on initial conditions and boundary limits to generate detailed results. Furthermore, mechanistic models require an excellent understanding of the system and forego the process of developing relationships, as would be the case with empirical models. Rather, these models describe changes likely to occur within the system (responses) and focus on how and why changes occur. Provided sufficient initial data and strong knowledge of the governing behavior within a system are available, a mechanistic model may be constructed to describe the unique responses of the system and generate output data that corresponds to internal stimuli.

The level of detail associated with a mechanistic model varies. Some models invoke a single ordinary differential equation within a highly constrained system whereas other models require a series of stochastic equations to capture the variability within a less-constrained model. Model development must match the nature of the initial data and boundary conditions, where models that are more complex result from large comprehensive data sets, explicit boundary conditions, and full mathematical representation of the intricacies of the system. The output of any model provides insight to the changes within the system, and the accuracy of this output is dependent on initial data and model construction.

Mechanistic models are used frequently in the realm of water quality. The majority of models found within the literature focus on transport of organic nutrients and sediment. Nutrient modeling includes detailed analysis of various transport mechanisms, and, to a lesser extent, transformation (fate) analysis. These analyses involve great knowledge surrounding physical, chemical, and biological processes, where responses of phosphorus and nitrogen to stimuli are better understood than responses of NOM. Input–output models are simple models constructed to describe cause and effect. These models operate through mass balance equations and equilibrium principles. Inherent in their name, input–output models describe fluxes into and out of control volumes. Changes between these fluxes are attributed to activity/reactions within the volume. The models offer great simplicity and serve as a foundation for more complex mechanistic modeling.

Lacking sufficient information to construct a detailed mechanistic model, an input–output model provides the opportunity to describe the behavior of TOC within a stream or river. Kinetic rate constants, absent from TOC literature, are crucial to the fate and transport of organic material. Input–output models applied to data collected by water utilities thus may provide insight into the behavior of TOC within various rivers across the country. This initial effort allows for the construction of mechanistic models with greater accuracy in the future.

The focus of this chapter is the development of a mechanistic model to describe TOC fate and transport in the Metedeconk watershed. The remainder of this chapter is organized as follows. First, the results of select past mechanistic modeling studies are summarized. Next, the



©2007 AwwaRF. All Rights Reserved.

mechanistic modeling methods used for this analysis are presented. Finally, the model results are presented and discussed.

## SELECT RESULTS OF PAST MECHANISTIC MODELING STUDIES

Several studies conducted in Europe and the United States have focused on the responses of phosphorus, nitrogen, and organic carbon to anthropogenic activity within various catchments. The level of detail incorporated in these studies varies. Each provides unique insight with regard to the effects of altered catchments on nutrient transport.

Hope et al. (1997) conducted a large investigation surrounding fluxes of organic carbon within British rivers. This study combined historical organic carbon data collected by the Scottish Environmental Protection Agency (formerly the River Purification Board) and the Environment Agency (formerly the National River Authorities) in England. The aim of the Hope study was twofold: (1) the study sought to describe riverine organic fluxes across the British river system, and (2) it sought to present gaps in organic carbon data from the two collection agencies. The study employed statistical analysis to relate carbon fluxes to several catchment characteristics. Linear relationships were developed among riverine DOC export, mean catchment soil organic carbon storage, annual precipitation, catchment area, and slope within 17 catchments (Hope et al. 1997). The modeling yielded results similar to other studies of organic carbon in British rivers.

In direct contrast to the linear regression employed by Hope et al. (1997) to describe carbon fluxes within British rivers, Hornberger, Bencala, and McKnight (1994) developed a mechanistic model to describe variation of organic carbon in headwater streams and interactions between organic carbon and other dissolved substances in the Snake River catchment in Colorado. Hornberger, Bencala, and McKnight (1994) obtained historical data previously collected by a local agency. The mechanistic model incorporated a simple chemical mixing model with the hydrological model called TOPMODEL. TOPMODEL operates through topographic indices that govern hydrological flow mechanisms, whereas the mixing model was based on a chemical mass balance within the soil structure. This model was used to simulate TOC concentrations during spring snowmelt conditions, when TOC occurrence is highest in the Colorado mountain region. The model was able to capture general trends in discharge and stream TOC levels, and provided the framework for a more refined model to better explain the behavior of organic carbon in this particular system. Future modeling efforts should aim to test the underlying assumptions of the model, refine kinetic processes, and improve the model through additional data (Hornberger, Bencala, and McKnight 1994; Boyer et al. 1996).

Mechanistic modeling is not limited to riparian systems; modeling efforts span lotic and marine systems as well. Marine environments do not often serve as water supply sources and are not included in the following discussion. Lotic systems, however, frequently serve as water supply sources. Mechanistic modeling is extensive within lakes and reservoirs and may include predation, nutrient/food-chain, nutrient loading, transformation, and microbial growth processes. A series of models was developed for the Cannonsville Reservoir* in New York. Several concurrent investigations developed phytoplankton growth and nutrient-phytoplankton models as well as models regarding fate and transport of DBP precursors, *Giardia* cysts, and *Cryptosporidium* oocysts (Stepczuk et al. 1998a, b, c; Auer and Forrer 1998a, b; Doerr et al. 1998). Chapra and

* The Cannonsville Reservoir is a water supply reservoir for New York City, located within the Delaware River watershed in southern New York state.



©2007 AwwaRF. All Rights Reserved.

Reckhow (1983a, b) present two mechanistic models: the first, a nutrient/food-chain model, and the second, a toxic substances model that involves fate, transport, and bioaccumulation. For brevity, the previously mentioned mechanistic models represent a fraction of the many models available for use within the realm of water quality analysis. The main focus of this study involves riparian systems and though several lakes and reservoirs serve as primary sources, greater emphasis is placed on riverine sources. As a result, a mechanistic model must be developed such that a unique model or its underlying structure may be applied to riparian systems across the United States to describe the fate and transport of organic carbon.

## METHODOLOGY

### Fibonacci Series

Circa 1200, Leonardo da Pisa—also known as Leonardo Fibonacci—sought to explain the reproduction phenomena observed in rabbits. His experiment followed the assumption that each pair of rabbits produces two successive generations (one pair per generation) before dying. The experiment began with one pair of rabbits. The next generation included a second pair of rabbits. The third generation included three pairs of rabbits, despite the loss of the original pair. The sequence—1, 1, 2, 3, 5, 8, 13…—summarizes the experiment and population growth of rabbits as observed by Leonardo da Pisa.

The Fibonacci series is both an additive and geometric series where each progressive number is the sum of the preceding two numbers ($f_{n+2} = f_{n+1} + f_n$) or a power of the previous number ($1 + \phi = \phi^2 \rightarrow 1, \phi, \phi^2, \phi^3, \phi^4, \ldots$) (Huntley 1970). This series not only quantifies rabbit populations but also represents numerous phenomena in nature, such as flower-petal arrangement, plant-bud development, and many similar types of phenomena (Conway and Guy 1996). Unknown to young Leonardo, the eloquent yet simple series held powerful applications that were better understood by other mathematicians.

Johannes Kepler (1571–1630), a gifted German mathematician and astrologist, observed that the ratio of consecutive terms ($f_{n+1}/f_n$) approached a value of 1.6180339… (Herz-Fischler 1987). Kepler was familiar with division in extreme and mean ratio (DEMR) and golden section geometry and recognized the numerical association between the Fibonacci series and DEMR. This value, 1.618, is commonly referred to as the "golden number." The golden number and golden section result from Euclidean geometry.

### Golden Section

Figure 11.1 illustrates golden section geometry, which is described as follows. There exists a point on every line segment that divides the segment such that the ratio of the two section lengths is equal to the ratio of the line segment length and the longer section length. Imagine that a line segment $AB$ of length $L_1$ is divided such that a point $C$, which is located on $AB$, separates the segment into $AC$ and $CB$ with respective lengths $a$ and $b$. If $C$ is positioned so that the ratios of $a{:}b$ and $L_1{:}a$ are equal, then $C$ is the golden section of the original segment $AB$.

A search algorithm developed by Johnson and Kiefer determines the golden section within segments of Euclidean geometry (Johnson 1956 and Kiefer 1953, as cited in Beveridge and Schechter 1970). The golden section search method allows optimization methods to determine the best solution without testing all possible solutions, by iterative numerical calculations using



©2007 AwwaRF. All Rights Reserved.

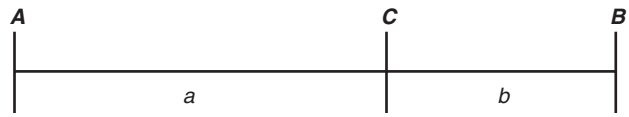



**Figure 11.1 Golden section geometry**

clearly defined logical procedures or algorithms (Reklaitis, Ravindran, and Ragsdell 1983). The search method, not unlike other optimization algorithms, requires well-defined boundary conditions as well as clearly defined performance criteria. Further, it is necessary to select performance evaluation criteria so that the "best" design or set of operating conditions can be identified (Reklaitis, Ravindran, and Ragsdell 1983). This optimization technique operates in an extremely efficient manner, eliminating the need for specific error criteria. As shown in Equations 11.1–11.8, the algorithm removes 38% of the search area with each iteration and quickly converges on the golden section. For the golden section algorithm, a maximum number of iterations opposed to a predetermined error value dictate final convergence. The keystone equations of the golden section search are summarized by Huntley (1970) and expand on the line segment in Figure 11.1. The ratios $L_1$:$a$ and $a$:$b$ are defined in Equations 11.1 and 11.2.

$$L_1 : \mathrm{a} \rightarrow \frac{\mathrm{a} + \mathrm{b}}{\mathrm{a}} \tag{11.1}$$

$$a : \mathrm{b} \rightarrow \frac{\mathrm{a}}{\mathrm{b}} \tag{11.2}$$

where $L_\mathrm{i}$ is the length of the line segment and $a$ and $b$ are as defined in Figure 11.1. The golden number is obtained by setting $b$ equal to a unit length and equating the two ratios, as shown in Equations 11.3 through 11.5.

$$\frac{\mathrm{a} + 1}{\mathrm{a}} = \frac{\mathrm{a}}{1} \tag{11.3}$$

$$a^2 - a - 1 = 0 \tag{11.4}$$

$$\mathrm{a} = \frac{1 + \sqrt{5}}{2} = 1.618 \quad \text{Solving for positive root} \tag{11.5}$$

The importance of the Fibonacci series is exemplified by Equations 11.3 and 11.5 in conjunction with the practice of DEMR. The golden number is obtained by setting $b$ equal to a unit length. However, if $a$ is set as a unit length instead of $b$, Equations 11.3 through 11.5 change, as shown in Equations 11.6 through 11.8.

$$\frac{1 + \mathrm{b}}{1} = \frac{1}{\mathrm{b}} \tag{11.6}$$



©2007 AwwaRF. All Rights Reserved.

$$b^2 + b - 1 = 0 \tag{11.7}$$

$$b = \frac{\sqrt{5} - 1}{2} = 0.618 \tag{11.8}$$

The golden section search algorithm utilizes this substitution and the resulting value for *b*. The search method is described in the following section.

**Fibonacci Algorithm**

The upper boundary $X_{MAX}$ and lower boundary $X_{MIN}$ define the search area for the algorithm. The golden section search requires intermediate values between $X_{MAX}$ and $X_{MIN}$. These two values are governed by the *b* value previously calculated. Intermediate minimum and maximum values may then be calculated, as shown in Equations 11.9 and 11.10.

$$X_1 = X_{MIN} + 0.382*(X_{MAX} - X_{MIN}) \tag{11.9}$$

$$X_2 = X_{MIN} + 0.618*(X_{MAX} - X_{MIN}) \tag{11.10}$$

With the two intermediate minimum and maximum values, $X_1$ and $X_2$, the optimization technique determines respective errors based on Equation 11.11 (Beveridge and Schechter 1970).

$$\delta = |TOC_{measured} - TOC_{predicted}| \tag{11.11}$$

where $\delta$ = error term
$TOC_{measured}$ = actual TOC recorded
$TOC_{predicted}$ = TOC predicted through the model

The respective errors for $X_1$ and $X_2$ are compared to one another. The larger of the two errors is removed and a single boundary condition is adjusted to the intermediate value. A new intermediate value is calculated and the algorithm proceeds. Each iteration of the golden section search removes 38% of the analysis region. Following the last iteration, the boundary limits, $X_{MAX}$ and $X_{MIN}$, are averaged to obtain the golden section, as shown in Equation 11.12.

$$X = \frac{X_{MAX} - X_{MIN}}{2} \tag{11.12}$$

The algorithm mechanics are further explained in Figure 11.2. The line segment is discarded for some line of the *n*th degree. The lower and upper boundary conditions are $X_{MIN}$ and $X_{MAX}$, respectively, and the intermediate values are $X_1$ and $X_2$. After completing one iteration, the algorithm determines that the error associated with $X_1$ is larger than that of $X_2$. Therefore, the lower 38% of the region (shaded area) is discarded. The new minimum boundary condition becomes $X_1$, and $X_1$ is recalculated for the new, smaller region. The algorithm then proceeds to the next iteration, as the upper boundary condition and $X_2$ remain unchanged.



©2007 AwwaRF. All Rights Reserved.

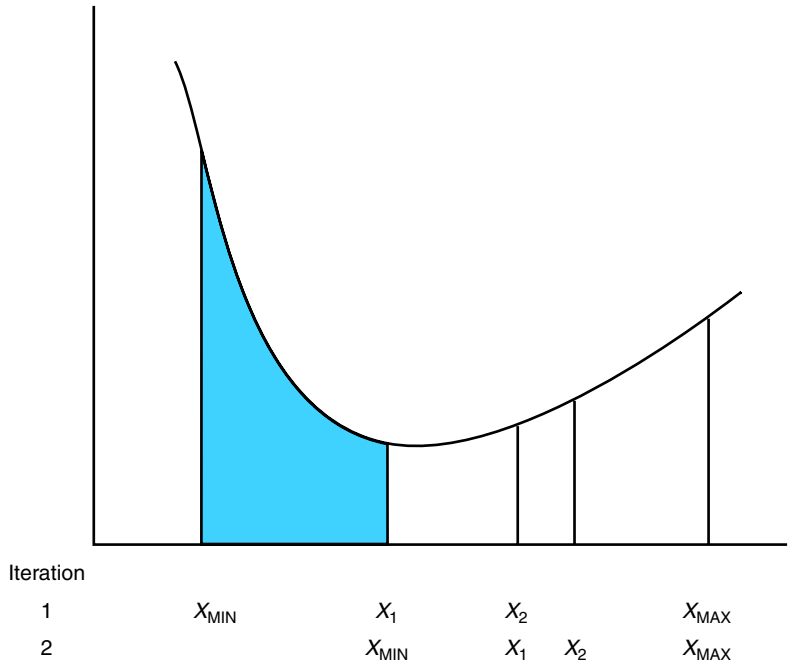



**Figure 11.2 Procedure of Fibonacci optimization technique**

## MODEL DEVELOPMENT

The Fibonacci optimization technique provides an efficient and thorough search routine. Moreover, the method allows multiple parameters to be simultaneously optimized with equal accuracy. This study made use of the golden section search to determine the fate of organic matter as the material traversed a riparian system. Uncertainties surrounding the fate of TOC during transport lead to construction of a simple input–output model. The kinetic reaction constants associated with degradation, sedimentation, and resuspension govern the behavior of the model. However, previous investigations of TOC provide little guidance as to the chemical and physical interactions of TOC within an aquatic system. As a result, a single kinetic reaction constant represented the various transformations associated with nutrient transport, and these reaction constants became the focal point of the optimization search.

Construction of the input–output model followed a single governing equation (Chapra 1997). The governing equation stems from a flow and mass balance of the river system. The general flow balance of the system is presented in Equation 11.13.

$$Q_{total} = Q_0 + Q_1 + \ldots + Q_n \tag{11.13}$$

where $Q_{total}$ = total discharge within a river
$Q_0$ = initial river discharge
$Q_1$ = discharge of the first tributary
$Q_n$ = discharge of the $n$th tributary



©2007 AwwaRF. All Rights Reserved.

