2004

A heuristic methodology for locating monitoring stations to detect contamination events in potable water distribution systems

James R. Chastain
University of South Florida

A Heuristic Methodology for Locating Monitoring Stations to Detect Contamination Events in Potable Water Distribution Systems

By

James R. Chastain, Jr.

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Environmental and Occupational Health
College of Public Health
University of South Florida

Major Professor: Yehia Y. Hammad, Sc.D.
Stuart M. Brooks, M.D.
Robert J. Murphy, Ph.D., P.E.
Yiliang Zhu, Ph.D.

Date of Approval:
October 14, 2004

Keywords: Sampling, Optimization, Drinking Water Protection, Vulnerability Assessment, Computer Modeling

Dedication

This work is dedicated to my wife, Nancy, who over the past twenty-eight years has provided encouragement, companionship and support while enduring a husband distracted by a seemingly endless stream of courses, assignments or readings.

To our children Lauren, Jay and Courtney, who have rarely known their father not studying alongside them, but who remain his constant source of joy and pride.

Acknowledgements

I would like to thank my committee members for their time and consideration during the development of my project. For a student working on a somewhat “non-standard” project, it was a great encouragement to have a supportive committee. I would like to especially thank Dr. Yehia Hammad for his guidance, insight and encouragement as I tried to formulate the approach and techniques used. I appreciate Dr. Yiliang Zhu’s attention to detail and the constructive comments which always helped me to think through my topic more completely. Dr. Jerry Murphy was always there with a smile and an encouraging word as he sought to help me focus on the overall objective. And special thanks to Dr. Stuart Brooks who first encouraged me to enter the doctoral program and to try to make a contribution which would combine my engineering and public health background and training.

I’d also like to thank Mrs. Lesa Richmond without whose spreadsheets and formatting wizardry this paper would have certainly taken much longer to complete. Mrs. Cindy Duhm was also a great help with her assistance at the office and her encouragement. To both of you thank you.

Anyone who has a family and has undertaken the rigors of a Ph.D. program knows that the family bears as much of a burden as the student. Accordingly, I must acknowledge my sincere appreciation and thanks to my wife, Nancy, and my children, Lauren, Jay and Courtney, for their indulgence during the years of studying, research and writing associated with this dissertation. I thankfully dedicate this work to you.

Table of Contents

List of Tables

List of Figures

List of Symbols and Abbreviations

Symbol	Description
A	Cross-sectional Area
ADD	Average Daily Demand
A_T	Cross-sectional Area of storage tank
AWWA	American Water Works Association
C	Concentration
C	Hazen Williams Constant
C*	Threshold concentration
CCL	Contaminant Candidate List
CCR	Consumer Confidence Rule
CFR	Code of Federal Regulations
cfs	Cubic feet per second
C_{m,i,t_s}	Contaminant Concentration measure at monitoring node m, generated from injection node i at time t since injection
CT	Concentration *Time value
CWS	Community Water System
d	Molecular diffusivity
D	Pipe diameter
DBPR	Disinfectant and Disinfection By-Products Rule
DBT	Design Basis Threat
DVM	Eulerian Discrete Volume Method
e	Pipe wall uniform roughness
EDM	Lagrangian Event-Driven Method
e_m	Number of times monitoring node, m, detects contamination over all injection nodes (over a specified time)
$e_{m,i}$	Number of times monitoring node, m, detects injection at node i, over a specified time
EPA	Environmental Protection Agency
f	Darcy-Weisbach friction factor
FAC	Florida Administrative Code
FDEP	Florida Department of Environmental Protection
FDM	Eulerian Finite Difference Method
FGN	Fixed Grade Node
f_{m,i,t_s}	Logical test comparing C* with value of concentration at monitoring node, m, resulting from injection at node, i, at time t_s
g	Gravitational constant
gpm	Gallons per minute
GWR	Ground Water Rule
H	Total Head
h_c	Pump Shutoff Head
h_f	Friction loss
ICR	Information Collection Rule
k	Decay (growth) constant
k_b	Bulk reaction coefficient at pipe wall
k_f	Mass transfer coefficient
k_w	Pipe wall reaction rate
L	Pipe length
LCCA	Lead Contamination Control Act of 1988

LCt_{50}	Lethal Concentration-time dose for 50% (median) of the subject population
LD_{50}	Lethal Dose for 50% (median) of the subject population
LOAEL	Lowest Observable Adverse Effect Limit
m	Number of monitoring nodes under consideration
m_b	Most efficient monitoring node for C*, t_s and specified time-of-day injection began
MCL	Maximum Contaminant Limit
MCLG	Maximum Contaminant Limit Goal
MDD	Maximum Daily Demand
mg/L	Milligrams per liter
MHD	Maximum Hourly Demand
m_i	Monitoring node efficiency for detecting injection at injection node i
n	Pump impeller rotational speed (rpm)
n	Number of nodes under consideration
NCWS	Non-community Water System
NOAEL	No Observable Adverse Effect Limit
NPDWR	National Primary Drinking Water Regulations
NSDWR	National Secondary Drinking Water Regulations
p	Pressure
P	Pump power
Q	Flow
q_s	External water supply or demand
r	Pipe wall reaction rate
r	Pipe radius
RfD	Reference Dose
R_H	Hydraulic Radius
SDWA	Safe Drinking Water Act
Sh	Sherwood Number
S_{m,i,t_s}	Sum of $f_{m,i,ts}$ for monitoring node, m, resulting from injection at node, i, from time t=0 to t=t_s
SWTR	Surface Water Treatment Rule
t	Time
TCR	Total Coliform Rule
t_d	Duration of the contamination event (time modeled)
TDM	Lagrangian Time-Driven Method
t_s	Time since injection began
V	Mean velocity
z	Height
ΔH_t	Change in tank water level at time t
δt	Change in time
γ	Specific weight of water
μ	Dynamic viscosity
ν	Kinematic viscosity
ρ	Fluid density

A Heuristic Methodology for Locating Monitoring Stations to Detect Contamination Events in Potable Water Distribution Systems

James R. Chastain, Jr.

ABSTRACT

The requirements to protect public water systems from intentional contamination have expanded in the years following September 11, 2001. The areal extent and non-linear nature of water demand and movement in the distribution system makes efficient location of sampling points difficult. This difficulty is compounded by the fact that contamination conceptually can occur at any point and at any time within the distribution system. Small to mid-sized water systems are especially at a disadvantage in addressing this issue due to limited resources available to them.

This paper proposes a heuristic methodology to identify strategic locations within the system that can be established as critical detection points for such occurrences. The process uses off-the-shelf software and is structured to be accessible to small and mid-sized water system managers. This methodology is different from others proposed in the literature in that it uses computer simulations to create a database of water system response to contamination at every node in the system. A process is developed to mine this database systematically after considering concentration thresholds and "time since injection" parameters. Finally, using pivot tables and graphs, a network of monitoring locations is identified to provide efficient coverage of the system under the conditions imposed.

Chapter 1 Problem Statement

Introduction

September 8, 1853 was a critical day for the water industry. On that day Dr. John Snow removed the handle from the Broad Street well which helped to stem the spread of cholera during the London epidemic (Aldrich, Griffith, & Cooke, 1993). In doing so he ushered in a new era not only in the field of analytic epidemiology but also in the role and responsibilities of public water supplies. A safe, reliable public water supply has always been a requirement for a sustainable community but from that point forward there was a new realization of the effect that water quality played in disease transmission and public health. Water supplies at the turn of the last century faced widespread fear of the effects of cholera, dysentery and other critical waterborne diseases. However, as the understanding of disease and its etiology improved, water utilities developed strategies to address those threats. The result was dramatic reductions in the morbidity and mortality of the population they served.

September 11, 2001 also will serve as a date which initiated a major paradigm shift in the water industry. Many of the design and operational practices incorporated throughout a water system are built on a foundation of public trust and the common good. However, now, in addition to accidental contamination, the possibility of intentional contamination of water supplies to disrupt society to achieve certain political or ideological goals must be considered. Water supplies have evolved into much more complex and sophisticated systems since Dr. Snow sought to protect consumers against the threat of

waterborne disease approximately 150 years ago. Now as a different dimension to an old problem asserts itself, it will be necessary to break down preconceived patterns of thinking and look carefully at each component of the infrastructure and examine its vulnerability, while keeping the overall framework in view.

A recent study proposed a helpful framework for addressing the water supply infrastructure as a whole (Haimes, Matalas, Lambert, Jackson, & Fellows, 1998). Their analysis recommends using a multivision risk identification method called Hierarchical Holographic Modeling (HHM) to "harden" individual water supplies against attack. This approach seeks to look at these complex systems in a holistic manner, as series of interconnected and overlapping sub-systems. Somewhat in the spirit of "a chain is no stronger than its weakest link", HHM seeks to identify and model all the components that compose the water system and affect its operation so the interactions and weaknesses may be exposed. Fifteen categories are proposed as comprising the basic decomposition of a water supply infrastructure. They are:

Category A: Physical Components

Category B: Scope of Impact (individual, plant, local, state, regional, national, or international)

Category C: Temporal (ability to detect, respond, and recover from a time perspective)

Category D: Maintenance (policies and capability)

Category E: Organizational (decision-making structure)

Category F: Management (security, short/long term emergency response)

Category G: Resource Allocation (prioritization of funds including system hardening)

Category H: SCADA (cyber-tampering, modeling accuracy, data management)

Category I: Systems Configuration (interconnection of physical, institutional, organizational and management configurations)

Category J: Hydrology (water sources and characteristics)

Category K: Geography/Physiography

Category L: External Ffactors (natural hazards, situations generating threats)

Category M: System Buffers (redundancy, over-design)

Category N: Contaminants (what range of contaminants can compromise a system)

Category O: Quality of Water (what safety and aesthetic characteristics must be met)

Whether or not one agrees with the system decomposition they propose, it is correct to point out that a wide range of factors contribute to the reliable delivery of water to the consumer with an acceptable quality level. And, while all can agree with the multifaceted approach to reducing system vulnerability, models of the individual sub-components must be defined with sufficient precision to adequately describe cause-effect responses for the system analyses.

When designing and operating a Public Water System many factors must be considered. However, none is more important than producing and maintaining an acceptable water quality. This is the unifying paradigm underlying the regulatory environment and the engineering rules of practice.

Following this then, one important aspect of system protection involves the capability to effectively monitor the water for constituents that could harm consumers. Obviously, without the ability to detect and quantify a contamination event, the ability to avert or effectively respond to it is severely constrained. As this study ranges over a myriad of

diverse topics, the issue of water quality and detection of deviances from acceptable values forms the central theme tying it together.

Literature Search

Water systems are generally divided into three major components: the water source or supply, the treatment plant and the distribution system. Engineering design of these components proceeds under the rubric of the "multiple barrier approach" in which a series of treatment processes or redundancies provide a safety net to assure downstream water quality. Water quality is typically monitored carefully at the source and during treatment stages of the system as required by operational and regulatory guidance. However, once the treated water enters the distribution system the level of monitoring effort is significantly reduced, reflecting an implicit assumption that the probability of water quality degradation in the distribution system is limited. This assumption can no longer be considered universally valid and in fact, because of the geographical extent and relatively direct access, it may be asserted that the distribution system presents itself as the most vulnerable component of the overall system.

Over the past decade studies have appeared which, apart from any terrorist activity, highlight the impact of the distribution system itself on water quality (Clark, Grayman, & Wymer, 1993; Craun & Calderon, 2001; Herwaldt, Craun, Stokes, & Juranek, 1992; Lindley & Buchberger, 2002). Given this set of conditions, improved monitoring of the distribution system is becoming an increasingly crucial task. However, there is surprisingly little in the literature relative to distribution system monitoring, characterization and protection from a water quality perspective. More specifically, the literature is sparse relative to quantitative methods to locate sampling stations within a

distribution system which increase the probability of detecting a contamination event (either accidental or intentional).

A few researchers have developed approaches which provide some guidance on the topic. One of the first attempts to objectively locate monitoring stations employed a method that used pathway analysis coupled with integer programming to identify which nodes have the maximum coverage of the distribution system (Lee & Deininger, 1992; Lee, Deininger, & Clark, 1991) . The methodology is helpful in situations with continuous, steady contamination in that the water quality effects are discerned by the hydraulic pathways. Intermittent or rapidly variable water quality swings, such as those that might be encountered in a terrorist event, weaken this approach. A later study (Kumar, Kansal, & Arora, 1997) refined the process of Lee et al. by proposing a methodology of re-ranking the pathway matrix to simplify its form thus making it more usable. However, the same base assumptions and weaknesses still apply. Kessler, Ostfeld & Sinai (1998) proposed a different, more novel approach by using the distribution system flows as a directed graph and incorporated an "all shortest paths" algorithm to find the minimum propagation times from any source node to other points in the system. Kessler et al.'s method is based on the hydraulics of the system and contamination is inferred only on the basis of water movement. It also assumes that contamination is continuous and any water passing through a contaminated node is considered contaminated regardless of the concentration. Non-conservative contaminants are not covered under this approach. Kessler et al. produces a pollution matrix for a level of service by which it is possible to estimate the volume of contaminant consumed prior to the first opportunity to detect under the stated assumptions. Ostfeld and Salomons (2003) build upon Kessler et al.'s methodology by randomizing the

pollution matrix and then using a genetic algorithm to seek the “most fit” set of monitoring points. Their approach expands Kessler et al. by allowing a limited number of events and defining a concentration hazard level. Berry, Fleisher, Hart & Phillips (2003) propose a mixed integer programming solution that follows from probability distributions of coupled population-weighted flows and contamination points. These risk points are developed through a Delphi type process. Bahadur, Samuels, Grayman, Amstutz & Pickus (2003) propose a GIS approach to the task in a computer program called PipelineNet. This program develops a detailed source prioritization ranking based upon a number of variables which are then coupled with a distribution system component scoring matrix and a population density and critical infrastructure matrix. Using a hierarchical approach the nodes are scored, ranked and related to the GIS map to identify the sites best suited to minimize system vulnerability.

Each of the approaches mentioned above has its own set of assumptions and applicability. From a conceptual standpoint a weakness asserted for these methodologies is that they do not capitalize on the power and flexibility of the extended period water quality models currently available. Consequently, significant assumptions regarding contamination modes and characteristics must be made rather than computing the effects directly. Most methodologies base their process on a limited number of contamination actualizations with an emphasis on the hydraulic portion of the model. This may not be reflective of likely contamination events, especially those that may occur as a pulse or discrete incident. Also from a pragmatic standpoint, these approaches have seen limited application because they tend to be mathematically complex and the programs and concepts are out of reach of most operators or consultants.

The purpose of this paper is to present a methodology for the selection of optimal monitoring locations within the distribution system. This methodology will be based upon a more stochastic approach than previously published studies. Furthermore, an approach is sought which can be tailored for use by small or mid-sized community water systems. It is asserted that these water systems are more vulnerable relatively speaking than larger distribution systems. Smaller systems typically have thinner management resources, smaller dilution effects and shorter critical response times. Certainly, subsets of larger systems can have similar characteristics, but again they can bring more sophisticated resources to bear on the problem.

Statement of Research Questions

It is posited, then, that the distribution system is a "weak link" in the monitoring chain and as such provides to those so motivated a potentially effective means of causing harm to the consumer. For reasons to be discussed later in this study, it is also believed that this is especially true in small to mid-sized water systems. The goal of this dissertation then is to outline a methodology which will guide more effective placement of water quality monitoring stations within distribution systems of small to mid-sized water systems. The questions to be addressed in this study are:

1. What methods have traditionally been referenced as a means to select monitoring station placement?
2. Are these methods adequate and easily accessible to small and mid-sized systems?
3. Can a methodology be developed, using commercially available "off the shelf" software, that would allow small and mid-sized operators and their consultants to

reliably analyze their systems to predict effective locations for water quality monitoring stations?

4. What are the key parameters to be considered when trying to establish a monitoring station network?

Document Organization

This dissertation is organized to proceed through the background information and findings of this study in the following fashion.

Chapter 2 provides an overview of the layout and design of a conventional public water supply and discusses the primary guidelines that govern the major components of a typical system.

Chapter 3 discusses the approaches that have been used in the past to establish water quality monitoring station locations within a distribution system. A discussion of the assumptions, strengths and weaknesses of these approaches will be discussed within the context of the focus of this study.

Chapter 4 discusses the contaminants that have been historically associated with waterborne diseases and contamination and provides a sketch of some of the agents that have been identified in the literature as candidates for concern.

Chapter 5 addresses the use of computer simulation as a tool to estimate the response of a water system to changes in operating conditions both from a hydraulic and water quality standpoint.

Chapter 6 combines these factors to present an approach that can be used to generate a database that will simulate an intentional attack upon a public water supply. Then an analysis and ranking algorithm will be proposed that may be used to predict an efficient set of monitoring station locations for the proposed contamination scenario.

Chapter 7 presents a sample application of the techniques and algorithms proposed and discuss the results.

Chapter 8 summarizes the findings of this study and lists the primary conclusions stemming from this research. Topics for additional research will also be proposed.

Chapter 2 Overview of the Design of a Conventional Water System

The fundamental purpose of a public water system is to deliver water to its customers in adequate quantities (volume and pressure) with acceptable quality and at a reasonable price. This implies two primary tasks. The first addresses the fact that physically a water system operates within a supply and demand framework. Water is essentially an incompressible fluid and because the demand can vary on an instantaneous basis, a modern water system must include a means of supplying, storing and pressuring the water to meet the range of demand reasonably anticipated. The second task recognizes that unless the quality of the water delivered is acceptable, the water cannot be used without incurring some level of harm or damage. All of this being accomplished within an economic framework that allows reasonable access to the resource.

From a conceptual point of view there are several "complicating" factors that create much of the difficulty associated with designing and operating a water system. They are:

1. Variability of the consumer demand
2. Incompressibility of water
3. Friction (energy) loss in transport systems

There are, of course, many other factors that present difficult technical and operational challenges, but these three in particular have widespread implications on the design of the physical system.

Consumer demand can vary significantly in time and space. Facility planning requires significant attention to the nature and location of users within the system as well as changes in environmental, economic and regulatory impacts. Regardless of changes in flow demand, the quality of the water produced must remain within regulatory guidelines and the volume and pressure must stay within acceptable levels. That is, increased demands in one part of the system should not adversely impact the level of service in other locations. This is complicated by the second factor in that efficiency of storage cannot be achieved in the same way that a compressible fluid can. Thus, storage is on a 1:1 basis and can require significant capital investment to provide sufficient volume to meet changing system demand. Storage and demand must be matched within reasonable limits because it is possible to build too much storage, in which case water can become stagnant, giving rise to water quality concerns.

Finally, the fact that water can have significant friction (energy) loss in its transmission to the consumer makes analysis, design and operation challenging. Considering the factors mentioned above, water will be demanded in varying quantities literally from moment to moment within the system. Energy loss through this variable demand is inherently *non-linear* and thus creates significant difficulties in developing analytical solutions to distribution system design.

This chapter will provide a brief overview of the means by which conventional water systems meet these requirements. A discussion of the historical context of public water supplies in this country will be presented followed by a brief outline of the regulatory framework within which water systems operate. Because of the nature of this study, more attention will be focused on the water quality aspects of the regulations than the

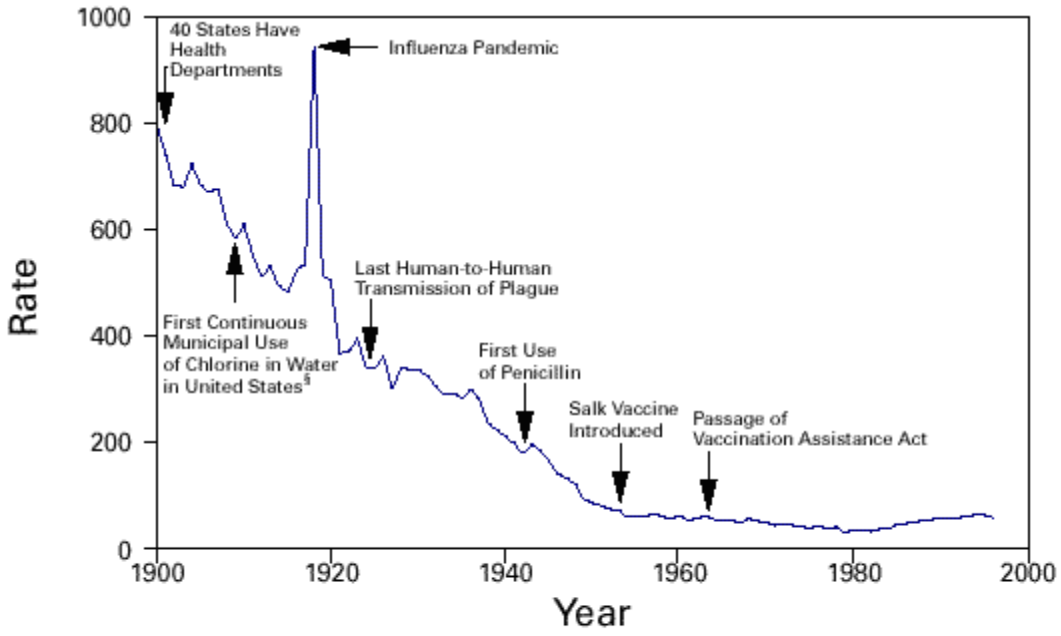
design aspects. An outline of the components and general processes used to meet water system demands and water quality is then presented.

Historical Context – the Rise of the Modern Water System

The design and operational standards of public water supplies has evolved dramatically over the past 100 years. Shortly after the turn of the century, Congress authorized the United States Public Health Service to develop regulations to minimize the spread of communicable diseases through public water supplies. One of its first tasks was to develop a means to distinguish a safe water from unsafe water. The relationship between acute waterborne disease and microbial activity had been recognized and, thus, attention was focused on improving the ability to reliably test for safe levels of pathogens. The initial standards identified coliform bacteria as an effective surrogate parameter for microbial contamination. The rationale being that, since coliform is an intestinal bacteria, a positive test would indicate that some level of fecal contamination must exist. Interestingly, almost 100 years from the time the test was proposed and implemented, the coliform test remains the primary routine method for evaluating water safety from a biological perspective (AWWA, 1990).

Concurrently, water treatment methods were being developed which could considerably improve water quality. It was observed that a reduction in turbidity through simple filtration provided significant beneficial effects, especially for those systems that depended on surface waters for their basic supply. Studies examining methods for disinfecting water demonstrated the efficacy of chlorine in reducing microbial levels. With the development of a means to safely apply chlorine at the water plant (i.e. the gas chlorinator), engineers began to include chlorine as a disinfectant in water supplies

beginning in 1908. With the establishment of these two processes a remarkable reduction in various waterborne diseases followed. Although other improvements in sanitation and medical treatment also contributed to this reduction, a significant portion of the seventy-five (75) per cent decline in the crude death rate for infectious disease from 1900 to 1940 can be attributed to these changes in water treatment (Armstrong, Conn, & Pinner, 1999). This is demonstrated graphically in Figure 1.



*Per 100,000 population per year.
†Adapted from Armstrong GL, Conn LA, Pinner RW. Trends in infectious disease mortality in the United States during the 20th century. JAMA 1999:281;61–6.
§American Water Works Association. Water chlorination principles and practices: AWWA manual M20. Denver, Colorado: American Water Works Association, 1973.

Figure 1 Crude Death Rate for Infectious Diseases

In the period between 1860 and 1960 the number of centralized municipal water systems had grown from 400 to 19,000 (AWWA 1990). Once the basic parameters of treatment were established, the rapid development of public supplies devoted much of its attention to the efficient and reliable delivery of water to its users. Advances in treatment process design, laboratory analyses, construction methods and materials, as

well as use of information technologies for modeling and control of system components have allowed greater sophistication and efficiency in the development of this country's water systems.

Regulatory framework – "What is a safe water?"

The design and operation of potable water systems in the United States is heavily regulated so as to provide a framework for the protection of public health and safety. Therefore, to accurately provide an overview of public water systems, the regulatory context within which they operate must be sketched. In fact the regulatory structure defines to a great extent what constitutes a "safe" or acceptable water quality, at least from a human consumption standpoint. These regulations codify much of the professional practice which over the years has created a network of water supplies that are unparalleled in history for their safety and reliability. Although taken for granted by most Americans, it is truly remarkable that one can travel from coast to coast drinking water from public supplies all along the way and not give any thought to the potential of contracting waterborne disease.

In general, public water supplies are governed by the Safe Drinking Water Act (SDWA) of 1974 and its amendments along with the Lead Contamination Control Act (LCCA) of 1988. At this point there have been five (5) major amendments to the SDWA which were promulgated in 1977, 1979, 1980, 1986, and 1996. These laws are codified in the *U.S. Code* at 42 U.S.C. s/s 300f *et seq.* which is also listed as Title XIV of the U.S. Public Health Services Act Section 1400 *et seq.* of the *United States Code Annotated*®. The Environmental Protection Agency (EPA) has federal responsibility for implementing the law. The regulations passed for this purpose are listed in the *Code of Federal*

Regulations at 40 CFR Parts 141 – 143 and 149. Responsibility for the operational management and enforcement for these regulations has for the most part been delegated to the states through the primacy process. In order to achieve primacy, states must stipulate to and demonstrate capacity to enforce the federal requirements as outlined in the law and the subsequent EPA regulations. Any state unwilling or unable to meet these requirements does not receive primacy and in that case the EPA assumes responsibility for regulation and enforcement in that state (Kucera, 2003). Florida has been granted primacy and enforces drinking water laws primarily through the Florida Department of Environmental Protection (FDEP). The FDEP by interdepartmental agreement has in turn delegated certain responsibilities to the Department of Health.

Consequently, every aspect of a water utility is governed by the overarching requirements of the SDWA as implemented through state laws and regulations. These regulations relate to water quality standards, testing methodologies, critical design criteria, source water protection, enforcement authority and consumer notification/awareness of violations. Among the more prominent rules issued by the EPA in conjunction with the SDWA, along with their primary references, are:

1. Surface Water Treatment Rule (SWTR):	40 CFR Part 141 Subpart H
2. Disinfectants/Disinfection By-Products Rule	40 CFR Part 141 Subpart L
3. Total Coliform Rule (TCR)	40 CFR Part 141.21
4. Lead and Copper Rule:	40 CFR Part 141 Subpart I
5. Information Collection Rule (ICR):	40 CFR Part 141 Subpart M
6. Consumer Confidence Rule (CCR):	40 CFR Part 141 Subpart O
7. Groundwater (Disinfection) Rule (Proposed):	65 FR 30194

These rules can have different applicability depending on the classification of the system. Public Water Systems are generally classified as Community Water Systems (CWSs) and Noncommunity Water Systems (NCWSs). This study addresses CWSs but does not explicitly consider NCWSs. A *Community Water System* is legally defined as a public water system which serves at least 15 service connections used by year-round residents or regularly serves at least 25 year-round residents (40 CFR 141.2). NCWSs are systems smaller than this and as such consist of components so small as to make monitoring system design trivial.

Water quality standards are among the most important sections of the regulations. They stipulate the chemical and biological concentrations and characteristics that constitute a "safe" or potable water. The most common standards are set forth in the National Primary Drinking Water Regulations (NPDWR) and the National Secondary Drinking Water Regulations (NSDWR).

The NPDWR, which currently consists of 87 chemical, microbiological and physical parameters, forms the basis of the regulatory examination of a water system's performance. These parameters are assigned enforceable criteria levels called Maximum Contaminant Levels (MCLs) that define a safe potable water. Facilities which exceed established MCLs are deemed in violation of their permit to operate and subject to enforcement action if not acceptably remedied. There are a few parameters in the NPDWR that have not been assigned an MCL because they cannot be feasibly measured or there is uncertainty about the appropriate limits. In those cases a *Treatment Technique* is assigned. Thus, in systems where a particular parameter has been demonstrated to exist or is likely to exist, the water utility must implement the

treatment technique associated with that parameter in order to be deemed in compliance with the regulations. In other words this is a presumptive criterion as explicit monitoring is not performed. In summary, MCLs are health based criteria, they are enforceable and they are developed with consideration to the cost-benefit associated with them.

NSDWRs differ from NPDWR in that they are non-enforceable guidelines that address contaminants that may have adverse cosmetic or aesthetic effects in drinking water. In other words these parameters may affect the palatability or cosmetic/staining characteristics of the water, but have no meaningful impact on its safety. While Secondary Maximum Contaminant Levels (SMCL) are not enforceable under federal law, it is noted that an individual state may elect to make a secondary parameter enforceable.

In addition the SDWA requires that every contaminant with an MCL have an associated Maximum Contaminant Level Goal (MCLG). This is a non-enforceable health-based criteria. These are based upon the National Research Council (NRC) risk assessment process and are formulated to be set at a level at which there are no known adverse effects and with an adequate safety factor (NRC, 1983). The MCLGs are set without regard to cost to achieve the stipulated concentration.

Historically, MCLGs for carcinogens have always been set at zero following theoretical and practical limits to determining the existence of a threshold of action. More recently the EPA has adopted a “weight of evidence” process which assigns a contaminant to one of three categories based on the knowledge base and potency of the carcinogen to determine whether a non-zero value for the MCLG may be assigned. While not

enforceable, the MCLGs provide valuable information as to treatment targets to be achieved and potential regulatory direction in future years (Kucera, 2003).

Of course, non-carcinogens also have MCLGs and are established based No Observable Adverse Effect Levels (NOAEL) or Lowest Observable Adverse Effect Levels (LOAEL) to determine a Reference Dose (RfD). This is then related to a Drinking Water Effect Level (DWEL) which is used to compute the MCLG.

The list of regulated contaminants continues to grow as the SDWA requires the EPA to monitor unregulated contaminants and add them to the NPDWR as appropriate. Potential contaminants are examined and monitored to determine whether they occur at a frequency and in concentrations that may warrant further study. If so they are placed on the Contaminant Candidate List (CCL). The CCL was initially published in 1998 and is reissued every five years. These listed contaminants are first assigned a proposed MCLG and then they proceed through scientific study and public notification and comment before a final decision is made on their regulatory status.

From this brief sketch it may be observed that Congress has established a broad and continually evolving framework to govern the design and operation of public water systems. Water quality relative to human consumption is paramount in the legislative and regulatory history. This responsibility is transferred to the design and operating professionals to develop the physical systems that will comply with these high standards. The following sections will provide an overview of the components of a water system and how a typical system sets about meeting the customers' demands while operating within the regulatory requirements.

Basic Water System Design Tenants

As briefly mentioned earlier widespread centralized urban water systems are a relatively new phenomenon. With the advent of basic treatment and disinfection techniques, the value of reliable and safe water supply began to manifest itself as a primary economic component in a community's development. While engineers continued to make improvements in various aspects of its treatment and delivery, a number of guiding design principles emerged. These have been formalized in professional technical guidelines and agency regulatory requirements.

Reliability

Among the most fundamental of guidelines for infrastructure design are the requirements for reliability and redundancy. *Reliability* or robustness refers to the dependability of the overall system and its component parts (AWWA & ASCE, 1998). It recognizes, for example, that variability of water demand is an inherent part of a water system operation. Estimation of water demand is an important task and influences the selection and sizing of all the downstream components. Because human consumption comprises only a very small part of the overall water demand it is necessary to carefully consider the composition of the community served. It was recognized early on that care must be taken to not only define average conditions, but also to gage the impact of extreme events. While there may be autocorrelation between day-to-day demands, seasonal and external events (ex. irrigation, fire flow, industrial user demands) can impose significant changes in demand patterns both on a short-term and long-term basis. Therefore, system requirements (Ysusi, 2000) are typically estimated for:

Average Daily Demand (ADD) = (total water consumed in one year) / 365

Maximum Daily Demand (MDD) = maximum 24-hour demand in one year

Maximum Hourly Demand (MHD) = maximum 1-hour demand in one year

In addition, fire flow demand and duration are estimated and generally superimposed upon the MDD or MHD, along with any other external demand requirements. In more complex systems, the analysis of demand requirements and patterns can become quite sophisticated, but the principle remains the same: water must be available for the user when required (for statistically reasonable demands), and the water must be delivered with acceptable quality while maintaining at least some given minimum residual line pressure.