The corresponding mass balance is presented in Equation 11.14.

$$0 = Q_0c_0 + Q_1c_1 + \ldots + Q_nc_n - Q_{total}c_{total} \tag{11.14}$$

where $Q_0c_0$ = first river loading
$Q_1c_1$ = second river loading
$Q_nc_n$ = $n$th river loading
$Q_{total}c_{total}$ = resulting river load

The downstream or outlet organic carbon concentration within the stream network is the dependent variable of the governing equation. The independent variables influence the dependent variable and include discharge, tributary loading (concentration multiplied by discharge), stream velocity, and channel length. The data collected by the water utility provided paired TOC concentration information along each branch of the Metedeconk River as well as at the treatment plant intake. USGS gauge stations located just upstream of the confluence provided discharge data for both the North and South branches. The TOC data for upstream stations were paired with spatially adjusted discharge values, derived from the USGS data, to represent tributary loadings to both branches.

The governing equation, based on the mass balance equation (Equation 11.14) for a two-branch river system, is shown in Equation 11.15 and predicts the outlet concentration of TOC based on the data collected within the watershed shown in Figure 11.3.

$$c_{final} = \left[\frac{\left(Q_1c_1e^{-\frac{k_1*x_1}{U_1}}\right) + \left(Q_2c_2e^{-\frac{k_2*x_2}{U_2}}\right)}{Q_1 + Q_2}\right]*e^{-\frac{k_{final}*x_3}{U_3}} \tag{11.15}$$

where $c_{final}$ = predicted TOC concentration at the intake location
$Q_1c_1$ = river loading for the first branch
$Q_2c_2$ = river loading for the second branch
$U$ = river velocity across each segment
$x$ = channel length across each segment
$Q$ = contributing discharge
$k_1$, $k_2$, $k_3$ = kinetic rate constants determined by the optimization algorithm

This governing equation may be altered to fit different riverine systems via alterations to number of stream branches, channel length, tributary loads, and stream velocity.

Equation 11.15 is a result of a combination of several similar equations because the North and South branch governing equations included respective tributary loadings and travel distances. The final TOC loads within each branch are represented by $Q_1c_1$ and $Q_2c_2$ for the North and South branches, respectively. The kinetic rate constants ($k_1$ and $k_2$) within Equation 11.15 are representative of the TOC losses or gains across the final segment of either branch. However, for



©2007 AwwaRF. All Rights Reserved.

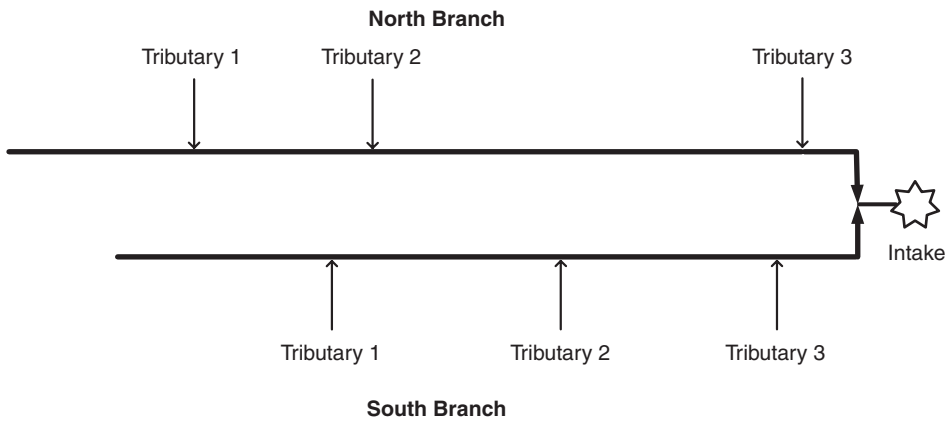



**Figure 11.3 Modeling simplification of Metedeconk watershed (to scale)**

simplicity, the model structure was such that these rate constants were assumed to be identical across the entire branch length; the losses or gains associated within an upstream segment are equivalent to losses and gains within a downstream segment. This simplification was responsible for some of the overall model error and is discussed further in a following section.

## Fibonacci Optimization

This study made use of the Fibonacci optimization technique to determine the appropriate values for the kinetic reaction constants within the governing equation. The results of the golden section search provide initial estimates for the behavior of TOC within the system. Furthermore, a compilation of kinetic constants from various water supply catchments affords investigators several degrees of freedom with which to draw tentative conclusions as to the fate of TOC during transport within riparian systems. The Fibonacci search sequence operates via a finite number of iterations and refines estimates of the target variables (kinetic constants) with each iteration. A single value for each parameter results following satisfactory convergence of the optimization routine.

The Fibonacci optimization sequence was conducted using VBA with Microsoft Excel. The code was developed to handle two-, three-, and four-parameter searches. Each search module (representative of an unknown parameter) operates across 15 iterations while a global error function forces the search routine to converge on an appropriate solution, as shown in Equation 11.11.

The optimization routine successfully functions without bias toward the location of the parameter module within the optimization algorithm. The results of the inner search routine (first parameter) yield answers of equivalent accuracy to that of a third (third parameter) or fourth (fourth parameter) module search routine. The governing equation for TOC within a riverine system may be altered to satisfy the physicality of the system within the realm of simple input–output models. Thus, the number of modules within the Fibonacci search sequence changes.

Literature pertaining to the fate and transport of nutrients indicates that most kinetic reaction constants fall between positive and negative 1. Therefore, the upper and lower boundary conditions for the search algorithm were set at 1 and –1, respectively. These boundary values allowed sufficient freedom for the algorithm to properly optimize the reaction constants and generate appropriate kinetic values. Without available literature that pertains to organic carbon kinetic reaction constants within riparian systems, the boundary conditions were quite broad. The



©2007 AwwaRF. All Rights Reserved.

**Table 11.1**
**Summary statistics for TOC data within the Metedeconk system**

| | Period of record | | Winter | Spring | Summer | Fall |
|---|---|---|---|---|---|---|
| Location | Median (mg/L) | SD* | Median (mg/L) | | | |
| Main branch intake | 5.2 | 3.1 | 4.4 | 5.8 | 5.5 | 4.6 |
| North Branch | | | | | | |
| NO | 9.9 | 6.7 | 7.1 | 11.15 | 15.9 | 9.6 |
| NL | 8.7 | 5.3 | 9.4 | 9.9 | 7.4 | 5.3 |
| NI | 6.3 | 3.3 | 5.9 | 7.55 | 5.95 | 5.55 |
| NB | 5.5 | 2.8 | 6.2 | 6.4 | 4.5 | 3.5 |
| South Branch | | | | | | |
| SL | 7.0 | 4.0 | 7.45 | 7.45 | 7.3 | 5.25 |
| SJ | 6.4 | 4.3 | 5.5 | 7.5 | 7.7 | 5.7 |
| SG | 5.0 | 2.5 | 4.5 | 6.2 | 5.0 | 4.45 |
| SB | 4.7 | 2.3 | 4.63 | 4.8 | 5.1 | 3.4 |

*SD = standard deviation.

remaining variables, excluding TOC, are system (physical) dependent and remain within the sphere of user influence. Furthermore, data collected by water utilities represent all variables within the governing equation except the kinetic reaction constants. Again, proper representation of the physical system is necessary in order to generate meaningful results from the input–output model and the golden section search method.

## Data Requirements

A detailed data set is required to determine kinetic rate constants from an input–output model; an optimization technique is also required. The data available for the input–output model was collected within the Metedeconk River watershed. The Metedeconk River is comprised of two branches, the North and the South, that converge upstream of the local water utility. The data collected from the North and South branches covered a period of 4 years (1999–2003), and TOC samples were collected approximately twice a month. Sampling locations along each branch provided several TOC measurements that capture tributary inputs and the influence of surrounding LULC characteristics via nonpoint loads. The variability associated with the TOC data for both the North and South branches is presented in Table 11.1.

## Results

The data collected within the Metedeconk watershed was divided into two subsets, water years 1999 and 2000 and water years 2001 and 2002, and these data were further divided into monthly median values. This second division lumped together the staggered biweekly samplings at



©2007 AwwaRF. All Rights Reserved.

each station and created uniform sampling frequencies within the watershed. The first subset was used for calibration of the model, and the second subset was used for validation of the model. The calibration data were input to the model, and the Fibonacci optimization algorithm determined optimal kinetic rate constants for each segment of the Metedeconk River system. The resulting rate constants for the North Branch, South Branch and main branch are 0.0729, $-1.6\times10^{-5}$, and 0.0707 $sec^{-1}$, respectively. The rate constants* of the North Branch and main branch indicate loss with travel time ($x/U$). Despite many wetland regions within the riparian zone, TOC losses were determined to most accurately fit the North Branch. Likewise, the optimization algorithm determined a similar rate constant along the main branch. The South Branch rate constant differed from the other two rate constants. The negative kinetic rate constant represents TOC gains across the length of the South Branch. The South Branch Metedeconk (Figure 11.3) passes through several lakes and wetlands areas in the lower third of the channel. These carbon sources—autochthonous and allochthonous, respectively—most likely contribute additional TOC and offset loss/decay.

The accuracy of the kinetic rate constants was determined through a comparison between actual (measured) TOC concentration at the water treatment facility (watershed outlet) and predicted TOC values resulting from the input–output model. The system or total error was an average of the monthly errors across the data set. Monthly errors were calculated by Equation 11.16.

$$\%\ \text{error} = \frac{\text{predicted TOC} - \text{measured TOC}}{\text{measured TOC}} \times 100 \tag{11.16}$$

where predicted TOC = value generated by the model
measured TOC = monthly average of actual TOC concentration collected within the watershed

## Calibration

The Metedeconk River system error was 18% with the optimal rate constants. Figure 11.4 presents a comparison between measured TOC concentration and predicted TOC values across the calibration period. Several erroneous predictions, such as in December 1999 and June 2000, inflate the system error. The level of inflation associated with each prediction was determined by the difference in system error, with and without the prediction present in the error calculation. The December 1999 prediction increased the total error by 2%, whereas the June and July predictions each increased the total by 1%. However, these three predictions in parallel increased the total system error by 5%. Table 11.2 presents monthly prediction errors for the model calibration period (Figure 11.4). Approximately half the predictions exceed the measured concentration of TOC. The average error associated with predictions above the measured TOC concentration was 22%, though the average error of predictions below the measured concentration was 16%. A probable cause of the erroneous predictions was seasonality. Hydrologic variability was not captured through monthly average values, and its affect on TOC likely caused erroneous predictions. The collected biweekly data (actual concentrations) more closely represented TOC fluctuations within the catchment than did the monthly averaged TOC concentration or the predicted TOC values.

As noted, the system error was largely influenced by the December 1999 prediction. A potential explanation for this is the effect of snowpack. The model incorporated TOC concentration via a load value (concentration * discharge). Therefore, variability of discharge significantly

* Positive is assumed to represent constituent loss/decay; negative is indicative of constituent gains.



©2007 AwwaRF. All Rights Reserved.

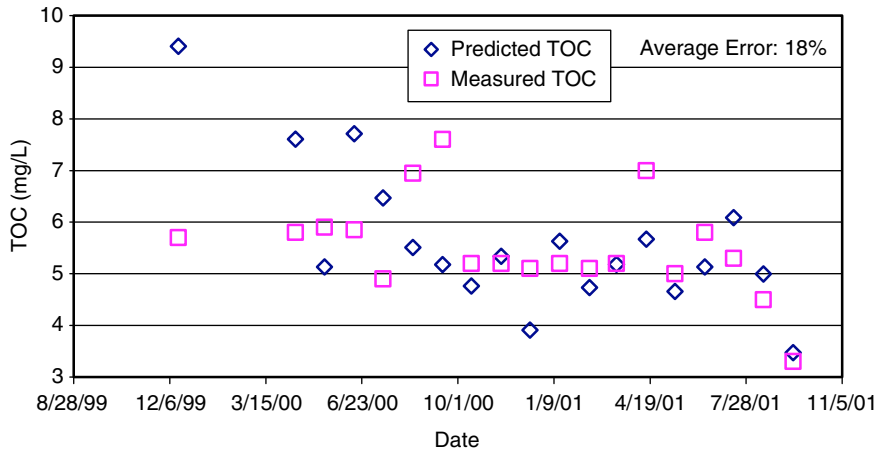



**Figure 11.4 Metedeconk calibration results**

**Table 11.2**
**Comparison between measured TOC concentration and predicted TOC values during the calibration period**

| Date | Measured TOC | Predicted TOC | % Error |
|---|---|---|---|
| 12/15/99 | 5.70 | 9.40 | 65 |
| 4/15/00 | 5.80 | 7.61 | 31 |
| 5/15/00 | 5.90 | 5.13 | 13 |
| 6/15/00 | 5.85 | 7.71 | 32 |
| 7/15/00 | 4.90 | 6.47 | 32 |
| 8/15/00 | 6.95 | 5.51 | 21 |
| 9/15/00 | 7.60 | 5.18 | 32 |
| 10/15/00 | 5.20 | 4.77 | 8 |
| 11/15/00 | 5.20 | 5.34 | 3 |
| 12/15/00 | 5.10 | 3.91 | 23 |
| 1/15/01 | 5.20 | 5.63 | 8 |
| 2/15/01 | 5.10 | 4.74 | 7 |
| 3/15/01 | 5.20 | 5.18 | 0 |
| 4/15/01 | 7.00 | 5.67 | 19 |
| 5/15/01 | 5.00 | 4.66 | 7 |
| 6/15/01 | 5.80 | 5.13 | 11 |
| 7/15/01 | 5.30 | 6.09 | 15 |
| 8/15/01 | 4.50 | 5.00 | 11 |
| 9/15/01 | 3.30 | 3.47 | 5 |
| Average | | | 18 |



©2007 AwwaRF. All Rights Reserved.

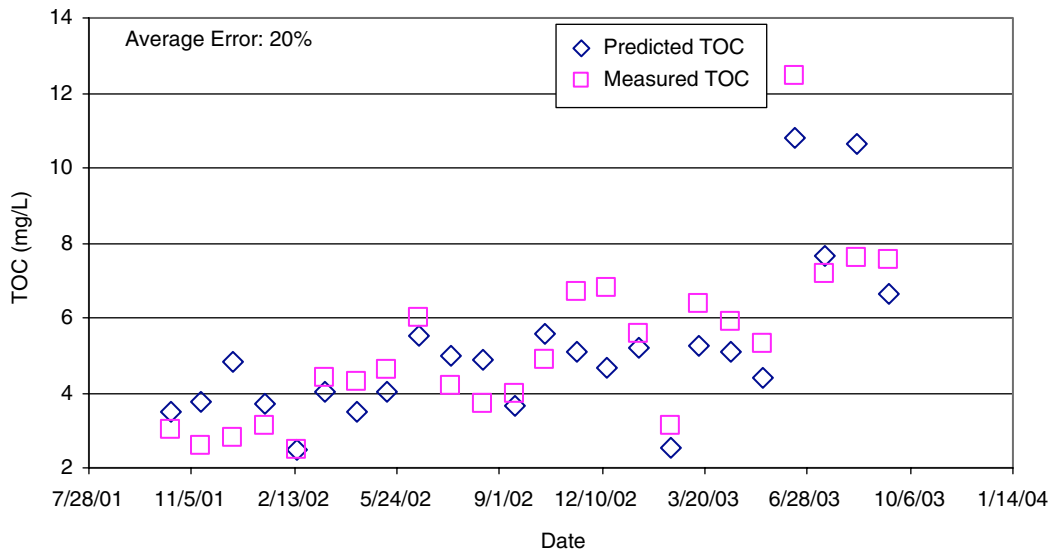



**Figure 11.5 Metedeconk validation results**

affected final TOC predictions. The discharge values in December 1999 were higher than values typically observed during spring months, whereas the amount of TOC available for transport was significantly reduced because of snow cover. The high discharge values generated a high load value, resulting in a falsely increased TOC concentration.

Similarly, low estimates of TOC may have been due to seasonality effects. The input–output model did not capture a "flush" within terrestrial and aquatic ecosystems. In June and July 2000, stream discharge was relatively low, but TOC concentrations were two and three times, respectively, the concentrations measured in December 1999. The resulting TOC load was lower than that in December 1999, and the model yielded an estimate of TOC considerably below the measured concentration.

### *Validation*

The kinetic rate constants generated from the first data set were applied to the second subset of TOC data to validate the model. The validation tested the rate constants for robustness and accuracy. The resulting error increased by 2% to a system error of 20%. Although the change in total error was slight, the Fibonacci optimization algorithm was employed to generate alternative kinetic rate constants. However, this effort to reduce the system error did not result in different kinetic rate constants. In comparing the two sets of constants, the difference between the original and alternative sets did not exceed 1%.

Figure 11.5 compares observed and predicted values based on the original kinetic rate constants applied across the second data set of TOC data. The majority of predictions related well to measured concentrations, but a few of the predictions negatively affected the system error. These differences were again attributed to the variability of collected TOC data due to hydrologic or seasonal differences not captured by the model. Table 11.3 presents monthly prediction errors in tabular format for the period of validation. The results indicate that the model performed reasonably over a range of hydrologic conditions, with the system error remaining under 20% as



©2007 AwwaRF. All Rights Reserved.

**Table 11.3**
**Comparison between measured TOC concentration and predicted TOC values during validation period**

| Date | TOC Measured | TOC Predicted | % Error |
|---|---|---|---|
| 10/15/01 | 3 | 3.49 | 16 |
| 11/15/01 | 2.6 | 3.75 | 44 |
| 12/15/01 | 2.8 | 4.84 | 73 |
| 1/15/02 | 3.1 | 3.70 | 19 |
| 2/15/02 | 2.5 | 2.46 | 2 |
| 3/15/02 | 4.4 | 4.02 | 9 |
| 4/15/02 | 4.3 | 3.48 | 19 |
| 5/15/02 | 4.6 | 4.02 | 13 |
| 6/15/02 | 6 | 5.50 | 8 |
| 7/15/02 | 4.2 | 4.98 | 19 |
| 8/15/02 | 3.7 | 4.87 | 32 |
| 9/15/02 | 4 | 3.65 | 9 |
| 10/15/02 | 4.9 | 5.59 | 14 |
| 11/15/02 | 6.7 | 5.07 | 24 |
| 12/15/02 | 6.8 | 4.67 | 31 |
| 1/15/03 | 5.55 | 5.18 | 7 |
| 2/15/03 | 3.1 | 2.51 | 19 |
| 3/15/03 | 6.4 | 5.25 | 18 |
| 4/15/03 | 5.9 | 5.10 | 14 |
| 5/15/03 | 5.3 | 4.37 | 17 |
| 6/15/03 | 12.45 | 10.80 | 13 |
| 7/15/03 | 7.2 | 7.65 | 6 |
| 8/15/03 | 7.6 | 10.64 | 40 |
| 9/15/03 | 7.55 | 6.64 | 12 |
| Average | | | 20 |

monthly differences between measured and predicted TOC concentrations fluctuated during the period of record. However, there was some seasonal variation in the accuracy of model predictions. Predictions during winter and summer months were consistently high. The structure of the input–output model was such that rapid variation within the watershed was not properly accounted for through the monthly data. However, in general, the resulting rate constants provided a good description of the fate and transport of TOC within the riparian system and indicated the ability of a Fibonacci optimization algorithm to successfully generate solutions for simple riverine input–output models.



©2007 AwwaRF. All Rights Reserved.

## SUMMARY AND CONCLUSIONS

The mechanistic model developed for the Metedeconk River system was a simple input–output model accounting for travel time and first-order reaction kinetics. The governing equations of the model were flow and mass balance (Equations 11.13 and 11.14) and did not explicitly include the effects of seasonality, precipitation, topography, and groundwater flow. Precipitation and groundwater flow indirectly affected the river discharge, included in the model, whereas topography indirectly affected time of transport, also included in the model. During validation, a 20% total error associated with the Metedeconk River system was found. This error was attributed to underlying hydrologic and TOC variability not captured by the model. Loss of variability was due in part to temporal smoothing (monthly averages) and lack of data (only biweekly measurements), resulting in a poor representation of certain characteristics and governing processes within the watershed, such as precipitation and snowmelt. It is evident that a more mechanistic model, which incorporates these parameters, may be necessary to improve model performance.

Each erroneous prediction exemplified the shortcoming of the model structure and indicated the attempt of the search algorithm to fit kinetic parameters to variable data. In general, the rate constant that describes a winter storm or snowmelt period will not properly describe conditions across a dry summer or wet fall. The Fibonacci algorithm minimized the system error (19%) by fitting a kinetic rate constant to the majority of the data as opposed to a single measurement. The three rate constants determined by the optimization routine were adequate and provided the input–output model with moderate accuracy; however, results could be improved by incorporation of additional variability, such as a dependence of reaction rate on temperature.

Ideally, more refined mechanistic models will be developed to describe not only TOC transport but also that of DBPs. A more detailed model would directly incorporate hydrologic impacts, soil moisture, groundwater flow contributions, and a host of other dynamic processes. The input–output model would also benefit from more data. A weekly or biweekly time interval might capture the system variability with more accuracy, provided sufficient data exist. Although more detailed discharge data are available for the Metedeconk River, their incorporation in the model is limited by the availability of paired TOC data. Nonetheless, this input–output model accurately described the general behavior of the Metedeconk River system. The kinetic rate constants resulted from a powerful optimization search algorithm and adequately represented the fate and transport of organic carbon within this system.



©2007 AwwaRF. All Rights Reserved.

# CHAPTER 12
# CONCLUSIONS

The major conclusions of this work can be summarized as follows.

## PROPERTIES OF NOM

THM precursor content of natural waters can vary sharply with changes in source type, location, and time of year. Although much of this variability can be attributed to differences in TOC, some cannot. By calculating the specific THMFP, one can explore the nature of the variability in this intensive property of NOM.

The formation of DBPs is heavily dependent on the characteristics of the precursor NOM. Although, SUVA has been successfully used as a measure of NOM reactivity with oxidants and disinfectants, it falls short of capturing the full complexity of these reactions. This study documents the range of variability in DBP yields from a broad spectrum of waters and NOM types. The major reason for higher THM levels from surface waters than groundwaters is due to higher TOC levels. On a per-carbon basis, surface waters produce only about one-third more THM than groundwaters. The median THM precursor level in raw waters (based on the UMass protocol) is 48 µg/mg C, and the bulk of the population falls within 30–80 µg/mg C (10th–90th percentiles). The average THM precursor level in hydrophobic acids closely mirrors the raw water database. The median THM precursor level (based on the UMass protocol) is 49 µg/mg C, and the bulk of the population falls within 20–80 µg/mg C (10th–90th percentiles). The median DHAA precursor level in hydrophobic acids is 21 µg/mg C, and the bulk of the population falls within 7–37 µg/mg C (10th–90th percentiles). The median THAA precursor level in hydrophobic acids is 24 µg/mg C, and the bulk of the population falls within 9–43 µg/mg C (10th–90th percentiles). Similar assessments have been made for precursor content of hydrophilic and mesophilic NOM.

Highly colored (i.e., high SUVA) organic matter has a much higher tendency to form THAAs. Hydrophobic and hydrophilic bases produce a disproportional amount of DHANs. Hydrophilic acids produce a higher level of DHAAs as compared to the other DBPs.

Both ozonation and chlorination resulted in increases in AOC and BDOC of all waters and NOM fractions. Ozonation generally resulted in the highest levels, but in several cases, chlorination caused increases that were greater than those observed in the ozonated samples.

The fractionation method used in this study allows recovery of most of the carbon and DBP precursors. However, freeze-drying seemed to enhance the production of DBPs, suggesting a change in chemical structure of NOM. Decarboxylation (loss of $CO_2$) and dehydration (loss of $H_2O$) probably occurs during freeze-drying, and these are likely to increase reactivity with an electrophilic substance such a hypochlorous acid.

Lignin was an important—though not major—contributor of DBP precursor structures to aquatic NOM. Using a standard aquatic fulvic acid, the contribution from lignin sources was estimated at 9% for THM precursors, 12% for DHAA precursors, and 25% for THAA precursors. Published abundance figures for natural tannins combined with estimated precursor yields suggest that these compounds may be the single most important contributors to the regulated DBPs in allochthonous material.



©2007 AwwaRF. All Rights Reserved.

## WATERSHED PROCESSES

Leaf litter will rapidly release DOM that is only moderately rich in DBP precursors. This NOM will immediately start to undergo biodegradation, which occurs on the order of days to weeks. As much as 55% of the fresh leached material is biodegradable, but older leaves may release NOM that is mostly recalcitrant (30% biodegradable). After degradation, the specific DBP precursor content of the NOM actually increases. This is thought to be a result of preferential loss of the most biodegradable fractions and the fractions that are also less reactive with chlorine.

There are indications that specific THM formation decreases as the TOC of surface water increases. Freshly released allochthonous NOM is low in precursor content. This is likely a result of the presence of polysaccharides and other biodegradable substances that have little reactivity with chlorine. Aging of this material leads to losses of DOC but increases in specific DBP precursor content. It is postulated that this is the end result of recalcitrant residues with high chlorine reactivity (e.g., lignin-type structures).

Comparison of the North American database with samples from a single intensively studied watershed shows remarkable agreement in median values and spread of precursor densities (THM, THAA, DHAA) as well as SUVA. The fact that this study of narrow geographic and climatic scope can produce an array of data that nearly matches the full continental database has many implications. The most significant pertain to the role of local versus broad geographical factors. It seems that time of year, as well as micro- and mesoscale processes (local precipitation, flow paths, riparian zone topography, sub-basin size) play a primary role in determining NOM character. Of lesser importance are the macroscale characteristics that only differ over broad geographic locations (e.g., ecoregion, dominant forest ecosystem, mean elevation, mean annual temperature, mean annual rainfall).

Simple export modeling based on land cover data supported the expectation that wetlands and urban lands result in higher levels of export per unit drainage area.

Long-term trends (1970–2005) tend to show decreases in TOC across the United States; however, increases were more common in a few regions, including the lower Mississippi River, the southern Appalachian area, and the Mid-Atlantic states. Care must be exercised in interpreting these data, as the quality of the earliest measurements may not meet modern standards. Near-term trends (1995–2005) in TOC show subtle increases along the Atlantic and Gulf coasts, and subtle decreases in the Pacific Northwest. Trends across the interior United States were less easily characterized.



©2007 AwwaRF. All Rights Reserved.

# CHAPTER 13
# RECOMMENDATIONS

## PRACTICAL APPLICATIONS AND RECOMMENDATIONS TO UTILITIES

Utilities in North America should plan for the possibility that TOC and DBP precursor levels in raw waters will increase as a result of long-term changes in land use and climate. The degree to which this occurs will depend on geographic location and land use changes in the watershed. Real changes over long periods of time will be slow, and can be easily masked by normal year-to-year variations in rainfall and other climatic factors.

Wetlands and urban areas contribute a disproportionally large amount of NOM and DBP precursors to surface waters. It is expected that expansion of urban area and associated impervious surfaces will result in higher levels of organic matter in raw waters derived from those watersheds. Changes in the amount of wetland area in a watershed can have an even greater impact. The distance (and water travel time) from these land areas to raw water intakes will be an important factor. Simple mechanistic models predict that wetlands and urban areas will have less of a negative impact if they are further from the stream channels or if a buffer zone exists. The authors judge that DBP precursors will be attenuated during additional travel time, although it was not possible to demonstrate this with the data collected in the current study.

NOM in runoff and small, low-order streams can be elevated and highly variable depending on the recent rainfall history. Similarly, the specific precursor content of raw water can rapidly change with varying precipitation.

Chlorination can cause as much general NOM oxidation and formation of biodegradable organic products as ozonation. This should be kept in mind as utilities weigh the relative merits of ozonation versus chlorination for primary disinfection and oxidation. Ozone is frequently linked with biological filtration. This is especially appropriate because ozone does not leave a lasting residual, which could inhibit biological growth in a downstream filter. Certain circumstances exist, however, where chlorine might be used to stimulate biodegradation in a filter (i.e., when practiced prior to filtration at doses low enough to be largely dissipated in the filter bed). The role of chlorine should also be considered in distribution systems prone to chlorine loss or those using postammoniation for maintaining chloramine residuals. The ability of chlorine to produce substantial amounts of AOC and BDOC will exacerbate the problem of biological growths in portions of the system with low residuals.

## RECOMMENDATIONS FOR FUTURE RESEARCH

Newly collected DBP precursor data should be compared to the population summaries based on the large literature databases presented in this document. This will help to place new data into context with prior experience. It will also allow the research community to identify anomalies more quickly, and thereby explore their origin in a timely fashion.

Future watershed studies should focus on microscale processes. It is important to understand and document changes in NOM and precursor flux with higher spatial and temporal resolution. This must include careful analysis of hydrologic and precipitation data.

An effort should be made to collect all historical and modern data on TOC for more extensive investigation of long-term trends. This should include the full USGS and USEPA databases,



©2007 AwwaRF. All Rights Reserved.

as well as the remaining data scattered among numerous local agencies, civic organizations, drinking water utilities, and research projects (published and gray literature). Future work of this type should make use of additional semi-empirical data, such as antecedent rainfall and soil types.

Much unpublished data exists in the databases of the utilities listed in chapter 3. Given their expressed interest in watershed management for precursor control, it is recommended that additional studies be conducted to better develop an understanding of relevant watershed processes. Any future studies on NOM in watersheds should make use of these utility lists and the accompanying database so as to best capture the full range of organic water qualities.

The role of tannins in THM and HAA formation should be explored. These compounds are expected to be major contributors in NOM from forested watersheds. Attempts should be made to identify plant-based sources, ultimately linking dominant flora to DBP export.



©2007 AwwaRF. All Rights Reserved.

# APPENDIX A
# MAIL SURVEY

**Watershed Sources of Organic Matter Utility Survey**

Department of Civil and Environmental Engineering
University of Massachusetts–Amherst
Amherst, MA 01003
(413) 545-5392

To the best of your knowledge, please provide the following information about water supply and quality *at your utility.*

1. The approximate total water production rate at your utility is: _____ mgd.

2. The fraction of water coming from surface supplies at your utility is: _____%.

3. We are interested in the total organic carbon (TOC) and/or the dissolved organic carbon (DOC) levels in the raw water source(s) for your utility. Please indicate either TOC or DOC levels below. If you have multiple raw water sources, please list each source separately.

| SOURCE | TOC (mg/L) | DOC (mg/L) |
|---|---|---|
| (1) _____ | _____ | _____ |
| (2) _____ | _____ | _____ |
| (3) _____ | _____ | _____ |

4. Aside from data reported to USEPA as part of the ICR, have any total organic halide (TOX) measurements been made on your *treated* water?

   1. YES
   2. NO

   Please provide the following information about disinfection by-product (DBP) precursor levels at your utility. We are interested in DBP precursor tests (e.g., formation potential and simulated distribution system tests) and NOT the finished DBP levels.

5. Does your utility have data on disinfection by-product (DBP) precursor levels for any of your *raw waters*?

   1. YES….. (if YES, proceed to question 6)
   2. NO….. (if NO, please skip to question 8)



©2007 AwwaRF. All Rights Reserved.

6. Which DBPs were measured as part of a precursor test?

Trihalomethane (THM) precursors ..... YES NO
Haloacetic acid (HAA) precursors ..... YES NO
Total organic halide (TOX) precursors ..... YES NO
Other (please specify):_____

7. What year were *any* of these DBP data *first* collected on raw water at your utility?

_____YEAR

8. Has *your utility* collected DBP precursor or other natural organic matter (NOM)-related data (e.g., TOC, DOC, UV absorbance) in the surface water samples collected within your watershed(s)?