Reliability concerns take into account all those resources necessary to properly operate the subject facility. In the event of disruption of normal service this would include adequate provision for auxiliary power, disinfectants, chemicals and other raw materials that are crucial to operating the plant as designed to meet the regulatory requirements.

Reliability also relates to delivery of water of safe and consistent quality to the consumer. Utilities that take their water from surface supplies (lakes and rivers) have more of a challenge in this regard than utilities with groundwater supplies. This is due to the influence of seasonal variations and the direct impact of surface features on the raw water quality. These water quality variations must be recognized and treated in real time before discharge to the system. Groundwater systems are usually more consistent with regard to water quality. Because they are not directly under the influence of surface conditions they are more stable, but many times will be more highly mineralized due to their long-term contact with aquifer materials. In either case, however, reliable delivery

of adequate volumes of water at sufficient pressure and quality is an essential feature of an efficient water utility (Harberg, 1997).

Redundancy

Related to reliability is the concept of *redundancy*. In any system where mechanical equipment and control systems exist, provision must be made for malfunctions; motors will burn out, lightning will strike control panels, pipes will rupture and the like. The system must be designed to continue to function, albeit at a lower level of service, even when these malfunctions occur. Therefore, all critical components are to be designed with duplicate units (as a minimum). Also, the system is to be capable of delivering the maximum daily demand, even when the largest unit is out of service.

The concept of redundancy extends to the power and control systems too. Auxiliary power units (with appropriate fuel resources) or redundant feed from separate electrical grids is necessary to assure sufficient capacity to operate critical equipment or treatment systems. As control systems become more integrated and sophisticated, care is necessary to assure that critical equipment, programs and data are duplicated or backed up in appropriate fashion (AWWA & ASCE, 1998).

Distribution systems have the additional requirement of maintaining a minimum residual pressure of 20 psig throughout the distribution system during the stress event. This is typically accomplished by on-line storage with auxiliary pumping systems or elevated storage tanks with adequate reserve capacity.

Resiliency

Unfortunately, no human system can be designed to account for all risks so one final aspect of good design involves the concept of system *resiliency*. Given that catastrophic events such as earthquakes, floods, and fires sometimes occur, thought should be given to 'how long will it take for the system to recover?'. By judicious planning, certain aspects of the system layout can minimize the amount of time that it takes to restore water service to the community. Many times this is as much a function of organizational and management structure as it is "pipes and pumps". As discussed above (Haimes et al., 1998) a system is composed of multiple aspects and the reliability and resilience are especially influenced by many aspects of the utility operation.

Water System Components – "How do we deliver a safe water?"

A water system is generally broken down into three primary components, the water supply, the treatment plant (including finished water storage), and the distribution system. This is shown schematically in Figure 2. In order to have an effective system each component must be matched to meet the requirements of the downstream element. Each constituent has its own set of design issues and its design has actually evolved into complete sub-disciplines.

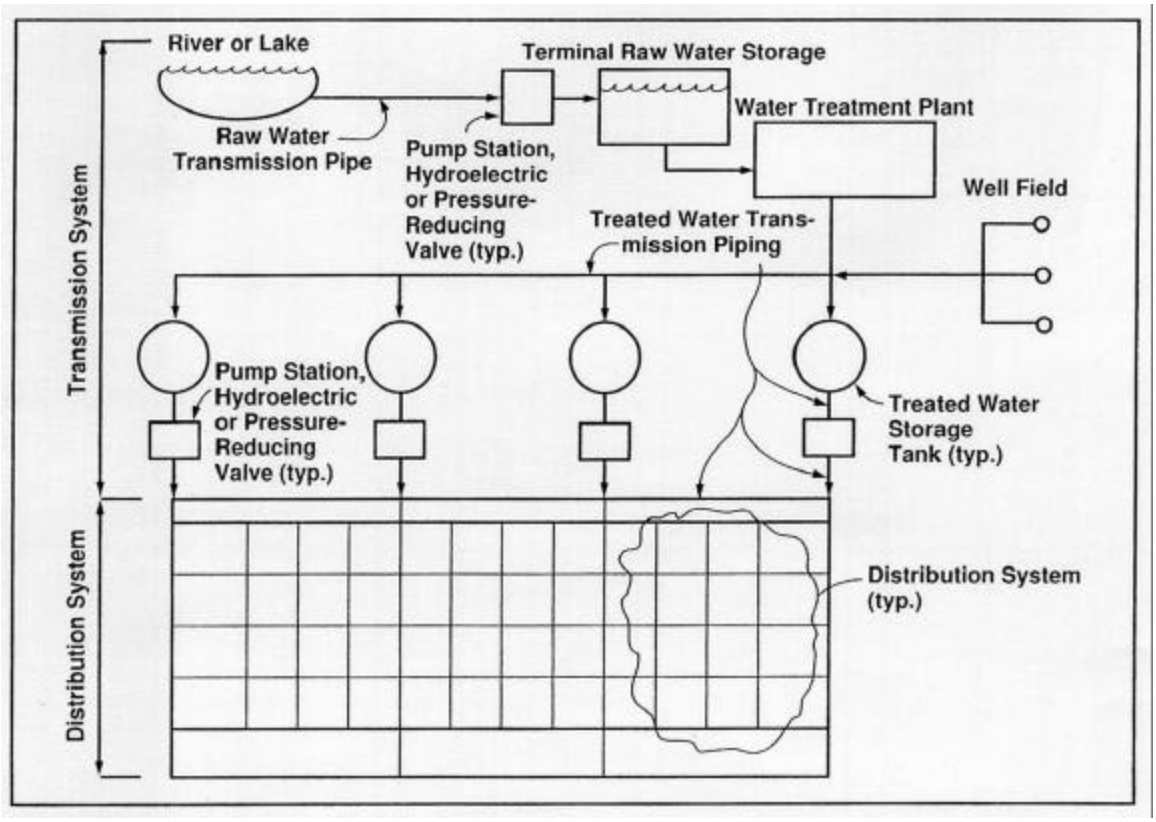


Figure 2 Generalized Layout of a Modern Potable Water System (Harburg, 1997)

It is the thesis of this study that the vast majority of attention, both in terms of regulations and research, has been focused upon the first two components (the water supply and the treatment systems) and that the distribution system has been largely neglected from the standpoint of its impact on water quality. This, of course, is to be expected in that the primary concern of a water utility is to find an adequate water supply with good raw water characteristics and then to treat it to a uniform quality that meets all health and regulatory requirements before release to the consumer. However, changes in water quality can occur within the distribution system for a variety of reasons, and it is here proposed that a more reliable means of detecting those changes is necessary to safeguard the consumer.

Water Supply

The first task of the water utility is to secure a reliable source of water with physical, chemical and biological properties that are within acceptable ranges. The nature of the source influences these properties considerably. Broadly classified, water supplies are listed as being a groundwater supply or a surface supply. Different regulatory requirements apply to a water system depending on the water source.

Typically groundwater supplies are preferable to surface supplies in that the water quality tends to be more stable and the volumetric yield is more predictable. The water may be more mineralized as a result of extended contact with the subsurface strata, but the water generally is relatively pure and not as prone to biological contamination. However, reliable ground water sources are not available in all parts of the country. This is a function of the geology of the area and whether suitable aquifers exist at a reasonable depth below ground. Florida is fortunate to have several high quality productive aquifers. Thus, most of the water supplied to its citizens comes from ground water sources.

Surface water sources are under the direct influence of stormwater runoff or snow melt and are tapped by a utility through river, reservoir or lake intake structures. Utilities using surface water sources must contend with daily variability in water quality (in that it is directly influenced by surface runoff) and seasonal volume availability which can dramatically affect public health issues (Craun, 1988). The SWTR, issued by the EPA is quite involved and has many stringent provisions which require water utilities to carefully consider and monitor the ways that source water is protected, the means by which it is

treated and what criteria must be met (and documented) prior to being made available to the public.

A special case of surface supply is brackish or ocean water. The ocean is the largest reservoir of water in the world but the adverse chemical composition, until recently, has precluded it as a viable source water. Interestingly, in many ways the ocean provides an ideal water supply, especially for coastal areas, in that the water supply is virtually inexhaustible and the water quality is consistently uniform. The primary problem has been that the technology to treat the water was inadequate or prohibitively expensive compared to other source options. Even in this case, consideration of reject water disposal must not be ignored.

Water Treatment

Once the water supply has been selected, a complete characterization of that water and its expected variability must be established. In many communities it is also not uncommon to have multiple water sources for reliability and redundancy purposes. This might also include mixed source combinations (ex. ground water and surface water supply components).

Using applicable state and federal regulations as the minimum finished water standard, the designer considers various treatment process trains to produce the required effluent quality. Careful consideration of the variability of different combinations of raw water characteristics, variation in consumer demand and cost of production is required. This can be especially complex when blending waters from multiple sources, due to chemical interactions that can occur when mixed that would not occur if they were separate.

Process design and treatment equipment have evolved significantly over the years and can effectively treat many waters that previously would have been deemed marginal, if not unacceptable. The most simple treatment process involves nothing more that chlorine disinfection before pumping it on to the consumer. This is normally only seen when a high quality ground water is the source water.

Source waters that are highly mineralized (e.g. hard waters or dissolved H_2S) or waters from surface supplies need more careful treatment and control. Water quality objectives are typically accomplished by using some form or combination of physical or chemical processes. Common component processes might include aeration, pH adjustment, sedimentation, coagulation/flocculation and clarification, lime softening, filtration, adsorption processes and more recently membrane processes.

In addition some form of disinfection will be applied. Historically, gas chlorination has been the primary means of providing microbial control of product water and has performed remarkably in terms of overall deactivation levels and long-term effectiveness. However, a change in design approach is underway as research has indicated that chlorination of certain waters can produce process by-products that increase the risk of bladder cancers and some level of adverse reproductive effects. Further research has indicated that all commonly used disinfectants have some level of by-product generation that must be controlled, with the exception of ultraviolet sterilization (AWWA, 1999; White, 1999).

Any cursory review of the regulatory requirements demonstrates that in-plant monitoring of incoming water quality and finished water quality can be quite thorough and

demanding. The laws and regulations are written to assure as much as practicable that the water produced by the CWS is safe to drink.

Distribution

Once the water has been treated it moves to the distribution network which is the final component of the water system. The distribution system is composed of pumps, the pipe grid, storage facilities and appurtenances (e.g. valves, hydrants, and service connections). Typically, the design of the distribution system has been considered primarily a hydraulic function. That is, the objective is to deliver the water from the treatment plant to the customer in sufficient volumes and at an acceptable pressure while minimizing the cost of delivery. This is accomplished by sophisticated computer analyses of pressure loss through pipes under varying conditions to meet regulatory criteria relative to residual pressure at all points in the system. In conjunction with this, sizing and placement of distribution system storage is an important factor in efficient design because it allows flow and pressure equalization which reduces energy loss and can reduce pipe stress. As mentioned earlier, because water is an incompressible fluid, clearwell storage at the water plant and distribution storage is absolutely necessary to handle the variation in pressure and demand experienced within the distribution grid. However, care must be taken not to oversize the pipes or storage because that can lead to unacceptable water age in the system.

The water quality emphasis has tended to be associated with maintenance of the integrity of system rather than the water quality itself. The implicit assumption being that if the water met compliance criteria when it left the treatment plant it will arrive at the customers tap in the same condition unless there is a breach in the structural integrity of

the system. For example, backflow prevention and cross-connection control are important programs in most water systems. The objective here is to protect the system from inadvertent connection of a non-potable water source with the potable water produced by the utility. There is also considerable effort to assure that when new water lines are added to the system or whenever the existing system is tapped for new connections that the pipe and water within the pipes are within standards. Again the focus of these important programs is to insure that the pipe grid integrity is uncompromised as opposed to detailed water quality monitoring within the network.

Water quality is monitored in distribution systems, but maybe not to the level that one might think. It is upon this portion of the system that this dissertation focuses attention asserting that the ability to detect contamination events, such as that imposed by a terrorist attack, is poorly developed in both regulatory guidance and operational practice. The focus of this study is to develop an approach to identify efficient monitoring locations with an emphasis on making those tools accessible to small to mid-sized water systems that in many ways represent the most vulnerable targets.

In order to understand the context from which this need arises, it will be necessary to briefly explore current practice in water quality monitoring and to identify weaknesses that characterize the process. This is discussed in Chapter 3.

One other means used as an indicator of the integrity of the network is the persistence of a chlorine residual in the system. Chapter 62-550.350 (6) FAC stipulates that the distribution system will maintain a 0.2 mg/L free chlorine residual or 0.6 mg/L combined chlorine residual at all times. Chlorine residual is useful in maintaining a germicidal

effect within the distribution network and can also provide an indication of a potential breach of the system. In other words if the residual becomes depressed in one section of the distribution system and not others, that would indicate a need to check the area to determine what is consuming the residual. While chlorine at that level is not a powerful disinfectant it is still useful to address many small microbiological events (VonHuben, 1999; White, 1999).

A summary of many of the most fundamental design requirements is indicated in Appendix 1 (Health Research, 1997).

Summary

The fundamental purpose of a public water system is to deliver water to its customers in acceptable quantities (volume and pressure) with an acceptable quality and at a reasonable price. In order to accomplish this, the system's characteristics must be quantified to adapt to highly non-linear factors such as demand variation and energy loss associated with bulk transport. The *sine qua non* of any potable water system, of course, is that the water is safe, palatable and suitable for use.

In the United States these goals are met within a regulatory framework which codifies much of the engineering, operational and public health experience developed over the years. The Safe Drinking Water Act, along with a number of implementation rules promulgated by the Environmental Protection Agency provides the basis for defining what a "safe" water is and how it should be treated and delivered to a system's customers.

Of fundamental importance are the 87 contaminants identified in the National Primary Drinking Water Regulations. These are the contaminants for which Maximum Contaminant Limits are set. These are the core parameters constituting the enforceable health based criteria which essentially circumscribe an acceptable water quality. The National Secondary Drinking Water Regulations are also desirable targets to meet aesthetic or functional goals but are not enforceable. An on-going process of scientific review and public scrutiny works to add to the list as additional "weight of evidence" data becomes available.

Because a community's water supply is crucial to its physical, social and economic well-being, care is taken to design, construct and operate a system that will be available to meet required demands under all reasonable conditions. To comply with that directive, water systems are designed with serious consideration given to reliability, redundancy and resiliency. Reliability relates to a system's capability of responding to the range of demands (from both a quantity and quality perspective) placed upon it. This typically involves various statistical projections relative to a community's current and future source and demand characteristics. Redundancy is somewhat related to reliability but focuses more upon the ability of a system to function acceptably when critical mechanical, electrical or control systems fail. A common example of this concept is the placement of multiple pumps in a station to allow continued pumping even when one or more pumps are taken out of service for one reason or another. Finally, water systems are designed with the concept of resiliency in mind. Emergency situations will occur occasionally which may require temporary redirections or other response actions to return service to an impaired system. By judicious placement of pipes, valves, pumps

and the like, a resilient system will allow at least nominal service to be returned in a minimum period of time.

To accomplish all these objectives, water systems are generally classified into three components: the water supply, water treatment systems and the distribution network. The water supply function is to identify, develop and protect the source system's water. Water supplies generally are classified as surface sources or groundwater sources. Different regulations apply to a system depending on that classification.

The water treatment system are those components that exist to alter the raw water characteristics of the water supply to conform to applicable regulations (as a minimum) to produce a safe, palatable and useful water. The water treatment plant(s) typically comprise the most complex operational portion of the water system and have heavy monitoring requirements relative to the other components.

Finally, the treated water is delivered to the customer via the water distribution network. This is composed of a pipe grid, appurtenances, storage, pumps and occasionally in-system treatment (ex. re-chlorination). Monitoring requirements are more rudimentary in many ways than those required in the supply and treatment segments of a water system. However, given the new post-9/11 realities, it is not prudent to allow network monitoring to remain tied to (blinded by) past practices. The focus of this study is to develop an approach to identifying efficient monitoring locations with an emphasis on accessibility of those tools to small to mid-sized water systems that in many ways represent the most vulnerable targets.

Chapter 3 Distribution System Monitoring

A key aspect of this study focuses on the capability of current monitoring practice to detect intentional contamination within the distribution system. Unless the monitoring system is able to efficiently detect the contamination event, its existence will be revealed by significant clinical response through the medical system. This, of course, is the effect that the monitoring effort endeavors to avoid. This dissertation, then, seeks to address the question of whether current regulatory and industry monitoring practice that has come to focus on minute chemical concentrations (e.g. parts per billion level) can effectively address events that lie on the other end of the spectrum, i.e. mass contamination in an acute event.

Before proposing modifications to distribution system sampling protocol, it is appropriate to examine the framework and practices underlying current monitoring practice. From a public health perspective, the US water industry has served its customers well. Major outbreaks are rare and, if an outbreak does occur, notification is prompt and specific (40 CFR Part 141, Subpart Q). The primary safeguards (Geldreich, 1996; VonHuben, 1999) for the distribution system are generally listed as being:

1. Continuous positive pressure in the water main
2. Maintenance of a minimum chlorine residual
3. Cross-connection and backflow prevention
4. Compliance monitoring
5. Corrosion control

As mentioned in Chapter 2, when one reviews the design, regulatory and operational literature it becomes apparent that until the passage of the SDWA in 1986, an operational "blind spot" existed. It seems that the generally held belief was that once water enters the distribution system, it would be delivered to the consumer with approximately the same quality it possessed when it left the treatment facility unless there was a structural break in the lines or some form of accidental intrusion occurred. Therefore, the operational and institutional focus of the water utilities tended to address construction monitoring, backflow and cross-connection programs, and customer complaint tracking.

Then in 1988 the issuance of the proposed Lead and Copper Rule required water system operators to test water "at the consumer tap" for compliance (40 CFR 141). As operators began to deal analytically with water quality issues within the distribution system itself an evolving view of water quality drivers began to occur. The subsequent Total Coliform Rule and Surface Water Treatment Rule reinforced the need to begin to address the fact that the distribution system itself may impact water quality.

This is not to say that the effects of the distribution system on water quality had gone completely unnoticed. In fact some researchers have described the distribution system more in terms of a living individual with its own dynamic characteristics than a network of pipes (Larson, 1966; Rossie, 1975). They use analogies in which the systems respond to stresses placed on them, just as a functioning entity would, resulting in changes in water quality and delivery capabilities. Also, epidemiological studies have provided ample documentation that finished water quality can deteriorate as a result of distribution system activities resulting in water-borne disease outbreaks (Besner, Gauthier, Servais,

& Camper, 2002; Clark et al., 1993; Craun & Calderon, 2001; Geldreich, 1996; Herwaldt et al., 1992; Logsdon, Schneider, & Budd, 2004). Thus, it is not the case that the literature had not documented the fact that the distribution system affects water quality. The distribution system has even been described as "the final barrier for the maintenance of water quality" (Clark et al., 1993).

However, over the past century the major battles in providing safe water have been associated with developing safe water supplies and providing adequate water treatment. If the distribution system was properly constructed and adequately maintained history generally confirmed that water quality in the distribution system remains similar to that which left the plant. Therefore, intensive monitoring for non-specific purposes tended to consistently prove negative and therefore has not been viewed as a prudent use of manpower and financial resources.

Even trying to determine what "current monitoring practice" entails is difficult. A recent AWWA Research Foundation report, commissioned to study distribution system monitoring practices noted, "Current guidelines on water quality monitoring are scattered throughout many sources and references and often are not specific to distribution systems. Written guidelines may be difficult to retrieve from utility archives and guidelines handed down verbally from past utility staff may be outdated or incomplete" (Kirmeyer et al., 2002). Thus coherent, structured distribution system monitoring plans are the exception rather than the norm in the water industry, especially in small to mid-sized systems.

This also highlights an interesting feature of water supplies. Unlike electric utilities, water systems are not connected in a global grid. Therefore, each water system has its own set of strengths and vulnerabilities. That is, not all systems are equally vulnerable. In a related fashion not all water systems have the same impact as "targets" to a terrorist group. For example, contamination of a water system serving a small mobile home park may not have the same "value" as an act that affects a major municipal system. However, one can't rule out an attack on a smaller system, because there may be localized motivations or possibly the intent to impose a psychological impact on other small system. In any event the point to be observed is that a multitude of independent water systems exist which serve the vast majority of the population. A coordinated attack on even a small proportion of the community systems could have widespread societal implications.

Although water distribution monitoring objectives have been classified in many different ways, ultimately they fall into one of two types of sampling programs, Compliance Monitoring and Special Purpose Sampling. These will be approaches will be outlined in the sections below with the intent of addressing the question posed at the outset - can monitoring systems that are established to deal with minute contaminant levels deal with massive contamination on an acute time scale?

Compliance Monitoring

Because waterborne disease posed the primary health risk during the formative years of large centralized water systems in this country, the most common monitoring activity, compliance monitoring, focused on managing that risk. Historically, that approach has been based upon the use of "indicator species" as a surrogate for overall water quality.

Indicator species have been used for that purpose for over 100 years in order to minimize the analyses required to confirm microbial contamination. The most common microorganisms used for this purpose are heterotrophic bacteria, total coliform and fecal coliform.

Over time there has been a subtle shift in emphasis from acute manifestations (ex. pathogenic organisms/diseases) to longer term or chronic outcomes (ex. carcinogens, Pb/Cu, etc). No doubt this has resulted from the marked success of improvements in water treatment and disinfection technology which dramatically reduced or eliminated waterborne disease outbreaks arising from CWSs. The implicit assumption underlying a compliance monitoring program is that the water is clean unless proven otherwise by sampling. Table 1 provides an overview of the primary regulations and requirements for sampling within the distribution system.

Table 1 Water Quality Parameters and Associated Regulations for Water Distribution Systems

Parameter	Sample Location	Regulatory Limit	Reference	Comments
Disinfectant residual*	Entry point to distribution system	Minimum 0.2 mg/L on a continuous basis	U.S. SWTR	Only applies to systems using surface water supplies. In U.S., Legionella is also regulated by a treatment technique.
Disinfectant residual	Distribution system	MRDL chlorine 4.0 mg/L MRDL chloramine 4.0 mg/L, running annual average	U.S. D/DBP Rule Stage 1	Surface water systems serving >10,000 people.
Disinfectant residual or HPC bacteria count*	Throughout distribution system	Detectable level of disinfectant residual or HPC bacteria count of 500 or less cfu/mL in 95 percent of samples collected each month for any two consecutive months	U.S. SWTR	Only applies to systems using surface water supplies.
Total trihalomethanes	Throughout distribution system	80 µg/L, running annual average based on quarterly samples	U.S. D/DBP Rule Stage 1	Surface water systems serving >10,000 people.
Haloacetic acids (HAA_5)	Throughout distribution system	60 µg/L, running annual average based on quarterly samples	U.S. D/DBP Rule Stage 1	Surface water systems serving >10,000 people.
Total coliform bacteria	Throughout distribution system	5% positive	U.S. Total Coliform Rule	Number of samples determined by population served.
Lead and copper	At customer's taps	Action levels: Lead 0.015 mg/L at 90% Copper 1.3 mg/L at 90%	U.S. Lead and Copper Rule	Number of samples determined by population served.
pH	Representative points in distribution system	Minimum of 7.0	U.S. Lead and Copper Rule	State primacy agency may waive requirement if the agency determines that it is not feasible or necessary to achieve pH 7.0

*Disinfectant residual may be regulated for some systems using groundwater supplies under the forthcoming Groundwater Rule.
Source: Kirmeyer et al, 2002.

Table 2 indicates the regulatory testing requirements for Coliform monitoring which is the test to be performed with the highest frequency (40 CFR 141.21; Chapter 62.550.518(2) FAC). The minimum number of samples required in the distribution system is typically established by state or federal regulatory standards. By and large it is left to the utility's discretion to locate the sampling points in areas which suitably characterize the system. Generally, sample points are selected in areas of reported or historical problems, different pressure zones, high risk areas or near interconnection points with adjacent

utilities. Some of the monitoring points might be permanent and some might be relocated from month to month (Geldreich, 1996; VonHuben, 1999).

Table 2 EPA Coliform Monitoring Frequency Requirements

POPULATION SERVED			Minimum number of samples per month	Population / Sample	
				Lowest Population	Highest Population
25	to	1,000	1	25	1,000
1,001	to	2,500	2	501	1,250
2,501	to	3,300	3	834	1,100
3,301	to	4,100	4	825	1,025
4,101	to	4,900	5	820	980
4,901	to	5,800	6	817	967
5,801	to	6,700	7	829	957
6,701	to	7,600	8	838	950
7,601	to	8,500	9	845	944
8,501	to	12,900	10	850	1,290
12,901	to	17,200	15	860	1,147
17,201	to	21,500	20	860	1,075
21,501	to	25,000	25	860	1,000
25,001	to	33,000	30	833	1,100
33,001	to	41,000	40	825	1,025
41,001	to	50,000	50	820	1,000
50,001	to	59,000	60	833	983
59,001	to	70,000	70	843	1,000
70,001	to	83,000	80	875	1,038
83,001	to	96,000	90	922	1,067
96,001	to	130,000	100	960	1,300
130,001	to	220,000	120	1,083	1,833
220,001	to	320,000	150	1,467	2,133
320,001	to	450,000	180	1,778	2,500
450,001	to	600,000	210	2,143	2,857
600,001	to	780,000	240	2,500	3,250
780,001	to	970,000	270	2,889	3,593
970,001	to	1,230,000	300	3,233	4,100
1,230,001	to	1,520,000	330	3,727	4,606
1,520,001	to	1,850,000	360	4,222	5,139
1,850,001	to	2,270,000	390	4,744	5,821
2,270,001	to	3,020,000	420	5,405	7,190
3,020,001	to	3,960,000	450	6,711	8,800
3,960,001	or	more	480	8,250	

Source: 40CFR141.21 (Sampling table)

By examining Table 2 and dividing the population served by the required samples, one develops an insight into the coverage each test provides. For all intents and purposes this ranges from 500 people covered per sample to 8800 people per sample. Recalling that these are *monthly* sampling requirements, it is easy to see that coverage for acute events is lacking in compliance monitoring. Most utilities will monitor in excess of the minimum number of samples required, but the essential pattern remains unchanged.

A slightly different way of looking at the monitoring practice is to examine the number of service connections covered by each biological test. This statistic is more reliable than population because water meters or connections are more easily and accurately determined by utility staff than population. Although some aspects of permit records are not available publicly subsequent to Homeland Security rules, a dataset was secured from the Florida Department of Environmental Protection (FDEP) during the early research for this paper in August, 2000. This data set lists all 2030 CWSs permitted in the state of Florida at that time. It addition it identified the population and service connections associated with the system along with the number of bacteriological tests required and the number actually performed. The number of bacteriological samples was correlated with the number of connections served. The bacteriological samples were chosen because they are the most frequently run tests. Examining water systems that serve at least 100 connections, the median number of connections covered by a bacteriological test is 99 connections per *month*. As the system size increases the unit coverage also increases. This is characteristic is indicated in Table 3.

Table 3 Average Number of Connections per Monthly Coliform Test Correlated to CWS Size

CWS with Connections Greater Than	Connections Covered per Monthly Bacteriological Test
100	99
500	194
1,000	212
2,000	228
5,000	253
10,000	261
20,000	303

Source: FDEP (Aug., 2000)

The point to notice is that both the spatial and temporal coverage of this testing is nominal. As one examines this data, however, it should become clear that the ability to detect an intentional, but unknown, contamination event within the distribution system would be limited. Since the laboratory turnaround time for coliform will be at least three days, even if the sample did detect the contamination the system managers wouldn't know in time to respond (assuming that the appropriate tests were being performed in the first place).

It should be noted that other more exhaustive distribution system compliance tests may be take once every one to three years depending on the parameter. For example, lead and copper are monitored at random consumer sites to verify compliance with regulations. However, compared with the sampling which is required for water plant operation (which varies from continuous to daily), the distribution system sampling scale is normally monthly. Obviously, any sort of acute event would not be picked up within the monitoring framework and would only serve as a confirmation test.

Special Purpose Sampling

The other type of distribution sampling can be classified as special purpose sampling. This type of monitoring is usually associated with a short-term intense sampling program to address a specific issue. This might include investigations related to water quality issues such as taste/odor complaints, compliance monitoring failure, or Pb/Cu studies. Upon recognition of an initiating event, a special purpose sampling program will be authorized to isolate the source, extent and magnitude of the contamination (Geldreich, 1996). Generally, the underlying assumption in this mode of sampling is that some level of contamination may exist in a specific area or region of the system, but the overall integrity of the system is intact.

Special purpose sampling is also used to meet operational needs which may not explicitly deal with public health issues. Monitoring for this purpose would include operational studies (ex. optimize pumping sequences or configuration), maintenance functions (identify pipeline replacement or flushing programs), or support for capital improvement programs (Kirmeyer et al., 2002).

For their purposes the foregoing methodologies have been adequate. However, when faced with detecting acute, fast-moving incidents, they are unsatisfactory. The feedback mechanism is much too slow (on the order of weeks or months) and in most cases the coverage of the system is severely limited. This paper proposes a change in sampling philosophy is now necessary. Using this paradigm the underlying assumption presumes that contamination exists somewhere in the system and must be detected. Accordingly, deployment of a set of continuous reading monitors within the network would be required. This, of course, assumes that instrumentation exists which can

detect contaminants of interest with sufficient speed and sensitivity to reliably assess the water status. Although such instrumentation for most contaminants does not exist or is prohibitively expensive, significant progress is being made in developing this capability.

This is not to imply that there is a wholesale decline in America's water supply capabilities, but rather it is a recognition of the need for a change in management culture which recognizes the multifaceted nature and the new vulnerability of the distribution system. On this basis, selection of the "best" set of monitoring points would be chosen with a slightly different set of criteria in mind. Water systems are designed and operated with an inherent level of trust that people will not intentionally contaminate a public water system resulting in harm to innocent people. However, once it is acknowledged that someone might be motivated to do such a thing, scenario development can quickly identify the weaknesses in the infrastructure, especially the distribution system, because *there is no routine short term, much less continuous, means of measuring water quality activity throughout the distribution system*.

If adopted, it is acknowledged that this approach will be expensive both in terms of capital and operating costs. Thus, the need to be able to strategically locate a limited number of effective stations is critical. It is further noted that small to mid-sized systems have even more pressure on them to allocate their limited resources in the best manner practicable, due to their limited financial and technical capacity. This is, in fact, the reason why the EPA has staggered implementation schedules for most rules, so that the smaller systems have more time to make the adjustments.

Summary

Utility designers and managers must concede that a new paradigm for network water quality surveillance now exists. Systems that have served well for over 100 years are now insufficient when one considers the possibility of direct and intentional contamination. Currently, water distribution system monitoring is generally categorized as “compliance monitoring” or “special purpose sampling”. Compliance monitoring is a routine process of sampling at various sites through the distribution system to confirm that water at those points meets specified regulatory standards. The time scale on this sampling approach (including analytical turnaround time) is generally on the order of weekly to monthly. Guidance on sampling location is limited. A summary of the major regulations driving compliance monitoring was presented in Table 1.

Special Purpose Sampling is associated with a short-term intense sampling program to address a specific issue. This is typically associated with a violation that occurs during the compliance monitoring or in response to complaints from system customers. This monitoring approach will be more thorough and have a shorter data turnaround time, however, it generally occurs only for a limited duration.

To establish a monitoring scheme that efficiently protects the network from all the potential levels of contamination is neither simple nor inexpensive…and probably, in an absolute sense, impossible. The burden on small to mid-sized water systems is especially burdensome due to their limited staff and budget constraints. A means of identifying and prioritizing a set of monitoring points that addresses this need would be a great help to utilities searching for a way to deal with the potential threat. That is the purpose of the remainder of this study.