1. YES
2. NO

9. Have any studies been done by *other groups* on DBP precursors or other NOM-related data in surface water samples collected within your watershed(s)?

1. YES
2. NO

To the best of your knowledge, please answer the following questions about hydrologic studies that may have been conducted in your region.

10. Have hydrologic studies been carried out in your watershed(s)?

1. YES, BY MY UTILITY
2. YES, BY OTHER GROUPS
3. NO

11. Have storm events in your watershed(s) ever been specially monitored *by your utility* in terms of the following?

Water quality ..... YES NO
Flow ..... YES NO
System response ..... YES NO



©2007 AwwaRF. All Rights Reserved.

12. Have storm events in your watershed(s) ever been specially monitored *by other groups* in terms of the following?

| | | |
|---|---|---|
| Water quality | YES | NO |
| Flow | YES | NO |
| System response | YES | NO |

Please indicate the level of interest that your utility may have in projects related to sources of organic matter and DBP precursors.

13. To what extent would the following issues be of interest to you?
1 = "Not at all interested," 2 = "Somewhat interested," 3 = "Moderately interested,"
4 = "Very interested," 5 = "Extremely interested"

| | | | | | |
|---|---|---|---|---|---|
| (a) DBP control | 1 | 2 | 3 | 4 | 5 |
| (b) Watershed management for control of organics | 1 | 2 | 3 | 4 | 5 |
| (c) Watershed management for control of pathogens | 1 | 2 | 3 | 4 | 5 |
| (d) Role of storm events in water quality | 1 | 2 | 3 | 4 | 5 |

14. Would your utility be willing to share existing water quality data with the University of Massachusetts?

1. YES
2. NO

15. Would your utility be willing to collect additional water quality data as part of an AwwaRF-sponsored project?

1. YES
2. NO



©2007 AwwaRF. All Rights Reserved.

©2007 AwwaRF. All Rights Reserved.

# APPENDIX B
# TELEPHONE SURVEY QUESTIONS

Telephone followups were performed by faculty and graduate students or postdoctoral scholars. The intent was to ask for some elaboration on questions asked in the written survey, as well as new topics not covered in the survey. This appendix includes the general guidelines and instructions given to University of Massachusetts (UMass) interviewers for conducting the followup telephone surveys.

## GENERAL INSTRUCTIONS

You'll be given telephone numbers and contact names in order to call a few utilities and perform telephone surveys. The numbers and names may or may not be up to date, so you may have to ask whomever answers the telephone whether there are other individuals you should talk to.

Please make careful notes of your discussions, and include these in a file about the utility. You are to start the file even before you make your first call, using information available at UMass and on the Web (see below). It is important to recognize that utility personnel are very busy and their time is valuable.

### Preliminary Work

It is important to do some preliminary research about the utilities you'll be calling. This will make your call more effective and will minimize the time that a utility person will have to spend on the interview.

1. Look in the compiled utility notebooks. There is some miscellaneous information about the water treatment plants, utilities, water quality, and watersheds. You'll also find copies of the returned surveys here. Look through the notebooks, and make notes or photocopies if there is some documentation you want to have in front of you when you call.
2. Perform a Google search on the water utility. There's no need to spend a lot of time doing this; just try a few phrases to see if there is anything on the Web that would be helpful to the project. Suggestions for search phrases include

   - Drinking water treatment for "city name"
   - Potable water for "city name"
   - Watershed for "city name"
   - Water supply for "city name"

   Please print out useful information, or print out a record of the links if there is a large document on the Web. Make a copy of this material (double sided, if possible) to give to Dave Reckhow for archiving; then keep the original for the utility file you're creating.



©2007 AwwaRF. All Rights Reserved.

3. Use the Web of Science to find literature about the watershed. Again, there's no need to spend a lot of time doing this. Run a quick general search using a characteristic watershed name (e.g., "Chickahominy"), city name (e.g., "water supply and Newport News"), or treatment plant name ("Randall Bold water treatment plant"). You are likely to find several papers, some of which may contain useful information for this project. Please print out a list (including abstracts) of papers that look at least remotely relevant. If some are clearly germane to this project, obtain full copies (from online sources, the UMass library, or interlibrary loan). As before, place all materials in the utility file you're creating. Make a duplicate copy (double sided) for Dave's archives.

**Suggested Telephone Etiquette**

> *"Hi, I'm Jane Doe from the University of Massachusetts, and I'm trying to reach the Water Quality Manager regarding an AwwaRF research project."*

When you connect with someone who can answer your questions, ask them whether they have a few minutes now or whether you should call back at a more convenient time. If this is a convenient time, proceed as follows:

> *"You may recall our mail survey from a few months ago. My call today is intended to follow up on that survey. We are in the process of narrowing down our list of candidate utilities for the project on watershed sources of natural organic matter, and I'm hoping that we can obtain a little more information about your system. I have several questions related to the watershed, water quality, and the water treatment system."*

***Water System***

1. What are the typical DBP levels in your system? THMs? HAAs? BDOM? Other?
2. Have any special DBP studies been conducted?
3. How many water treatment systems do you have? (Obtain names of systems, typical flow [million gallons per day], and general types [conventional, ozone, etc.]).
4. Do you have a brochure on the water treatment system(s) that you could send us (e.g., something you hand out to groups that visit the plant)?

***Water Sources and Watersheds***

1. What are your water sources (verify and expand on data from written survey)?
2. Describe the watershed(s). (Ask first if there are written summaries or studies that we could obtain. There's no need to ask them to describe something that already exists on paper.) Include the following information:
   - Size
   - Lakes or impoundments
   - Residence times
   - Trophic state
   - Presence of wetlands, swamps, marshes
   - Sources of pollution
   - Protection and level of development



©2007 AwwaRF. All Rights Reserved.

3. Is there a general watershed protection plan or other similar document that we might benefit from seeing?
4. Are there known major influences on water quality, or issues that the utility is particularly concerned about?

NOTE: Because many utilities will have multiple surface water sources feeding multiple plants, it is at this point that you may need to consult Dave Reckhow about how to select the raw water source that is to be the focus of future study by UMass. In some cases, it will be easiest to collect data on all sources and decide on the most promising one later; and in other cases, it will be important to focus on a single source first.

***Water Quality Data***

1. Is there a general impression at your utility that the raw water quality is slowly improving/deteriorating with time? If so, do you have evidence for this or is it anecdotal?
2. Do you know why the changes, if any, are occurring?
3. What is the nature of your historical raw water TOC or DOC data?
   - How many measurements per year?
   - When (what year) did you start collecting it?
   - Note the agreement or lack thereof with the survey question on this issue. Follow up to clarify whether there is a substantial discrepancy.
   - How many locations in the watershed, if any?
4. Were DBP formation potentials (or precursor tests) run on your raw water? If so when was this first done?
5. What other water quality information has been collected/studied in the watershed?
6. If there were studies of DBP precursors done by others on your watershed, what was the nature of these studies? Who did them, when were they done, and how might we obtain these data?

***Hydrology Data***

1. What types of hydrologic data are collected for your watershed, at how many locations, at what frequency (hourly, daily, weekly), and by whom (the utility, or are there known U.S. Geological Survey gauges that the utility relies on)?
2. Do you have the locations (latitude/longitude) of the sampling sites?
3. For how long have data been collected?
4. Have more detailed hydrologic studies been carried out for your watershed(s)?
5. Have storm events ever been specially monitored in terms of water quality, flow, or system response?



©2007 AwwaRF. All Rights Reserved.

### *Meteorological Data*

1. What types of meteorological data are collected for your watershed, at how many locations, at what frequency (hourly, daily, weekly), and by whom (the utility, or are there known National Weather Service gauges that the utility relies upon)?
2. Do you have the locations (latitude/longitude) of the sampling sites?
3. For how long have data been collected?

### *Ancillary Data*

1. What types of additional data are available for your watershed(s), either internally or from external sources, particularly state and local agencies?
2. When were these data collected?
3. Are there historical data?
4. Are these data available in hard copy and/or geographic information system format?
5. Are there delineated watershed boundaries and sampling locations? If so, who has this information?
6. What is the scale or level of detail of the data?
7. Is detailed additional information available about city services such as the sewer network (storm drain locations and pipe locations), combined sewer overflows, the road network, high-resolution topography data, etc.?
8. Is there land use and land cover data? Historical data? Is high-resolution data available, and does it contain imperviousness information (particularly for cities)? Do these data exist in geographic information system format?

### *Microbial Monitoring*

1. What microbial parameters do you monitor for? Do you do monitoring within your watershed? At how many locations? What is the frequency of monitoring and the length of record?
2. Are latitude/longitude coordinates available for your sampling locations?
3. Is detailed land use around your sampling locations available?

### *Protozoan Monitoring*

1. Do you monitor for *Cryptosporidium* and/or *Giardia* within your watershed?
2. Do you monitor just at the intake and in finished water?

## ADDITIONAL DOCUMENTATION

The interviewer may wish to use the following table to document type and frequency of current and historic data collection. Add as much additional information as possible about the types of locations sampled (e.g., major tributaries, storage reservoirs, point sources of pollution).



©2007 AwwaRF. All Rights Reserved.

| Parameter | Number of locations | Frequency or year | Length of record | Collecting agency |
|---|---|---|---|---|
| **Plant Water Quality Data** | | | | |
| TOC or DOC | | | | |
| UV absorbance | | | | |
| THMs | | | | |
| HAAs | | | | |
| TOX | | | | |
| Fe | | | | |
| Mn | | | | |
| Residual disinfectant | | | | |
| **Watershed Water Quality Data** | | | | |
| TOC or DOC | | | | |
| UV absorbance | | | | |
| THMFP | | | | |
| HAAFP | | | | |
| TOXFP | | | | |
| Chlorophyll | | | | |
| Algal counts | | | | |
| Phosphorus species | | | | |
| Nitrogen species | | | | |
| TDS or conductivity | | | | |
| Turbidity | | | | |
| **Hydrology Data** | | | | |
| Snow depth | | | | |
| Stage | | | | |
| S-Q rating curves | | | | |
| Velocity data | | | | |
| Time of concentration | | | | |
| Wet weather travel times | | | | |
| Dry weather travel times | | | | |
| Cross-section data | | | | |
| River profile data | | | | |



©2007 AwwaRF. All Rights Reserved.

| Parameter | Number of locations | Frequency or year | Length of record | Collecting agency |
|---|---|---|---|---|
| **Meteorological Data** | | | | |
| Rainfall | | | | |
| Snowfall | | | | |
| Temperature | | | | |
| Dew-point temperature | | | | |
| Wind speed | | | | |
| Wind direction | | | | |
| Cloud cover | | | | |
| Evapotranspiration | | | | |
| Other | | | | |
| **Ancillary Data** | | | | |
| Land use | | | | |
| Geology | | | | |
| Soils | | | | |
| Zoning | | | | |
| Utilities (sewer, septic) | | | | |
| Wetlands | | | | |
| Detailed vegetation (species) | | | | |
| Topography | | | | |
| River reaches | | | | |
| Roads | | | | |
| **Microbial Data** | | | | |
| Fecal coliform | | | | |
| Total coliform | | | | |
| *E. Coli* | | | | |
| Total heterotrophic plate counts | | | | |
| Source-indicators | | | | |
| **Protozoan Data** | | | | |
| *Cryptosporidium* | | | | |
| *Giardia* | | | | |



©2007 AwwaRF. All Rights Reserved.

# REFERENCES


Aiken, G., and E. Cotsaris. 1995. Soil and Hydrology: Their Effect on NOM. *Journal AWWA*, 87(1):3645.

Aiken, G.R., and J.A. Leenheer. 2004. Removal of Dissolved Organic Matter in a Subsurface Recharge System. In *Natural Organic Material Research: Innovations and Applications for Drinking Water.* Edited by G. Newcombe and L. Ho. London: IWA Publishers.

Aitkenhead-Peterson, J.A. 2000. Source, Production and Export of Dissolved Organic Carbon and Nitrogen. Doctoral diss., University of New Hampshire, Durham.

Allgeier, S.C., R.S. Summers, J.G. Jacangelo, V.A. Hatcher, D.M. Moll, S.M. Hooper, J.W. Swertfeger, and R.B. Green. 1996. A Simplified and Rapid Method for Biodegradable Dissolved Organic Carbon Measurement. In *Proc. of the Water Quality Technology Conference*. Denver, Colo.: AWWA.

Amy, G.L., P.A. Chadik, and Z.K. Chowdhury. 1987. Developing Models for Predicting Trihalomethane Formation Potential and Kinetics. *Journal AWWA*, 79(7):89–97.

Amy, G.L., M.R. Collins, C.J. Kuo, and P.H. King. 1987. Comparing Gel Permeation Chromatography and Ultrafiltration for the Molecular Weight Characterization of Aquatic Organic Matter. *Journal AWWA*, 79(1):43–49.

Amy, G.L., R.A. Minear, and W.J. Cooper. 1987. Testing and Validation of a Multiple Nonlinear-Regression Model for Predicting Trihalomethane Formation Potential. *Water Research*, 21:649–659.

Amy, G.L., M. Siddiqui, K. Ozekan, H.W. Zhu, and C. Wang. 1998. *Empirically Based Models for Predicting Chlorination and Ozonation By-products: Trihalomethanes, Haloacetic Acids, Chloral Hydrate, and Bromate*. USEPA Report. Washington, D.C.: U.S. Government Printing Office.

Amy, G.L., R.A. Sierka, J. Bedessem, D. Price, and L. Tan. 1992. Molecular-Size Distributions of Dissolved Organic Matter. *Journal AWWA*, 84:67–75.

Amy, G.L., J.M. Thompson, L. Tan, M.K. Davis, and S.W. Krasner. 1990. Evaluation of THM Precursor Contributions From Agricultural Drains. *Journal AWWA*, 82:57–64.

APHA, AWWA, and WEF (American Public Health Association, American Water Works Association, and Water Environment Federation). 1998. *Standard Methods for the Examination of Water and Wastewater*, 20th ed. Washington, D.C.: APHA.

Arana, I., P. Santorum, A. Muela, and I. Barcina. 2000. Effect of Disinfection Upon Dissolved Organic Carbon (DOC) in Wastewater: Bacterial Bioassays. *Letters in Applied Microbiology*, 31(2):157–162.

Ashbolt, N.J. 2004. Risk Analysis of Drinking Water Microbial Contamination Versus Disinfection By-products (DBPs). *Toxicology*, 198(1–3):255–262.

Auer, M.T., and B.E. Forrer. 1998a. Development and Parameterization of a Kinetic Framework for Modeling Light and Phosphorus Limited Phytoplankton Growth in Cannonsville Reservoir. *Journal of Lake and Reservoir Management*, 14(2-3):290–300.

Auer, M.T., and B.E. Forrer. 1998b. Phosphorus Bioavailability and P-Cycling in Cannonsville Reservoir. *Journal of Lake and Reservoir Management*, 14(2-3):278–289.

Benner, R., J.D. Pakuluski, M. McCarthy, J.I. Hedges, and P.G. Hatcher. 1992. Bulk Chemical Characteristics of Dissolved Organic Matter in the Ocean. *Science*, 255:1561–1564.




©2007 AwwaRF. All Rights Reserved.

Besner, M.-C., V. Gauthier, P. Servais, and A. Camper. 2002. Explaining the Occurrence of Coliforms in Distribution Systems. *Journal AWWA*, 94:95–109.

Beveridge, G.S.G., and R.S. Schechter. 1970. *Optimization: Theory and Practice*. New York: McGraw-Hill.

Bezbarua, B.K. 1997. Modeling Reactions of Ozone With Natural Organic Matter. Ph.D. diss., University of Massachusetts, Amherst.

Bissett, A., J. Bowman, and C. Burke. 2006. Bacterial Diversity in Organically-Enriched Fish Farm Sediments. *FEMS Microbiology Ecology*, 55(1):48–56.

Block, J.C., L. Mathieu, P. Servais, D. Fontvieille, and P. Werner. 1992. Indigenous Bacterial Inocula for Measuring the Biodegradable Dissolved Organic Carbon (BDOC) in Waters. *Water Research*, 26(4):481–486.