Chapter 4 Potential Contamination Agents

Introduction

Because this study is focusing on the possibility of individuals or groups attempting to intentionally contaminate a public water supply, one of the first questions that might be asked is, 'what contaminants might be used to cause mass contamination of a water distribution system?' In other words if the utility is trying to monitor to detect something, what would that be? This is not a trivial matter in that, at first glance, it would seem that thousands of potential contaminants exist. That being the case it is appropriate to establish a framework for examining potential agents and then attempt to evaluate their potential for effective contamination on a public health basis.

Parenthetically, it is noted that while assembling the data gathered for this paper, it became evident that the documented effects of various treatment processes from this perspective are not well developed. In fact, it is not uncommon to find diverging, if not conflicting, data. This is probably not surprising in that (fortunately) these agents are rarely found in water supplies, so few have ever had to be concerned with them. However, given the new world realities, we can expect that future research will accelerate in this area.

In organizing the research and presentation of results, some framework needed to be set. A pattern of analysis has been established for analyzing environmental contamination which has applicability to this problem, i.e. the Risk Assessment process.

The challenge with intentional contamination events is that they do not follow traditional probability distributions and tend to be extreme (low probability) events. However, the assessment of risk can still follow the same general procedure.

The format of a Risk Assessment, generally used for chemical exposures, is broken into four basic steps which are to be rooted in and focused on the scientific aspects of the process. This has been defined (NRC, 1983) to include:

1. Hazard identification, i.e. what agent is present that will cause adverse effects?
2. Toxicity (dose-response) assessment, i.e. what is the relationship between an exposure dose and an adverse health effect in humans?
3. Exposure assessment, i.e. what exposures are currently experienced or likely to occur under different conditions?
4. Risk characterization, i.e. what is the estimated incidence of health impairment to a given population?

Augmenting this process Haas, Rose and Gerba (1999) demonstrate that microbial exposures can be examined in a quantitative or semi-quantitative manner by adapting the NRC methodology. By the nature of infectious microbial agents, the process invariably has data quality and management issues. However, research is proceeding to attempt to enhance traditional epidemiological tools in managing this type of risk.

Routes of Exposure

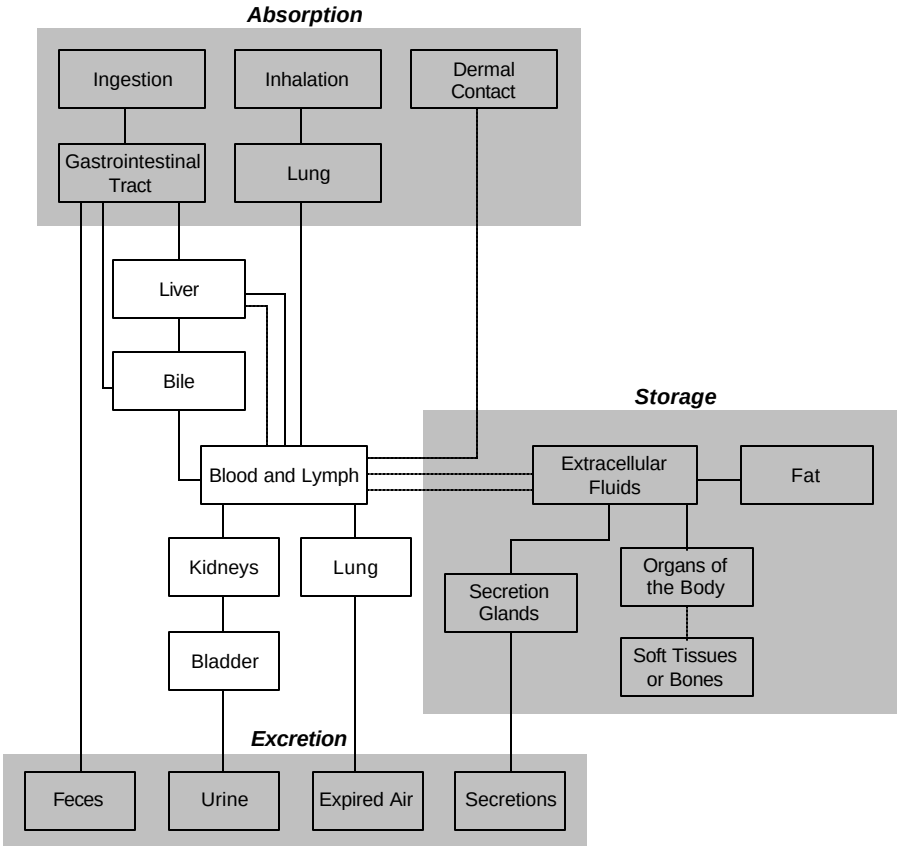
Exposures occur primarily through three routes: inhalation, ingestion and dermal contact. Dermal contact is sometimes further divided to distinguish between percutaneous

absorption and percutaneous (puncture) wounds. The latter is not normally a significant route in environmental assessments and is many times ignored. However, it is a major concern in medical settings where “needle sticks” can transfer microbial agents. It is important to note that chemical toxicity is greatly influenced by the route and method of exposure, so the dose-response relationships used in an analysis must be related to the appropriate route. To illustrate this intuitively, Figure 3 outlines schematically the interrelationships between the major physiological systems.

Because the focus of this study is drinking water, the primary route of interest is *ingestion*. It is possible that inhalation or dermal exposures can occur from activities such as one showering, cooking, irrigating the lawn, but these are deemed to be secondary routes. While these routes exist and should not be completely ignored, their effect should be minor.

Clear definition of routes of exposure and exposure protocols is important because the agent will impact the body in different ways and at different rates depending on the method of absorption. Toxicity tests and development of uptake rates must, therefore, differentiate between exposure mechanisms. This is even more difficult, because the body has defense mechanisms that try to neutralize xenobiotic (foreign) substances before they damage internal organs. Therefore, an applied dose is not the same as an absorbed dose. Even the absorbed dose (which enters the blood stream) is not the same as the “delivered” dose to a target organ. From the delivered dose, the information truly desired is the “biologically effective” dose, which is the quantity of chemical that the organ incorporates into the organ function (Chastain, 1998; Klaasen, 2001; USEPA, 1990).

From the standpoint of planning for a bioterrorism emergency, event planners typically assume that the primary method of agent dissemination for mass casualties will be through the *inhalation* route. It has been shown that aerosol dispersion of appropriate substances in highly populated areas is more effective than other means of application. This, among other reasons, is why many of the substances deemed critical were developed for aerosol dispersion or are focused on adapting the substance for exposure via the inhalation route. Accordingly, most of the toxicological data is listed for inhalation instead of ingestion. As mentioned earlier, the body's reaction to an agent can be greatly different depending on how the material is presented to the victim.



after USEPA (1990)

Figure 3 Key Routes of Chemical Absorption, Distribution, and Excretion in Humans

Because critical doses/exposures to biological agents and chemical agents occur on different scales, a brief note about the measures of morbidity and mortality is appropriate. When examining the toxicity or virulence of a given substance/organism, there are a number of ways in which the effect can be measured or reported (Klaasen, 2001).

LD_{50}: This is an abbreviation for Lethal Dose which results in the death of 50% (median) of the subject population. This is a statistically derived measure of the single dose of substance that can be expected to cause death in 50% of the test animals. Although it is treated as a biological constant, it is not. It is subject to many variables.

LCt_{50}: Because the actual applied dose is a function of both the concentration (C) of the agent and the length of time (t) exposed, many toxicologists are moving to a weighted scale that shows the interaction of the effects. The result is the LCt_{50} which is the Lethal Concentration-time dose in which 50% of the population dies. This measure is primarily seen in inhalation studies, but is also being recognized in other route exposures.

NOAEL: No Observable Adverse Effect Level. This is the highest dosage of a given test which does not result in a measurable or observed adverse effect on the test animals (population).

RfD: The Reference Dose is the dose that estimates the level of a substance that can be safely consumed by the general population. It is calculated by dividing the NOAEL by Uncertainty factors and Modifying factors. It is essentially equivalent to the Average Daily Intake (ADI) factor which has become outmoded. In general, the lower the RfD, the higher the toxicity.

Infectious dose: The dose at which an infectious agent (microorganism) enters the body and begins to multiply. It may or may not result in a clinical manifestation of the disease. The period between the exposure and the first shedding or excretion of the agent, i.e., manifestation is called the Latent period. The period of time which may exist between the time the infectious agent enters the body and the time it begins to multiply is called the Lag period.

Other measures of toxicity exist; however, these are among the most common and can provide at least a qualitative means of comparing risks between various agents. Also, having some quantification of the toxicity values allows one to begin to estimate that

amount of a specific agent that must be present to contaminate a water system to produce a lethal effect.

Assumptions Regarding System Contamination

There are many contaminants that can pollute a water supply. Of specific interest to this study are those contaminants that are not readily detectable by the consumer and which can cause some degree of morbidity or mortality upon exposure. It is also assumed that the agent is capable of mass contamination, otherwise known as "*weapons of mass destruction*". In other words, the intent would be to corrupt a large volume of water as opposed to service to a few houses or a small isolated subdivision.

Focus is also limited to those agents that cause a major *acute* response contra a *chronic* response. In other words, carcinogenic agents are not considered in the screening process, because the time of action is too long. This is not to say that they might not have serious consequences, but they don't typically fit within a terrorist agenda and thus present poor agent choices from that perspective.

For a contaminant to effectively harm a population by contaminating the water, it would need, at a minimum, to have the following qualities:

1. Highly soluble or finely divided and suspended in water
2. Colorless
3. Odorless
4. Tasteless
5. Highly toxic

In other words, to achieve the goals of the offending entity, the water would need to be contaminated in such a way as to be *undetectable* when consumed, so that exposure to the contaminant would be unrestrained. Given these criteria, one is able to then begin to sort through a list of potential candidates to identify possible contaminants for use in situations as posed here. Additional pragmatic concerns also come into play, such as the ease of acquiring and the cost of the material as well as the ability to transport and disperse the contaminant.

The contaminant can be drawn from several different categories: Chemical, Microbial or Physical (Radiation). A number of studies have sought to screen various databases to select likely candidates that fit the specified profile. As an example, Burrows and Renner (1999) did a careful study of biological agents and their potential as threats to potable water. That work examined 27 organisms or biotoxins that are likely to be used as contamination agents. While the work was prepared by the U.S. Army with a focus on field water treatment units, it has applicability to this study.

Assuming chemicals formed the primary threat, the University of Michigan's Studies in Urban Security Group (1997) prepared a report for the U.S. Environmental Protection Agency that examined 79,000 potential contaminants taken from the Registry of Toxic Effects of Chemical Substances. They identified 35 chemicals that met the criteria of being tasteless, odorless, colorless and had an LD_{50} of 10 mg/kg. Another study prepared at the University of Michigan (Deininger, 2000) conceptually widened the scope to consider biological agents and "designer drugs" as a threat, although no specific analysis was provided. The authors did note, however, that the use of "designer

drugs" could be particularly difficult to identify because of the ability of a terrorist to develop an agent that does not exist in any tracking database.

An interesting feature of each of these reports is that they all assume contamination ahead of the treatment plant, thus allowing the treatment process to reduce the efficacy of the agent. While this is valid if the contamination of the supply occurs, it leaves many questions unanswered if the application occurred downstream. Although that perspective will be changing as utilities begin to analyze their systems from a true vulnerability perspective, it demonstrates the potential blind spot that has existed in the past.

Another scenario that does not seem to be considered is the mixture of contaminants. Should several different contaminants be combined into one injection incidence, it could cause a more serious response in the population. The individual immune system could be challenged on several fronts resulting in a weakening of the body's ability to resist. For example, if *Cryptosporidium* and *Giardia* were mixed, the resulting incident could cause an epidemic of severe gastric distress that could result in additional secondary infections.

Classes of Contamination Agents

Given the preceding discussion, a number of military, academic and public health institutions have developed lists of agents to be considered as "weapons of mass destruction". There are a number of different ways to classify them but generally, they can be organized as follows. (More expansive lists also include radioactive materials, but those are not considered sufficiently viable to include in this analysis.)

1. Biological Agents

    - Bacteria/Rickettsiae
    - Protozoa
    - Toxins
    - Viruses

2. Chemical Agents

    - Warfare Agents (Casualty agents)
    - Nerve Agents
    - Blood Agents
    - Choking Agents
    - Vesicants (Blister Agents)
    - Toxic Industrial Chemicals

Biological Agents

In many ways, the biological agents possess the most potential for system contamination. Conceptually, any of the waterborne pathogens could be used as an agent. The difference between the pathogens would be their virulence, ability to survive in the distribution system environment and the ultimate effect. Many waterborne diseases cause gastric illnesses, but have no long term effect, while others can result in protracted illness or death. The key to selection lies in the intent of the perpetrator (i.e. morbidity/mortality, disruption, psychological trauma, etc.), ease of access to the materials and ability to deliver the agent. While the issue is new to this generation of water system designers and managers, it is certainly not the first time that it has been

considered. Berger and Stevenson (1955), in fact, presented an analysis of biological contamination following the initiation of the Cold War. Although the equipment and technology have changed significantly, that article still contains pertinent information.

The Center for Disease Control and Prevention (CDC) has attempted to prioritize different biological agents by classifying them as Class A, B, or C. The class codes relate to their virulence and likelihood of use with Class A being the most damaging. The listing for the bio-agents is as follows (CDC, 2000; Kahn, Morse, & Lillibridge, 2000; Rotz, Kahn, Lillibridge, Ostroff, & Hughes, 2002). It should be noted that most of the agents are established using the inhalational exposure route because that is deemed to present the most vulnerable scenario.

Category A.

High-priority agents include organisms that pose a risk to national security because they:

- can be easily disseminated or transmitted from person to person;
- cause high mortality, and have the potential for major public health impact;
- might cause public panic and social disruption; and
- require special action for public health preparedness.

Agents currently classified as *Category A Disease/Agents*

- Anthrax (*Bacillus anthracis*)
- Botulism (*Clostridium botulinum* toxin)
- Plague (*Yersinia pestis*)
- Smallpox (*Variola major*)
- Tularemia (*Francisella tularensis*)

- Viral hemorrhagic fevers (Arenaviruses, Filoviruses, Bunyaviruses, and Flaviviruses)

Category B.

Second highest priority agents include those that:

- are moderately easy to disseminate;
- cause moderate morbidity and low mortality; and
- require specific enhancements of CDC's diagnostic capacity and enhanced disease surveillance.

Agents currently classified as *Category B Disease/Agents*

- Brucellosis (*Brucella* species)
- Epsilon toxin of *Clostridium perfringens*
- Food safety threats (*Salmonella, E. coli O157:H7, Shigella*)
- Glanders (*Burkholderia mallei*)
- Melioidosis (*Burkholderia pseudomallei*)
- Psittacosis (*Chlamydia psittaci*)
- Q fever (*Coxiella burnetti*)
- Ricin toxin from Ricinus communis (castor beans)
- Staphylococcus enterotoxin B
- Typhus fever (*Rickettsia prowazekii*)
- Viral encephalitis
- Water Safety Threats (*Cryptosporidium parvum, Vibrio Cholerae*)

Category C.

Third highest priority agents include emerging pathogens that could be engineered for mass dissemination in the future because of:

- availability;
- ease of production and dissemination; and
- potential for high morbidity and mortality and major health impact.

Agents currently classified as *Category C Disease/Agents*

- Hantaviruses
- Multidrug-resistant tuberculosis
- Nipah virus
- Tickborne encephalitis viruses
- Tickborne hemorrhagic fever viruses
- Yellow fever

In this context Biological agents can be conveniently classified into four different categories: Bacteria/Rickettsiae, Protozoa, Toxins and Viruses. For the purpose of this analysis, an attempt to summarize the findings of numerous reports has been listed in Table 4, Table 5 and Table 6. Because of the differences in the modes of operation of the different organisms (as well as information available), each table is arranged a little differently. The primary intent of the tables is to list in a compact fashion the nature of the threat, the number of organisms necessary to cause an infection and whether or not the organism is sensitive to inactivation by chlorine. Obviously, the fewer organisms required to infect, i.e., the more virulent the agent, the more effective it could be in meeting a terrorist objective.

From a public health perspective, it is important to know if these agents are replicating (able to reproduce themselves) and what their environmental (extra-host) survivability is likely to be. The danger of certain pathogenic organisms can be magnified because, given appropriate growing conditions, they can continue to increase in concentration after the initial contamination event. Other substances such as toxins are derivative substances, i.e. poisonous substances of plant or animal origin. A toxin is dependent on a predecessor organism (plant or animal) producing it; therefore, it is not able to replicate itself. This is important in that the concentration of the substance once in place will not increase as is possible with replicating agents. Accordingly, an attempt has been made to segregate the agents in terms of their ability to reproduce.

Table 4 Replicating Agents

Disease	Agent	Agent Type	Weaponized	Water threat	Infective dose[a]	Stable in water	Chlorine tolerance[b]
Anthrax	*Bacillus anthracis*	Bacteria	Yes	Yes	6,000 spores (inh)	2 years (spores)	Spores resistant
Brucellosis	*Brucella melitensis*	Bacteria	Yes	Probable	10,000 organisms (uns)	20-72 days	Unknown
Cholera	*Vibrio cholerae*	Bacteria	Unknown	Yes	1,000 organisms (ing)	Survives well	Easily killed
Gas Gangrene	*Clostridium perfringens*	Bacteria	Probable	Probable	10^8 organisms (ing)	Common in sewage	Resistant
Glanders	*Burkholderia mallei*	Bacteria	Probable	Unlikely	3.2×10^6 organisms (uns)	Up to 30 days	Unknown
Melioidosis	*Pseudomonas pseudomallei*	Bacteria	Possible	Unlikely	Unknown	Unknown	Unknown
Plague (Bubonic)	*Yersinia pestis*	Bacteria	Probable	Yes	500 organisms (inh)	16 days	Unknown
Psittacosis	*Clamydia psitacci*	Parasite	Possible	Possible	Unknown	18-24 hr, seawater	Unknown
Q fever	*Coxiella burnetti*	Rickettsia	Yes	Possible	25 organisms (uns)	Unknown	Unknown
Salmonellosis	*Salmonella sp.*	Bacteria	Unknown	Yes	10^4 organisms (ing)	8 days, fresh water	Inactivated
Shigellosis	*Shigella sp.*	Bacteria	Unknown	Yes	10^4 organisms (ing)	2-3 days	Inactivated, 0.05 ppm, 10 min
Tularemia	*Francisella tularensis*	Bacteria	Yes	Yes	10^8 organisms (ing)	Up to 90 days	Inactivated, 1 ppm, 5 min
Typhus	*Rickettsia prowazekii*	Rickettsia	Probable	Unlikely	10 organisms (uns)	Unknown	Unknown
Encephalomyelitis	*VEE*	Virus	Probable	Unlikely	25 particles (aer)	Unknown	Unknown
Enteric viruses		Virus	Unknown	Yes	6 particles (ing)	8-32 days	Readily inactivated (rotavirus)
Hemorrhagic fever	VHF, Rotavirus, Norwalk & Ebola virus	Virus	Probable	Unlikely	10^5 particles (ing)	Unknown	Unknown
Smallpox	*Variola major*	Virus	Possible	Possible	10 particles (uns)	Unknown	Unknown
Cryptosporidiosis	*Cryptosporidum parvum*	Protozoa	Unknown	Yes	132 oocysts (ing)	Stable days or more	Resistant

Abbreviations: aer-aerosol; ing-ingestion; inh-inhalation; uns-unspecified.
[a] Total infective dose used to calculate water values. [b] Ambient temperature, ≤ 1 ppm free available chlorine, 30 min or as indicated.

Source: Burrows and Renner, 1999

Toxins have been proven to be potent poisons in a number of settings. The range of plant and animal based toxins is very large and the use of these substances to contaminate water supplies is somewhat unconventional. To provide an indication of the relative potency of the biotoxins, the following table is provided. (A few warfare agents

are also listed below because they are, in essence, toxins. Note that the data is based on the effects on laboratory mice.)

Table 5 Comparative Lethality of Selected Biotoxins and Chemical Agents in Laboratory Mice

AGENT	LD50 (µ G/KG)	MOLECULAR WEIGHT	SOURCE
Botulinum toxin	0.001	150,000	Bacterium
Shiga toxin	0.002	55,000	Bacterium
Tetanus toxin	0.002	150,000	Bacterium
Abrin	0.04	65,000	Plant (Rosary Pea)
Diphtheria toxin	0.10	62,000	Bacterium
Mainotoxin	0.10	3,400	Marine Dinoflagellate
Palytoxin	0.15	2,700	Maine Soft Coral
Ciguatoxin	0.40	1,000	Marine Dinoflagellate
Textilotoxin	0.60	80,000	Elapid Snake
C. perfringens toxins	0.1 – 5.0	35-40,000	Bacterium
Batrachotoxin	2.0	539	Arrow-Poison Frog
Ricin	3.0	64,000	Plant (Castor Bean)
alpha-Conotoxin	5.0	1,500	Cone Snail
Taipoxin	5.0	46,000	Elapid Snake
Tetrodotoxin	8.0	319	Puffer Fish
alpha-Tityustoxin	9.0	8,000	Scorpion
Saxitoxin	10.0 (Inhal 2.0)	299	Marine Dinoflagellate
VX	15.0	267	Chemical Agent
SEB (Rhesus/Aerosol)	27.0 (ED50~pg)	28,494	Bacterium
Anatoxin-A(s)	50.0	500	Blue-Green Algae
Microcystin	50.0	994	Blue-Green Algae
Soman (GD)	64.0	182	Chemical Agent
Sarin (GB)	100.0	140	Chemical Agent
Aconitine	100.0	647	Plant (Monkshood)
T-2 Toxin	1,210.0	466	Fungal Mycotoxin

From: Medical Management of Biological Casualties, U.S. AMRIID, July 1998, Appendix C

Table 5 clearly shows that the Botulinum toxin is a thousand times more effective than Ricin, and many of the other agents. However, Botulinum will be deactivated by chlorine if a water system has reasonable CT, whereas many of the others are insensitive to chlorine.

Other routes of exposure are typically more efficient methods of delivery, but some can be adapted to a water based delivery method. Table 6 was developed by Burrows and Renner (1999) to show the biotoxins that are most likely to be used in a water contamination event. (Notice that, in general, the toxins are relatively resistant to chlorine at the levels commonly found in potable water systems.)

Table 6 Selected Biotoxins

Biotoxin	**Weaponized**	**Water threat**	**NOAEL, 2 L/day[a]**	**Stable in water**	**Chlorine tolerance [b]**
Aflatoxin	Yes	Yes	75 µg/L	Probably stable	Probably tolerant
Anatoxin A	Unknown	Probable	Unknown	Inactivated in days	Probably tolerant
Botulinum toxins	Yes	Yes	0.0004 µg/L	Stable	Inactivated, 6 ppm, 20 min
Microcystins	Possible	Yes	1.0 µg/L[c]	Probably stable	Resistant at 100 ppm
Ricin	Yes	Yes	15 µg/L	Stable	Resistant at 10 ppm
Saxitoxin	Possible	Yes	0.4 µg/L	Stable	Resistant at 10 ppm
Staphylococcal enterotoxins	Probable	Yes	0.1 µg/L	Probably stable	Unknown
T-2 mycotoxin	Probable	Yes	65 µg/L[d]	Stable	Resistant
Tetrodotoxin	Possible	Yes	1 µg/L	Probably stable	Inactivated, 50 ppm

NOAEL, no-observed-adverse-effect level.

[a]Estimated as 7.5 times the NOAEL calculated for consumption of 15 L/day. [b]Ambient temperature, < 1 ppm free available chlorine, 30 min or as indicated. [c]World Health Organization drinking water standard. [d]Derived from short-term U.S. Department of Defense Tri-Service standard (77).

Source: Burrows and Renner, 1999

It should be noted that a number of these agents have not been thoroughly evaluated in terms of removal or disinfection efficacy, because they are not of routine concern to water utilities. Therefore, they have not had the level of research associated with them that the more common indicator species have had. In light of that fact, it is not surprising that there is disagreement in the literature about the characteristics of the organisms (infective dose, survivability and efficacious chlorine dose). This is no doubt due to the different conditions under which the studies in the literature were performed.

Variables Affecting Disinfection

Disinfection processes are typically viewed as being the primary treatment barrier for reducing or eliminating the threat of biological pathogens. Chlorine (HOCl and OCl^- species) has been the primary agent of disinfection for roughly 100 years. However, in recent years alternative disinfectants/oxidants are being used in specific situations. Examples would include ozone, ultraviolet light, chlorine dioxide and gamma radiation.

The effectiveness of the disinfection or neutralization process is governed by more variables than merely the mass of disinfectant/oxidant applied. One thing noted while reviewing the literature was the fact that many of the key variables were not listed in the articles. Therefore, it is possible that the dosages or effectiveness ratings of disinfectant (chlorine specifically) may not be exactly as listed. Generally speaking, in order to fairly estimate the efficacy of a disinfectant, the temperature, pH, contact time (assuming adequate mixing), dosage, and presence of interfering chemicals/turbidity must be known. Then, given the initial concentrations of the microorganisms, a reasonable estimate of the disinfection process can be made.

Portions of the Surface Water Treatment Rule (SWTR) were promulgated by the USEPA to account for variations in temperature, pH, dosage and contact time for various disinfectants when designing disinfection systems. The SWTR established what are known as the CT tables for a given organism and a given disinfectant at specific pH and temperatures. CT is shorthand notation for (disinfectant) Concentration (C) times Contact time (T). The concentration is measured in mg/L and the time is measured in minutes.

It is reasonable to assume that the higher the disinfectant concentration, the higher the rate of destruction. Again, the actual CT values depend on the microorganism, the pH of the water, the temperature, and the turbidity of the water. The turbidity is a measurement of the amount of material suspended in the water that influences the transmission of light. While turbidity itself is not a significant variable, studies have shown that bacteria attach themselves to the surface of particles. If the material is organic, bacteria may hide inside the particle and may be protected from the disinfectant (White, 1999).

Generally speaking, chemical reactions proceed faster at higher temperatures. Chlorine reacting with an organic compound will be faster at higher temperatures. Thus, the inactivation (destruction) of viruses and protozoa is slower at lower temperatures than at higher ones. For example, tables in the SWTR show that the required CT times for inactivation at low temperatures (5° Celsius) are about 4 times higher than at 25° Celsius. This phenomenon holds for all disinfectants, except for UV radiation. Thus, the higher the CT value, the more disinfectant needed and/or the longer the detention time required. This is clearly shown in Table 7 and Table 8 (AWWA, 1991).

The guidance manual does not list any CT values for bacteria since they are so much lower than the CT values for viruses and protozoa. Since chlorine is the most prevalent disinfectant used in the water industry (80%), a comparison is instructive. At a temperature of 15° C and an inactivation by 3 logs (99.9%), the CT value for a representative protozoa is 75 and for viruses 3 (see the Tables 7 and 8). By contrast, the CT value for coliforms is about 0.5. Deininger (2000) notes that the CT values for protozoa are an order of magnitude higher than the CT values for viruses, and the CT for

viruses are an order of magnitude higher than those for bacteria. One exception was noted. Spores of *Bacillus anthracis* are very resistant to chlorine and will not be inactivated by chlorine at the normal concentrations and contact times used in water treatment plants. While most attention is focused on anthrax as an inhalational agent, its use through the ingestion route can likewise be a potent agent (Berger & Stevenson, 1955; Inglesby et al., 1999). Other microorganisms are resistant to chlorine (ex. *giardia* and *cryptosporidium*), but none has presented themselves to date that have the virulence of *B. anthracis*.

Table 7 CT Tables for Inactivation of Giardia Cysts (at pH 7)

Disinfectant		**Temperature ºC**		
	Inactivation*	5	15	25
	1	.63	.32	.16
	2	1.3	.63	.32
	3	1.9	.95	.48
	1	50	25	12
Chlorine	2	99	50	25
	3	169	75	37
	1	8.7	6.3	3.7
Chlorine Dioxide	2	17	13	7.3
	3	26	19	11
	1	735	500	250
Chloramine	2	1,470	1,000	500
	3	2,200	1,500	750

*Inactivation is in log units. A 2 means a 100 fold reduction, or a 99% inactivation.

The numbers are CT values (mg/L * minutes)

Source: AWWA, 1991

Table 8 CT Tables for Inactivation of Viruses (pH 6-9)

Disinfectant	Temperature ºC			
	Inactivation*	5	15	25
Ozone	2	.6	.3	.15
	3	.9	.5	.25
	4	1.2	.6	.3
Chlorine	2	4	2	1
	3	6	3	1
	4	8	4	2
Chlorine Dioxide	2	5.6	2.8	1.4
	3	17.1	8.6	4.3
	4	33	16.7	8.4
Chloramine	2	857	428	214
	3	1,423	712	356
	4	1,988	994	497
UV Irradiation	2	21	21	21
	3	36	36	36

The numbers are CT values – (mg/L * minutes)
UV Irradiation: Numbers are in mWsec/cm
Source: AWWA, 1991

Chemical Agents

Advances in chemistry have driven many of the technological, economic and social changes in developed society. While much of this advancement has been for the good, the large number of new chemicals with dangerous characteristics has caused much environmental damage and now threatens to be used in more sinister ways. Although chemical warfare has been used at various times throughout history, it was first utilized on a large scale during World War I. Although banned by various conventions, research and development within the military has continued and development of chemicals specifically as warfare agents has occurred.

As mentioned, these agents are classified by their physiological effect. Chemicals which were developed to have high mortality (death) levels are called *Casualty Agents*. They

are sub-classified as Blood Agents, Choking Agents, Nerve Agents and Vesicants (Munro et al., 1999; Reutter, 1999). Other agents have been developed whose purpose is not necessarily to cause death, but rather to incapacitate the target in some manner. These are called *Harassing Agents*. Among this classification are the Lachrymators (Tear agents), Sternutators (Vomiting agents), Depressants, Psychedelic agents, and Stimulants.

With regard to the specific warfare agents, they have been developed almost exclusively as inhalational or dermal agents. The toxicity of ingesting the material is less well documented but, in most cases, the agents are impractical because of solubility limitations. If there were to be a threat, it would most likely be caused by volatilization of the agent out of the water and inhaled.

Most of the critical agents from an ingestion standpoint are likely to be some form of Organophosphate (Nerve agents), Arsenic (Vesicant) or Cyanide (Blood). The organophosphates (OP) are related to the OP pesticides but have much higher mammalian acute toxicity (particularly through the percutaneous route). Therefore, some toxicity extrapolation may be inferred from that. These agents attack the central nervous system in one way or another and generally are related to the interruption of acetylcholinesterase activity. The OP agents are generally divided into the G-group and V-group. The G agents are more volatile than the V agents, but all exist as colorless and, for the most part, odorless liquids.

A brief description of the primary unclassified warfare agents listed in the public domain follows in Table 9 (Marrs, Maynard, & Sidell, 1996). A more detailed discussion of the

toxicology and fate is provided in Munro et al. (1999), which also includes some data on environmental fate issues. It is also interesting to note that much of the development of chemical warfare agents centered on causing maximum casualties for the least sensitive male soldier. When considering the general population, lethal dosages may be need to be adjusted given the fact that women, children, elderly and other sub-populations can be affected. Apart from the acute event there is also the issue of the potential for longer term low dose effects (Reutter, 1999).

Other agents that fall into the narcotic or drug classification probably should be considered too, but it is difficult at this point to screen them. For example, certain hormone, heart stimulants/depressants, or psychedelic drugs could effect different populations in different ways. These agent classes would not be likely to show up on most of the search routines reflected in these results. Accordingly, additional research may be required to cover pharmaceutical class substances. Many of these chemicals are sensitive to the disinfectants in common use, so if they were added to the water upstream of treatment, their effectiveness is assumed to be significantly reduced.

Table 9 Chemical Warfare Agents

Agent Type	Agent	Solubility	Water threat	RfD (mg/L)	LD_{50} (µG/KG)	Molecular Weight	Stable in water	Chlorine tolerance[b]	CAS No.
Blood Agents	Arsine	Sparingly sol.	Unlikely			77.9			7784-42-1
	Cyanogen Chloride (CK)	Soluble	Possible	750		61.5			506-77-4
	Hydrogen Cyanide (AC)	Miscible	Yes	750		27.0			74-90-8
Choking Agents	*Chlorine*		No			70.9	yes		7782-50-5
	Diphosgene								
	PFIB								
	Phosgene	Reactive	Unlikely			98.9			75-44-5
Vesciants	*Sulfur mustard (HT)*	Insoluable	Unlikely	140				degrades	6392-89-8
	Distilled mustard (HD)	Sparingly sol.	Unlikely	7				degrades	505-60-2
	Ethyldichoroarsine			80					
	Lewisite 1			80					
	Lewisite 2								
	Lewisite 3								
	Methyldichloroarsine								
	Mustard-Lewisite mixture (HL)	Insoluable	Unlikely			186.4			
	Mustard-T mixture								
	Nitrogen Mustard 1 (HN-1)	Sparingly sol.	Unlikely			170.1			538-07-8
	Nitrogen Mustard 2 (HN-2)	Sparingly sol.	Unlikely			156.1			51-75-2
	Nitrogen Mustard 3 (HN-3)	Insoluable	No			204.5			555-77-1
	Phenyldichloroarsine								
	Phosgene oxime (CX)	Insoluable	No			113.9			1794-86-1
	Sesqui mustard								
Nerve Agents	*GA (Tabun)*	Miscible	Unlikely	70		162.3			77-81-6
	GB (Sarin)	Miscible	Unlikely	13.8	100	140.1			107-44-8
	GD (Soman)	Slightly sol.	Unlikely	6	64	182.2			96-64-0
	GE								
	GF					180.1			
	VE								
	VG								
	VM								
	VX	Moderate	Unlikely	7.5	15	267.0			50782-69-9

[a] Total infective dose used to calculate water values. [b] Ambient temperature, < 1 ppm free available chlorine, 30 min or as indicated.
Source: Marrs, Maynard, & Sidell (1997)

In addition to warfare agents, Toxic Industrial Chemicals (TICs) could represent a threat for use in a contamination event. Again, if one assumes that an aesthetic change (taste, odor, color) will deter individuals from consuming the product, the number of toxic chemicals can be significantly reduced. Deininger (2000) reports that the EPA provided a list of chemicals for analysis that was produced by a screening on some unknown basis. The list included methyl mercaptan, arsine, dimethyl sulfate, acrolein, toluene 2, 4-diisocyanate, bromine, fluorine, methyl hydrazine, phosgene, stibine, chlorine, nickel carbonyl, hydrogen sulfide, hydrogen chloride, hydrogen fluoride, triethylamine,

gluteraldehyde, nitric acid vapors, furan, methylethyl ketone, methylvinylketone, osmium tetroxide aerosol and chromic acid aerosol.

Although most of the chemicals are certainly toxic, a review of this list reveals that none of these chemicals presents a credible threat in this context. For example, chlorine is listed as one of the TICs. Chlorine is without a doubt toxic and must be monitored for an aerosol release. In fact, chlorine is listed in the Clean Air Act Amendments as a chemical to be modeled in Risk Management Plans (RMP). However, in order to be toxic from an ingestion standpoint, the taste and odor would be so pungent that the palatability of the water would render it undrinkable by most people. Therefore, it is not considered a credible threat, because it could be too easily detected. Many of the other chemicals on the list have a similar weakness from the standpoint of a lethal contaminant via the ingestion route.

Additional research is necessary to more fully develop a viable list of candidate agents in the TIC category. At this point, it would appear that the primary TICs would be related to some of the warfare agents, such as cyanides, orthophosphate pesticides and arsenicals. Again, this issue of viability of the toxic agents via the ingestion route (as opposed to inhalational or dermal) while being undetectable by the senses is the major problem.

Survival and Deactivation Rates of Agents

If a contamination event were to occur, a question of critical importance relates to the persistence of the agent. With regard to bacteria, few have been studied in detail (in this context), except the die-off of coliform organisms. The EPA Qual II model has a

provision for the modeling of these organisms which occur in sewage effluents and surface runoff. They are normally believed to follow the common exponential decay form:

$$y(t)=y(0)*e^{(-k*t)} \qquad \text{(Equation 1)}$$

The quoted rates for decay constant, k, vary between 0.5 and 4.0 (Deininger, 2000; USEPA, 1985).

A textbook on water microbiology (McFeters, 1990) discusses a number of infectious microorganisms. In the absence of more detailed studies, it may be prudent to assume a worst case scenario, i.e. neglect the die-off and assume that there would be no decrease in the agent concentration. As an indication of relative survivability, however, McFeters quotes *half-lives* of 7 to 16 hours for *Vibrio cholerae*, *Shigella* spp., Enterotoxigenic *E.coli,* and *Salmonella*. *Campylobacter jejuni* has a half-life of 3 days, and *Yersinia enterocolitica* has a half-life of more than a year. *B. anthracis* has a half-life of several years. Viruses are reported to survive well in river settings; they die-off faster at warm temperatures than at cold ones. Survival times for Echo 7 virus, Echo 12 virus, Cocksackie A9 virus and Poliomyelitis virus at 4° C are: 99.9% are dead after 20 days, 99% are dead after 14 days, and 90% are dead after 7 days (Deininger, 2000).

Information on this aspect of toxins was difficult to find, but they are probably very stable in water. Burrows and Renner (1999) speculate on the most likely biotoxins to be used for water contamination and conclude that for the most part they seem to be stable in water and more or less resistant to chlorine levels typically found in water systems.

Therefore, the default assumption would be that they would remain potent in whatever concentration they were applied.

Microbiological organisms require nominal food sources and nutrients to survive. Therefore, many biological agents may not survive as well as the literature values indicate, because it is noted that the data are taken from stream or wastewater effluent sources. However, at specific stages of the life cycle (ex. Cysts), certain organisms are well protected from environmental factors and remain dormant. This fact makes those organisms attractive for use in surreptitious attacks.

The Toxic Industrial Chemicals will most likely be persistent and will leave the water based on their partial pressures or fugacity. In turbulent situations (ex. pouring or shaking), there will be a higher transfer to the atmosphere than in tranquil situations (storage). Obviously, there would need to be exposure to the atmosphere for this to occur.

General Discussion of the Effect of Treatment Processes on Contamination

The 1990 edition of *Water Quality and Treatment* (4th ed.) provides a table that shows the effectiveness of treatment processes in terms of poor, fair, good, and excellent. This table was developed by summarizing forty-five research reports in existence at that time and is reproduced as Table 10 below (AWWA, 1990). Taking midpoints of the sometimes wide ranges, leads to Table 11 below with numeric fractions. This table should be used with extreme caution. The data need to be verified in a pilot treatment plant. It is interesting to note that the 5th edition of *Water Quality and Treatment* (1999)

has dropped this estimate of efficiency for each process, and has replaced it by only a table with an “X” when the process is deemed to be effective (AWWA, 1999).

Table 10 Effectiveness of Processes for Contaminant Removal

	Bacteria	Viruses	Protozoa	VOC	SOC	TOC	Taste/Odor
Aeration, Air Stripping	P	P	P	G-E	P-F	F	F-E
Coagulation Sediment/Filtration	G-E	G-E	G-E	G-E	P	P-G	P-G
Lime Softening	G-E	G-E	G-E	P-F	P-F	G	P-F
Ion Exchange	P	P	P	P	P	G-E	-
Reverse Osmosis	E	E	E	F-E	F-E	G	-
Ultra Filtration	E	E	E	F-E	F-E	G	-
Disinfection	E	E	E	P-G	P-G	G-E	P-E
Granular Activ. Carbon	F	F	F	F-E	F-E	F	G-E
Powdered Activ. Carb.	P	P	P	P-G	P-E	F-G	G-E
UV Irradiation	E	E	E	G	G	G	G

P–Poor (0-20% removal); F–Fair (20-60%); G–Good (60-90%); E–Excellent (90-100%); NA –Insufficient Data

Source: Water Quality & Treatment, 4th edition, 1990

If one takes the midpoints of the categories of the above table, it is possible to generate the following table. Obviously, each source water is different, so actual removals will show a significant variance, but this table provides a relative indication of the results.

Table 11 Percentage Removal of Contaminants

	Bacteria	Viruses	Protozoa	VOC	SOC	TOC	Taste/Odor
Aeration, Air Stripping	10	10	10	85	25	40	70
Coagulation Sedimentation	85	85	85	10	40	25	25
Lime Softening	85	85	85	25	25	75	25
Ion Exchange	10	10	10	10	10	85	-
Reverse Osmosis	95	95	95	70	70	75	-
Ultra Filtration	95	95	95	40	40	85	-
Disinfection	95	95	95	40	40	85	22
Granular Activ. Carbon	40	40	40	70	70	40	85
Powdered Activ. Carb.	10	10	10	40	50	55	75
UV Irradiation	95	95	95	75	75	75	75

Adapted from: Water Quality & Treatment, 4th edition, 1990

Summary

In the post-September 11 world, all utilities are examining their vulnerability to attack. This especially includes bio-chemical contamination, although other facets of operation are being considered too (cyber/control, disabling critical structures, critical personnel backup, etc). An enormous number of microbiological and chemical substances exist which could cause harm if consumed. However, the list of probable candidates can be screened by looking at the virulence or toxicity of a substance and its ability to be undetected (ex. no taste, odor, color). This reduces the number of viable candidates to a more manageable number.

With regard to bacterial agents, *Shigella*, *Salmonella*, *Vibrio cholerae* are viable threats; however, a reasonable chlorine residual will destroy them. The greatest threat seems to be the spores of *B. anthracis*. They are not affected by the normal disinfectant concentrations used in water treatment. For the same reason, *Cryptosporidium* or *Giardia* are also worrisome due to its low LD_{50} and its resistance to disinfectants.

Further examination of toxins is probably in order. Tables 2 and 3 are helpful as guides. It appears that *Botulinum*, Aflatoxin and Ricin are the most likely to have been weaponized.

Chemical agents need additional examination for viable candidates. The current thinking in many venues is that some form of cyanide, arsenic or orthophosphates holds the most potential as a contamination agent. The issue of detectability for some of these agents could reduce their utility, however, these certainly can have a significant effect and should be placed on a priority list. It seems highly likely that other chemicals

(including pharmaceuticals or TIC) will expand the list, so periodic review of potential candidates is warranted.

It should be noted that no judgment has been made as to the logistics of feeding an adequate dose of toxicant into the water. Further screening could be accomplished by looking at the quantity of the agent required to result in a toxic dosage in the water. As additional LD_{50} or other dose-response data becomes available on an ingestion basis, this computation may be helpful.

Chapter 5 Computer Simulation of Distribution System Water Quality

Because it is impracticable to physically test a wide range of contamination events, it becomes immediately apparent that any reasonable approach seeking to address the problem must rely upon a workable extended period simulation of the distribution network in question. Although there are certainly many factors, assumptions and sources of errors in developing an extended period water quality model (see Appendix 2), with time and effort the network simulations can be refined and calibrated to yield a reasonably accurate representation of reality. With this in place it is possible to investigate various system events and observe the results.

Because the computer simulations generate the data to be analyzed in this study, it may be useful to survey the basic features of the mathematical models that replicate the physical system. The modeling of a water distribution system is difficult because it is governed by complex, nonlinear, non-convex and discontinuous hydraulic and water quality equations. To realistically simulate the operation of a water system one must consider the effects of pipes, valves, pumps (and their controls), storage tanks, spatial and temporal variations in water demand, and variations in water quality. Fortunately, a number of computer packages have been developed to handle the technical mathematical aspects of network modeling. A few of the network solvers available include the EPA's freeware EPANET, Haestad Methods WaterCAD®/WaterGEMS®, University of Kentucky's KYPIPE, MWH's H_2ONET®, and Advantica's Stoner SynerGEE® Water to name a few. These packages are more affordable and accessible now than in

the past and for many reasons every community water system should have a working simulation of their distribution network.

Hydraulic Analysis

The first set of physical relationships that must be understood and converted to mathematical constructs relate to basic fluid flow. Pipe flow and especially networked pipe flow, is complex because of its nonlinearity which arises primarily because of pipe friction. In addition, the effects of water storage facilities and pump systems create a discontinuous set of boundary conditions that must be managed. The primary components of a distribution system network are links which represent pipes, pumps, valves and nodes which correspond to junctions, tanks and reservoirs. The pipes and junction nodes are connected in loops which provide a means of distributing water efficiently with reduced energy requirements and improved operational redundancy. This section will discuss the fundamental fluid relationships that form the foundation of network simulation.

Characteristics of Water

Water from a hydraulic standpoint has a number of features that must be considered in any simulation effort. For example, it is essentially, although not absolutely, incompressible. It has a specific weight (γ) of 62.4 pounds per cubic foot. The specific gravity is usually defined as 1.000 at 60^{o}F, however, some references list it as 1.000 at 39.2^{o}F, which is the point of maximum water density. The kinematic viscosity (ν) is 0.00001216 f^{2}/sec at 60^{o}F. Like all materials its viscosity varies with temperature and so may be a factor in computing flow characteristics. It is considered the “universal acid”

and as such can be reactive to piping and other material in the system which over time can change the flow characteristics of the conduits (Camp & Meserve, 1974; Westaway & Loomis, 1977).

Basic Fluid Equations

The hydraulic analysis of pipe systems rest primarily upon two basic laws of physics: the law of conservation of mass and the law of conservation of energy. Thus, when examining any control volume of water or any change of state, equations must be written which conserve mass and energy. This is the starting point for network models.

The *law of conservation of mass* is understood to mean that storage is equal to inflow minus outflow. In pressure pipe networks there is no storage within the pipe system although tank storage may change over time. Thus, in a pipe or junction node, the inflow must equal the outflow. Mathematically, then:

$$Q_i = Q_e \qquad \text{(Equation 2)}$$

or

$$\sum Q_i = \sum Q_e \qquad \text{(Equation 3)}$$

where: Q_i = inflow (cfs)

Q_e = outflow (cfs)

When considering pressure flow (no free water surface) in a conduit, the flow is also related to the velocity of the water in accordance with the wetted cross-sectional area of the conduit. Thus,

$$Q = V * A \quad \text{(Equation 4)}$$

where: V = velocity (f/s)

A = cross-sectional area (f^2)

and $$V_1 * A_1 = V_2 * A_2 \quad \text{(Equation 5)}$$

When a storage tank is considered, it is possible to accumulate or expel water during a time increment and thus the equation becomes:

$$\sum Q_i - \sum Q_e = q_s \quad \text{(Equation 6)}$$

where: q_s = external demand or supply (cfs)

If the elevation and configuration of the storage tank is known (ex. tank diameter and minimum and maximum pool elevations), an equation for head and storage can be written. This, in conjunction with the conservation equation above, can provide an estimate of the effect of the tank on the water system. As mentioned, distribution storage is essential to provide flow and head equalization during demand stress periods. Therefore, it is critical to have the ability to estimate analytically the response of the tank on the system.

Likewise, the *law of conservation of energy* asserts that the energy state (potential and kinetic energy) between two points is constant. Of course, useful energy is converted or "lost" due to a number of factors including friction and momentum. Thus, the term "friction loss", which will be estimated in following sections, refers to the energy that is associated with water movement and converted to energy that is no longer available for system use.

Water system energy is commonly referred to as *pressure* or *head*. In an ideal state head is the sum of three components:

1. Elevation head (potential energy)
2. Pressure head (potential energy)
3. Velocity head (kinetic energy)

The Bernoulli equation relates the total head at a point as:

$$H = z + \frac{p}{?} + \frac{V^2}{2g} \qquad \text{(Equation 7)}$$

where: H = total head (f)

z = elevation above datum (f)

p = pressure (lb/f^2)

V = pipe velocity (f^2/s)

g = gravitational constant (32.2 f/s^2)

γ = water specific weight (62.4 lb/f^3)

Also, from the conservation of energy for two points in a frictionless environment:

$$H_1 = H_2 \qquad \text{(Equation 8)}$$

thus,
$$z_1 + \frac{p_1}{?} + \frac{V_1^2}{2g} = z_2 + \frac{p_2}{?} + \frac{V_2^2}{2g} \qquad \text{(Equation 9)}$$

When friction loss is considered:

$$H_1 = H_2 + h_f \qquad \text{(Equation 10)}$$

where: h_f = friction loss (f)

so $$z_1 + \frac{p_1}{?} + \frac{V_1^2}{2g} = z_2 + \frac{p_2}{?} + \frac{V_2^2}{2g} + h_f$$ (Equation 11)

Now given the fundamental framework for flow and energy related to pressure pipe flow, the next task is to sketch the equations which are commonly used to estimate flow, velocity and headloss. These equations will form the basis for describing the mass and energy effects as water moves through the distribution system (Mays, 1999; Mays, 2000).

Friction (Energy) Loss Equations

The study of fluid movement has produced a number of categories for classifying flow. At low flows (velocities) there is relatively little wall interaction relative to the bulk flow and so the regime is deemed to be in laminar mode. This occurs when the Reynold's number (R_e) is less than 2000; the Reynold's number, of course, being:

$$R_e = \frac{V D ?}{\mu}$$ (Equation 12)

where: V = mean velocity

D = Diameter

ρ = fluid density

μ = dynamic viscosity

As the flow increases the regime moves through a transition phase before becoming fully turbulent. Turbulent flow is generally defined as a flow that occurs with a Reynold's number greater than 4000 (Westaway & Loomis, 1977).

In addition flow systems can be further classified as steady or unsteady. Most typical computations assume steady flow. Steady flow is defined as the condition where the flow rate at a specific point is constant with respect to time. Otherwise the flow is classified as unsteady or transient. By this definition most engineered systems, technically speaking, are unsteady. However, computation of flow properties under unsteady conditions is extremely difficult, if not intractable, so steady flow is assumed in most cases. This assumption appears to be acceptable so long as the temporal mean velocity does not change over brief periods. Unsteady flow then, for engineering design, is associated with rapidly varying velocity or pressure gradients over a short period of time (Mays, 2000).

Because water systems are just designed to convey water that is fit for human consumption, certain assumptions can be reasonably made. These assumptions include modeling a single phase, Newtonian fluid (i.e. water) flowing in a full pipe. This means that the computations will assume pressure flow as opposed to open channel flow. The only place a free water surface should occur in a water distribution system is a storage tank or reservoir.

Darcy-Weisbach Equation

In 1845 Julius Weisbach published his three volume set on engineering mechanics which included his theory of pressure loss through pipe. Over the next 10 to 15 years he refined his approach, which is still widely used today. In fact it is still viewed as superior to most other systems in that it is applicable to any fluid, is dimensionally correct and takes into account temperature of the fluid (Brater & King, 1976; Haestaed et al., 2003). The common forms of his equation are:

$$h_L = f\frac{LV^2}{D2g} = \frac{8fLQ^2}{gD^5p^2} = 0.02517\frac{fLQ^2}{D^5} \qquad \text{(Equation 13)}$$

where: h_L = Friction loss

f = D-W friction factor

L = Length of pipe (f)

D = Diameter of pipe (f)

The Darcy-Weisbach equation can also be expressed in common units of D (in) and Q (gpm), in which case the head loss is:

$$h_L = 0.031\frac{fLQ^2}{D^5} \qquad \text{(Equation 14)}$$

The friction factor, f, is approximated by the following relationship in turbulent flow (R_e > 4000):

$$\frac{1}{\sqrt{f}} = 1.14 - 2\log\left(\frac{e}{D}\right) \qquad \text{(Equation 15)}$$

where: e = pipe wall uniform roughness (f)

For flow regimes in the transition zone (2000 > R_e > 4000) the friction factor is approximated by the Colebrook-White equation:

$$\frac{1}{\sqrt{f}} = -2\log\left(\frac{e/D}{3.7} + \frac{2.51}{R_e\sqrt{f}}\right) \qquad \text{(Equation 16)}$$

A key aspect to notice with the Darcy-Weisbach friction factor in the transition zone is that it is not constant for a given pipe. It will change in relationship to the Reynold's number over the range of flows encountered in the field (Brater & King, 1976).

Computer power has improved the convenience of using the Darcy-Weisbach equation in large scale hydraulic computations, but most of the friction loss computations in water systems rely on the following equation set.

Hazen-Williams Equation

Although the Darcy-Weisbach equation offers a rational approach to friction loss analysis, over the years it has been avoided in complex computations because the friction factor, f, cannot be directly computed. Rather, an iterative approach is necessary and, until recent advances in computer power, the time and effort required to perform all of the computations made it undesirable.

The Hazen-Williams equation has been widely used in the water industry especially in America. It is valid only for water with a turbulent flow regime and a fluid temperature of 60°F. Because there is an implicit kinematic viscosity assumption of $\nu = 0.00001216$ f^2/sec, as the temperature approaches the freezing or boiling point, the equation can yield results with variances of as much as 40% (Westaway & Loomis, 1977). However, water distribution systems tend to stay within the valid range in all but the coldest climates. The mathematical representation of pipe friction loss according to the Hazen-Williams equation is as follows:

$$h_L = \frac{4.73L}{D^{4.8655}}\left(\frac{Q}{C}\right)^{1.852} \qquad \text{(Equation 17)}$$

where: L = Length (f)

D = Diameter (f)

Q = Flow (f^3/s)

C = HW Coefficient

The equation, expressed in common units of L (f), D (f) and Q (gpm), V (f/s) is:

$$h_L = 0.00208L\frac{Q^{1.852}}{D^{4.8655}}\left(\frac{100}{C}\right)^{1.852} = 3.0226\frac{L}{D^{1.167}}\left(\frac{V}{C}\right)^{1.852} \qquad \text{(Equation 18)}$$

Although the formula is incorrect from a theoretical point of view, it still yields acceptable results in practice. This is probably due to the built-in uncertainty implicit in the C coefficient. The HW coefficient is generally obtained by using published tables for pipes and appurtenances, although it is occasionally checked in the field by measuring flow and pressure drop between hydrants (known pipe length) and back-computing the C-value.

Minor Losses

In addition to energy consumed as the water traverses the pipe route, it is also consumed when the flow line changes. This typically occurs in flow through valves, reducers, tees, bends, and other appurtenances in the piping system. The losses, generally called minor or local losses, are due to the turbulence in the bulk flow as it moves through the fittings. Thus, the energy loss tends to be proportional to the velocity

of the water flowing through the fittings. The headloss through these appurtenances is generally estimated by the following equation:

$$h_L = K \frac{V^2}{2g} \qquad \text{(Equation 19)}$$

where: K = minor loss coefficient

The minor loss coefficient is usually determined experimentally by the manufacturer of the equipment or material in question and is provided in tabular form.

Energy Inputs: Pumps

Much of the discussion thus far has dealt with energy loss during delivery of water from source to use. The supply of water and energy imparted to the system is due primarily to pumping systems. While there are many types of pumps, the vast majority of pumps used in the water industry are centrifugal pumps. Centrifugal pumps impart energy to the water volume by a rotating impeller located in an eccentric casing. The pump inlet characteristics (impeller design and rotational speed, outlet features and power applied) form the specifics of each pump and its efficiency. Because of the physics of the pump, the pressure (head) and outlet volume are inversely related so that as the flow increases the pressure decreases and vice versa. This can correspond nicely with varying water system demands in that pump combinations can be selected that naturally adjust to the demand changes by moving up and down the demand curves (Karassik, Krutzsch, Fraser, & Messina, 1976).

The output of each pump design is different so manufacturers publish the pumping characteristics of each of their particular pumps in Head-Capacity or H-Q curves. Horizontal split-case centrifugal pumps are used in most applications because the split case allows removal of the impeller without disturbing the inlet or outlet piping plus the operating efficiency of this type of pump is quite high. Figure 4 illustrates a typical split case centrifugal pump curve showing operating curves for five different impeller options that can be used in the 5x6x11 pump operating at 1750 rpm (Aurora Pump, 1989). The operating head (Y-axis) and the pump discharge (X-axis) stipulates the pump response to different demand requirements. The point at which the impeller curve intersects the Y-axis is called the cutoff head. This represents the pressure at which the pump is unable to pump water (Q = 0) into the system. Thus, if this condition is reached the pump impeller can spin but is unable to overcome the pressure imposed upon it.

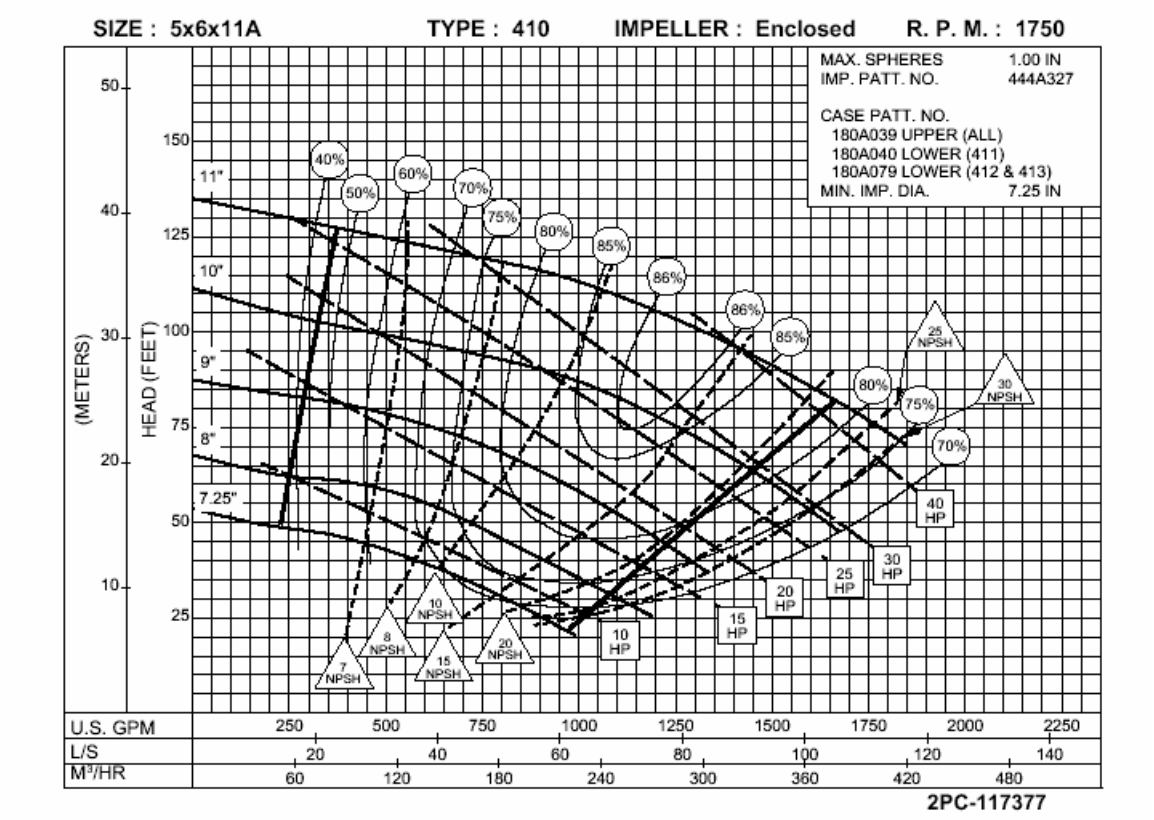


Figure 4 Typical Centrifugal Pump Curve (Source: Aurora Pump, 1989)

Pump design is complex, but fortunately most single speed centrifugal pump curves (H-Q curves) can be approximated by a quadratic equation of the form:

$$h_p = AQ^2 + BQ + h_c \qquad \text{(Equation 20)}$$

where: h_p = pump head (f)

A, B = pump coefficients

Q = pump discharge (gpm)

h_c = pump cutoff head (f)

Data points are picked off the manufacturer's published curves and fit to the equation by regression methods. This allows a mathematical expression for pump activity that can be programmed when solving network problems.

Over the past 15 to 20 years improvement in equipment has led to increased use of variable speed centrifugal pumps instead of constant (single) speed pumps. The pumps provide more flexibility to an operating system, but are more difficult to incorporate in design and modeling efforts. Variable speed operations can be predicted from single speed data by using pump affinity laws. They are:

$$\frac{Q_1}{Q_2} = \frac{n_1}{n_2} \qquad \text{(Equation 21)}$$

$$\frac{H_1}{H_2} = \left(\frac{n_1}{n_2}\right)^2 \qquad \text{(Equation 22)}$$

$$\frac{P_1}{P_2} = \left(\frac{n_1}{n_2}\right)^3 \qquad \text{(Equation 23)}$$

where: Q = pump discharge (gpm)

H = pump head (f)

P = pump power (hp)

n = impeller rotational speed (rpm)

With these equations and the concept of a Fixed Grade Node (FGN), the energy inputs to the water distribution system can be described mathematically. This will be necessary for solving the simultaneous equations describing the system. An FGN is a reservoir of

assumed infinite volume operating at a fixed head (Boulos, Lansey, & Karney, 2004; Haestaed et al., 2003; Mays, 2000).

Network Simulation

The preceding sections have described the major relationships that must be considered when developing an overall model of the hydraulic aspects of a distribution system. The law of conservation of mass states that at every node within the system, the flow entering that node must equal the flow leaving it. Storage does not exist within the system except for storage tanks and reservoirs where there is a free water surface and sufficient freeboard to accommodate the changes in water elevation.

The conservation of energy principle states that the energy between any two nodes must balance regardless of the path taken. This brings into play the concept of friction loss through pipes, valves and other appurtenances within the system. The headloss equations listed above are used to estimate the energy loss as a function of flow and system characteristics (pipe materials and configuration, elevation, control valves, etc). In addition, energy inputs into the system are considered by modeling pumping systems or FGN.

As a matter in passing, because distribution networks can become unwieldy if all small diameter pipes are included in the model, many times the modeler will consolidate small elements of the network into hydraulically equivalent pipe configurations. This process, called *skeletonization*, is a means simplifying or removing elements with small or inconsequential effects. The data associated with removed elements is not typically discarded, but is included with adjacent larger features, using conservation of mass and

energy principles. A frequently used technique of converting flow and headloss for small pipes, valves, bends and other appurtenances into the larger components is the *equivalent pipe method*.

Steady State Model

Once the distribution system has been configured it will have N nodes (junctions) and L links (pipes). Thus, there will be N conservation of mass equations and L non-linear conservation of energy equations that must be solved simultaneously. These are configured according to the physical layout of the pipe network and its loops or topology. Accordingly, the system cannot be addressed by analytical (closed form) methods. Rather, the problem must be solved iteratively (numerically). A number of approaches have been proposed as computer power and mathematical sophistication has increased.

The simplest model to develop is the static or steady state model. In this configuration all the input values are set and the computer iteratively steps to the solution that would exist if the system were allowed to exist in equilibrium. In other words the dynamic features of the system (ex. tank levels, nodal demands, pump output (curve)) are specified and the computer iteratively solves the descriptive equations (Boulos et al., 2004; Haestaed et al., 2003; Mays, 2000).

Extended Period Model

Developmentally, once system programmers had developed the ability to simulate steady state conditions, the opportunity to move forward to a more dynamic representation of a system was possible. This was conceptualized as being a time variation problem, in that engineers had the ability to create steady state configurations

of the network and, if a means of updating the model with new system status could be developed, the model would begin to predict actual system changes. In essence an extended period simulation may be envisioned as a string of steady state models incrementing forward by an arbitrary time step, while updating tank levels, system demands and operational changes.

In order to accomplish this effectively, additional criteria and relationships must be defined to instruct the model how these factors will change from step to step. Of fundamental importance is the ability to reliably define or predict:

- Change in demand both spatially and temporally
- Change in control mechanisms (pumps, control valves and system constraints)
- Change in tank storage volumes and levels

Although demands can theoretically change in almost an infinite number of ways in space and time, for each particular water system certain patterns tend to emerge after study of daily and seasonal water trends. From that information diurnal relationships can be structured by using an assumed baseline flow. If it is assumed that these relationships apply uniformly across the network (spatial variance), one is able to establish a procedure for modifying system demands over time that are imposed on the pumping system.

Given new demand levels, the supply system must respond by either increasing or decreasing the pumping rate or utilizing storage capacity or both. This involves some means of simulating the control system governing the network. Most current software

incorporates some type of scripting or rule-based logic that allows an analyst to describe the conditions under which a pump will switch on or off, PRV's to operate, check valves to open or close, etc. As the new demands are imposed, the control system along with branching logic relating tank storage to the new demands defines how the system adjusts to the new demands.