Bose, P., D.A. Reckhow, and B.K. Bezbarua. 1994. Effect of Ozonation on Some Physical and Chemical Properties of Aquatic Natural Organic Matter. *Ozone Science and Engineering*, 16:89–112.

Bourdeau, D.H. 2004. *Effect of Beaver (*Castor canadensis*) on Stream Water Quality: A Paired Watershed Study.* M.S. Environmental Engineering Project Report. Department of Civil and Environmental Engineering. Amherst: University of Massachusetts.

Boyce, S.D., and J.F. Hornig. 1983. Reaction Pathways of Trihalomethane Formation From the Halogenation of Dihydroxyaromatic Model Compounds for Humic-Acid. *Environmental Science and Technology*, 17:202–211.

Boyer, E.W., G.M. Hornberger, K.E. Bencala, and D. McKnight. 1996. Overview of a Simple Model Describing Variation of Dissolved Carbon in an Upland Catchment. *Ecological Modelling*, 86:183–188.

Bradford, S.M., C.J. Palmer, and B.H. Olson. 1994. Assimilable Organic Carbon Concentrations in Southern California Surface and Groundwater. *Water Research*, 28(2):427–435.

Camper, A.K. 2004. Involvement of Humic Substances in Regrowth. *International Journal of Food and Microbiology*, 92(3):355–364.

Chadik, P.A., and G.L. Amy. 1983. Removing Trihalomethane Precursors From Various Natural Waters by Metal Coagulants. *Journal AWWA*, 75:532–536.

Chadik, P.A., and G.L. Amy. 1987. Molecular Weight Effects on THM Control by Coagulation and Adsorption. *Journal of Environment Engineering*, 113(6):1234–1248.

Chang, S.D., and P.C. Singer. 1991. The Impact of Ozonation on Particle Stability and the Removal of TOC and THM Precursors. *Journal AWWA*, 83:71–79.

Chapra, S.C. 1997. *Surface Water Quality Modeling*. Boston: McGraw-Hill.

Chapra, S.C., R.P. Canale, and G.L. Amy. 1997. Empirical Models for Disinfection By-products in Lakes and Reservoirs. *Journal of Environmental Engineering*, 123(7):714–715.

Chapra, S.C., and K.H. Reckhow. 1983a. *Engineering Approaches for Lake Management—Volume 1: Data Analysis and Empirical Modeling*. Boston: Butterworths.

Chapra, S.C., and K.H. Reckhow. 1983b. *Engineering Approaches for Lake Management—Volume 2: Mechanistic Modeling*. Boston: Butterworths.

Charnock, C., and O. Kjønno. 2000. Assimilable Organic Carbon and Biodegradable Dissolved Organic Carbon in Norwegian Raw and Drinking Waters. *Water Research*, 34(10):2629–2642.

Charriere, B., F. Gadel, and L. Serve. 1991. Nature and Distribution of Phenolic Compounds in Water and Sediments From Mediterranean Deltaic and Lagunal Environments. *Hydrobiologia*, 222:89–100.



©2007 AwwaRF. All Rights Reserved.

CheckLight Company. 2007. Company Web site [Online]. Available: <http://www.checklight.co.il/>. [cited January 27, 2007]

Cheek, K. 1998. Water Velocities and Dispersion in the Winsor Dam Basin of the Quabbin Reservoir. M.S. thesis, University of Massachusetts, Amherst.

Childress, A.E., E.M. Vrijenhoek, M. Elimelech, T.S. Tanaka, and M.D. Beuhler. 1999. Particulate and THM Precursor Removal With Ferric Chloride. *Journal of Environmental Engineering*, 125:1054–1061.

Christman, R.F., D.L. Norwood, D.S. Millington, J.D. Johnson, and A.A. Stevens. 1983. Identity and Yields of Major Halogenated Products of Aquatic Fulvic-Acid Chlorination. *Environmental Science and Technology*, 17:625–628.

Christman, R.F., J. Shi, D. Wagoner, C. Sharpless, E. Fisher, and J. Schupback. 1998. *A New Method for Characterizing Aquatic Organic Matter*. Raleigh, N.C.: Water Resources Research Institute of the University of North Carolina.

Cipparone, L.A., A.C. Diehl, and G.E. Speitel Jr. 1997. Ozonation and BDOC Removal: Effect on Water Quality. *Journal AWWA*, 89(2):84–97.

Clair, T.A., and J.M. Ehrman. 1996. Variations in Discharge and Dissolved Organic Carbon and Nitrogen Export From Terrestrial Basins With Changes in Climate: A Neural Network Approach. *Limnology and Oceanography*, 41:921–927.

Cleveland, C.C., J.C. Neff, A.R. Townsend, and E. Hood. 2004. Composition, Dynamics, and Fate of Leached Dissolved Organic Matter in Terrestrial Ecosystems: Results From a Decomposition Experiment. *Ecosystems*, 7(3):275–285.

Conway, J.H., and R.K. Guy. 1996. *The Book of Numbers*. New York: Copernicus.

Cooke, G.D., and R.E. Carlson. 1989. *Reservoir Management for Water Quality and THM Precursor Control*. Denver, Colo.: AwwaRF.

Corpe, W.A., and T.E. Jensen. 1996. The Diversity of Bacteria, Eukaryotic Cells and Viruses in an Oligotrophic Lake. *Applied Microbiology and Biotechnology*, 46:622–630.

Cronan, C.S., and G.R. Aiken. 1985. Chemistry and Transport of Soluble Humic Substances in Forested Watersheds of the Adirondack Park, New York. *Geochimica et Cosmochimica Acta*, 49:1697–1705.

Croué, J.P., D. Violleau, and L. Labouyrie. 2000. Disinfection By-product Formation Potentials of Hydrophobic and Hydrophilic Natural Organic Matter Fractions: A Comparison Between a Low- and a High-Humic Water. In *Natural Organic Matter and Disinfection By-Products: Characterization and Control in Drinking Water*. Edited by S.E. Barrett, S.W. Krasner, and G.L. Amy. Washington, D.C.: American Chemical Society.

Currie, W.S., and J.D. Aber. 1997. Modeling Leaching as a Decomposition Process in Humid Montane Forests. *Ecology*, 78:1844–1860.

Da Cunha, L.C., L. Serve, F. Gadel, and J.L. Blazi. 2001. Lignin-Derived Phenolic Compounds in the Particulate Organic Matter of a French Mediterranean River: Seasonal and Spatial Variations. *Organic Geochemistry*, 32:305–320.

De Laat, J. 1981. Contribution à l'Étude du Mécanisme de Formation des Trihalométhanes: Incidence de l'Azote Ammoniacal et des Traitements de Préoxydation. Thesis, Universite de Poitiers, France.

Dilling, J., and K. Kaiser. 2002. Estimation of the Hydrophobic Fraction of Dissolved Organic Matter in Water Samples Using UV Photometry. *Water Research*, 36:5037–5044.

Dillman, D.A. 1978. *Mail and Telephone Surveys*. New York: John Wiley and Sons.



©2007 AwwaRF. All Rights Reserved.

Ditton, R.B., D.K. Loomis, and S. Choi. 1992. Recreation Specialization: Reconceptualization From a Social Worlds Perspective. *Journal of Leisure Research*, 24:33–51.

Doerr, S.M., E.M. Owens, M.G. Gelda, M.T. Auer, and S.W. Effler. 1998. Nutrient-Phytoplankton Model for Cannonsville Reservoir. *Journal of Lake and Reservoir Management*, 14(2–3):301–321.

Edzwald, J.K., W.C. Becker, and K.L. Wattier. 1985. Surrogate Parameters for Monitoring Organic Matter and THM Precursors. *Journal AWWA*, 77(4):122–131.

Edzwald, J.K., B. Gong, and C.A. McGowan. 1981. Trihalomethane Formation and Precursor Removal. In *Proc. of the ASCE Environmental Engineering Division Specialty Conference*. Edited by F.M. Saunders. New York: American Society of Civil Engineers.

Ertel, J.R., J.M. Caine, and E.M. Thurman. 1993. Biomarker Compounds as Source Indicators for Dissolved Fulvic-Acids in a Bog. *Biogeochemistry*, 22:195–212.

Ertel, J.R., and J.I. Hedges. 1984. The Lignin Component of Humic Substances: Distribution Among Soil, and Sedimentary Humic, Fulvic, and Base-Insoluble Fractions. *Geochimica et Cosmochimica Acta*, 48(10):2065–2074.

Ertel, J.R., J.I. Hedges, and E.M. Perdue. 1984. Lignin Signature of Aquatic Humic Substances. *Science*, 223:485–487.

Escobar, I.C., S.K. Hong, and A.A. Randall. 2000. Removal of Assimilable Organic Carbon and Biodegradable Dissolved Organic Carbon by Reverse Osmosis and Nanofiltration Membranes. *Journal of Membrane Science*, 175(1):1–17.

Escobar, I.C., and A.A. Randall. 2001. Assimilable Organic Carbon (AOC) and Biodegradable Dissolved Organic Carbon (BDOC): Complementary Measurements. *Water Research*, 35(18):4444–4454.

Farella, N., M. Lucotte, P. Louchouarn, and M. Roulet. 2001. Deforestation Modifying Terrestrial Organic Transport in the Rio Tapajos, Brazilian Amazon. *Organic Geochemistry,* 32:1443–1458.

Fengel, D., and G. Wegener. 1984. *Wood: Chemistry, Ultrastructure, Reactions*. New York: Walter de Gruyter.

Fields, B.S. 1996. The Molecular Ecology of *Legionellae*. *Trends in Microbiology*, 4:286–290.

Findlay, S., J.M. Quinn, C.W. Hickey, G. Burrell, and M. Downes. 2001. Effects of Land Use and Riparian Flowpath on Delivery of Dissolved Organic Carbon to Streams. *Limnology and Oceanography*, 46(2):345–355.

Fram, M.S., R. Fujii, J.L. Weishaar, B.A. Bergamaschi, and G.R. Aiken. 1999. *How DOC Composition May Explain the Poor Correlation Between Specific Trihalomethane Formation Potential and Specific UV Absorbance*. U.S. Geological Survey Water Resources Investigation Report 99-4018B. Washington, D.C.: U.S. Government Printing Office.

Freeman, C., C.D. Evans, D.T. Monteith, B. Reynolds, and N. Fenner. 2001. Export of Organic Carbon From Peat Soils. *Nature*, 412:785.

Frias, J., F. Ribas, and F. Lucena. 1992. A Method for the Measurement of Biodegradable Organic Carbon in Waters. *Water Research*, 26(2):255–258.

Frias, J., F. Ribas, and F. Lucena. 1994. Critical Study of the Use of *Pseudomonas fluorescens* P17 to Determine Assimilable Organic Carbon (AOC). *Water Research*, 28(6):1463–1469.

Frias, J., F. Ribas, and F. Lucena. 1995. Comparison of Methods for the Measurement of Biodegradable Organic Carbon and Assimilable Organic Carbon in Water. *Water Research*, 29(12):2785–2788.



©2007 AwwaRF. All Rights Reserved.

Fujii, R., A.J. Ranalli, G.R. Aiken, and B.A. Bergamaschi. 1998. *Dissolved Organic Carbon Concentrations and Compositions, and Trihalomethane Formation Potentials in Waters From Agricultural Peat Soils, Sacramento–San Joaquin Delta, California: Implications for Drinking-Water Quality.* Sacramento, Calif.: U.S. Geological Survey.

Galapate, R.P., A. Kitanaka, K. Iro, T. Mukai, E. Shoto, and M. Okada, M. 1997. Origin of Trihalomethane Precursors in Kurose River, Hiroshima, Japan. *Water Science and Technology*, 35(8):15–20.

Garvey, E., J.E. Tobiason, M. Hayes, E. Wolfram, D.A. Reckhow, and J.W. Male. 1998. Coliform Transport in a Pristine Reservoir: Modeling and Field Studies. *Water Science and Technology*, 37(2):137–144.

Garvey, E.A. 2000. Natural Organic Matter Fate and Transport in Quabbin Reservoir. Doctoral diss., Department of Civil and Environmental Engineering, University of Massachusetts, Amherst.

Garvey, E.A., and J.E. Tobiason. 2003a. Assessment of Natural Organic Matter in Quabbin Reservoir. *Journal of Water Supply Research and Technology–Aqua*, 52:19–36.

Garvey, E.A., and J.E. Tobiason. 2003b. Relationships Between Measures of NOM in Quabbin Watershed. *Journal AWWA*, 95(11):73–84.

Garvey, E.A., J.E. Tobiason, D.A. Reckhow, and J.K. Edzwald. 2001. Natural Organic Matter Fate and Transport in Quabbin Reservoir. Unpublished report to Metropolitan District Commission. Amherst: University of Massachusetts.

Gergel, S.E., M.G. Turner, and T.K. Kratz. 1999. Dissolved Organic Carbon as an Indicator of the Scale of Watershed Influence on Lake and Rivers. *Ecological Applications*, 9(4):1377–1390.

Gergel, S.E., M.G. Turner, J.R. Miller, J.M. Melack, and E.H. Stanley. 2002. Landscape Indicators of Human Impacts to Riverine Systems. *Aquatic Sciences*, 64:118–128.

Gessner, M. 1991. Differences in Processing Dynamics of Fresh and Air-Dried Leaf Litter in a Stream Ecosystem. *Freshwater Biology*, 26:387–398.

Gibbs, R.A., J.E. Scutt, and B.T. Croll. 1993. AOC Concentrations and Bacterial Numbers in a Water Distribution System. *Water Science and Technology*, 27(3–4):159–166.

Glaze, W.H., and J.L. Wallace. 1984. Control of Trihalomethane Precursors in Drinking Water—Granular Activated Carbon With and Without Preozonation. *Journal AWWA*, 76:68–75.

Goni, M.A., and J.I. Hedges. 1992. Lignin Dimers—Structures, Distribution, and Potential Geochemical Applications. *Geochimica et Cosmochimica Acta*, 56:4025–4043.

Goni, M.A., and S. Montgomery. 2000. Alkaline CuO Oxidation With a Microwave Digestion System: Lignin Analyses of Geochemical Samples. *Analytical Chemistry*, 72:3116–3121.

Goni, M.A., K.C. Ruttenberg, and T.I. Eglinton. 1998. A Reassessment of the Sources and Importance of Land-Derived Organic Matter in Surface Sediments From the Gulf of Mexico. *Geochimica et Cosmochimica Acta,* 62(18):3055–3075.

Goni, M.A., M.J. Teixeira, and D.W. Perkey. 2003. Sources and Distribution of Organic Matter in a River-Dominated Estuary (Winyah Bay, South Carolina, USA). *Estuarine Coastal and Shelf Science*, 57:1–26.

Gonzalez, C., C. Gutierrez, and T. Grande. 1987. Bacterial Flora in Bottled Uncarbonated Drinking Water. *Canadian Journal of Microbiology*, 33:1120–1125.

Goslan, E.H., D.A. Fearing, J. Banks, D. Wilson, P. Hillis, A.T. Campbell, and S.A. Parsons. 2001. Assessing Seasonal Variations in the DBP Precursor Profile of a Reservoir Water. In *Proc. of the AWWA 2001 Water Quality Technology Conference.* Denver Colo.: AWWA.



©2007 AwwaRF. All Rights Reserved.

Goslan, E.H., S. Voros, J. Banks, D. Wilson, P. Hillis, A.T. Campbell, and S.A. Parsons. 2004. A Model for Predicting Dissolved Organic Carbon Distribution in a Reservoir Water Using Fluorescence Spectroscopy. *Water Research*, 38(3):783–791.

Gruden, C., S. Skerlos, and P. Adriaens. 2004. Flow Cytometry for Microbial Sensing in Environmental Sustainability Applications: Current Status and Future Prospects. *FEMS Microbiology Ecology*, 49(1):37–49.

Haddix, P.L., N.J. Shaw, and M.W. LeChevallier. 2004. Characterization of Bioluminescent Derivatives of Assimilable Organic Carbon Test Bacteria. *Applied and Environment Microbiology*, 70(2):850–854.