Since the storage tanks have predefined configurations the water levels can be updated by the conservation of mass and energy relationships. Given the updated demand data the adjustment in water level can be estimated by (Mays, 2000):

$$? H_t = \frac{V_t}{A_T} = \frac{Q_t ? t}{A_T} \qquad \text{(Equation 24)}$$

where: ΔH_t = Change in tank water level at time t

V_t = Volume of water in tank at time t

A_T = Cross sectional area of tank

Δt = Time increment

A_T generally assumes a constant section, but the same concept can be used to predict changing tank levels and volumes on more complex tank shapes.

Dynamic Water Quality Models

The next step in computer simulation of distribution systems is the development of mechanisms to compute concentrations and movement of water quality parameters though the network. At the outset it important to understand the nature of the contaminant of interest. Generally, contaminants may be classified in one of two ways.

They are either conservative or non-conservative. Conservative constituents are those that do not degrade. Constituents in this category tend to be inorganic ionic species like chlorides, fluorides, nitrates, etc. Their concentration is attenuated only by dilution effects. Non-conservative or reactive constituents are those whose concentration are a function of concentration and time and would be represented by microbiological contaminants, THMs, chlorine and the like. A water quality model, then, must factor these concepts into an analysis that can be divided into the following considerations:

- Hydraulics
- Transport Mechanisms
- Bulk Reactions
- Wall Reactions
- Tank Hydrodynamics

Hydraulics.

A dynamic water quality model is dependant upon an extended period hydraulic simulation in order to provide the basic "input" of flow, velocities, direction of movement and storage tank levels as well as fundamental network descriptive data (ex. pipe diameter and length). In essence then, the overall water quality model is actually composed of two component models: the extended period hydraulic simulation and the associated water quality solution algorithm. The extended period hydraulic simulation has been discussed above and will not be repeated here other than to emphasize the importance of developing and maintaining a calibrated simulation of the distribution system in question. Because the water quality model depends on the hydraulic simulation to make its projections, poorly developed hydraulic models will result in

propagation of errors through to the water quality output. This can result in a faulty analysis and improper conclusions.

Transport mechanisms.

Conceptually, water contaminants can be modeled by three mechanisms as they move along a pipe route. These transport models are generally denominated as being advective, dispersive or diffusive (actually, radiation is a fourth mechanism, but it doesn't have applicability in this context). Advective transport assumes that the constituent moves down the pipe axis at a velocity equivalent to the mean velocity of the bulk water column. Dispersive transport in this context is generally defined as radial movement, i.e. from the axis to pipe wall. Diffusion is descriptive of interface transport which occurs at a molecular level and in this context would occur at the leading or trailing edge of a concentration gradient. Because of the increased complexity associated with a multi-dimensional analysis, most popular current analytical schemes uniformly assume that only advective transport is significant. Most packages use a mathematical representation similar to EPANET (Clark & Grayman, 1998; Rossman, 1994), which is:

$$\frac{\partial C_{ij}}{\partial t} = -v_{ij}\frac{\partial C_{ij}}{\partial x} + k_{ij}\left(C_{ij}\right) \qquad \text{(Equation 25)}$$

where: C_{ij} = concentration at position x at time t in the link between nodes i and j

v_{ij} = velocity in link between nodes i and j

k_{ij} = substance reaction rate in link between nodes i and j

It is also important to note the means by which the concentration changes occur as various pipes converge on a junction node. At any given node, multiple pipes can enter

the node and multiple pipes can exit the node. In addition, it is possible for an external demand or supply to exist at the node. Because nodes are a fundamental feature of the pipe network, the computation of the situation should be rational and consistent. Computational models uniformly assume that mixing of flows in a node is instantaneous and complete and that no storage occurs within a node. It is further assumed that conservation of mass principles apply. This leads to the following relationship to define mixing at pipe junctions (Rossman, 1994):

$$C_{i|x=0} = \frac{\sum_{j \in I_k} Q_j C_{j|x=L_j} + Q_{k,ext} C_{k,ext}}{\sum_{j \in I_k} Q_j + Q_{k,ext}} \qquad \text{(Equation 26)}$$

where: $C_{i|x=0}$ = Concentration at start of node i

i = Link with flow leaving node k

I_k = Set of pipes with flow into node k

Q_j = Flow in link j

L_j = Length of link j

$Q_{k,ext}$ = External source flow entering network at k

$C_{k,ext}$ = Concentration of external flow entering at k

Bulk reactions.

As the water moves through the pipe network it is subject to reactive changes. Bulk reactions refer to the changes that occur in the water column itself. The nature of the changes is a function of the contaminant in question. In reality chemical components may change oxidation state or react with other chemicals in the water to form new species or compounds. This is especially prevalent when a community supplies water to

the distribution system from several different sources. These waters mix in the distribution system and can change the character of both waters as a new chemical equilibrium is established. At this point network analysis software only handles biological or chemical reactions as a single species decay (or growth) function. In other words multi-species decay or interaction processes cannot be directly processed. Accordingly, there is no direct way, for instance, for the software to estimate the effect of a specified chlorine level in the water on a specified or computed concentration of a contaminant. Work is being performed on developing toolkits for that purpose but as of this point they do not exist in production form (Uber, Shang, & Rossman, 2004).

Biological contaminants on the other hand may grow or decay depending on the nature of the biological species and the availability of sustaining nutrients in the water. As mentioned in Chapter 4, from a monitoring standpoint it is important to know if a monitored microorganism can replicate outside a host and what its environmental survivability characteristics are. Water systems themselves can impact the environmental survivability. For instance, it is not uncommon in systems that have multiple water sources to create zones within the system that are conducive to growth of microbiological colonies as the waters mix.

Under the assumptions of this paper, the focus is upon contaminants that are specifically injected into the water column and the resultant fate of those contaminants. Certainly there are dilution effects but, in addition, other changes may occur as the water moves through the system. The simulation software must have a means of approximating these changes.

The means used to accomplish this draws from fundamental biological and chemical relationships and the kinetic transformation models used in chemical engineering. These models can become quite sophisticated but generally are unwarranted because of the uncertainty associated with reaction factors in highly dilute and spatially variable environments such as water systems.

Reaction rates are generally classified by reaction order. In simple reactions the order is equal to the sum of the exponents in the rate equation. A rate equation states the relationship between the reactants and the products and is determined experimentally. The general form is patterned as follows:

$$Rate = k[A]^n \qquad \text{(Equation 27)}$$

where: k = equilibrium constant

A = reactant

n = reaction order

Simple reaction relationships used in water quality modeling include zero order, first order and second order equations. Accordingly, then a zero order reaction would indicate that the reaction is independent of constituent concentration, and a second order equation would be proportional to the product or square of the concentration (Williams & Williams, 1967).

The most commonly used model is the first order single species decay model which states that the constituent decays at a rate proportional to initial concentration. Chlorine

is commonly assumed to decay within a distribution system following this pattern unless consumed by an unexpected demand within the system. Powell, West, Hallam, Forester, and Simms (2000) explore this simplifying assumption and the variability possible by using it. This can be represented mathematically by the following expression:

$$C_t = C_0 \exp^{-kt} \qquad \text{(Equation 28)}$$

where C_t = Concentration at time t

C_o = Original concentration

k = decay (growth) constant

t = time

Note that growth of a constituent can also be represented by this relationship if the growth constant is expressed as a negative. The growth indicated would be unconstrained and increase at a rate determined by k.

Biologicals on the other hand may exhibit growth potential up to some maximum or limiting concentration due to limitation of critical nutrients in a potable water system. This can be modeled by a first order growth to equilibrium. This has the same form as that above except that it must consider the constraint posed by the limiting concentration. This is represented by:

$$C_t = C_{max} - \left(C_{max} - C_0\, e^{-kt}\right) \qquad \text{(Equation 29)}$$

where: C_{max} = Maximum concentration

Other reaction relationships can be modeled but this is only productive if information regarding the fundamental reaction mechanisms is known along with the boundary conditions in which they operate. Even the most straightforward modeling efforts require a number of assumptions so, unless specific data is available upon which to base the effort, a more complex analysis may be misguided (Weber & DiGiano, 1996; Williams & Williams, 1967).

Wall reactions.

Wall reactions historically have been deemed to be secondary or tertiary effects with respect to water quality except in cases where corrosion or physical damage occurs. More recently however the implications of wall effects are being reconsidered especially in pipes with high surface area to volume ratios. For example, it has been discovered that the interior of water distribution systems can have thin but significant colonies of microbiological species called biofilms. These colonies have existed inside the networks for years but the impact on water quality (if any) is not well understood or documented. Dissolved species flowing in the water column can contact and react with pipe wall materials or these biofilms attached to the pipe wall and may contribute to changes in the water quality (Camper et al., 2003; Camper, Warnecke, Jones, & McFeters, 1998; Donlan, 2002; Geldreich, 1996).

The following relationship has been proposed as a means of quantifying wall effects on water quality (Rossman, 1994). Assuming a first order reaction, the key unknown is the decay or rate coefficient. This is computed as:

$$r = \frac{k_w k_f}{R_H (k_w + k_f)} \qquad \text{(Equation 30)}$$

where: r = pipe wall reaction rate

k_w = pipe wall reaction rate constant

k_f = mass transfer coefficient

R_H = Hydraulic radius

and

$$k_f = Sh\frac{d}{D} = \left(0.0149 (Re)^{0.88} \left(\frac{n}{d} \right)^{1/3} \right) \frac{d}{D} \qquad \text{(Equation 31)}$$

where Sh = Sherwood number

d = molecular diffusivity of transported species

D = pipe diameter

This is then combined with the bulk reaction rate constant (k_b) to yield an overall decay or growth response as water moves through the pipes. The bulk flow coefficient, k_b, representing the reaction in the water column is normally found experimentally for the particular constituent under consideration.

Thus, the complete rate constant will be computed as:

$$K = k_b + \frac{k_w k_f}{R_H (k_w + k_f)} \qquad \text{(Equation 32)}$$

It then follows that the overall concentration impact can be computed as:

$$C_t = C_0 \exp^{-Kt} \qquad \text{(Equation 33)}$$

Tank hydrodynamics.

As discussed in Chapter 2, storage tanks are important elements in a properly functioning distribution system. They provide the buffering capacity within the system to cope with variations in supply and demand, which reduces stress on the treatment, pump and piping components. Accordingly, the volume of these tanks or reservoirs can have an impact on the movement of a contaminant through the distribution system.

In its most simple form the tank behavior is approached assuming its contents are homogeneous and completely mixed. Thus, incremental changes in incoming water quality are blended with existing tank water in using a simple conservation of mass approach. Also, since these tanks can have significant volume, it is recognized that some level of constituent growth or decay can occur given the detention time afforded. This is represented as follows (Clark & Grayman, 1998):

$$\frac{d(V_s C_s)}{dt} = \sum_k Q_{ks} C_{ks|k=L} - \sum_i Q_{sj} C_s + k_{ij}(C_s) \qquad \text{(Equation 34)}$$

where: C_s = Concentration for tank s

V_s = Volume in tank at node s

dt = Change in time

Q_{ks} = Flow from node k to s

Q_{sj} = Flow from node s to j

C_{ks} = Concentration at end of links

k_{ij} = Decay coefficient between node i and j

Under field conditions, however, it is not unusual to find that the water in a tank is not of homogeneous quality. Short-circuiting and temperature stratification can create significant concentration gradients that clearly impact water quality. These tanks can manifest "dead" zones in which there is very little water movement or exchange. This can result in microbial growth, chlorine destruction, chemical precipitation and other adverse consequences. Current software updates are using research into these issues to refine the simulation capabilities to more closely match field results. This is currently being done by the two and three compartment models that are beginning to be incorporated into the software. Use of these features though requires additional base data to accurately define the functional description and boundary conditions under which the tank operates (Grayman et al., 2000).

Dynamic Water Quality Solution Algorithms

Now that the descriptive components of the network are available, an algorithm can be defined to solve the contaminant propagation equations. There are, in fact, four different algorithms that have evolved over recent years and have come to provide the primary "accounting" options for keeping track of the disparate components in water quality simulation. These are classified by the numerical method used to segment the links and then the manner by which these segments are tracked through the time steps.

Before proceeding with this discussion, it is important to point out the distinction between the "hydraulic time step" used in computation of bulk flow movement and the "water quality time step" that is associated with the segmentation of the links. Generally, the hydraulic time step is measured in hours and is associated with the extended period hydraulic simulation that computes the macro-scale parameters of the system operation.

The “water quality time step” is established by the computational schemes below and typically is measured in hundredths of an hour. This allows the algorithms to establish small completely mixed segments that then move down the pipe and decay (or increase) in accordance with the equations mentioned in the sections above. The means by which this segmentation and tracking occurs is at the heart of the four solution methods below.

The first two methods may be classified as Eulerian after the Swiss mathematician, Leonard Euler, who pioneered the basic numerical approach to solving differential equations. The second two methods are Lagrangian following Joseph-Louis Lagrange’s approach to solving large systems of simultaneous equations. These solution algorithms are briefly described below (Boulos, Altman, Jarrige, & Collevati, 1995; Clark & Grayman, 1998; Mays, 2000; Rossman & Boulos, 1996):

Eulerian Finite-Difference Method (FDM).

This approach seeks to solve the advective transport model (Eq.25) by using variations of Eulerian numerical methods to approximate the derivatives with their finite difference equivalents along a fixed grid of point in time and space (Islam, Chaudhry, & Clark, 1997). The end result is a series of algebraic equations for the entire network that are solved by marching forward in time and down the length of each pipe. At the start of each new hydraulic time step a new grid spacing is chosen in each link so that the number of intervals is as large as possible while staying less than a scaling ratio ($L/V\Delta t$). The process is then repeated until the specified time steps have been completed.

Eulerian Discrete Volume Method (DVM).

In this method each link is divided into a series of equally spaced, completely mixed segment volumes. At the end of each successive water quality time step, concentration within each volume is "reacted" (using the relative equations above) and then transferred to the adjacent downstream segment. When the adjacent segment is a junction node, the resultant (outgoing) concentration for this and all segments entering that node is computed using Eq. 26. This sequence is repeated until a new hydraulic condition is encountered. At that point the network is resegmented to reflect changes in link travel time and the contaminant mass are reassigned to the new segments, at which time the computations continue. The number of volume segments in a pipe is the same as the FDM method, i.e. largest integer less than or equal to the travel time divided by the water quality time step.

Lagrangian Time-Driven Method (TDM).

Instead of dividing each link into equally sized segments and react and transport the water between segments at fixed time intervals as the Eulerian methods do, the Lagrangian methods track the position of variably-sized segments in each link. The difference between the two Lagrangian methods has to do with when the segment conditions are updated. In the TDM method the conditions are at fixed time intervals. The segment so determined is then reacted and the results reestablished as the segment moves on through the link and/or node.

This method uses the water quality timestep (as opposed to the hydraulic timestep) to organize each segment reaction, keeping a cumulative account of the total mass and

flow entering each node. The water quality timestep is normally much shorter than the hydraulic timestep to accommodate the short travel times that can occur within pipes.

Because this is the system algorithm incorporated by the software used in Chapter 7, a more detailed description of the sequence is listed below. The description is taken from the WaterCAD User's Manual:

The following steps occur at the end of each water quality time step.

1. "The water quality in each segment is updated to reflect any reaction that may have occurred over the time step.
2. The water from the leading segments of pipes with flow into each junction is blended together to compute a new water quality value at the junction. The volume contributed from each segment equals the product of its pipe's flow rate and the time step. If this volume exceeds that of the segment, then the segment is destroyed and the next one in line behind it begins to contribute its volume.
3. Contributions from outside sources are added to the quality values at the junctions. The quality in storage tanks is updated depending on the method used to model mixing in the tank.
4. New segments are created in pipes with flow out of each junction, reservoir, and tank. The segment volume equals the product of the pipe flow and the time step. The segment's water quality equals the new quality value computed for the node.
5. To cut down on the number of segments, this step is only carried out if the new node quality differs by a user-specified tolerance from that of the last segment in the outflow pipe. If the difference in quality is below the tolerance, then the size

of the current last segment in the outflow pipe is increased by the volume flowing into the pipe over the time step." (Haestaed Methods, 2002)

The segment positions are also updated at each timestep. New nodal concentrations are computed and new segments are created when the downstream segment is outside the concentration tolerance. The process is repeated for each water quality timestep. At the start of the next hydraulic timestep, the order of the segments is reversed for any link experiencing a flow reversal. Otherwise no ordering adjustment is necessary.

Lagrangian Event-Driven Method (EDM).

As mentioned above this Lagrangian method establishes variably-sized segments based on pipe volume that has the same concentration (within a stated tolerance). In the EDM approach the update computations occur only at times when the leading segment in a link completely disappears through its downstream node. This requires that the program keep an ordered list of the projected "lifetime" of the leading edge of each link. The next "event" occurs for the segment at the head of the list as it intersects and exists through the next node. When the "event" occurs, the following actions occur: 1) the "event" segment is destroyed and the simulation clock is updated, 2) a new concentration is recorded at the node receiving the "event" in accordance with Eq 26 as it mixes with other water entering the node, 3) if the concentration at the "event" node is above a specified tolerance, a new segment is generated at the start of all links leaving the node (with the concentration of each of those links being equal to the updated concentration), and 4) the projected lifetimes of all the leading segments are adjusted and the event is reordered accordingly.

The process continues until the end of the current hydraulic time step. At that time all the segments' positions and concentrations are updated. As new hydraulic time steps increment, directions of flow are recorded, a new ordered event list is generated, and the event processing is continued through the system.

Rossman and Boulos (1996) provide a description and comparison of each of these methods and conclude that all of them are capable of adequately representing observed water quality behavior in actual water distribution systems. They did observe, however, that the Lagrangian time-driven method (TDM) was the most versatile of the methods tested, but did require more computer memory when water age was being computed. If computer memory is at a premium given the size of the system in question the Eulerian methods might be better suited when trying to model water age.

Summary

This chapter provides an overview of the fundamental factors associated with developing a computer simulation of an operating water distribution system. Water distribution systems can be described in terms of links (pipes), nodes (interconnection junctions), and pipe loops. The movement within water systems is simulated by application of the laws of conservation of mass and conservation of energy. The hydraulic movement of water is inherently non-linear because friction (energy) loss within the pipes is highly variable based on flow rate and other physical factors. Fluid formulas exist (Darcy-Weisbach and Hazen-Williams) which can reasonably predict the flow and energy levels (pressure or head) associated with the water at any particular point in a pipeline.

Because of the interconnection of pipes in loops, the analysis of fluid movement becomes quite complex as the water seeks to supply demand through the system with minimum energy. The resulting flow and energy results are again highly non-linear and in many cases non-intuitive. Equations balancing mass at each node and energy along each link must all be solved simultaneously. A number of numerical processes have been developed to solve these equations.

A steady state analysis is the most fundamental investigation of a system as it addresses the response at a single point in time while all external forces are constant and the system is in equilibrium. From that extended period simulation, processes have been developed that define changes in external forces (included description of pump, valve and tank control sequences) that allow a series of steady state models to step through time (hydraulic time step) and simulate actual hydraulic response to the changes imposed. By calibrating the hydraulic model to actual field data and making appropriate adjustments in the model assumptions the computer simulation can provide a reliable proxy for the physical reality.

After the calibrated extended period simulation for the distribution network has been developed, it is possible to expand the computer model to form a representation of water quality movement and fate in the system. This is accomplished by using the extended period hydraulic simulation to compute the bulk movement components of the water, define the advective transport features, bulk and wall decay reactions as well as tank hydrodynamic characteristics. As these parameters and their associated relationships are defined, the water quality simulation takes the hydraulic model output data and using one of four solution algorithms, computes the changes in water quality resulting from the

external forces applied. While the water quality simulations need to be subjected to calibration exercises also, they have been found to provide a reasonable surrogate for actual field studies. In fact they allow exploration of system response to atypical events that otherwise would not be subject to analysis. That is the sense in which the computer simulations will be used in later sections of this study.

For reference, a list of assumptions and potential error sources associated with extended period dynamic water quality simulations is presented in Appendix 2. The purpose of the list is to provide a sober assessment of the data needs necessary to develop a reliable and workable model. Fortunately, computer simulations seem to be fairly robust and can provide valuable insights into system operation if a conscientious effort is made to calibrate the initial model. As George Box, the statistician, once stated so well, “all models are wrong, but some models are useful.” That is precisely the view that should be maintained with water network simulations. They are incredibly powerful and useful tools that can provide system designers and operators a richness of system understanding that could not be gained by any other means.

Chapter 6 An Algorithm to Identify Efficient Monitoring Station Locations

As this study moves forward, it might be helpful to recap the ground that has been covered. Up to this point the primary effort has been to develop the context and major factors of the research effort. A statement of the research objective was presented and an overview of current water system design and distribution system monitoring practice was discussed to establish the framework within which the research will be conducted. Then a discussion of potential contaminants followed to illustrate the range of characteristics, effects and fate they present in water systems. Finally, a discussion of the major features of water quality simulation was presented, to identify the source and basis of the data to be used to analyze a specific system. At this point the study will proceed with developing the approach and algorithm used to decompose the distribution system and organize the data to reach the stated objectives.

Algorithm Setting

As mentioned earlier, because the nature of the contamination event and the operational state of the subject distribution system has an infinite number of configurations, there is in an absolute sense no one true optimum set of sampling locations, and variation in the "optimum" set is to be expected given different base assumptions. However, it is posited that there does exist a cluster of points that can serve as robust sentinels for detection of a wide range of events.

For a methodology to be productive, the procedure must define a feasible contamination event and make realistic assumptions regarding its occurrence and characteristics. This

assumption set will be applied to a computer simulation of the physical system to produce an assessment of the system's response to the critical event.

These assumptions may be considered to exist at two primary levels. The first involves the physical reality of the event followed by a definition of the simplifying assumptions related to the modeling endeavor. A reasonable set of physical assumptions would include the following:

1. Contamination can occur at any point in the distribution system at any time and for any duration.
2. The contaminant can be injected in either a continuous or discrete mode.
3. The contaminant can be chemical, biological or physical.
4. The contamination event can consist of a single agent or multiple agents injected simultaneously at single or multiple points.
5. The contaminant is assumed to be sufficiently toxic to adversely affect consumers at some arbitrary threshold concentration and is undetectable by the consumer's senses. The adverse endpoint can be defined in terms of morbidity, mortality, psychological trauma or economic effects.
6. A distribution system monitoring point is assumed to continuously monitor for the presence and concentration of the contaminant of interest with real-time feedback.
7. Economic and/or technical factors constrain the number of sampling points to a small number relative to the total nodes in the system.

Beginning with these assumptions, the following sections develop a methodology to identify, in a heuristic manner, a set of monitoring points that efficiently ranks and covers the distribution system from contamination at any point within the system. At this point the approach is presented in a "proof of concept" form which will be subject to refinements and open areas of future research identified in the final chapter. As a final note it is again mentioned that the methodology should be within the reach of its target audience, small to med-sized water systems. Heuristic approaches have historically been more effective in that practice segment.

Develop a Computer Simulation of the Distribution System

Because it is impracticable to physically test a wide range of contamination events, it becomes immediately apparent that any reasonable approach seeking to address the problem must rely upon a workable, extended period of computer simulation of the distribution network in question. Thus, the first step in the process is to develop an extended period water quality model that simulates the normally functioning water system. This process was discussed in Chapter 5. In that chapter, there are certainly many assumptions and sources of errors associated with an extended period water quality model. However, a network simulation can be refined and calibrated to yield a sufficiently accurate representation of reality. The simulation packages are more affordable and accessible now than in the past and for many reasons every community water system should have a working simulation of their distribution network.

Develop a Contamination Scenario

Once the basic extended period model describing the distribution system is in place, the next step in the process is to define a contamination incident description or scenario.

Vulnerability analysis training identifies this as the Design Basis Threat (DBT). The DBT should be reflective of a realistic scenario that could be feasibly imposed upon the system resulting in a defined adverse effect.

The first issue in developing the DBT is selecting the contaminating agent. Many contaminants can pollute a water supply but, for this purpose, the agent is assumed capable of mass contamination. In other words, the intent of contamination would be to corrupt a large volume of water as opposed to a few connections. To achieve the goals of the offending entity, the water would need to be contaminated in such a way as to be undetectable when consumed, so that exposure to the contaminant would be unconstrained. Given these criteria, Chapter 4 outlines a list of potential candidates which can be considered for use in situations proposed here. Additional pragmatic concerns also come into play, such as the ease of acquiring and the cost of the material as well as the ability to transport and disperse the contaminant.

At a minimum the following information should be delineated to properly model the event:

1. Identification of the contaminant: biological, chemical or physical, along with toxicological properties, and detection methods or characteristics.
2. The mass of the contaminant to be injected into the system.
3. The duration or injection rate that the contaminant will be introduced into the distribution system.
4. Time-of-day that the injection begins.

General Modeling Assumptions and Analysis Concepts

Once the basic distribution system simulation has been developed and the DBT has been delineated, it is possible to predict the impact of the contamination event on the distribution system. Given the fundamental set of physical assumptions presented above, a corresponding set of primary modeling assumptions must be stated. Most distribution system software provides a wide range of customizable attributes (with associated assumptions) which allows simulation of complex and compounded events that can approximate the physical assumptions. For purposes of this paper the following represents the *general modeling assumptions* used in the development of the contamination event data.

1. Contamination can occur at any point in the distribution system at any time and for any duration.
2. The contamination event is assumed to occur at only one location in the distribution system at a time;
3. Contamination is assumed to occur only at the junction nodes.
4. All nodes are equally important.
5. The contamination event will consist of a single agent.
6. The contamination event will be assumed to occur in a single discrete event as opposed to continuous injection. (Continuous injection is permitted by the algorithm, but is not as realistic for terrorist activity.)
7. Monitoring points are assumed to be located at the junction nodes.
8. The contaminant is carried with the bulk flow of water. The contaminant can be treated as observing conservative or non-conservative particle assumptions depending on the parameter and capability of the modeling software.

9. Biological measurement assumes that count (i.e. cfu, oocysts, etc.) per million is equivalent to parts per million (ppm). Strictly speaking this is incorrect because ppm is a weight/weight relationship, but this assumption is sufficiently accurate for simulation purposes.
10. A monitoring point that detects, above a given threshold, contamination originated from multiple points is considered superior to a monitoring point that detects contamination from fewer points.
11. Given a fixed number of monitoring points, M, the "best" location set shall be those M points that detect contamination from the largest number of unique origination nodes with the greatest redundancy.

The procedure proposed here differs from others cited in that, rather than utilizing a representative hydraulic analysis and/or water quality simulation and then identifying a monitoring set using various optimization techniques based on that one actualization, a separate extended period water quality analysis is performed with each node in turn being used as the contamination injection node all of which are then used to find the optimum set. Thus if there are **n** nodes in the system there will be **n** extended period simulations performed. In this way any combination of injection events conceptually can be assumed for consideration. Using the DBT, the **n** extended period simulations will be run for a duration necessary for the entire contamination mass to be consumed. Noting that the concentration of contamination will be computed for each node along the selected timestep, **t**, it is observed that all other nodes in the system can be considered monitoring nodes. Recall that the principal of superposition is a useful tool in handling problems of this type (Boccelli et al., 1998; Tryby, Boccelli, Uber, & Rossman, 2002). By combining all **n** analyses for all **t**-timesteps it is possible by linear superposition to

examine and rank each node for its efficiency in detecting contamination occurring at any point in the system.

Until recently such an approach would be prohibitively burdensome computationally except for systems that had been dramatically skeletonized. However the power, efficiency and cost of computers and analysis software have improved to the point that this procedure is in fact feasible for small to mid-sized water systems.

It is now possible to develop an analysis methodology which heuristically identifies a ranked set of monitoring points to guide the establishment of a monitoring network. In a generic sense the process proceeds as follows:

1. Develop a calibrated extended period hydraulic model of the system in question.
2. Develop a prototype contamination event.
3. Develop a contamination event database as follows.
4. Beginning with the first node in the distribution system, assume that the contamination event occurs at that node, run the extended period water quality model and record the computed concentrations for each node at each timestep of the analysis.
5. Repeat step (4) for each node in the system in turn, applying the contamination criteria to that node.
6. Analyze the resulting dataset for patterns which identify the nodes that have a higher probability of detecting contamination events occurring in the system than other nodes according to the criteria outlined in modeling assumption 10 above.

The following sections of this paper will outline issues that arise when implementing this strategy and provide a means of organizing the data to identify the desired monitoring points.

Scenario Development and Analysis Issues

Water Distribution System Model and Development of Contamination Database

Time-of-day that injection begins.

The DBT should stipulate at what time of the day the injection begins and how long it occurs. While a time-of-day can be arbitrarily selected, it may be prudent to run a sensitivity analysis by starting the injection at several different times during the day. Depending on the size of the system, this can create a significant amount of additional work because it means that the entire modeling and analysis effort must be rerun for each new starting time. It may be worth the additional effort, however, because the selected monitoring locations may shift somewhat depending on the time that the injection began. For example, the shift in water patterns between a contamination event that begins at midnight and one that begins at noon may be sufficient to cause a re-weighting of the monitoring nodes. Examination of several analysis sets using different injection initiation times can help to create a more realistic portrayal of the range of contamination events resulting in a more robust network of monitoring points.

Duration of analysis.

An extended period simulation requires selection of a hydraulic time step. This is the time between each quasi-steady state analysis through the model duration which is the

total time that the analysis runs. In other words the analysis will begin at t=0 and proceed incrementally by the hydraulic time step (Δt) until the duration specified (t_d) has been reached. Because a contaminant will move through (and exit) the system at differing rates depending on when and where the injection occurs, one cannot necessarily select an arbitrary duration of analysis, such as 24 hours. In order to avoid comparing scenarios that may have different levels of total contaminant consumption, it is necessary to estimate the time it takes for all of the contaminant to clear the system. For example, injection of a contaminant near a highly active node may result in a short duration uptake of the total mass whereas injection near a storage reservoir may cause a lower concentration, longer duration feed. One convenient way to estimate the appropriate duration of analysis is to select a number of diverse nodes and run simulations at varying durations to see how long it takes for the contaminant to be consumed at each. The intent is to set a duration that is long enough to provide a comparable mass consumption for all permutations and yet as short as possible to minimize the data to be processed.

Functionally, some level of judgment may be required. In situations with long concentration tails, it may be necessary to set the duration at a level shorter than required to completely clear the system. This entails a judgment call considering the amount of extraneous data collected in the overall analysis along with the reduced health implications at those low tail-levels.

Development of contamination database.

Once the DBT scenario has been defined and the distribution system modeled and calibrated, the contamination database is then created. As mentioned above an injection "pattern" corresponding to the DBT is defined to include a contamination initiation time and duration. The basic network model parameters have been set which will include the selected hydraulic time step (Δt) and duration of the analysis (t_d). Apply the DBT pattern to the first junction node and run the extended period water quality simulation. The simulation data for all nodes is generated and stored, taking care to note which node was the contamination node. The process is then repeated for all nodes in turn. The resulting data set represents the response of the distribution system to all nodes serving as the contamination source, for the assumptions associated with the DBT pattern.

Manipulation and analysis of contamination database.

Once the full set of junction nodes have been simulated with each node serving as the injection node, the resulting database provides a wealth of information about the system and its response. In order to mine the data effectively, several factors need to be addressed to improve the flexibility necessary to define the dominant patterns along with their stability.

Threshold of concentration.

In one sense any node with a positive computed contaminant value should be considered worthy of note and counted in the assessment process. However,

pragmatically there is a difference between a computed simulation value and a measured lab value, not to mention an actual infectious dose or NOAEL. It is not unusual, for example, to observe a computed value that may have a magnitude of 0.001 mg/L. Depending upon on the agent, this may be neither significant nor detectable in a field setting and thus is more of a computational artifact than a useful detection point. Thus, when devising a ranking system, establishing a threshold of concentration provides the analyst flexibility in exploring whether small computational values skew the result. Also, allowing for a concentration threshold can demonstrate the sensitivity of the monitoring points to changes in detection levels in the field.

Time since injection versus consumption volume.

There are two fundamental approaches to establishing the criteria for evaluating the most desirable set of monitoring points. These approaches are related but can yield results that may differ. One approach seeks to optimize the set of monitoring points by minimizing the time to detection while the other approach focuses on establishing a system that minimizes consumption of contaminated water. Both approaches have merit depending on the emphasis of the program. It would seem intuitive that the basic objective of a monitoring system would be to detect contamination as quickly as possible after it occurs; however, the ultimate goal of monitoring is to minimize exposure of customers to contamination. A brief discussion of these perspectives is given in Kumar, Kansal and Arora (1999).

The method proposed here presents a methodology which seeks to utilize the time since injection as a major variable, but is balanced against the stability of the monitoring set

over time. By extension the proposed approach can also be used to estimate the "volume consumed" and compare that with monitoring set stability. The basic thought under consideration in this paper, is that for any given system, the "optimum" monitoring set may vary for a given elapsed time or volume consumed and it is instructive to develop a sense of the magnitude of the set variation prior to forming a conclusion.

Detection count versus average contaminant concentration.

Once the database is populated with concentration values for all contamination injection possibilities over all time steps, it is necessary to provide a criterion for evaluation and ranking. For a given concentration threshold, two straightforward means of accomplishing this are (1) averaging the computed concentrations at each monitoring node over time **t** or (2) summing the number of injection nodes that the response node detects along with the frequency of detection. These statistics (means or totals) will be cumulatively computed for each timestep at each response node.

This process (successive cumulative computations from **t**=0) tends to have more impact on the evaluation of nodes by averages than by totals/frequency, so the counts tend to be more stable, if not informative, than the averages. This follows from the fact that the range of values encountered during a contamination event will be large and the use of averages may overwhelm small but significant signals at some nodes. The counts will not be as sensitive to extreme values, and, if a concentration threshold is considered, counts will provide a reliable indication of the frequency at which each response node can detect contamination from other nodes in the system.

Ranking algorithm.

At this point it is possible to present a step by step summary of a proposed algorithm that can be used to isolate nodes within the distribution system that are more efficient than others in detecting external contamination.

1. Apply the injection pattern to each node in turn and record all computed concentrations for all nodes for all timesteps through the modeled duration (t_d). This generates a matrix populated with C_{m,i,t_s}, the detected concentration at the monitoring node m measured at time t_s since injection at node i.
2. Proceed through the data matrix generated above and, for each injection node, compare all computed timesteps with a preselected threshold concentration (C*). In a parallel Evaluation matrix, if the computed concentration is greater than the threshold concentration, place "1" in the cell; otherwise insert "0" in the cell. In other words the matrix will be populated by values computed as follows: $f_{m,i,t_s} = 1$ if $C_{m,i,t_s} > C^*$ or 0 if $C_{m,i,t_s} \leq C^*$. This indicates whether the monitoring node m detects injection at node i at timestep t_s.
3. Using the Evaluation matrix, increment through the data and record in a Summary matrix the number of times that Monitoring node, m, detects injection by node i for all timesteps for $0 = t = t_s$. This sum will be recorded by timestep t_s.

$$s_{m,i,ts} = \sum_{t=0}^{t_s} f_{m,i,t}$$

4. For simplification of presentation steps 4 through 9 assume a stipulated C* and t_s. Using the stated C* and t_s, compute the efficacy ($e_{m,i}$) of each Monitoring node

m for injection at node i by noting if it detects any contamination for all timesteps from t=0 to t_s . In other words $e_{m,i} = 1$ if $s_{m,i,t_s} > 0$ or 0 if $s_{m,i,t_s} = 0$.

5. Compute the efficiency with which each monitoring node detects contamination (e_m) by summing the number of injection nodes for which Monitoring node m has an $e_{m,i}$ value equal to 1. Hence,

$$e_m = \sum_{i=1}^{n} e_{m,i}$$

6. From the Summary matrix find the most efficient monitoring node (m_b). This is the monitoring node that detects the most injection nodes for $0 = t = t_s$ and with $C > C^*$. Select node m_b such that $e_{m_b} = \max(e_m)$. Thus, by ranking all monitoring nodes by e_m in descending order, the monitoring nodes are ranked by the number of injection nodes they detect. If there is a tie for m_b (i.e. several monitoring nodes have the same max(e_m)), rank the tied monitoring nodes by the number of monitoring events that each node detects. This corresponds to a secondary sort of values based on total detection events by the monitoring node over t_s. Symbolically then, chose from the set of tied nodes, the one node with the largest $\sum_{i=1}^{n} s_{m,i,t_s}$. This says that, in the event that several monitoring nodes detect the same number of injection nodes, the most efficient node is the one that has the most individual detection events through t_s.
7. Selection of the most efficient node based on m_b, however, does not assure complete coverage of the system as injection at some nodes may not be

detected by the selected node. Additional monitoring nodes, then, are required by selecting all injection nodes i not detected by m_b (i.e. m_b with $s_{m_b i,t_s} = 0$).

8. For each injection node, i, that is not detected by m_b (i.e. $s_{m_b i,t_s} = 0$), select other monitoring nodes, m, for *that* injection node i for which $s_{m_b,i,t_s} > 0$. Record the s_{m,i,t_s} counts for each of those nodes. Select the m with maximum counts, i.e. $\max(s_{m,i,t_s})$. If there is a tie, select the m that has the greatest redundancy for that t_s, i.e. the node with the largest e_m among the remaining nodes. This identifies the monitoring node that provides the best opportunity to detect the injection node under consideration along with the greatest redundancy for detecting other injection nodes that are covered by other monitoring nodes.
9. To prioritize additional monitoring nodes not covered by m_b, analyze and sort in the same manner as above.
10. Perform sensitivity analysis by varying "time since injection" (t_s) and "threshold concentration level" (C*). Typically, this task is accomplished by systematically varying t_s from some minimum value to t_d, observing the stability of the selected monitoring set(s). Likewise C* should be varied through a reasonable range to observe the effect that C* has on the selection of the monitoring set.
11. Using the information from step 10, develop contingency tables and charts to assist in decision making. This can take a number of forms depending on the technical ability of the analyst and degree of sophistication required. At a minimum, one should plot and/or tabulate the selected monitoring sets with t_s and C* and note the magnitude of change in selected nodes. Decision rules for acceptable stability will be system specific.

The data set developed above has an explicit assumption regarding the time-of-day that the DBT pattern was applied. To validate the selected monitoring points, consideration should be given to repeating this process after setting the DBT pattern to initiate contamination at a different time-of-day. This allows the examination of effects caused by variable water demand experienced over the modeling duration. A separate set of data will result reflecting the different initiation time of contamination. Additional sets may be run reflecting other initiation times if there are discrepancies in the optimal sampling points.

Summary

This chapter provides the context and assumptions used to develop an algorithm that can be applied to locate a robust set of monitoring stations within a potable water distribution system. First, a database is developed by computing concentration values assuming all nodes are potential injection points for a stipulated Design Basis Threat. This database is analyzed using the algorithm presented to select in ranked order the nodes that are robust and effective in detecting the contamination event when considering the time since injection and a threshold concentration level. The next chapter will seek to illustrate the process on a small test water distribution system that is frequently used in the literature.

Chapter 7 An Illustrative Example Using the Ranking Algorithm

As a means of demonstrating the procedure described above, it will be applied to the Anytown, USA model (Kessler et al., 1998; Walski et al., 1987). This test set is frequently used in the literature and should provide a convenient means of setting up the program and experimenting with results. The system consists of 34 pipes, 16 nodes and 17 loops. It also has two elevated storage tanks, a single well and pumping station. A schematic drawing of the system is provided in Figure 5. The specific characteristics of the system are delineated in Table 12 through Table 15. The assumed demand flow pattern is listed on Table 16.

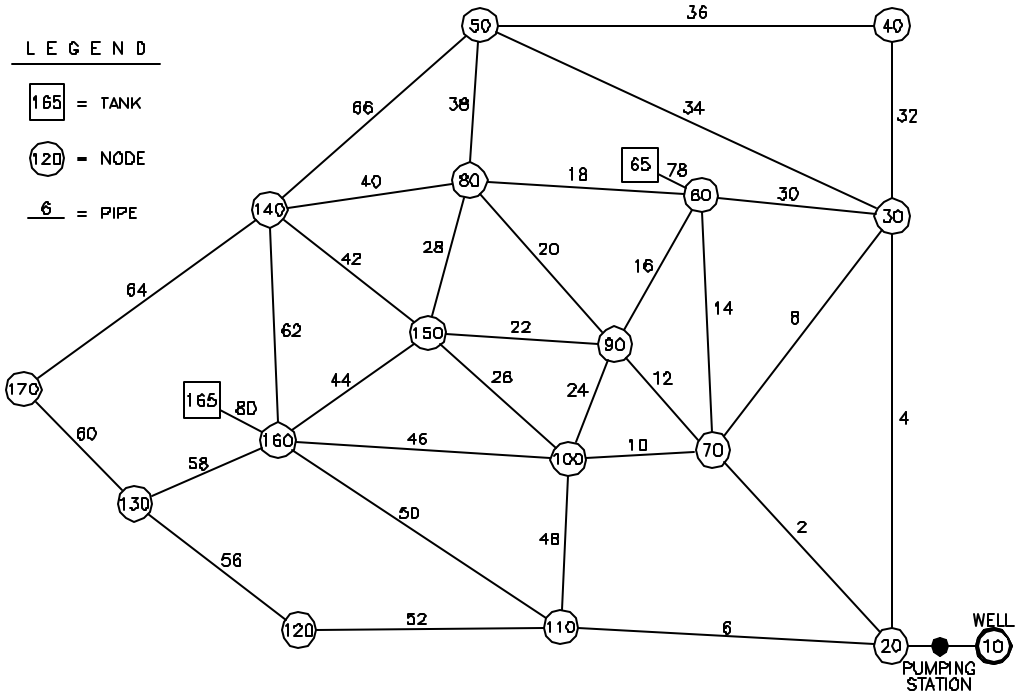


Figure 5 Schematic Drawing of the System

Table 12 Pipe Characteristics

Pipe Number	Length (ft)	Diameter (in.)	C-factor
2	12,000	16	120
4	12,000	12	120
6	12,000	12	120
8	9,000	12	120
10	6,000	12	120
12	6,000	10	120
14	6,000	12	120
16	6,000	10	120
18	6,000	12	120
20	6,000	10	120
22	6,000	10	120
24	6,000	10	120
26	6,000	12	120
28	6,000	10	120
30	6,000	10	120
32	6,000	10	120
34	9,000	10	120
36	6,000	10	120
38	6,000	10	120
40	6,000	10	120
42	6,000	8	120
44	6,000	8	120
46	6,000	8	120
48	6,000	8	120
50	6,000	10	120
52	6,000	8	120
56	6,000	8	120
58	6,000	10	120
60	6,000	8	120
62	6,000	8	120
64	12,000	8	120
66	12,000	8	120
78	100	12	120
80	100	12	120

Table 13 Node Characteristics

Node Number	Elevation (ft)	Average Demand (gal./min.)
20	20	500
30	50	200
40	50	200
50	50	200
60	50	500
70	50	500
80	50	500
90	50	1,000
100	50	500
110	50	500
120	120	200
130	120	200
140	80	200
150	120	200
160	120	800
170	120	200

Table 14 Tank Characteristics

Tank Number	Elevation (ft)	Initial level (ft)	Minimum level (ft)	Maximum level (ft)	Diameter (ft)
65	215	35	10	35	53.8
165	215	35	10	35	53.8

Table 15 Pump Characteristics

Discharge (gal./min)	Head (ft)
0	358
12,000	270
24,000	190
33,000	0

Table 16 Demand Characteristics

Time of Day	Multiplier of average demand
06-09	1.2
09-12	1.3
12-15	1.2
15-18	1.1
18-21	1.0
21-24	0.9
24-03	0.7
03-06	0.6

Water Distribution System Simulation

An extended period water quality simulation for Anytown was prepared using Haestad Methods WaterCAD® v. 6.01. After configuring the model using the components as listed above, the simulation was run several times under different initial hydraulic conditions (ex. elevated tank water levels, pump control settings, etc) with a one week duration. This helped to establish stable values for the pump controls to match the water levels stipulated in the Anytown problem statement. These were then used to set the initial conditions for the simulation used in this example. A one hour hydraulic time step was used as the default increment between simulations.

Design Basis Threat Definition

For analysis purposes it will be assumed that the distribution system will be subject to a biological contamination event. The agent chosen will be the parasite *Cryptosporidium parvum*, which is the same organism involved in the accidental contamination of the Milwaukee, Wisconsin water system in 1993. This agent, while typically having a low mortality rate, has many characteristics of an ideal contaminant. It has a low infective

dose, is highly resistant to normal disinfection procedures, can be significantly concentrated and is easily handled (Burrows & Renner, 1999; Teunis, Chappell, & Okhuysen, 2002). It can also give rise to subsequent secondary infections. Using this agent the contamination scenario is assumed to unfold as follows.

It is noticed that large road or construction projects use water tankers as a means of dust control. These tankers are frequently observed connecting to fire hydrants to replenish the water in the tanker. The terrorists reason that it would be a fairly straightforward task to fit the tanker with a pump such that rather than receiving water the tanker could overcome the backpressure and pump water back into the distribution system. Cryptosporidium oocysts can be cultivated and concentrated (Teunis et al., 2002) to the point that a 3000 gallon tanker can be brought to a concentration of 300,000 oocysts/L. For planning purposes it is assumed that the tanker is fitted with a 50 gpm pump which would allow a full discharge of the tanker in 60 minutes. As a simplifying assumption it is assumed that each node in the distribution system computer model represents a fire hydrant (i.e. injection node). Thus, the DBT envisions an event that discharges 3000 gallons of contaminant into the distribution system at a rate of 50 gpm for one hour with a concentration of 300,000 oocysts/L (assume this is equivalent to 300,000 mg/L for modeling purposes). Although pragmatically unlikely, for the initial scenario it is assumed that the injection occurs at midnight.

Development of Contamination Database

The DBT scenario summary information is now developed into an injection pattern which will be applied to each node of the distribution system. First, however, a few of the nodes are sampled to establish the duration of the extended period water quality

simulation of the contamination event. For this project, it is determined that a 24 hour duration from the time of the injection should be sufficient for the majority of the contamination to exit the system. The short duration is probably due to the fact that the elevated storage tanks are small relative to the system demands which results in high turnover rates. This in turn reduces the dampening effects of the storage tank. In any event this step provides guidance on how to set the simulation parameters for the extended period water quality model which will generate the fundamental data set.

Now, applying the injection pattern to each of the **n** nodes in turn, the simulation is run **n** times. The time increment for the extended period model (Δt) in this example is 1 hour and the total duration of the run (t_d) is 24 hours. Therefore, the analysis data set is based upon 24 time steps ($t_d / \Delta t = 24 / 1 = 24$) for each of the **n** simulations. Since there are 16 nodes which can serve as monitoring points (including the contamination node), this yields 384 data points for each contamination scenario. The simulation is then repeated for each of the 16 contamination nodes, which yields a total of 6144 points populating the database for consideration. A concentration matrix results with Monitoring Nodes (m = 1, 2,…, 16) forming an X-axis, Injection Nodes (j = 1, 2,…,16) defining the Y-axis and the Time-steps (t_0 through t_d) as the Z-axis. By examining each slice through the Y-axis (injection at that node), a record of contaminant movement through the system over time is observed (in the X-Z plane) and the nodes which detected that movement over the course of the event may be identified.

Analysis of Contamination Database

Now that the contamination matrix is defined it is possible to develop different queries to test the efficiency of different nodes with respect to detecting contamination over a range

of contamination injection points. For purposes of this analysis, the efficiency of the monitoring nodes will be based upon the total counts (observations) recorded for each monitoring node over the different scenarios. Because an objective of this study is to use tools that are commonly available to operational personnel, Microsoft Excel® is used for this illustration to process the data. As the distribution system increases in size it is necessary to use a relational database, however by using Excel for this example the process is more transparent.

Preprocess Raw Data

One of the limitations of Excel is that it can only accept 256 columns of data. Because there are 256 node points (16 nodes * 16 nodes) to consider plus the need for some labeling and computational columns, the first task is to import the database and transpose the time and node axes. Once this is accomplished the analysis can proceed for projects in which the timesteps under consideration are less than 256, which is typically the case. However, as will been seen in the tasks to follow, this column constraint will limit the use of Excel alone to systems with no more than 250 modeled nodes.

Another comment on setting up the analysis database is to note that a simulation will likely have many intermediate or fractional timesteps. This results from "change of state" operations in the system (e.g. a pump turns on or off, a tank switches from filling to discharging, etc.) These timesteps may or may not have concentration data associated with the operations, but to keep from "oversampling" during fractional timesteps it is normally best to eliminate these points from the analysis and just consider status at each standard hydraulic timestep. This essentially forms a systematic sampling of the data

and prevents the database from expanding unnecessarily. The initial state (t = 0) is also included in the event that an initial or background concentration needed to be considered (which in this case it did not).

Having imported and transposed the simulation results, a two dimensional spreadsheet exists with evenly incremented timesteps across the top and monitoring-injection nodes down the side. Table 17 provides a partial view of this information. Reading this chart for example, 9 hours after the contamination event, monitoring node 80 detects a contaminant level of 2313.849 mg/L (oocysts/L) for contamination that originated at node 60.

Table 17 Transposed Simulation Results (Partial)

Time (hr)	6	7	8	9	10	11	12
80 J20 Inj. Conc. (mg/l)	216.6066	1181.621	267.7076	170.9058	341.2332	565.2613	425.2074
80 J30 Inj. Conc. (mg/l)	43796.71	11069.39	154.569	297.9569	478.0985	783.9133	724.057
80 J40 Inj. Conc. (mg/l)	0	0	0	0	0	0	0
80 J50 Inj. Conc. (mg/l)	0	0	0	0	0	0	0
80 J60 Inj. Conc. (mg/l)	112787	10982.69	1200.131	2313.849	3713.53	6088.093	5623.952
80 J70 Inj. Conc. (mg/l)	469.0788	1539.53	191.9643	370.0469	593.7739	973.5806	899.2426

Develop a Count Evaluation Matrix

Next, construct a parallel spreadsheet with the same headings as above. Using a Threshold Concentration (C*) established by the analyst, compare the value in each cell of the raw data database with that threshold concentration. If the cell value is greater than to the threshold concentration, a “1” is placed in the cell of the parallel spreadsheet. Otherwise a “0” is placed in the cell. Upon completion this spreadsheet constitutes a count evaluation matrix that is populated with a record of the number of times a particular (monitoring) node detected contamination greater than C* from any injection point in the system measured from the stated time.

Prepare Count Summary Table

Using the parallel spreadsheet (Count Evaluation Matrix), develop a Count Summary Table. This table has the Injection nodes listed across the top and the Monitoring nodes listed down the side. Select a timestep value, then sum the number of "1"s in each Monitoring node cell associated with each Injection node from the Counts worksheet. Sum only through the time step indicated in the timestep cell.

Now it is possible to count across the table to determine the number of Injection nodes that each monitoring node detects through the time period indicated in the timestep cell. Note that this count is independent of the number of times that the monitoring point detected contamination for an Injection node – only that it detected contamination at least once during the period indicated. However, along with that count, another column will be inserted that sums the total number of times that the monitoring node recorded a detection over all nodes during the period indicated. With this summary information in hand it is now possible to refine the search for the most robust set of monitoring points. A partial view of the Count Summary Table is presented in Table 18.

Table 18 Count Summary Table (Partial) – At 12 Hours With 100 Mg/L Threshold

Greater than variable:

Node	J20	J30	J40	J50	J60	J70	J150	J160	J170	COUNT	Ave	Total
20	1									1	1	1
30		1								1	1	1
40		1	1							2	1	2
50	3	4		2	5	2				5	3	16
60	8	8			10	8				4	9	34
70						1				1	1	1
80	7	8		1	11	7				6	6	38
90	5	5			6	6		1		7	4	26
100	2									2	2	3
110										1	1	1
120										1	1	1
130	6				5	5	6	8		8	5	38
140	5	5		1	7	5	2	6		10	4	38
150	5	3			4	4	6	7		9	4	35
160	8	1			8	9	7	10		7	7	50
170	4	2			3	4	1	6	1	11	3	33

Develop a Count Sorting Table

Using the same configuration as the Count Summary Table, a parallel table is constructed. See Table 19. Using the data from the Count Summary Table, this new table is sorted using the Count column as the primary key and the Count Total as the secondary sort key. The top ranked node is the monitoring node that is the most efficient in detecting contamination over the time and concentration criteria that have been stipulated. In other words this node detects contamination from more Injection nodes than any other Monitoring node in the system and, in the event that several monitoring nodes detect the same number of Injection nodes, the top ranked node has the largest number of total detection events. This will be called the Primary Monitoring Node. In most cases though, this monitoring node will not detect contamination at all nodes. In order to identify the additional monitoring points necessary to cover these undetected nodes, an additional table is required.

Table 19 Sorted Summary Table (Partial) – At 12 Hours With 100 Mg/L Threshold

Sorted values:

Node	J20	J30	J40	J50	J60	J70	J150	J160	J170	COUNT	Ave	Total
170	4	2			3	4	1	6	1	11	3	33
140	5	5		1	7	5	2	6		10	4	38
150	5	3			4	4	6	7		9	4	35
130	6				5	5	6	8		8	5	38
160	8	1			8	9	7	10		7	7	50
90	5	5			6	6		1		7	4	26
80	7	8		1	11	7				6	6	38
50	3	4		2	5	2				5	3	16
60	8	8			10	8				4	9	34
100	2									2	2	3
40		1	1							2	1	2
20	1									1	1	1
30		1								1	1	1
70						1				1	1	1
110										1	1	1
120						1				1	1	1

Develop Supplemental Node Search Table

To identify and rank the additional monitoring nodes necessary to cover undetected nodes in the distribution system, set up a table with headings as shown Table 20. This table will list the nodes not covered by the Primary Monitoring Node identified in the Count Sorting Table above. A computer search populates this table as follows.

Using the Count Sorting Table, proceed across the top row (the Primary Monitoring Node) and identify any injection node that has no detections noted. Then drop down that Injection node column to locate the maximum number of detections. From that row, record the monitoring node along with the total number of injection nodes that the monitoring node detected as well as the total number (i.e. sum) of detections. See Table 20. Record in the "Nodes Not Covered" Table along with the number of other Injection nodes detected and also the total number of times that node detected.

Table 20 Nodes Not Covered

Summary of filtered values:

Col Node	Mon Node	COUNT	Total
-			
-			
J40	40	2	2
J50	50	5	16
-			
-			
-			
J90	90	7	26
J100	150	9	35
-			
J120	120	1	1
-			
-			
-			
-			
-			

In order to identify other undetected nodes, return to the Primary Monitoring Node and continue across the top row and repeat the column lookup process described above for any remaining injection nodes that had no detections. This is repeated across the top row until no Injection node is left undetected. An Injection node will detect itself so there will always be at least one monitoring node to detect every Injection node. At the completion of this effort a table will be populated that has all Injection nodes not covered by the Primary Monitoring Node identified along with the Monitoring nodes that do detect them.

At this point the table searches for the monitoring node that most effectively detects each Injection node. This is accomplished by sorting and ranking the nodes. It will be recalled that one of the decision criteria for determining the most efficient set of monitoring nodes will be those nodes that detect the most injection nodes over the system and with the greatest redundancy. Since the Monitoring Nodes listed in the "Nodes Not Covered" Table represent the largest (highest) number of detections for

each Injection node not detected by the Primary Monitoring Node, the table can be used to identify the nodes that have the greatest redundancy for detecting other Injection nodes. Thus, as the nodes in this table are sorted, the resulting monitoring node ranking will be the one that detects an Injection node not covered by the Primary Monitoring Node, and the most other Injection nodes. If there is a tie at this level, the node is selected which has the highest total number of total detects during the period.

At this point the process has identified and ranked all of the monitoring nodes necessary to detect the DBT for the time and concentration threshold specified. It is quite possible that a number of Injection nodes only detect themselves which of course means that the contamination stays localized and does not spread to other parts of the system. At this point the water system managers must make a value judgment as to how many monitoring stations to deploy and what portions of the system are covered and which sections are still vulnerable.

Sensitivity Analysis Based on Time Since Injection and Concentration Threshold

While the process above identifies the set of monitoring points which covers the system for the event duration, it is intuitively obvious that at earlier points in the contamination event the dispersion will not be as extensive as at later times and thus may have an effect on the monitoring set. Also, if different threshold contamination levels are selected, what impact would that have on the best set of monitoring nodes? It is reasonable then to explore the effects of changes in time and threshold concentration on the location of the best set of monitoring nodes.

In order to accomplish this, a system must be devised to repeat the process above in a systematic way while varying the "Time Since Injection" (t_s) and the threshold concentration. This summarizing and sorting of data, based on the new threshold, can be accomplished with Excel by using the Visual Basic for Applications (VBA) scripting language which is a part of the Excel package.

After writing the appropriate script, the Anytown database was queried using Threshold Concentrations of 0, 100, 200, 300 and 400 mg/L. Since some reports (Rose, 1988; Teunis et al., 2002) show that the doses between 10 and 500 oocyst will provide infection of the majority of subjects and if the population drinks 1 to 2 liters per day, this range presents the approximate coverage of a minimum infective zone. Also, the "time since injection" factor was considered by examining the data in 2 hour increments from 6 to 24 hours. It was observed that durations less than 6 hours after injection were not useful because the movement of the contaminant through the system was too limited. This procedure along with the VBA code is listed in Appendix 3.

Interpretation of Results

The analyst now has a means of organizing and ranking monitoring systems that provide coverage of the entire network with varying levels of efficiency. At this point it is now possible to develop additional procedures to assist in the selection of the set of monitoring stations that provide the most robust coverage for the range of assumptions that the analyst believes to be most appropriate.

Number of Monitoring Stations Required

Probably the first question that an operations manager would want to know is 'how many monitoring stations are required to cover the entire system?' As the sensitivity analysis under the varying t_s and C* assumptions indicates, the answer to this question is dependent on both variables. Therefore, the analyst must frame the response by organizing the data into a form that incorporates the impacts of each function. This may be accomplished by counting the number of stations determined in the t_s and C* sensitivity analysis. In other words as the t_s is iterated from t=0 to t=t_d for each C* value, a defined number of monitoring stations is counted to cover the system. By listing the number of stations required for each combination of t_s and C*, a table can be developed to summarize data. This data can then be plotted for convenient comparison of the impact of each variable.

For the Anytown example, Figure 6 is plotted which illustrates the outcome. It shows that as time increases and threshold concentration decreases, the number of monitoring stations decreases from 8 stations to 4 stations with a zero concentration threshold. On the other hand with C* equal to 0, 100, 200 or 300 mg/L, the number of monitoring stations decreases from 9 to 5 stations necessary to cover all nodes. With a 400 mg/L threshold the number of stations required does not fall below 6 stations.

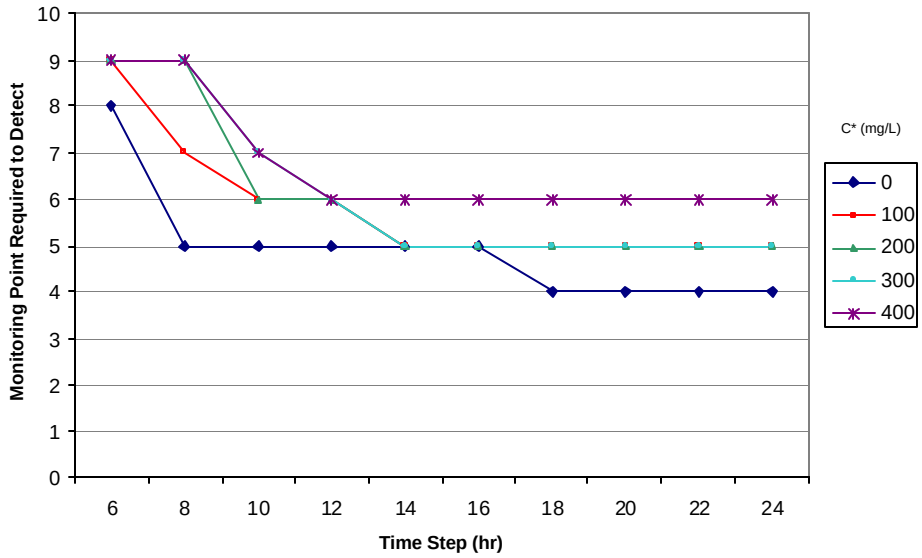


Figure 6 Number of Monitoring Stations Required vs. Time to Detect: Injection at Midnight

With this information one can develop a sense of the resources required to develop a system that will detect contamination introduced at any point in the system. In most real water systems the number of stations required will greatly exceed the capital and human resources available to cover the whole network. Factors such as dead-end lines, pressure zones, lines with high water age and the like, will create areas that have low efficacy from a system monitoring perspective and, thus, will require their own sampling station to detect contamination for that small area. This issue will become more apparent as the methodology seeks to identify and rank the effectiveness of various monitoring nodes.

So, having developed an understanding of the number of nodes required to cover the system, the core issues are now ready to be addressed, i.e.

1. Where are those nodes to be located?

2. If all of the nodes necessary to cover the system cannot be implemented (ex. resource constraints), what nodes should be selected for that smaller subset?
3. If a smaller subset is chosen, what portion of the system is covered; and conversely, what portion will not be covered?

The next section develops the strategy for locating and ranking monitoring nodes in terms of their suitability to detect contamination at other nodes within the pipe network.

Location of Monitoring Stations

Using the information from the sensitivity iterations, it is possible to construct contingency tables to summarize the data and put it into a useful form to aid in evaluation of the options. Excel has an internal "pivot table" function which accomplishes this task and allows convenient plotting of the results. The contingency table allows one to view the ranking of the Primary Monitoring Node and the associated secondary monitoring nodes over time and at varying threshold concentration levels. In this way the stability of the monitoring node set may be observed as the conditions vary. This allows a realistic assessment of the functionality of a particular location sequence which is invaluable when making risk tradeoffs or economic judgments. A few examples illustrating the use of these tables and charts are shown below.

Suppose that the utility operating budget will allow no more than four monitoring stations to cover Anytown. First, it is necessary to determine how many monitoring stations are necessary to cover Anytown and then to locate those stations within the distribution system. Using the sorting and supplemental node search tables it is also possible to determine what sections of town are vulnerable if four monitoring stations are not

sufficient. Recalling that the data is organized using cumulative time (i.e. time since the contamination event), the best set of monitoring stations is determined for the time and threshold concentration selected. For example, the chart above indicates that four monitoring stations will be sufficient to cover the entire town if it is believed that no threshold concentration is applicable. It also shows that five stations will be required if 100 mg/L must be observed before a detection is counted.

An interesting note here is that it is commonly assumed that using a zero threshold for detection is the most conservative assumption in establishing the monitoring network. Actually this is not the case because, from the network design perspective, the contamination "signal" must be strong enough to be confirmed (threshold) and should reject spurious events. Thus, this tends to require more monitoring stations, which yields a more conservative design. Once the system is in place, of course, any detection will be conservatively treated as an event, but that assumption should be used on the operations side, not necessarily on the optimization side.

One way to develop a good overview of the optimum station locations is to look at all times and all thresholds at once. For the midnight time of injection the monitoring combinations are shown in Figure 7.