Hautala, K., J. Peuravuori, and K. Pihlaja. 1997. Estimation of Origin of Lignin in Humic DOM by CuO-Oxidation. *Chemosphere*, 35(4):809–817.

Hautala, K., J. Peuravuori, and K. Pihlaja. 1998. Organic Compounds Formed by Chemical Degradation of Lake Aquatic Humic Matter. *Environment International*, 24(5–6):527–536.

Heathwaite, A.L, P.J. Johnes, and N.E. Peters. 1996. Trends in Nutrients. *Hydrological Processes*, 10:263–293.

Hedges, J.I., R.A. Blanchette, K. Weliky, and A.H. Devol. 1988. Effects of Fungal Degradation on the CuO Oxidation Products of Lignin: A Controlled Laboratory Study. *Geochimica et Cosmochimica Acta*, 52:2717–2726.

Hedges, J.I., and J.R. Ertel. 1982. Characterization of Lignin by Gas Capillary Chromatography of Cupric Oxide Oxidation Products. *Analytical Chemistry*, 54:174–178.

Helsel, D.R., and R.M. Hirsch. 1992. *Statistical Methods in Water Resources*. New York: Elsevier.

Hem, L.J., and H. Efraimsen. 2001. Assimilable Organic Carbon in Molecular Weight Fractions of Natural Organic Matter. *Water Research*, 35(4):1106–1110.

Hemond, H.F. 1990. Wetlands as the Source of Dissolved Organic Carbon to Surface Waters. In *Organic Acids in Aquatic Ecosystems*. Edited by E.M. Perdue and E.T. Gjessing. New York: John Wiley and Sons.

Herz-Fischler, R. 1987. *A Mathematical History of Division in Extreme and Mean Ratio*. Waterloo, Ont.: Wilfrid Laurier University Press.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. Techniques of Trend Analysis for Monthly Water Quality Data. *Water Resources Research*, 18(1):107–121.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1991. Selection of Methods for the Detection and Estimation of Trends in Water Quality. *Water Resources Research*, 27(5):803–813.

Hoben, H.J., and P. Somasegaran. 1982. Comparisons of the Pour, Spread, and Drop Plate Method for Enumeration of *Rhizobium* spp. in Inoculants Made From Presterilized Peat. *Applied and Environment Microbiology*, 44:1246–1247.

Hodgkins, L. 1999. Measurement and Characterization of Biodegradable Dissolved Organic Carbon in MWRA's Drinking Water Supply and Distribution System. M.S. thesis, University of Massachusetts, Amherst.

Hoefel, D., W.L. Grooby, P.T. Monis, S. Andrews, and C.P. Saint. 2003. Enumeration of Waterborne Bacteria Using Viability Assays and Flow Cytometry: A Comparison to Culture-based Techniques. *Journal of Microbiological Methods*, 55(3):585–597.

Hoehn, R.C., D.B. Barnes, B.C. Thompson, C.W. Randall, T.J. Grizzard, and P.T.B. Shaffer. 1980. Algae as Sources of Trihalomethane Precursors. *Journal AWWA*, 72(6):344–350.

Hongve, D. 1999. Production of Dissolved Organic Carbon in Forested Catchments. *Journal of Hydrology*, 224:91–99.



©2007 AwwaRF. All Rights Reserved.

Hongve, D., G. Riise, and J.F. Kristiansen. 2004. Increased Colour and Organic Acid Concentrations in Norwegian Forests Lakes and Drinking Water—A Result of Increased Precipitation? *Aquatic Sciences*, 66(2):231–238.

Hood, E.W., and M.W. Williams. 2001. Seasonal Changes in the Character and Nitrogen Content of Dissolved Organic Material in an Alpine/Subalpine Headwater Catchment. Optimizing Nitrogen Management in Food and Energy Production and Environmental Protection: Proceedings of the Second International Nitrogen Conference on Science and Policy. *The Scientific World*, 1:556–563.

Hope, D., M.F. Billett, R. Milne, and T.A.W. Brown. 1997. Exports of Organic Carbon in British Rivers. *Hydrological Processes*, 11:325–344.

Hornberger, G.M., K.E. Bencala, and D.M. McKnight. 1994. Hydrological Controls on Dissolved Organic Carbon During Snowmelt in the Snake River Near Montezuma, Colorado. *Biogeochemistry*, 25:147–165.

Houghton, R.A., J.L. Hackler, and K.T. Lawrence. 1999. The U.S. Carbon Budget: Contributions From Land-Use Change. *Science*, 285:574–578.

Houghton, R.A., and J.L. Hackler. 2000. Changes in Terrestrial Carbon Storage in the United States. 1: The Roles of Agriculture and Forestry. *Global Ecology and Biogeography*, 9(2):125–144.

Huck, P.M. 1990. Measurement of Biodegradable Organic Matter and Bacterial Growth Potential in Drinking Water. *Journal AWWA*, 82(7):78–86.

Huck, P.M., P.M. Fedorak, and W.B. Anderson. 1990. Methods for Determining Assimilable Organic Carbon and Some Factors Affecting the Van Der Kooij Method. *Ozone: Science and Engineering*, 12(4):377–392.

Huntley, H.E. 1970. *The Divine Proportion: A Study in Mathematical Beauty.* New York: Dover.

Hyotylainen, J., J. Knuutinen, P. Malkavaara, and J. Siltala. 1998. Pyrolysis-GC-MS and CuO-Oxidation-HPLC in the Characterization of HMMs From Sediments and Surface Waters Downstream of a Pulp Mill. *Chemosphere*, 36(2):297–314.

Hyotylainen, J., J. Knuutinen, and E. Vilen. 1995. Characterization of High Molecular Mass Fractions of Receiving Waters and Sediments of a Pulp Mill by CuO-Oxidation and HPLC. *Chemosphere*, 30(5):891–906.

Jago, P.H. 1993. *Review of Biodegradable Organic Matter Determination Techniques*. Marlow, United Kingdom: Foundation for Water Research.

Jordan, M., and G.E. Likens. 1975. An Organic Carbon Budget for an Oligotrophic Lake in New Hampshire, U.S.A. *Verhandlungen Internationale Vereinigung Limnologie*, 19:994–1003.

Joret, J.C., Y. Levi, and C. Volk. 1991. Biodegradable Dissolved Organic Carbon (BDOC) Content of Drinking Water and Potential Regrowth of Bacteria. *Water Science and Technology*, 24(2):95–101.

Kalbitz, K., D. Schwesig, J. Schmerwitz, K. Kaiser, L. Haumaier, B. Glaser, R. Ellerbrock, and P. Leinweber. 1970. Changes in Properties of Soil-Derived Dissolved Organic Matter Induced by Biodegradation. *Soil Biology and Biochemistry*, 35:1129–1142.

Kaplan, L.A., T.L. Bott, and D.J. Reasoner. 1993. Evaluation and Simplification of the Assimilable Organic Carbon Nutrient Bioassay for Bacterial Growth in Drinking Water. *Applied and Environmental Microbiology*, 59(5):1532–1539.

Kaplan, L.A., and M.W. LeChevallier. 1993. *Assimilable Organic Carbon Measurement Techniques*. Denver, Colo.: AwwaRF.



©2007 AwwaRF. All Rights Reserved.

Karimi, A.A., and P.C. Singer. 1991. Trihalomethane Formation in Open Reservoirs. *Journal AWWA*, 83(3):84–88.

Kaushik, N. 1971. The Fate of the Dead Leaves That Fall Into Streams. *Archiv fur Hydrobiologie*, 68:465–515.

Kemmy, F.A., J.C. Fry, and R.A. Breach. 1989. Development and Operational Implementation of a Modified and Simplified Method for Determination of AOC in Drinking Water. *Water Science and Technology*, 21(3):155–159.

Khan, E., R.W. Babcock, S. Jongskul, F.A. Devadason, and S. Tuprakay. 2003. Determination of Biodegradable Dissolved Organic Carbon Using Entrapped Mixed Microbial Cells. *Water Research*, 37(20):4981–4991.

Khan, E., R.W. Babcock, M.H. Suffet, and M.K. Stenstrom. 1998. Method Development for Measuring Biodegradable Organic Carbon in Reclaimed and Treated Wastewaters. *Water Environment Research*, 70(5):1025–1032.

Kim, J., and D. Reckhow. 2006. Disinfection Byproduct Formation From Selected Model Compounds Representing Natural Plant Product Substructures. Unpublished. Amherst: University of Massachusetts.

Kitis, M., T. Karanfil, and J.E. Kilduff. 2004. The Reactivity of Dissolved Organic Matter for Disinfection By-product Formation. *Turkish Journal of Engineering Environmental Science*, 28:167–179.

Kitis, M., T. Karanfil, A. Wigton, and J.E. Kilduff. 2002. Probing Reactivity of Dissolved Organic Matter for Disinfection By-product Formation Using XAD-8 Resin Adsorption and Ultrafiltration Fractionation. *Water Research*, 36:3834–3848.

Krasner, S.W. 1999. Chemistry of Disinfection By-product Formation. In *Formation and Control of Disinfection By-products in Drinking Water*. Edited by P.C. Singer. Denver, Colo.: AWWA.

Krasner, S.W., J.P. Croué, J. Buffle, and E.M. Perdue. 1996 Three Approaches for Characterizing Natural Organic Matter. *Journal AWWA*, 88(6):66–79.

Kuenzler, E.J. 1979. Organic Carbon Export From Upland and Forested Wetland Watersheds. *Limnology and Oceanography*, 24(5):960–966.

Langmark, J., M.V. Storey, N.J. Ashbolt, and T.A. Stenstrom. 2004. Artificial Groundwater Treatment: Biofilm Activity and Organic Carbon Removal Performance. *Water Research*, 38(3):740–748.

Larson, R.A., and A.L. Rockwell. 1979. Chloroform and Chlorophenol Production by Decarboxylation of Natural Acids During Aqueous Chlorination. *Environmental Science and Technology*, 13(3):325–329.

LeChevallier, M.W., N.E. Shaw, L.A. Kaplan, and T.L. Bott. 1993. Development of a Rapid Assimilable Organic Carbon Method for Water. *Applied and Environment Microbiology*, 59(5):1526–1531.

Leenheer, J.A., and J.P. Croué. 2003. Characterizing Aquatic Dissolved Organic Matter. *Environment and Science Technology*, 37(1):18A–26A.

Leenheer, J.A., and T.I. Noyes. 1984. A Filtration and Column Adsorption System for On-site Concentration and Fractionation of Organic Substances From Large Volumes of Water. U.S. Geological Survey Water-Supply Paper 2230. Washington, D.C.: U.S. Government Printing Office.

Leff, L.G., and J.L. Meyer. 1991. Biological Availability of Dissolved Organic Carbon Along the Ogeechee River. *Limnology and Oceanography*, 36(2):315–323.



©2007 AwwaRF. All Rights Reserved.

Lehtola, M.J., I.T. Miettinen, T. Vartiainen, and P.J. Martikainen. 1999. A New Sensitive Bioassay for Determination of Microbially Available Phosphorus in Water. *Applied and Environmental Microbiology*, 65(5):2032–2034.

Lehtola, M.J., I.T. Miettinen, T. Vartiainen, P. Rantakokko, A. Hirvonen, and P.J. Martikainen. 2003. Impact of UV Disinfection on Microbially Available Phosphorus, Organic Carbon, and Microbial Growth in Drinking Water. *Water Research*, 37(5):1064–1070.

Lettenmaier, D.P. 1988. Multivariate Nonparametric Tests for Trend in Water Quality. *Water Resources Bulletin*, 24(3):505–512.

Lettenmaier, D.P., E.R. Hooper, C. Waginer, and K.B. Faris. 1991. Trends in Stream Quality in the Continental United States, 1978–1987. *Water Resources Research*, 27(3):327–339.

Li, S.Y., X.H. Yang, R.L. Qiu, and P. Wang. 2003. Contents and Leaching of Trihalomethane Precursors in Soils. *Water, Air, and Soil Pollution*, 145:35–52.

Liang, L., and P.C. Singer. 2001. Factors Influencing the Formation and Relative Distribution of Haloacetic Acids and Trihalomethanes Under Controlled Chlorination Conditions. In *Proc. of the AWWA 2001 Water Quality Technology Conference*. Denver, Colo.: AWWA.

Liang, L., and P.C. Singer. 2003. Factors Influencing the Formation and Relative Distribution of Haloacetic Acids and Trihalomethanes in Drinking Water. *Environmental Science and Technology*, 37:2920–2928.

Liao, W., R.F. Christman, D.J. Johnson, D.S. Millington, and J.R. Hass. 1982. Structural Characterization of Aquatic Humic Material, Humic Degradation Products. *Environmental Science and Technology*, 16(7):402.

Liu, W., H. Wu, Z. Wang, S.L. Ong, J.Y. Hu, and W.J. Ng. 2002. Investigation of Assimilable Organic Carbon (AOC) and Bacterial Regrowth in Drinking Water Distribution System. *Water Research*, 36(4):891–898.

Lobbes, J.M., H.P. Fitznar, and G. Kattner. 1999. High-Performance Liquid Chromatography of Lignin-Derived Phenols in Environmental Samples With Diode Array Detection. *Analytical Chemistry*, 71:3008–3012.

Loomis, D.K. 1996. *New Hampshire Anglers' and Hunters' Awareness of Resource Management Agencies, Programs and Funding Sources*. Washington, D.C.: U.S. Fish and Wildlife Service.

Louchouarn, P., S. Opsahl, and R. Benner. 2000. Isolation and Quantification of Dissolved Lignin From Natural Waters Using Solid-Phase Extraction and GC/MS. *Analytical Chemistry*, 72:2780–2797.

MacCarthy, P. 2001. *The Principles of Humic Substances: An Introduction to the First Principle*. Edited by E.A. Ghabbour and G. Davies. Cambridge: The Royal Society of Chemistry.

Magill, A.H., and J.D. Aber. 2000. Dissolved Organic Carbon and Nitrogen Relationships in Forest Litter as Affected by Nitrogen Deposition. *Soil Biology and Biochemistry*, 32:603–613.

Malafronte, J.A. 2003. *Disinfection By-product Models for the Stamford, CT, Water System*. M.S. Environmental Engineering Project Report. Department of Civil and Environmental Engineering. Amherst: University of Massachusetts.

McArthur, M.D., and J.S. Richardson. 2002. Microbial Utilization of Dissolved Organic Carbon Leached From Riparian Litterfall. *Canadian Journal of Fisheries and Aquatic Sciences*, 59(10):1668–1672.



©2007 AwwaRF. All Rights Reserved.

McClellan, J.N., D.A. Reckhow, J.E. Tobiason, J.K. Edzwald, and A.F. Hess. 1996. Empirical Models for Chlorination By-products: Four Years of Pilot Experience in Southern Connecticut. In *Water Disinfection and Natural Organic Matter: Characterization and Control.* Edited by R.A. Minear and G.L. Amy. Washington, D.C.: American Chemical Society.

McClellan, J.N., D.A. Reckhow, J.E. Tobiason, J.K. Edzwald, and D.B. Smith. 2000. *A Comprehensive Kinetic Model for Chlorine Decay and Chlorination Byproduct Formation. Natural Organic Matter and Disinfection By-products: Characterization and Control in Drinking Water: Anaheim, CA*. Washington, D.C.: American Chemical Society.

Meyer, J.L. 1990. Production and Utilization of Dissolved Organic Carbon in Riverine Ecosystems. In *Organic Acids in Aquatic Ecosystems*. Edited by E.M. Perdue and E.T. Gjessing. New York: John Wiley and Sons.

Meyer, J.L., and C.M. Tate. 1983. The Effects of Watershed Disturbances of Dissolved Organic Carbon Dynamics of a Stream. *Ecology*, 64(1):33–44.

Miettinen, I.T., T. Vartiainen, and P.J. Martikainen. 1999. Determination of Assimilable Organic Carbon in Humus-Rich Drinking Waters. *Water Research*, 33(10):2277–2282.

Miettinen, I.T., T. Vartiainen, T. Nissinen, T. Tuhkanen, and P.J. Martikainen. 1998. Microbial Growth in Drinking Water Treated With Ozone, Ozone/Hydrogen Peroxide or Chlorine. *Ozone Science and Engineering*, 20:303–315.

Millard, S.P., and N.K. Neerchal. 2001. *Environmental Statistics With S-Plus.* Boca Raton, Fla.: CRC Press.

Miller, R.E., S.J. Randtke, L.R. Hathaway, and J.E. Denne. 1990. Organic Carbon and THM Formation Potential in Kansas Groundwaters. *Journal AWWA*, 82(3):49–62.

Mogren, E.M., P. Scarpino, and R.S. Summers. 1990. Measurement of Biodegradable Dissolved Organic Carbon in Drinking Water. In *Proc. of the 1990 Annual Conference*. Denver, Colo.: AWWA.

Mulholland, P.J. 1981. Organic Carbon Flow in a Swamp–Stream Ecosystem. *Ecological Monographs*, 51(3):307–322.

Mulholland, P.J., and E.J. Kuenzler. 1979. Organic Carbon Export From Upland and Forested Wetland Watersheds. *Limnology and Oceanography*, 24(5):960–966.

Neff, J.C., and G.P. Asner. 2001. Dissolved Organic Carbon in Terrestrial Ecosystems: Synthesis and a Model. *Ecosystems*, 4(1):29–48.

Norton, C.D., and M.W. LeChevallier. 2000. A Pilot Study of Bacteriological Population Changes Through Potable Water Treatment and Distribution. *Applied and Environment Microbiology*, 66:268–276.

Nwachcuku, N., and C.P. Gerba. 2004. Emerging Waterborne Pathogens: Can We Kill Them All? *Current Opinion in Biotechnology*, 15(3):175–180.

Nykvist, N. 1963. Leaching and Decomposition of Water-Soluble Organic Substances From Different Types of Leaf and Needle Litter. *Studia Forestalia Suecica*, 3:3–39.

Oliver, B.G., and J. Lawrence. 1979. Haloforms in Drinking Water: A Study of Precursors and Precursor Removal. *Journal AWWA*, 71:161–163.

Omernik, J.M., and R.G. Bailey. 1997. Distinguishing Between Watersheds and Ecoregions. *Journal AWWA*, 33(5):935–949.

Owen, D.M., G.L. Amy, and Z.K. Chowdhury. 1993. *Characterization of Natural Organic Matter and Its Relationship to Treatability.* Denver, Colo.: AwwaRF.

Owen, D.M., G.L. Amy, Z.K. Chowdhury, R. Paode, G. McCoy, and K. Viscosil. 1995. NOM Characterization and Treatability. *Journal AWWA*, 87(1):46–63.



©2007 AwwaRF. All Rights Reserved.

Page, D.W., J.A. Van Leeuwen, K.M. Spark, M. Drikas, N. Withers, and D.E. Mulcahy. 2002. Effect of Alum Treatment on the Trihalomethane Formation and Bacterial Regrowth Potential of Natural and Synthetic Waters. *Water Research*, 36:4884–4892.

Park, S.K., K.R. Pak, S.C. Choi, and Y.K. Kim. 2004. Evaluation of Bioassays for Analyzing Biodegradable Dissolved Organic Carbon in Drinking Water. *Journal of Environment Science and Health, Part A—Toxic/Hazardous Substances and Environmental Engineering*, 39(1):103–113.

Petersen, R., and K. Cummins. 1974. Leaf Processing in a Woodland Stream. *Freshwater Biology*, 4:343–368.

Plummer, J.D., and J.K. Edzwald. 2001. Effect of Ozone on Algae as Precursors for Trihalomethane and Haloacetic Acid Production. *Environmental Science and Technology*, 35:3661–3668.

Polanska, M., K. Huysman, and C. van Keer. 2005. Investigation of Assimilable Organic Carbon (AOC) in Flemish Drinking Water. *Water Research*, 39:2259–2266.

Randtke, S.J., F. deNoyelles Jr., C.E. Burkhead, R.E. Miller, E. Rachel, J.E. Denne, L.R. Hathaway, and A.S. Melia. 1988. Trihalomethane Precursors in Kansas Water Supplies: Occurrence, Source Control Measures, and Impacts on Drinking Water Treatment. In *Proc. of the 38th Annual Environmental Engineering Conference.* Lawrence: University of Kansas.

Randtke, S.J., F. deNoyelles Jr., J.E. Denne, L.R. Hathaway, R.E. Miller, A.S. Melia, and C.E. Burkhead. 1987. Source Control of THM Precursors. In *Proc. of the 1987 Annual Conference: Our Water Makes the Difference*. Denver, Colo.: AWWA.

Rathbun, R.E. 1996. Regression Equations for Disinfection By-products for the Mississippi, Ohio and Missouri Rivers. *Science of the Total Environment*, 191:235–244.

Reckhow, D.A., G. Hua, J. Kim, P.G. Hatcher, S.A.L. Caccamise, and R. Sachdeva. 2007. *Characterization of TOX Produced During Disinfection Processes.* Final Report. Denver, Colo.: AwwaRF. Forthcoming.

Reckhow, D.A., and G. Makdissy. 2007. Formation of Disinfection By-products From Isolated Aquatic Organic Matter. Unpublished. Amherst: University of Massachusetts.

Reckhow, D.A., J.N. McClellan, B. Mousset, and B.K. Bezbarua. 1999. A Comprehensive Look at DBP Precursors and Their Chemical Properties. American Chemical Society, *Preprint of Extended Abstracts*, 39(1):217–218.

Reckhow, D.A., P.L. Rees, and D. Bryan. 2004. Watershed Sources of Disinfection Byproduct Precursors. *Water Supply,* 4(4):61–69.

Reckhow, D.A., and P.C. Singer. 1985. Mechanisms of Organic Halide Formation During Fulvic Acid Chlorination and Implications With Respect to Preozonation. In *Water Chlorination: Environmental Impact and Health Effects.* Edited by R.L. Jolley, R.J. Bull, W.P. David, S. Katz, M.H. Roberts Jr., and V.A. Jacobs. Chelsea, Mich.: Lewis Publishers.

Reckhow, D.A., and P.C. Singer. 1990. Chlorination By-Products in Drinking Waters: From Formation Potentials to Finished Water Concentrations. *Journal AWWA*, 82:173–180.

Reckhow, D.A., P.C. Singer, and R.L. Malcolm. 1990. Chlorination of Humic Materials: Byproduct Formation and Chemical Interpretations. *Environmental Science and Technology*, 24:1655–1664.

Reckhow, D.A., J.E. Tobiason, and J.K. Edzwald 1999. *Evaluation of Organic Matter in the MWRA Drinking Water Distribution System*. Final project report to Massachusetts Water Resources Authority. Amherst: University of Massachusetts.



©2007 AwwaRF. All Rights Reserved.

Reklaitis, G.V., A. Ravindran, and K.M. Ragsdell. 1983. *Engineering Optimization: Methods and Applications.* New York: John Wiley and Sons.

Repeta, D.J., T.M. Quan, L.I. Aluwihare, and A.M. Accardi. 2002. Chemical Characterization of High Molecular Weight Dissolved Organic Matter in Fresh and Marine Waters. *Geochimica et Cosmochimica Acta*, 66(6):955–962.

Ribas, F., J. Frias, and F. Lucena 1991. A New Dynamic Method for the Rapid Determination of the BDOC in Drinking Water. *Journal of Applied Bacteriology*, 71:371–378.

Ribas, F., J. Perramon, A. Terradillos, J. Frias, and F. Lucena. 2000. The *Pseudomonas* Group as an Indicator of Potential Regrowth in Water Distribution Systems. *Journal of Applied Microbiology*, 88(4):704–710.

Rice, E.W., P.V. Scarpino, G.S. Logsdon, D.J. Reasoner, P.J. Mason, and J.C. Blannon. 1990. Bioassay Procedure for Predicting Coliform Bacterial Growth in Drinking Water. *Environmental Technology*, 11:821–828.

Richardson, S.D. 2003a. Disinfection By-products and Other Emerging Contaminants in Drinking Water. *Trends in Analytical Chemistry*, 22(10):666–684.

Richardson, S.D. 2003b. Water Analysis: Emerging Contaminants and Current Issues. *Analytical Chemistry*, 75:2831–2857.

Roberts, M. 2003. Impacts of Urbanization on Riparian Leaf Inputs: Initial Leaf Processing Results [Online]. Center for Water and Watershed Studies, Annual Review of Research. Feb. 6, 2003. University of Washington, Seattle. Available: <http://depts.washington.edu/cws/Outreach/Events/AnnualReview/Abstracts/2003abstracts.pdf> [cited January 14, 2007]

Roberts, M.G., P.C. Singer, and A. Obolensky. 2002. Comparing Total HAA and Total THM Concentrations Using ICR Data. *Journal AWWA*, 94:103–114.

Robertson, A. 1988. Decomposition of Mangrove Leaf Litter in Australia. *Journal of Experimental Marine Biology and Ecology*, 116:235–247.

Robinson, C., and M. Gessner. 2000. Nutrient Addition Accelerates Leaf Breakdown in an Alpine Springbrook. *Oecologia*, 122:258–263.

Robinson, C., M. Gessner, and J. Ward. 1998. Leaf Breakdown and Associated Macroinvertebrates in Alpine Glacial Streams. *Freshwater Biology*, 40:215–228.

Robinson, T. 1991. *The Organic Constituents of Higher Plants.* North Amherst, Mass.: Cordus Press.

Rodriguez, M.J., and J.B. Serodes. 2001. Spatial and Temporal Evolution of Trihalomethanes in Three Water Distribution Systems. *Water Research*, 35(6):1572–1586.

Rostad, C.E., J.A. Leenheer, B. Katz, B.S. Martin, and T.I. Noyes. 2000. Characterization and Disinfection By-product Formation Potential of Natural Organic Matter in Surface and Ground Waters From Northern Florida. In *Natural Organic Matter and Disinfection By-products: Characterization and Control in Drinking Water*. Edited by S.E. Barrett, S.W. Krasner, and G.L. Amy. Washington, D.C.: American Chemical Society.

Saka, S. 2001. Chemical Composition and Distribution. In *Wood and Cellulosic Chemistry.* Edited by D.N.S. Hon and N. Shiraishi. New York: Marcel Dekker.

Sakakibara, A., and Y. Sano. 2001. Chemistry of Lignin. In *Wood and Cellulosic Chemistry.* Edited by D.N.S. Hon and N. Shiraishi. New York: Marcel Dekker.

Schnoor, J.L., J.L. Nitzschke, R.D. Lucas, and J.N. Veenstra. 1979. Trihalomethanes Yields as a Function of Precursor Molecular Weight. *Environment and Science Technology*, 13:1134–1138.



©2007 AwwaRF. All Rights Reserved.

Schonter, R., and V. Novotny. 1993. Predicting Attainable Water Quality Using the Ecoregional Approach: Proceedings of the First International Conference on Diffuse Nonpoint Pollution. *Water Science and Technology*, 28(3-5)149–158.

Sen, P.K. 1968. Estimates of the Regression Coefficient Based on Kendall's Tau. *Journal of the American Statistical Association*, 63:1379–1389.

Servais, P., A. Anzil, and C. Ventresque. 1989. Simple Method for Determination of BDOC in Water. *Applied and Environmental Microbiology*, 55(10):2732–2734.

Servais, P., G. Billen, and P. Bouillot. 1994. Biological Colonization of Granular Activated Carbon Filters in Drinking-Water Treatment. *Journal of Environmental Engineering*, 120(4):888–899.

Servais, P., G. Billen, and M.-C. Hascoet. 1987. Determination of the Biodegradable Fraction of Dissolved Organic Matter in Waters. *Water Research*, 21(4):445–450.

Servais, P., P. Laurent, and G. Randon. 1993. Impact of the Biodegradable Dissolved Organic Carbon (BDOC) on Bacterial Dynamics in Distribution Systems. In *Proc. of the AWWA Water Quality Technology Conference*. Denver, Colo.: AWWA.

Shukairy, H.M., R.J. Miltner, and R.S. Summers. 1994. Bromide's Effect on DBP Formation, Speciation, and Control. Part 1, Ozonation. *Journal AWWA*, 86:72–87.

Siddiqui, M., G.L. Amy, K. Ozekin, and P. Westerhoff. 1994. Empirically and Theoretically-based Models for Predicting Brominated Ozonated By-products. *Ozone: Science and Engineering*, 116:157–178.

Singer, P.C. 1994. Control of Disinfection By-products in Drinking Water. *Journal of Environmental Engineering*, 120(4):727–744.

Singh, R., O.C. Stine, D.L. Smith, J.K. Spitznagel Jr., M.E. Labib, and H.N. Williams. 2003. Microbial Diversity of Biofilms in Dental Unit Water Systems. *Applied and Environmental Microbiology*, 69:3412–3420.

Sinsabaugh, R.L., R.C. Hoehn, W.R. Knocke, and A.E. Linkins. 1986. Precursor Size and Organic Halide Formation Rates in Raw and Coagulated Surface Waters. *Journal of Environmental Engineering*, 112(1):139–153.

Smith, R.A., R.B. Alexander, and M.G. Wolman. 1987. Water-Quality Trends in the Nation's Rivers. *Science*, 235:1607–1615.

Smolander, A., and V. Kitunen. 2002. Soil Microbial Activities and Characteristics of Dissolved Organic C and N in Relation to Tree Species. *Soil Biology and Biochemistry*, 34:651–660.

Søndergaard, M., and J.W. Worm. 2001. Measurement of Biodegradable Dissolved Organic Carbon (BDOC) in Lake Water With a Bioreactor. *Water Research*, 35(10):2505–2513.

Speiran, G.K. 2000. Dissolved Organic Carbon and Disinfection By-product Precursors in Waters of the Chickahominy River Basin, Virginia, and Implications for Public Supply. Water-Resources Investigations Report 00-4175. Reston, Va.: U.S. Geological Survey.

Sprott, D.G., S.F. Koval, and C.A. Schnaitman. 1993. Cell Fractionation. In *Methods for General and Molecular Bacteriology*. Edited by P. Gerhardt. Washington, D.C.: American Society of Microbiology.

Steinberg, C., and U. Muenster. 1985. Geochemistry and Ecological Role of Humic Substances in Lakewater. In *Humic Substances in Soil, Sediment and Water: Geochemistry, Isolation and Characterization*. Edited by G.R. Aiken, D.M. McKnight, R.L. Wershaw, and P. MacCarthy. New York: John Wiley and Sons.

Steinberg, C.E.W. 2003. *Ecology of Humic Substances in Freshwaters*. New York: Springer-Verlag.



©2007 AwwaRF. All Rights Reserved.

Stepczuk, C.L., A.B. Martin, S.W. Effler, J.A. Bloomfield, and M.T. Auer. 1998a. Spatial and Temporal Patterns of THM Precursors for a Eutrophic Reservoir. *Journal of Lake and Reservoir Management*, 14(2-3):356–366.

Stepczuk, C.L., A.B. Martin, P. Longabucco, J.A. Bloomfield, and S.W. Effler. 1998b. Allochthonous Contributions of THM Precursors in a Eutrophic Reservoir. *Journal of Lake and Reservoir Management*, 14(2-3):344–355.

Stepczuk, C.L., E.M. Owens, S.W. Effler, J.A. Bloomfield, and M.T. Auer. 1998c. A Modeling Analysis of THM Precursors for a Eutrophic Reservoir. *Journal of Lake and Reservoir Management*, 14(2-3):367–378.

Stout, L.M., and K. Nüsslein. 2005. Shifts in Rhizoplane Communities of Aquatic Plants After Cadmium Exposure. *Applied and Environmental Microbiology*, 71:2484–2492.

Strauss, E.A., and G.A. Lamberti. 2002. Effect of Dissolved Organic Carbon Quality on Microbial Decomposition and Nitrification Rates in Stream Sediments. *Freshwater Biology*, 47:65–74.