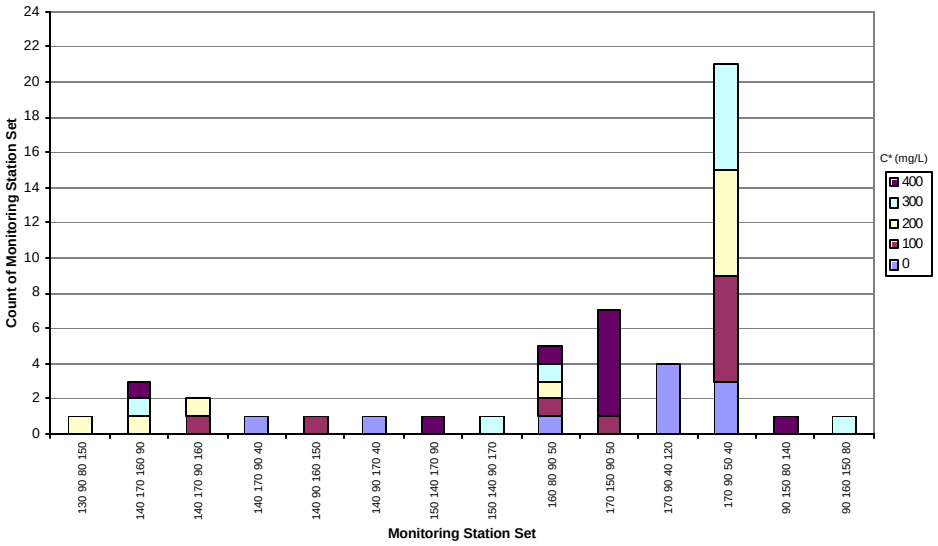


Figure 7 Monitoring Station Configuration: Midnight Contamination Event (Permutation Order)

Node Ranking in Permutation Order

Again remembering that time is cumulative (t=0 to t=t_s) and therefore longer time periods overlap shorter periods, this gives a first cut ordering of the ranked monitoring station locations. The numbers along the X-axis, of course, being the Monitoring Nodes in permutation order (Primary node first followed by ranked nodes). The Y-axis represents the number of injection nodes that the indicated monitoring node combination detects for the given t_s and C*.

The purpose of this graph is to quickly scan the range of data to observe patterns that may exist or to identify any dominant combinations that may exist. For the DBT with a midnight injection at some point within the distribution system, the node sequence of 170/90/50/40 seems to provide the best coverage over all the range of data assumptions.

Node Ranking in Combination Order

It is noticed that the data presentation above has a “permutation” characteristic in that each monitoring set is based on the Primary Monitoring Node holding the key position and other nodes listed in ranked order. This is important when developing the node list, however, when evaluating the efficacy of different monitoring node sets, the internal order doesn’t matter. In other words once a node set is selected for consideration, it is the efficacy of the group that matters since at that point there is no value to the order. So in the example above if monitoring equipment was deployed at nodes 170/90/50/40, it inconsequential whether the computed priority order is 170/90/50/40, 90/50/40/170, 50/40/170/90 or the like. Once those four nodes are selected it is the overall efficiency of the group that matters. For that reason the node sets should be evaluated in “combination” order, as opposed to “permutation” order, when selecting the most robust set of monitoring nodes.

The same information viewed in combination order (equivalent node sets regardless of ranking) is shown in Figure 8. It is noted that the number of columns is reduced from 14 to 11 as some of the nodes sets were combined when priority order was not considered. This sorting operation is done through Excel using the VBA script indicated in Appendix 3.

From this chart it is seen that the node set 40/50/90/170 provides the most robust coverage under all conditions. By incrementing through the time scale it is seen that this sequence stabilizes as the dominant set 12 hours after the contamination event is initiated. Also by observing the supplemental node tables at threshold concentrations above 100 mg/L, it is observed that the area around node 120 will not be covered by this

sequence (see Table 21). Table 21 is also useful in observing the time variance (or stability) of the monitoring sets through the analysis period.

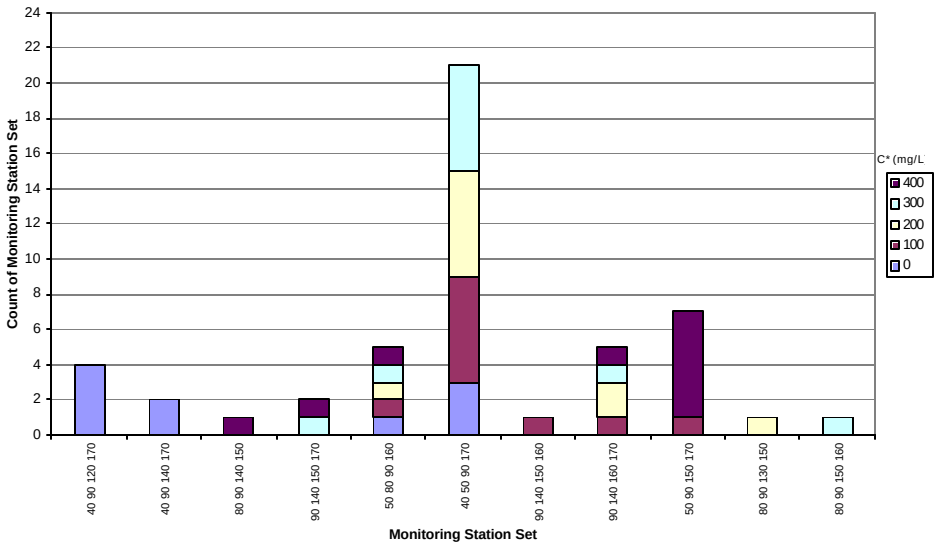


Figure 8 Monitoring Station Configuration: Midnight Contamination Event (Combination Order)

Table 21 Node Ranking for Midnight Injection Scenario

		Rank							
Threshold	Hour	1st	2nd	3rd	4th	5th	6th	7th	8th
0	6	160	80	90	50	140	170	40	120
0	8	140	90	170	40	12			
0	10	140	170	90	40	120			
0	12	170	90	50	40	120			
0	14	170	90	50	40	120			
0	16	170	90	50	40	120			
0	18	170	90	40	120				
0	20	170	90	40	120				
0	22	170	90	40	120				
0	24	170	90	40	120				
100	6	160	80	90	50	140	150	170	40
100	8	140	90	160	150	170	40	120	
100	10	140	170	90	160	40	120		
100	12	170	150	90	50	40	120		
100	14	170	90	50	40	120			
100	16	170	90	50	40	120			
100	18	170	90	50	40	120			
100	20	170	90	50	40	120			
100	22	170	90	50	40	120			
100	24	170	90	50	40	120			
200	6	160	80	90	50	140	150	170	40
200	8	130	90	80	150	140	50	170	40
200	10	140	170	90	160	40	120		
200	12	140	170	160	90	40	120		
200	14	170	90	50	40	120			
200	16	170	90	50	40	120			
200	18	170	90	50	40	120			
200	20	170	90	50	40	120			
200	22	170	90	50	40	120			
200	24	170	90	50	40	120			
300	6	160	80	90	50	140	150	170	40
300	8	90	160	150	80	140	50	170	40
300	10	150	140	90	170	50	40	120	
300	12	140	170	160	90	40	120		
300	14	170	90	50	40	120			
300	16	170	90	50	40	120			
300	18	170	90	50	40	120			
300	20	170	90	50	40	120			
300	22	170	90	50	40	120			
300	24	170	90	50	40	120			
400	6	160	80	90	50	140	150	170	40
400	8	90	150	80	140	160	50	170	40
400	10	150	140	170	90	50	40	120	
400	12	140	170	160	90	40	120		
400	14	170	150	90	50	40	120		
400	16	170	150	90	50	40	120		
400	18	170	150	90	50	40	120		
400	20	170	150	90	50	40	120		
400	22	170	150	90	50	40	120		
400	24	170	150	90	50	40	120		

Sensitivity Analysis Based on Injection Time-of-Day

Because the water demand throughout the system drives the contamination, there is reason to believe that monitoring node selection may be sensitive to fluctuations caused by the diurnal pattern. Therefore, it would seem prudent to explore the implications of this factor on the stability of the monitoring node set.

To illustrate the effect of "Time-of-Day" that the injection occurs, the entire methodology is performed again by setting up a new database for applying the DBT to scenarios with injection beginning at 6 a.m., Noon and 6 p.m. The three new databases allow the development of three new monitoring networks which use the same assumptions as stipulated in the example problem, except that the injection time-of-day has been changed. By comparing the "optimum" monitoring networks for each of the three additional injection "Time-of-Day" injection scenarios, one can explore the impact of the "Time-of-Day" variable and the stability of the midnight injection monitoring node set.

Monitoring Locations Based Upon 6 a.m. Injection

Using the DBT contamination event pattern, each node was modeled with a 6 a.m. injection. The resulting information was used to populate a database which was then analyzed as described above. The results of the analysis are interesting in that they affirm the fact that, as hypothesized, the time-of-day that the injection occurs does in fact have an impact on the monitoring set.

Fortunately, contamination in this instance can ultimately be detected with two monitoring points as shown in Figure 9. Figure 10 and Figure 11 show that Monitoring

nodes 170 and 90 will cover the system effectively. Table 22 illustrates that this pattern stabilizes at hour 16 for all concentration thresholds.

It should be remembered that the time scale remains the same on all “Time-of-Day” scenarios. Because this scenario dataset was generated from a contamination event that started at 6 a.m., the pattern actually stabilized 10 hours after the event initiation.

Table 22 Node Ranking for 6 a.m. Injection Scenario

		Rank							
Threshold	Hour	1st	2nd	3rd	4th	5th	6th	7th	8th
0	12	160	140	90	130	170			
0	14	170	160	90					
0	16	170	90						
0	18	170	90						
0	20	170	90						
0	22	170	90						
0	24	170	90						
0	26	170	90						
0	28	170	90						
0	30	170	90						
100	12	160	140	90	130	170			
100	14	170	160	90					
100	16	170	90						
100	18	170	90						
100	20	170	90						
100	22	170	90						
100	24	170	90						
100	26	170	90						
100	28	170	90						
100	30	170	90						
200	12	160	140	90	130	170			
200	14	170	160	90					
200	16	170	90						
200	18	170	90						
200	20	170	90						
200	22	170	90						
200	24	170	90						
200	26	170	90						
200	28	170	90						
200	30	170	90						
300	12	160	140	90	130	170			
300	14	170	160	90					
300	16	170	90						
300	18	170	90						
300	20	170	90						
300	22	170	90						
300	24	170	90						
300	26	170	90						
300	28	170	90						
300	30	170	90						
400	12	160	140	90	130	170			
400	14	170	160	90					
400	16	170	90						
400	18	170	90						
400	20	170	90						
400	22	170	90						
400	24	170	90						
400	26	170	90						
400	28	170	90						
400	30	170	90						

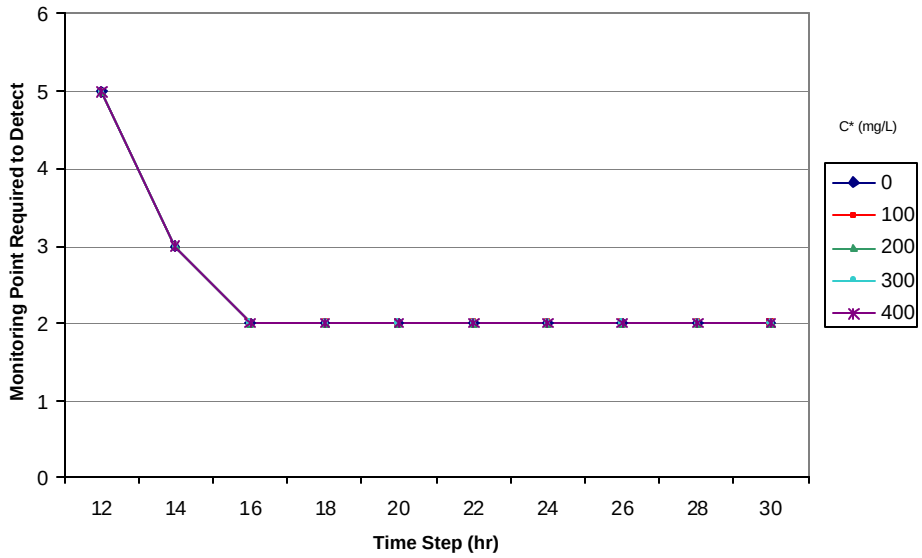


Figure 9 Number of Monitoring Stations Required vs. Time to Detect: Injection at 6 a.m.

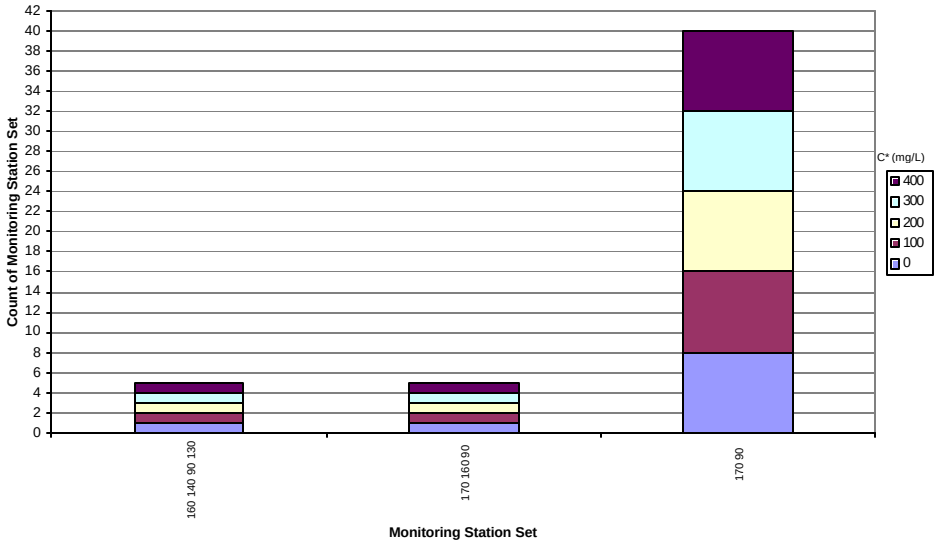


Figure 10 Monitoring Station Configuration: 6 a.m. Contamination Event (Permutation Order)

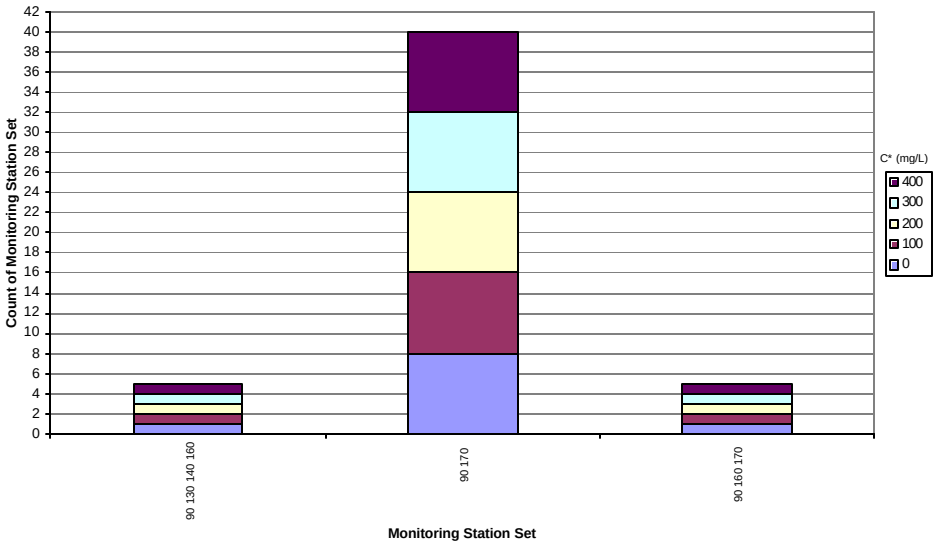


Figure 11 Monitoring Station Configuration: 6 a.m. Contamination Event (Combination Order)

Monitoring Locations Based Upon Noon Injection

Stepping the time of injection forward another six hours, a new database was created with contamination beginning at noon. A set of exhibits similar to those above are developed and presented in Figure 12, Figure 13 and Figure 14 along with Table 23.

Again, it is interesting to see the shift in the monitoring station patterns. In this case it is observed that the distribution system can be covered with a single monitoring node after 10 hours from injection. As before this is a subset of the set selected for the midnight event, but indicates rather dramatically the affect that the diurnal patterns can have on the analysis.

Table 23 Node Ranking for Noon Injection Scenario

		Rank							
Threshold	Hour	1st	2nd	3rd	4th	5th	6th	7th	8th
0	18	160	130	170					
0	20	130	170						
0	22	170							
0	24	170							
0	26	170							
0	28	170							
0	30	170							
0	32	170							
0	34	170							
0	36	170							
100	18	160	130	170					
100	20	170	160						
100	22	170							
100	24	170							
100	26	170							
100	28	170							
100	30	170							
100	32	170							
100	34	170							
100	36	170							
200	18	160	130	170					
200	20	170	160						
200	22	170							
200	24	170							
200	26	170							
200	28	170							
200	30	170							
200	32	170							
200	34	170							
200	36	170							
300	18	160	130	170					
300	20	170	160						
300	22	170							
300	24	170							
300	26	170							
300	28	170							
300	30	170							
300	32	170							
300	34	170							
300	36	170							
400	18	160	130	170					
400	20	170	160						
400	22	170							
400	24	170							
400	26	170							
400	28	170							
400	30	170							
400	32	170							
400	34	170							
400	36	170							

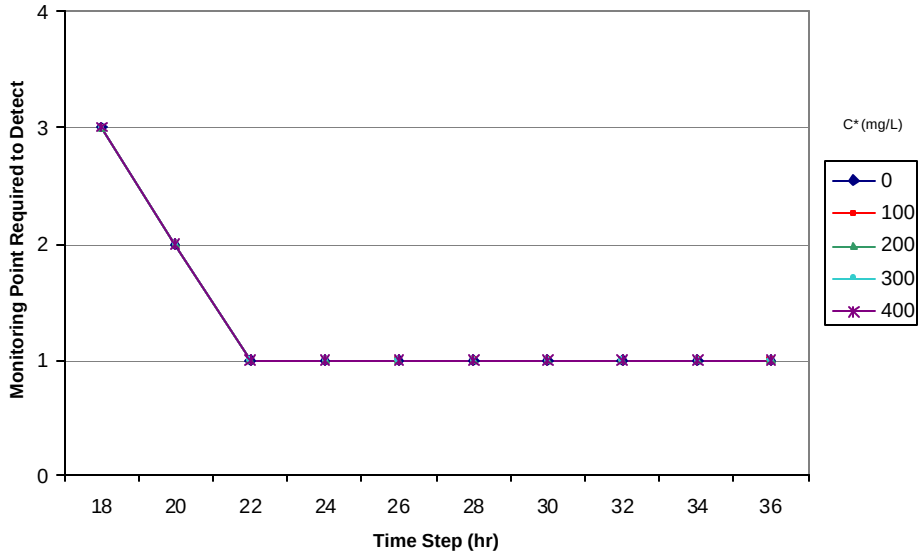


Figure 12 Number of Monitoring Stations Required vs. Time to Detect: Injection at Noon

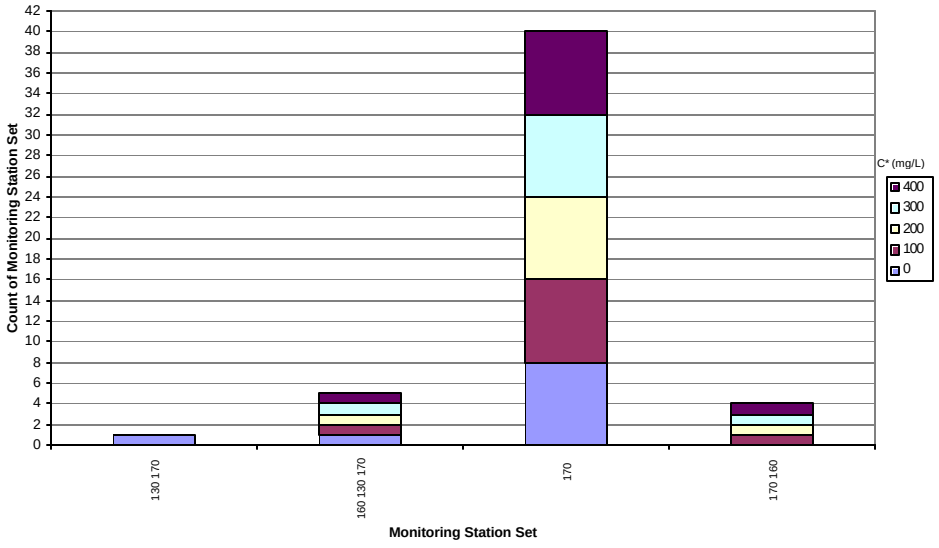


Figure 13 Monitoring Station Configuration: Noon Contamination Event (Permutation Order)

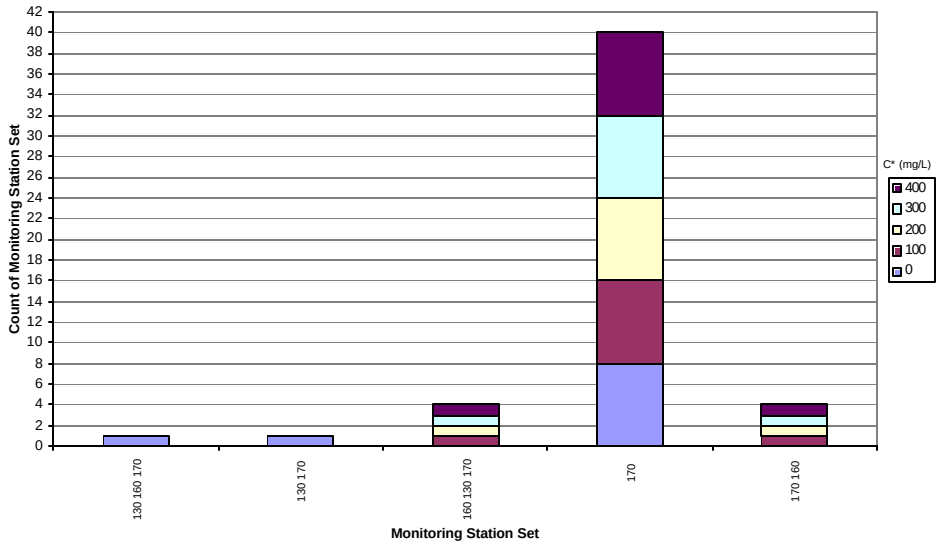


Figure 14 Monitoring Station Configuration: Noon Contamination Event (Combination Order)

Monitoring Locations Based Upon 6 p.m. Injection

Finally, the initiation time is moved forward another six hours and the full set of injection nodes replicated sequentially beginning at 6 p.m. The new database is generated and the sorting analysis recompiled. As with the 6 a.m. and noon analyses the optimum monitoring set differs from the midnight contamination scenario. However, as with the other initiation times, the selected set for the 6 p.m. event remains subsumed within the midnight set. The ranking results for the scenario are present in Table 24.

Table 24 Node Ranking for 6 p.m. Injection Scenario

		Rank							
Threshold	Hour	1st	2nd	3rd	4th	5th	6th	7th	8th
0	24	160	130	80	170				
0	26	160	130	170					
0	28	170	160						
0	30	170							
0	32	170							
0	34	170							
0	36	170							
0	38	170							
0	40	170							
0	42	170							
100	24	160	140	130	80	170			
100	26	160	130	170					
100	28	170	160						
100	30	170							
100	32	170							
100	34	170							
100	36	170							
100	38	170							
100	40	170							
100	42	170							
200	24	160	140	130	80	170			
200	26	160	130	170					
200	28	170	160						
200	30	170							
200	32	170							
200	34	170							
200	36	170							
200	38	170							
200	40	170							
200	42	170							
300	24	160	140	130	80	170			
300	26	160	130	170					
300	28	170	160						
300	30	170							
300	32	170							
300	34	170							
300	36	170							
300	38	170							
300	40	170							
300	42	170							
400	24	160	140	130	80	170			
400	26	160	130	170					
400	28	170	160						
400	30	170							
400	32	170							
400	34	170							
400	36	170							
400	38	170							
400	40	170							
400	42	170							

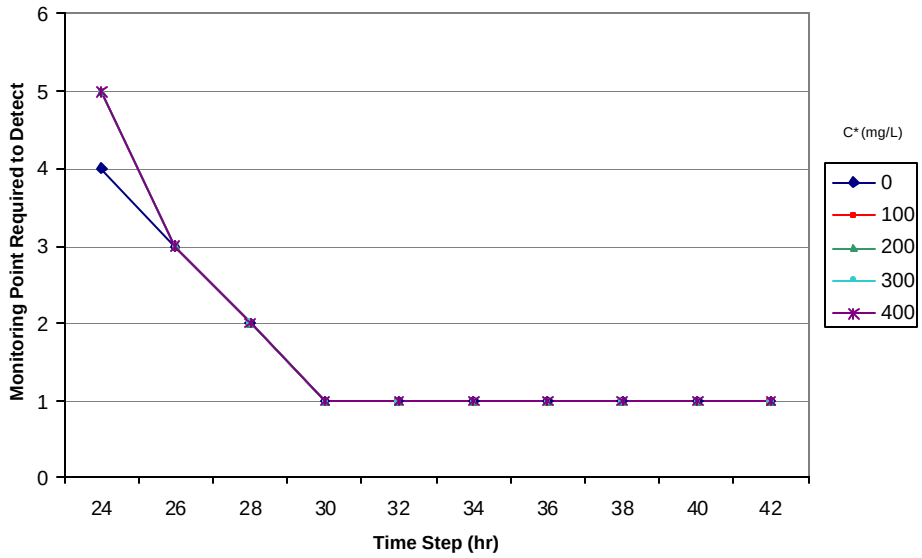


Figure 15 Number of Monitoring Stations Required vs. Time to Detect: Injection at 6 p.m.

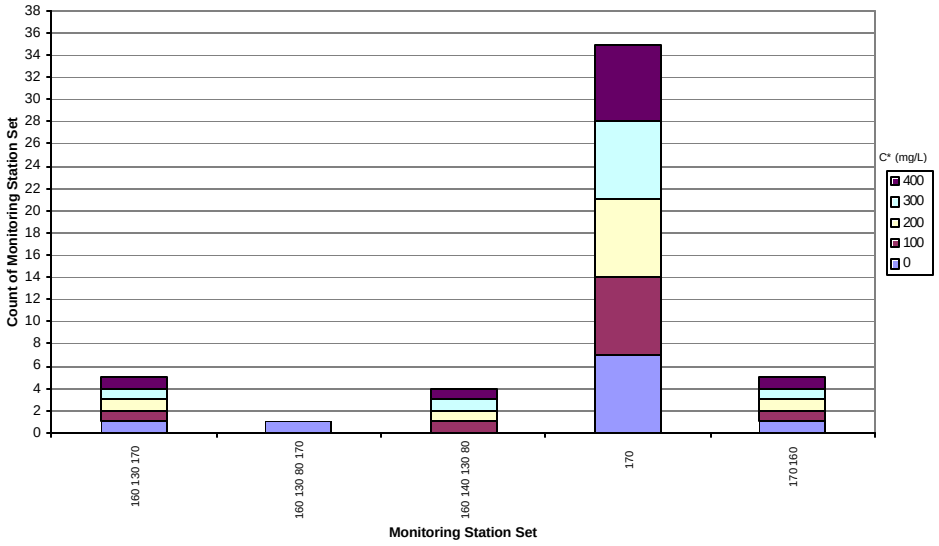


Figure 16 Monitoring Station Configuration: 6 p.m. Contamination Event (Permutation Order)

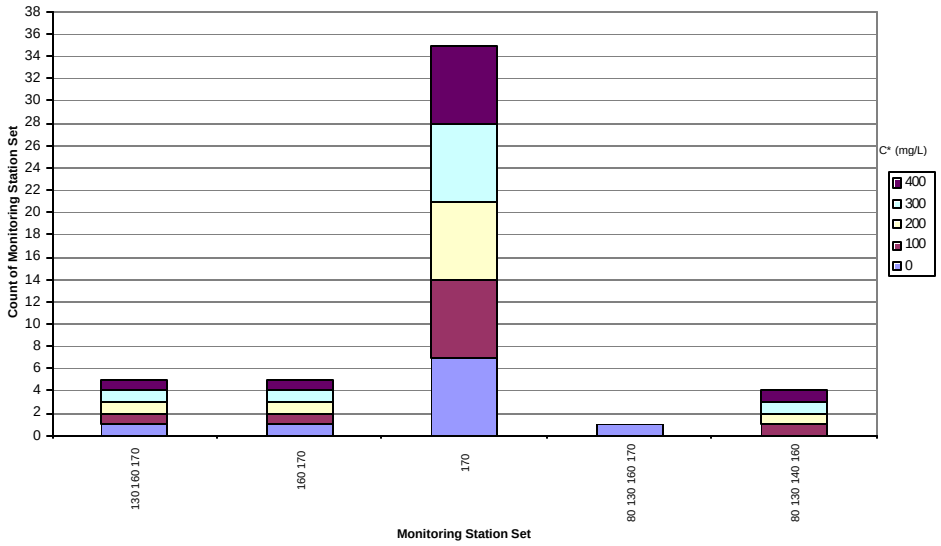


Figure 17 Monitoring Station Configuration: 6 p.m. Contamination Event (Combination Order)

Discussion

Examination of Figure 8, Figure 11, Figure 14 and Figure 17 shows that the time-of-day at which injection occurs can be a significant factor in the monitoring set analysis. Diurnal hydraulic patterns cause changes in flow velocity and direction which can result in differences in timing and extent of contaminant plumes. However, in this system the vulnerability is not quite as severe as might have been indicated by just an analysis of the midnight injection alone. Accordingly, the vulnerability of a system should be measured not only in a space dimension, but also in a time dimension. The implication here being that the level of monitoring resources required may not be the same through time.

It is also observed that some nodes may predominate in detection efficiency during the day. In this illustration Node 170 is clearly the most important node in the monitoring network. Aside from being the primary node for the Midnight injection, Node 170

stabilizes as the only node necessary to detect contamination for the whole system 10 hours after a noon injection with approximately the same results for injection occurring at 6 am and 6 pm.

An application of this result would be that particular care should be taken to protect this node and keep it operating at peak performance. While it is highly unlikely that this level of dominance will be evident in larger networks, certain nodes will be more efficient, and therefore, more critical than others in managing risks.

While this small illustration will not necessarily characterize every system in that all monitoring sets will be subsumed in the "worst case" or master cluster of monitoring points, it is hypothesized that there will be a tendency for this to occur in systems that have stable operation and regular diurnal patterns. Systems with more substantial temporal discontinuities such as significant seasonal or tourist impacts, large industrial demands with varying production schedules, etc. are more likely to result in varying focus points.

The analysis also shows when the water system is the most vulnerable. Because more nodes are required for detection during the midnight event and because it takes longer for the selected monitoring set to detect the range of events, the utility manager is given an indication of the time of day and areas that are most vulnerable. This information can be used to inform police or other security services of necessary patrol patterns which will provide additional protection for their system.

Summary

An illustrative example has been developed which constructs an analysis of a small water system. The example uses commercially available off-the-shelf software (WaterCAD and Microsoft Excel) to perform the analysis. The process is not dependant on any proprietary software and can be adapted to any small system.

The sample problem steps through a heuristic process that organizes the extended period water quality model data into a form that can be manipulated to tease out sets of monitoring stations that outperform others in terms of detecting potential contamination occurring at any node in the system.

The process demonstrates and provides a means of dealing with the variables of Concentration threshold (C*), Time Since Injection (t_s), and Time of Day of Injection. Each of these variables should be explored in any analysis to determine the magnitude of its influence on the monitoring set. While in some cases the selected monitoring set may fluctuate given the weight assigned to each variable, this should not necessarily be viewed as a negative consequence.

Recall that the overall objective is, given a fixed number of monitoring stations, to select a set of monitoring nodes that most efficiently detects contamination which can occur anywhere in the distribution system at any time. In most real situations it is not likely that the selected set will provide complete coverage of the distributions system. This analysis also provides direction to utility managers regarding susceptible zones and times of day that the system has higher risk than others. This information can in turn be

used to implement institutional or operational control which can further reduce overall system vulnerability.