Summers, R.S., S.M. Hooper, H.M. Shukairy, G. Solarik, and D. Owen. 1996. Assessing the DBP Yield: Uniform Formation Conditions. *Journal AWWA*, 88:80–93.

Sun, L., E.M. Perdue, J.L. Meyer, and L. Weis. 1997. Use of Elemental Composition to Predict Bioavailability of Dissolved Organic Matter in a Georgia River. *Limnology and Oceanography*, 42(4):714–721.

Sutton, R., and G. Sposito. 2005. Molecular Structure in Soil Humic Substances: The New View. *Environmental Science and Technology*, 39:9009–9015.

Swietlik, J., A. Dabrowska, U. Raczyk-Stanislawiak, and J. Nawrocki. 2004. Reactivity of Natural Organic Matter Fractions With Chlorine Dioxide and Ozone. *Water Research*, 38(3):547–558.

Symons, J.M. 1999. DBPs: A Historical Perspective. In *Formation and Control of Disinfection By-products in Drinking Water.* Edited by P.C. Singer. Denver, Colo.: AWWA.

Symons, J.M., G.E. Speitel Jr., A.C. Diehl, and H.W. Sorenson Jr. 1994. Precursor Control in Waters Containing Bromide. *Journal AWWA*, 86:48–60.

Takiar, S.M. 2001. *Characterization of Natural Organic Matter in the MDC/MWRA Wachusett Water Supply.* M.S. Environmental Engineering Project Report. Department of Civil and Environmental Engineering. Amherst: University of Massachusetts.

Tao, S. 1998. Spatial and Temporal Variation in DOC in the Yichun River, China. *Water Research*, 37(7):2205–2210.

Tao, S., and B. Lin. 2000. Water Soluble Organic Carbon and Its Measurement in Soil and Sediment. *Water Research*, 34:1751–1755.

Tauscher, T.L. 2005. *Microbial Source Tracking as Part of an Integrated Approach to the Assessment of Wachusett Reservoir Watershed Hotspots*. M.S. Environmental Engineering Project Report. Department of Civil and Environmental Engineering. Amherst: University of Massachusetts.

Taylor, J.K. 1987. *Quality Assurance of Chemical Measurements*. Chelsea, Mich.: Lewis Publishers.

Theil, H. 1950. A Rank-Invariant Method of Linear and Polynomial Regression Analysis. *Nederlandse Akademie Wetenchappen Series A* 53:386–392.

Thompson, J.M., L. Tan, G.L. Amy, and M.K. Davis. 1988. Agricultural Drains in the Sacramento River Delta: Trihalomethane Precursor Contributions to California State Project Water. In *Proc. of the 1988 Annual Conference*. Denver, Colo.: AWWA.



©2007 AwwaRF. All Rights Reserved.

Thurman, E.M. 1985. *Developments in Biochemistry: Organic Geochemistry of Natural Waters*. Dordrecht, Netherlands: M. Nijhff, Dr. W. Junk Publishers.

Uhl, W., and G. Schaule. 2004. Establishment of HPC(R2A) for Regrowth Control in Non-Chlorinated Distribution Systems. *International Journal of Food Microbiology*, 92(3):317–325.

USEPA (U.S. Environmental Protection Agency). 1998. National Primary Drinking Water Regulations, Disinfectants and Disinfection By-products, Final Rule. *Federal Register*, 63(241):69390 (Dec. 16, 1998).

USEPA (U.S. Environmental Protection Agency). 2005. Appendices to the Economic Analysis for the Final Stage 2 Disinfectants and Disinfection Byproducts Rule, Volume I (A-E2). Washington, D.C.: USEPA.

USEPA (U.S. Environmental Protection Agency). 2006. Ecoregions of the United States [Online]. Available: <http://www.epa.gov/bioindicators/html/usecoregions.html> [cited September 6, 2006]

USGS (U.S. Geological Survey). 2006. National Water Information System Web site [Online]. Available: <http://waterdata.usgs.gov/nwis> [cited October 15, 2006]

van der Kooij, D. 1990. Assimilable Organic Carbon (AOC) in Drinking Water. In *Drinking Water Microbiology*. Edited by G. McFeters. New York: Springer-Verlag.

van der Kooij, D. 1992. Assimilable Organic Carbon as an Indicator of Bacterial Regrowth. *Journal AWWA*, 84:57–65.

van der Kooij, D., W. Hijnen, and J. Kruithof. 1989. The Effects of Ozonation, Biological Filtration and Distribution on the Concentration of Easily Assimilable Organic Carbon (AOC) in Drinking Water. *Ozone Science and Engineering*, 11:297–311.

van der Kooij, D., A. Visser, and W.A.M. Hijnen.1982. Determining the Concentration of Easily Assimilable Organic Carbon in Drinking Water. *Journal AWWA*, 74(10):540–545.

van der Kooij, H. 1984. Biofilm Formation on Surfaces of Glass and Teflon Exposed to Treated Water. *Applied and Environmental Microbiology*, 47(3):551–559.

van der Mee-Marquet, N., A. Domelier, L. Arnault, D. Bloc, P. Laudat, P. Hartemann, and R. Quentin. 2006. *Legionella anisa*, a Possible Indicator of Water Contamination by *Legionella pneumophila*. *Journal of Clinical Microbiology*, 44:56–59.

Volk, C., C. Renner, C. Robert, and J.C. Joret. 1994. Comparison of Two Techniques for Measuring Biodegradable Dissolved Organic Carbon in Water. *Environmental Technology*, 15(6):545–556.

Volk, C.J., and M.W. LeChevallier. 1999. Impacts of the Reduction of Nutrient Levels on Bacterial Water Quality in Distribution Systems. *Applied and Environmental Microbiology*, 65(11):4957–4966.

Volk, C.J., and M.W. LeChevallier. 2002. Effects of Conventional Treatment on AOC and BDOC Levels. *Journal AWWA*, 94(6):112–123.

Waldron, M.C., and G.C. Bent. 2001. *Factors Affecting Reservoir and Stream-Water Quality in the Cambridge, Massachusetts, Drinking-Water Source Area and Implications for Source-Water Protection*. Water-Resources Investigations Report 00-4262. Northborough, Mass.: U.S. Geological Survey.

Walker, W.W., Jr. 1983. Significance of Eutrophication in Water Supply Reservoirs. *Journal AWWA*, 75(1):38–42.



©2007 AwwaRF. All Rights Reserved.

Weinberg, H.S., S.W. Krasner, S.D. Richardson, and A.D. Thruston Jr. 2002. *The Occurrence of Disinfection By-products (DBPs) of Health Concern in Drinking Water: Results of a Nationwide DBP Occurrence Study.* Athens, Ga.: U.S. Environmental Protection Agency.

Weishaar, J.L., G.R. Aiken, B.A. Bergamaschi, M.S. Fran, R. Fujii, and K. Mapper. 2003. Evaluation of Specific Ultraviolet Absorbance as an Indicator of the Chemical Composition and Reactivity of Dissolved Organic Carbon. *Environmental Science and Technology*, 37(20):4702–4708.

Wetzel, R.G. 2001. Detritus: Organic Carbon Cycling and Ecosystem Metabolism. In *Limnology: Lake and River Ecosystems*, 3rd ed. New York: Academic Press.

Wetzel, R.G. 2003. Dissolved Organic Carbon: Detrital Energetics, Metabolic Regulators, and Drivers of Ecosystem Stability of Aquatic Ecosystems. In *Aquatic Ecosystems: Interactivity of Dissolved Organic Matter*. Boston: Academic Press.

Williams, M.M., J.W.S. Domingo, M.C. Meckes, C.A. Kelty, and H.S. Rochon. 2004. Phylogenetic Diversity of Drinking Water Bacteria in a Distribution System Simulator. *Journal of Applied Microbiology*, 96(5):954–964.

Withers, N., and M. Drikas. 1998. Impact of Water Treatment Processes on Assimilable Organic Carbon. Fifth Germany–Australia Workshop on Water and Environmental Research and Development, May 1998, at University of Karlsruhe, Germany.

Wolfram, E. 1996. Determination of the Decay Rate for Indicator Bacteria Introduced by Sea Gulls to an Oligotrophic Drinking Water Reservoir. M.S. thesis, University of Massachusetts, Amherst.

Wolman, M.G. 1971. The Nation's Rivers. *Science*, 174:905–917.

Woolschlager, J., and B. Rittmann. 1995. Evaluating What Is Measured by BDOC and AOC Tests. *Revue des Sciences de L'Eau*, 8(3):371–385.

Xie, Y.F. 2003. Disinfection Byproducts in Drinking Water: Formation, Analysis and Control. Boca Raton, Fla.: CRC Press.

Yavich, A.A., K.H. Lee, K.C. Chen, L. Pape, and S.J. Masten. 2004. Evaluation of Biodegradability of NOM After Ozonation. *Water Research*, 38(12):2839–2846.

Zavaleta, J.O., F.S. Hauchman, and M.W. Cox.1999. Epidemiology and Toxicology of Disinfection By-products. In *Formation and Control of Disinfection By-products in Drinking Water*. Edited by P.C. Singer. Denver, Colo.: AWWA.



©2007 AwwaRF. All Rights Reserved.

# ABBREVIATIONS

| | |
|---|---|
| AcS | acetosyringone |
| AcV | acetovanillone |
| $(Ac/Al)_v$ | acid-to-aldehyde ratio of vanillyl phenols |
| AMU | atomic mass unit |
| AOC | assimilable organic carbon |
| APHA | American Public Health Association |
| APS | acylheteropolysaccharides |
| ATP | adenosine triphosphate |
| AWWA | American Water Works Association |
| | |
| BAS | biologically active sand |
| BCAA | bromochloroacetic acid |
| BDCAA | bromodichloroacetic acid |
| BDOC | biodegradable dissolved organic carbon |
| BDOM | biodegradable dissolved organic matter |
| BIF | bromine incorporation factor |
| BOD | biochemical oxygen demand |
| BOM | biodegradable organic matter |
| BSF | bromine substitution factor |
| | |
| CAD | *p*-coumaric acid |
| CDBAA | chlorodibromoacetic acid |
| cfu | colony-forming unit |
| CGR | column growth response |
| CIAD | cinnamic acid |
| C/V | cinnamyl to vanillyl phenols (ratio) |
| | |
| DA | drainage area |
| DBAA | dibromoacetic acid |
| DBP | disinfection by-product |
| DBPFP | disinfection by-product formation potential |
| DCAA | dichloroacetic acid |
| DEMR | division in extreme and mean ratio |
| DHAA | dihaloacetic acid |
| DHAAFP | dihaloacetic acid formation potential (a measure of DHAA precursors) |
| DHAN | dihaloacetonitrile |
| DHANFP | dihaloacetonitrile formation potential (a measure of DHAN precursors) |
| DI | deionized |
| DNA | deoxyribonucleic acid |
| DO | dissolved oxygen |
| DOC | dissolved organic carbon |
| DOM | dissolved organic matter |
| DPD | *N*,*N*-diethyl-*p*-phenylenediamine |



©2007 AwwaRF. All Rights Reserved.

| | |
|---|---|
| EVA | ethyl vanillin |
| | |
| FA | fulvic acid |
| FAD | ferulic acid |
| FCM | flow cytometry |
| FD | freeze-dried, freeze-drying |
| FISH | fluorescent in situ hybridization |
| | |
| GAC | granular activated carbon |
| GC | gas chromatography |
| GC/ECD | gas chromatography/electron capture detector |
| GC–MS | gas chromatography–mass spectrometry |
| GF/F | glass microfiber filter |
| GIS | geographic information system |
| | |
| HAA | haloacetic acid |
| HAAFP | haloacetic acid formation potential (a measure of HAA precursors) |
| HAA6 | the sum of six haloacetic acids |
| HAA9 | the sum of nine haloacetic acids |
| HAN | haloacetonitrile |
| HK | haloketone |
| HMW | high molecular weight (NOM fraction) |
| HPI | hydrophilic (NOM fraction) |
| HPLC | high-performance liquid chromatography |
| HPO | hydrophobic (NOM fraction) |
| HUC | hydrologic unit code |
| | |
| ICR | Information Collection Rule |
| IHSS | International Humic Substances Society |
| | |
| kD | kilodalton |
| | |
| LC–MS | liquid chromatography–mass spectrometry |
| LDC | Legacy Data Center |
| LULC | land use/land cover |
| | |
| MCAA | monochloroacetic acid |
| MHAA | monohaloacetic acid |
| MLR | multiple linear regression |
| MN | number-averaged molecular weight |
| MTBE | methyl-*tert*-butyl-ether |
| MW | molecular weight |
| MWCO | molecular weight cutoff |
| MUA | Municipal Utilities Authority |
| | |
| NA | not applicable |
| NBTOC | North Branch total organic carbon |



©2007 AwwaRF. All Rights Reserved.

| | |
|---|---|
| ND | not determined; no data |
| NG | no growth |
| NOM | natural organic matter |
| NOX | *Spirillum NOX* |
| ntu | nephelometric turbidity unit |
| NWIS | National Water Information System |
| | |
| PAD | *p*-hydroxybenzoic acid |
| PAL | *p*-hydroxybenzaldehyde |
| PON | *p*-hydroxyacetophenone |
| POC | particulate organic carbon |
| PTFE | polytetrafluoroethylene |
| PWSID | public water system identification |
| | |
| rRNA | ribosomal ribonucleic acid |
| RVC | rotavap (rotary evaporator) concentrate |
| RW | raw water |
| RWF | raw water, filtered |
| | |
| SAD | syringic acid |
| SAL | syringaldehyde |
| SBOD | soluble biochemical oxygen demand |
| SBTOC | South Branch total organic carbon |
| SD | standard deviation |
| SDS | simulated distribution system |
| SN | syringaldehyde |
| SON | acetosyringone |
| sp. | species (singular) |
| spp. | species (plural) |
| SPP | simple plant products |
| std. | standard |
| std.sp. | standard specific |
| STORET | STOrage and RETrieval |
| SUVA | specific ultraviolet absorbance |
| S/V | syringyl to vanillyl phenols (ratio) |
| | |
| TBAA | tribromoacetic acid |
| TCAA | trichloroacetic acid |
| THAA | trihaloacetic acid |
| THAAFP | trihaloacetic acid formation potential (a measure of THAA precursors) |
| THM | trihalomethane |
| THMFP | trihalomethane formation potential (a measure of THM precursors) |
| TOC | total organic carbon |
| TOX | total organic halide |
| TOXFP | total organic halide formation potential |
| TPL | Trust for Public Land |



©2007 AwwaRF. All Rights Reserved.

| | |
|---|---|
| TTHM | total trihalomethanes |
| TTHMFP | total trihalomethane formation potential |
| | |
| UF | ultrafiltration |
| UFC | uniform formation conditions (test) |
| UFP | ultrafiltration permeate |
| UFR | ultrafiltration retentate |
| UFR-FD | ultrafiltration retentate, freeze-dried |
| UMass | University of Massachusetts |
| USEPA | United States Environmental Protection Agency |
| USGS | United States Geological Survey |
| UV | ultraviolet |
| $UV_{254}$ | ultraviolet absorbance at 254 nm |
| | |
| VAD | vanillic acid |
| VAL | vanillin |
| VBA | Visual Basic for Applications (Microsoft) |
| VN | vanillin |
| VON | acetovanillone |
| | |
| WEF | Water Environment Federation |
| | |
| XADEF | XAD effluent |
| XADEF-FD | XAD effluent, freeze-dried |
| XADEL | XAD eluate |
| XADELC | XAD eluate concentrate |
| XADEL-FD | XAD eluate concentrate, freeze-dried |



©2007 AwwaRF. All Rights Reserved.

Awwa Research Foundation

6666 West Quincy Avenue
Denver, CO 80235-3098 USA
P 303.347.6100
www.awwarf.org
email: info@awwarf.org

**Sponsors Research**

**Develops Knowledge**

**Promotes Collaboration**

1P-2.5C-91186-07/07-NH