Chapter 8 Conclusions and Recommendations for Future Research

Summary and Conclusions

Water is a fundamental resource, not only for life, but for a community's social and economic structure. Therefore, protection of a community's water system is crucial to its well-being. Over the years regulations, design procedures and operational practices have evolved in this country to safeguard the water delivered to each customer. By and large this effort has been successful given the reduction in waterborne diseases and reliability of delivery and overall water quality. However, representative sampling of water systems to ensure adequate water quality has always been complex and technically challenging. With the paradigm shift that occurred subsequent to the events of September 11, 2001, it is necessary to reevaluate the processes and procedures used to detect contamination in our public water systems.

It is the thesis of this study that the distribution system itself is particularly vulnerable to contamination, especially intentional contamination, which has not generally been recognized in the water industry. The procedures that have been sufficient to monitor systems in the past are not adequate to detect, respond and recover from significant mass contamination events that can occur within the distribution system itself. This results primarily from the following general beliefs prevalent in the water industry.

First, the long term experience in the water industry has generally proven that if the structural integrity of the pipe grid is maintained, drastic changes in plant effluent water

quality would not occur. Consequently, regulatory and operations management primarily has been focused on cross-connection, backflow prevention and corrosion control, as well as strict standards for disinfecting and testing new pipe connections. This belief was founded on the facts that the pipe network (1) by and large is buried and difficult to access, (2) it has a positive backpressure (20 psig minimum) which makes it difficult for contaminants to enter the system and (3) with backflow and cross-contamination regulations and programs in place have been effective in protecting network integrity.

In addition, a chlorine residual is maintained within the distribution system which provides nominal disinfectant capability throughout the pipe network. The residual also serves as a sentinel for detecting contaminant intrusion. Finally, monitoring in accordance with regulatory and "professional practice" guidelines was felt to provide adequate quantification of water quality within the system.

The documented epidemiological record of waterborne illnesses in the United States has shown a high level of progress and success over the past century. While periodic events will occur, the overall record speaks for itself. This applies to both chemical and biological etiologies.

However, as this study has probed the regulatory requirements and general technical literature which provides the guidance by which distribution system sampling is structured, a void is identified relative to intentional contamination. This study has illustrated that the programs that can function acceptably when considering normal system failures or even significant naturally occurring disasters, are ill-equipped to deal with intentional mass contamination events. This results from several factors:

1. Most routine monitoring is accomplished primarily through indicator or surrogate parameters and cannot detect specific contaminants in real time, and in most cases the results will not be known for at least three to five days.
2. If the technology does exist to detect a set of contaminants within one hour or less, the financial and operational ability to deploy that technology makes it impracticable, especially in small to mid-sized systems.
3. Guidance on location of monitoring stations is not focused on searching for optimum locations based on flow and water quality patterns.

Recognizing the deficiency of current monitoring practice to quickly detect an acute, massive contamination event, water system managers need assistance in developing an approach to identify a robust, if not optimized, set of monitoring stations. When one examines the regulations and literature there is little, if any, explicit guidance instructing water system managers on where to locate water quality monitoring stations to detect mass contamination within a distribution system.

Locating a robust set of monitoring locations is important because monitoring stations are expensive to purchase and operate. Plus, a large number of stations generating streams of real-time information can produce a massive volume of data that could create data management and interpretation problems. Thus, the objective of this study is to maximize coverage of the network with the minimum number of stations, again remembering that sampling for this purpose has a different focus than for the TCR.

The characteristics of an ideal contaminant were defined as those that are essentially undetectable by the senses (e.g. colorless, tasteless, and odorless), highly soluble or

finely divided and highly toxic. The contaminant can be either biological or chemical. A review of pertinent literature has been performed to identify agents that have a reasonable likelihood of being used to contaminate a public water supply. Logistical, toxicity and/or infectivity considerations reduce the number of potential candidates to a more or less manageable number, if the purpose of the contamination event is to impose a health risk to a large number of people. However, it should be pointed out that the literature does not explicitly address fate of most agents under conditions found in a potable water distribution system. Considerable research is necessary to address the paucity of data along those lines. Having said that, it would seem that *Shigella, Salmonella*, and *Vibrio cholerae* are viable bacterial threats although they do have susceptibility to reasonable chlorine levels.

It also appears that *B. anthracis* could be a very potent and difficult agent to eradicate within a water system. While the vast majority of attention is focused on its inhalational risk, the threat from ingestion may also present a worrisome risk. Additional research on this agent and this route of exposure is encouraged.

Cryptosporidium or *Giardia* should also be considered because of their chlorine resistance and low LD_{50} levels. Additional research with toxins in water systems is warranted, but it appears that *Botulinum*, Aflatoxin and Ricin are the most likely to be weaponized at this point.

Likewise, chemical agents have considerable gaps with regard to the ingestion route. The current thinking in many venues is that some form of cyanide, arsenic or orthophosphates holds the most potential as viable contamination agents. The issue of

detectability for some of these agents could reduce their utility. However, these certainly have characteristics that can manifest adverse health effects and should be considered for monitoring.

A few caveats are worthy of note. First, the literature sources used in this study are in the public domain. Other classified agents that are not searchable are likely to exist. Also, this study does not consider pharmaceuticals or explicit hallucinogenic drugs which may result in an effect desired by terrorists. While this means that additional agents could exist, the knowledge of and accessibility to such agents, as well as their logistical constraints, hopefully, would significantly limit their utility.

Due to the complexity of water distribution systems there is no practicable way to physically test a system to isolate the effective monitoring nodes. However, the speed and power of computers and network analysis software have developed to the point that it is now possible to develop mathematical models of those systems. They are also capable of managing the large datasets that are produced in ways that were not practicable in the past. This coupled with the fact that the cost of these tools has dropped to the point that it is accessible to the target audience, i.e. small to mid-sized water systems. Accordingly, the tools now exist to begin to address the problem in a more scientific fashion.

This study has proposed a heuristic algorithm targeted for small to mid-sized water systems which identifies a robust set of monitoring station locations capable of detecting mass contamination occurring within the distribution system itself. The algorithm is structured to treat every node in turn as a contaminant source and model the fate of the

contaminant through the system over time. The results of these simulations are then organized and analyzed to search for the number of nodes required to cover the system and identify which node sets demonstrate stability over a range of detection times and concentration thresholds. This process is then applied to a small test distribution system using off-the-shelf commercially available software to illustrate its applicability to small or mid-sized water systems.

Following the pattern used by other studies in the literature search, initial configurations of the algorithm sought to optimize the monitoring node set by using a fixed analysis time. Usually, this was tied to a diurnal sequence and, given the outcome of that one set of conditions, the monitoring locations were established. When the early versions of this algorithm were developed, it solved the problem in the same way. However, one of the advantages of structuring the problem in the manner discussed in this paper is that one is able to examine and manipulate the raw data directly. This allows full use of all of the modeling software to describe an event, or combination of events, of interest to the analyst without being constrained by the optimization process. While this can create additional work in the modeling effort, current computer power and software routines can manage those requirements for the size of system targeted in this study.

Using this approach and examining the results raised the question as to whether an arbitrarily set time of analysis was warranted. Although some discussion occurs in the literature regarding minimum time to detect contamination, none of the other papers have suggested an approach that explicitly considers “Time Since Injection” (t_s) as a variable for analysis. However, after looking at the raw data, and upon further consideration of the context, it seemed that expanding the algorithm to include this factor

would be prudent. The results did, in fact, confirm that t_s is important in developing a sense of the stability of the monitoring set. Accordingly, this study recommends that future optimization processes consider t_s when establishing the monitoring set.

Again, analysis of the raw data set and consideration of the context suggested that it might be appropriate to consider a Threshold Concentration (C*) factor in the analysis. This resulted from two observations. First, the mathematical computations would at times produce very small values for some nodes. These values, though greater than zero, seemed to be more computational artifacts than significant concentrations. Thus, if a selection rule was established that noted any cell with a value greater than zero, these cells would be selected and could bias the results with insignificant readings. Accordingly, the algorithm was expanded to allow the analyst to set a C* value such that any cell whose computed concentration was less than C* would be rejected from the standpoint of influencing the monitoring station set. There could, of course, be situations where an analyst wished to consider all data in which case one would set C* = 0.

Another reason for establishing a C* variable is to permit future use of a NOAEL or other threshold concentration effect level (See Appendix 2). Thus, if the examination of a system included some assumption of an acceptable concentration level the analysis could accommodate such a consideration. A cautionary note in this regard is warranted. While the computer simulation of water networks has developed considerably, there are still many assumptions and sources of error which must be recognized. Therefore, as with many epidemiological studies, the C* consideration should be viewed more as an "indication" of effect rather than an absolute value. In other words depending on the nature of the contaminant and the C* selected, a water system should not necessarily be

considered “clean” if all concentration values are below C*. Rather, C* provides a means of examining the sensitivity of an optimum monitoring station set to small concentration values.

The implications of this note are interesting. The inability to compute and evaluate concentrations at absolute levels is not necessary to effectively use the process described in this paper to locate the monitoring stations. While experimenting with different DBT scenarios during the initial phases of this study, it was realized that the relationships influencing node selections are relative. The DBT selected influences the magnitude of the concentration spikes, and thus has an impact from a NOAEL or action level point of view, but the locations of the monitoring stations remain more or less insensitive to that variable. Thus, the monitoring stations can be selected using any hypothetical event while remaining more or less insensitive to the nature of the contamination event itself. Of course, the more that contaminants exhibit significant growth or decay characteristics relative to each other, an effect may be noticed. However, modeling that type of behavior is imprecise in its own right and has its own set of instability that must be evaluated.

Use of C* as an analysis variable also pointed out a fact that the current literature does not seem to adequately consider. Most studies use the assumption that the most conservative form of analysis must consider any positive concentration value at any node to define the optimum monitoring set. Actually, this is not true. When defining a more rigorous monitoring set, using a C* reduces the number of detection data points and requires a “stronger” data signal to be considered. Consequently the analysis relies on fewer data points. As one considers this implication, using a C* will tend to require

more sampling stations. More sampling stations in this context are more conservative than fewer sampling stations. The reason this point is overlooked is that this assumption (consideration of all positive concentration values) is the most conservative assumption when actually monitoring the distribution system, i.e. as an operational rule. It is not, however, the most conservative assumption when designing the monitoring set.

Finally, an outcome of this research emphasized the need to consider "Time of Day of Injection" (TDI) as a variable for consideration. Because diurnal variations exist within all community water systems, the direction and magnitude of flow within the pipes will vary throughout the day. The literature to this point in time has tended to assume a time of injection and compute a monitoring set using the approach proposed in that paper. This study used the algorithm proposed in Chapter 6 and went a step farther by examining the effect of TDI on the proposed monitoring set. It was determined that TDI can, in fact, have a significant impact on the monitoring set and should be considered in any methodology that seeks to truly define the rigor, if not the validity, of a specified monitoring location set.

Areas of Future Research

This dissertation has provided an approach and solution to the research questions identified in Chapter 1. However, while providing a direct response to the questions, further research is warranted to enhance the approach to improve its utility and expand the theoretical basis underlying portions of the algorithm. More specifically, these recommendations are as follows.

It is clear that additional research into the impact and fate of various contaminants within water systems is necessary. As this study has indicated the discussion of fate must also include the conditions existent within the distribution system itself. It is, again, the thesis of this dissertation that the distribution network is actually the most vulnerable component of the water system and must be considered as a potential point of contaminant injection. The limited studies to date have focused on the contaminant removals or inactivation by water treatment processes. With consideration of the nature of the distribution system environment, guidance relative to fate, survivability, toxicity or infectivity of the agent, especially via the ingestion route, is necessary.

Related to this, continued research is necessary to develop the sensing devices that can detect concentration variances at acceptable cost, turnaround time, durability, and operational stability for the parameters in question. Much work is proceeding in this area and with advances in nanotechnology, miniaturization and biotechnology instruments to meet the assumptions of this paper seem within reach. Obviously, the more parameters that a single instrument can detect the more desirable it is. Interestingly, many of the issues seem to be related as much to power (battery life) as to technical issues of detecting a contaminant.

One assumption implicit in this methodology is that each node is assigned equal important in the screening process. From a pragmatic standpoint this is not true. For example, a node demand which is used primarily for non-potable uses (ex. irrigation) may not be as critical as one that is sensitive for water quality (ex. hospital, key government buildings, etc.) Therefore, an expansion of this model may be to develop

and apply a means of weighting node sensitivity and importance in terms of stipulated outcomes.

This study used Microsoft Excel® as the primary tool to manipulate the data, because it is more transparent than a relational database and, therefore, more suitable for examination in a study such as this. However, because Excel® has inherent data size limitations, the next phase in developing this algorithm will be to improve the data handling characteristics by use of more sophisticated database structures and automation features.

Associated with the improved database structure would be the extension of the algorithm to consider monitoring station location on the basis of "Volume Consumed" as well as "Time Since Injection". This dissertation used t_s as the initial basis of location because it is more straightforward and is an important criterion in detection efficiency. However, if the ultimate objective of monitoring is to minimize adverse health effects, then the volume consumed is a more critical parameter than t_s. While the two are related they are not necessarily equivalent and thus consideration should be given to adapting the algorithm to weighting the selection process to volume rather than time. With improved data storage and management this task should be feasible.

With regard to possibilities for improving the underlying theoretical basis for the algorithm, it seems possible to treat the outcome data as a Markov Chain process. The data, especially using C* to create a binomial response, seems amenable to processing using that mathematical approach. This could lead to analysis and decision processes that would be more efficient and rigorous than those proposed here. Genetic algorithms

are being proposed in some of the more recent papers (Ostfeld & Salomons, 2003) and have shown some potential for searching for optimum solutions. However, the Markov Chain approach has not been proposed in any of the papers viewed to date.

Finally, research should be initiated relative to incorporating this work into a public health surveillance program. The purpose of water quality monitoring is to provide warning of anomalous events as early as possible in order to protect public health and property. While this study, hopefully, makes a contribution in terms of assisting system managers in strategically placing the sensing devises, there is much work that needs to be done to set this in an operational framework.

The SDWA has detailed requirements regarding consumer notification of MCL exceedances. However, as the frequency and number of parameters monitored increases, the likelihood of false positives can create nuisance, in not disruptive, notifications. Decision rules and processes must be developed to give guidance to water system operators to adequately balance these conditions.

In the past there has been little actual operational communications between the utilities and the public health/medical community. With the passage of the Public Health Security and Bioterrorism Preparedness and Response Act of 2002, this deficiency is beginning to lessen through the Vulnerability Assessment process, however, there is still much work to be done. Just as water system monitoring is not sufficiently responsive for acute massive events, the public health network is believed to suffer from the same deficiencies.

Future research directed at developing a surveillance and communications model with an emphasis on two-way communication would seem to be useful. In other words the model would not just promote communication from the public health/medical community to the utility by identifying waterborne health effects, but via improved short-term monitoring the utility could report waterborne anomalies to the public health and medical communities that could be used in patient notification and care.

This point is believed to be of importance because it is possible that the engineering aspects of detection may be more effective than discovery via the clinical perspective, especially for biological agents. This follows from the fact that many biological agents have an incubation period measured in days if not weeks. If an efficient water system monitoring program was established, detection could be noted in less than a day (at least for covered nodes). If identified agents were communicated to the medical community, this information could provide valuable insight into diagnosis and treatment as well as provide critical response time for public notification, supplemental supplies and support, if necessary. The primary issue being highlighted here is whether there exists a functional partnership between these important social entities that is sufficient to defeat or mitigate a massive contamination event.

References

Aldrich, T., Griffith, J., & Cooke, C. (1993). *Environmental epidemiology and risk assessment*. New York, NY: Van Nostrand Reinhold.

Armstrong, G. L., Conn, L. A., & Pinner, R. W. (1999). Trends in infectious disease mortality in the United States during the 20th century. *JAMA, 281*(1), 61-66.

Aurora Pump. (1989). AP-3eng product catalog (p. 410). North Aurora, IL: Pentair Pump Group.

AWWA. (1990). *Water quality and treatment: A handbook of community water supplies* (4th ed.). New York, N.Y.: McGraw-Hill, Inc.

AWWA. (1991). *Guidance manual for compliance with the filtration and disinfection requirements for public water systems using surface water sources*. Denver, CO: American Water Works.

AWWA. (1999). *Water quality and treatment: A handbook of community water supplies* (5th ed.): McGraw-Hill.

AWWA, & ASCE. (1998). *Water treatment plant design* (3rd ed.). New York, NY: McGraw-Hill.

Bahadur, R., Samuels, W. B., Grayman, W., Amstutz, D., & Pickus, J. (2003). *Pipelinenet: A model for monitoring introduced contaminants in a distribution system.* Paper presented at the World Water & Environmental Resources Congress, Philadelpia, PA.

Berger, B. B., & Stevenson, A. H. (1955). Feasibility of biological warfare against public water supplies. *J. AWWA, 47*(2), 101-110.

Berry, J., Fleischer, L., Hart, W., & Phillips, C. (2003, July 23-26, 2003). *Sensor placement in municipal water networks.* Paper presented at the Water World & Environmental Resources Congress, Philadelphia, PA.

Besner, M.-C., Gauthier, V., Servais, P., & Camper, A. (2002). Explaining the occurrence of coliforms in distribution systems. *J. AWWA, 94*(8), 95-109.

Boccelli, D. L., Tryby, M. E., Uber, J. G., Rossman, L. A., Zierolf, M. L., & Polycarpou, M. M. (1998). Optimal scheduling of booster disinfection in water distribution systems. *Journal of Water Resources Planning and Management, 124*(2), 99-111.

Boulos, P. F., Altman, T., Jarrige, P.-A., & Collevati, F. (1995). Discrete simulation approach for network-water-quality models. *Journal of Water Resources Planning and Management, 121*(1), 49-60.

Boulos, P. F., Lansey, K. E., & Karney, B. W. (2004). *Comprehensive water distribution systems analysis handbook for engineers and planners* (1st ed.). Pasadena, CA: MWH Soft, Inc.

Brater, E. F., & King, H. W. (1976). *Handbook of hydraulics for the solution of hydraulic engineering problems* (6th ed.). New York, NY: McGraw-Hill Book Company.

Burrows, W. D., & Renner, S. E. (1999). Biological warfare agents as threats to potable water. *Environmental Health Perspectives, 107*(12), 975-984.

Camp, T. R., & Meserve, R. L. (1974). *Water and its impurities* (2nd ed.). Stroudsburg, PA: Dowden, Hutchinson & Ross, Inc.

Camper, A. K., Brastrup, K., Sandvig, A., Clement, J., Spencer, C., & Capuzzi, A. J. (2003). Effect of distribution system materials on bacterial regrowth. *J. AWWA, 95*(7), 107-121.

Camper, A. K., Warnecke, M., Jones, W. L., & McFeters, G. A. (1998). *Pathogens in model distribution system biofilms*. Denver, CO: AWWARf and AWWA.

CDC. (2000). MMWR: Biological and chemical terrorism: Strategic plan for preparedness and response (Vol. 49 RR-4, pp. 1-14).

Chastain, J. R., Jr. (1998). Orientation to health risk assessment for practicing engineers. *Practice Periodical of Hazardous, Toxic and Radioactive Waste Management, 2*(1), 22-32.

Clark, R. M., & Grayman, W. M. (1998). *Modeling water quality in drinking water distribution systems*. Denver, CO: American Water Works Association.

Clark, R. M., Grayman, W. M., & Wymer, L. J. (1993). Effect of the distribution system on drinking water quality. *J. Water Supply Research and Technology-AQUA, 42*(1), 30-38.

Craun, G. F. (1988). Surface water supplies and health. *J. AWWA, 80*(2), 40-52.

Craun, G. F., & Calderon, R. L. (2001). Waterborne disease outbreaks caused by distribution system deficiencies. *J. AWWA, 93*(9), 64-75.

Deininger, R. A. (2000). *The threat of chemical and biological agents to public water supply systems* (No. TSWG T-1211). Ann Arbor, MI: The University of Michigan.

Donlan, R. M. (2002). Biofilms: Microbial life on surfaces. *Emerging Infectious Diseases, 8*(9), 881-890.

Geldreich, E. E. (1996). *Microbial quality of water supply in distribution systems*. Boca Raton, FL: CRC Press, Inc.

Grayman, W. M., Rossman, L. A., Arnold, C., Deininger, R. A., Smith, C., Smith, J. F., et al. (2000). *Water quality modeling of distribution system storage facilities* (No. 90774). Denver, CO: AWWARf and AWWA.

Haas, C. N., Rose, J. B., & Gerba, C. P. (1999). *Quantitative microbial risk assessment*. New York, NY: John Wiley & Sons, Inc.

Haestaed, Walski, T., Chase, D., Savic, D., Grayman, W., Beckwith, S., et al. (2003). *Advanced water distribution modeling and management* (1st ed.). Waterbury, CT: Haestad Methods, Inc.

Haestaed Methods. (2002). *WaterCAD v6 user's guide*. Waterbury, CT: Haestad Methods, Inc.

Haimes, Y. Y., Matalas, N. C., Lambert, J. H., Jackson, B. A., & Fellows, J. F. (1998). Reducing vulnerability of water supply systems to attack. *Journal of Infrastructure Systems, 4*(4), 164-177.

Harberg, R. J. (1997). *Planning and managing reliable urban water systems*. Denver, CO: American Water Works Association.

Health Research, I. (1997). *Recommended standards for water works*. Albany, NY: Health Education Services.

Herwaldt, B. L., Craun, G. G., Stokes, S. L., & Juranek, D. D. (1992). Outbreaks of waterborne disease in the United States: 1989-90. *J. AWWA, 84*(4), 129-135.

Inglesby, T. V., Henderson, D. A., Bartlett, J. G., Ascher, M. S., Eitzen, E., Friedlander, A. M., et al. (1999). Anthrax as a biological weapon: Medical and public health management. *JAMA, 281*(10), 1735-1745.

Islam, M. R., Chaudhry, M. H., & Clark, R. M. (1997). Inverse modeling of chlorine concentration in pipe networks under dynamic conditions. *Journal of Environmental Engineering, 123*(10), 1033-1040.

Kahn, A. S., Morse, S., & Lillibridge, S. (2000). Public health preparedness for biological terrorism in the USA. *The Lancet, 356,* 1179-1182.

Karassik, I. J., Krutzsch, W. C., Fraser, W. H., & Messina, J. P. (Eds.). (1976). *Pump handbook*. New York, NY: McGraw-Hill, Inc.

Kessler, A., Ostfeld, A., & Sinai, G. (1998). Detecting accidental contamination in municipal water networks. *Journal of Water Resources Planning and Management, 124*(4), 192-198.

Kirmeyer, G. J., Friedman, M., Martel, K., Thompson, G., Sandvig, A., Clement, J., et al. (2002). *Guidance manual for monitoring distribution system water quality*. Denver, CO: AWWA and AWWARf.

Klaasen, C. D. (Ed.). (2001). *Cassarett and Doull's toxicology: The basic science of poisons* (6th ed.). New York, NY: McGraw-Hill, Inc.

Kucera, D. J. (2003). Safe drinking water act. In T. F. P. Sullivan (Ed.), *Environmental law handbook* (17th ed.). Rockville, MD: Government Institutes.

Kumar, A., Kansal, M. L., & Arora, G. (1997). Identification of monitoring stations in water distribution system. *Journal of Environmental Engineering, 123*(8), 746-752.

Kumar, A., Kansal, M. L., & Arora, G. (1999). Discussion of 'detecting accidental contamination in municipal water networks' by A. Kessler and A. Ostfeld. *Journal of Water Resources Planning and Management, 125*(5), 308-310.

Larson, T. E. (1966). Deterioration of water quality in distribution systems. *J.AWWA, 58*(10), 1307-1316.

Lee, B. H., & Deininger, R. A. (1992). Optimal locations of monitoring stations in water distribution systems. *Journal of Environmental Engineering, 118*(1), 4-17.

Lee, B. H., Deininger, R. A., & Clark, R. M. (1991). Locating monitoring stations in water distribution systems. *J.AWWA, 83*(7), 60-66.

Lindley, T. R., & Buchberger, S. G. (2002). Assessing intrusion susceptibility in distribution systems. *J. AWWA, 94*(6), 66-79.

Logsdon, G. S., Schneider, O. D., & Budd, G. C. (2004). Hindsight is 20/20: Using history to avoid waterborne disease outbreaks. *J.AWWA, 96*(7), 66-74.

Marrs, T. C., Maynard, R. L., & Sidell, F. R. (1996). *Chemical warfare agents: Toxicology and treatment*. West Sussex, England: John Wiley & Sons, Ltd.

Mays, L. W. (Ed.). (1999). *Hydraulic design handbook*. New York, NY: McGraw-Hill.

Mays, L. W. (Ed.). (2000). *Water distribution systems handbook*. New York, NY: McGraw-Hill, Inc.

McFeters, G. A. (1990). *Drinking water microbiology: Progress and recent developments*. New York, NY: Springer-Verlag.

Munro, N. B., Talmage, S. S., Griffin, G. D., Waters, L. C., Watson, A. P., King, J. F., et al. (1999). The sources, fate and toxicity of chemical warfare agent degradation products. *Environmental Health Perspectives, 107*(12), 933-974.

NRC. (1983). *Risk assessment in the federal government: Managing the process*. Washington, D.C.: National Academy Press.

Ostfeld, A., & Salomons, E. (2003). *An early warning detection system (EWDS) for drinking water distribution systems security.* Paper presented at the World Water & Environmental Resources Congress, Philadelphia, PA.

Powell, J. C., West, J. R., Hallam, N. B., Forester, C. F., & Simms, J. (2000). Performance of various kinetic models for chlorine decay. *Journal of Water Resources Planning and Management, 126*(1), 13-20.

Reutter, S. (1999). Hazards of chemical weapons release during war: New perspectives. *Environmental Health Perspectives, 107*(12), 985-990.

Rose, J. B. (1988). Occurrence and significance of cryptosporidium in water. *J. AWWA, 80*(2), 53-58.

Rossie, W. L. (1975). Control of water quality in transmission and distribution systems. *J.AWWA, 67*(8), 425-427.

Rossman, L. A. (1994). *EPANET user's manual* (version 1.1). Cincinnati, OH: Risk Reduction Engineering Laboratory, Office of Research and Development, United States Environmental Protection Agency.

Rossman, L. A., & Boulos, P. F. (1996). Numerical methods for modeling water quality in distribution systems: A comparison. *Journal of Water Resources Planning and Management, 122*(2), 137-146.

Rotz, L. D., Kahn, A. S., Lillibridge, S. R., Ostroff, S. M., & Hughes, J. M. (2002). Public health assessment of potential biological terrorism agents. *Emerging Infectious Diseases, 8*(2), 225-230.

SUSG. (1997). *Protection of public water supplies*. Ann Arbor, MI: The University of Michigan.

Teunis, P. F. M., Chappell, C. L., & Okhuysen, P. C. (2002). Cryptosporidium dose response studies: Variation between isolates. *Risk Analysis, 22*(1), 175-183.

Tryby, M. E., Boccelli, D. L., Uber, J. G., & Rossman, L. A. (2002). Facility location model for booster disinfection of water supply networks. *Journal of Water Resources Planning and Management, 128*(5), 322-333.

Uber, J. G., Shang, F., & Rossman, L. (2004). *Incorporating general chemical reaction dynamics in water distribution system water quality modes: Extensions to EPANET.* Paper presented at the AWWA Annual Conference, Orlando, FL.

USEPA. (1985). *Rates, constants and kinetic formulations in surface water quality modeling* (No. EPA 600/3-85/040). Athens, GA: USEPA Environmental Research Laboratory.

USEPA. (1990). *Risk assessment, management and communication of drinking water communication* (No. EPA/625/4-89/024). Washington, DC: Office of Research and Development.

VonHuben, T. (Ed.). (1999). *Water distribution operator training handbook* (2nd ed.). Denver, CO: American Water Works Association.

Walski, T. M., Brill, E. D., Gessler, J., Goulter, I. C., Jeppson, R. M., Lansey, K., et al. (1987). Battle of the network models: Epilogue. *Journal of Water Resources Planning and Management, 113*(2), 191-203.

Weber, W. J., Jr., & DiGiano, F. A. (1996). *Process dynamic in environmental systems*. New York, NY: Wiley-Interscience.

Westaway, C. R., & Loomis, A. W. (Eds.). (1977). *Cameron hydraulic data* (15th ed.). Woodcliff Lake, NJ: Ingersoll-Rand.

White, G. C. (1999). *Handbook of chlorination and alternative disinfectants*. New York, NY: John Wiley & Sons, Inc.

Williams, V. R., & Williams, H. B. (1967). *Basic physical chemistry for the life sciences*. San Francisco: W.H. Freeman and Company.

Ysusi, M. A. (2000). System design: An overview. In L. W. Mays (Ed.), *Water distribution systems handbook* (pp. 3.1-3.49). New York, NY: McGraw-Hill.

Appendices

Appendix 1: Selected General Distribution System Design Guidelines

Key Water Demand Criteria:

1. Average Day: total year's pumpage divided by 365 (expressed in MG/day)
2. Maximum Day: maximum 24-hour demand in one year
3. Maximum Hour: maximum 1-hour demand in one year
4. Fire Flow Demand: flow to provide for fire defense; generally computed by the Insurance Services Office (ISO) criteria in the Fire Suppression Rating Schedule.
5. Present and Future
    a. Type: Residential, Commercial, Industrial, Irrigation
    b. Location: Direction/areas of growth

Redundancy:

1. At least two pumping units to be provided. With any pump out of service the remaining pumps shall be capable of providing the maximum pumping demand of the system.
2. Auxiliary Power provided from two separate feed sources or have auxiliary power generator capable of operating critical components to service 50% of Maximum Daily Flow

Pressure:

1. Minimum: 20 psig at ground level
2. Normal: 40-70 psig (35-60 psig RSWW recommended)

Diameter:

1. Minimum: 6” for lines providing fire service; larger if required to maintain minimum pressure during fire event.
2. Smaller mains (3” recommended min.) are allowed only in special circumstances based on land use and hydraulic analysis. No fire service required.

Fire Protection:

3. Design system (supply, storage, pumps and mains) to meet fire demands of State Insurance Services Office or other jurisdictional agencies.
4. Account shall be taken for hourly and season fluctuations in water demand when designing for fire flow.

Distribution Mains:

1. Dead end mains minimized to reduce head loss and improve reliability of service.
2. Hydrants should be placed at street intersections and at intervals of 350-600’. Hydrants not connected to mains that don’t support minimum fire flow (500 gpm)
3. Valves should be placed at minimum intervals of 500’ in industrial areas, 800’ in residential areas and one mile in rural areas.

(adapted from: Recommended Standards for Water Works, 1997)

Appendix 2: Major Variables in Modeling Water Distribution System Response to Terrorist Attack

Contaminant Characterization

1. Nature of the Contaminant
    a. Biological, chemical, or physical
    b. Infective/toxic dose characteristics (dose-response information)
        i. NOAEL
        ii. Regulatory default values
2. Dose Characterization
    a. How much injected into system
        i. Volume injected
        ii. Initial Concentration of injection
        iii. Duration on injection
    b. Location of injection
    c. Date effects (ex. weekend/holiday/max. usage)
    d. Time of Injection
    e. Growth/Decay of contaminant in the system
3. Analytical Techniques Available
    a. Accuracy (sensitivity, specificity)
    b. Turnaround time
    c. Sampling characteristics
        i. Volume of sample required
        ii. Time to sample
        iii. Skill/effort of sampling staff

iv. Cost (equipment/expendables)

Distribution System Characterization

1. Hydraulic characterization
    a. Distribution system geometry
        i. Actual vs. Skeletonized
        ii. Pipe information
            (a) Diameter
            (b) Length
            (c) Friction factor (function of Age, Material, water quality)
        iii. Valve information (type, location, settings (open/closed/fractional))
        iv. Storage tanks
            (a) Elevated vs. Ground
            (b) Geometric configuration (Diameter/Height/Volume; feed method)
            (c) Operational Control (control equipment and setpoint/levels)
            (d) Functional characteristics
                (i) Short-circuiting
                (ii) Diurnal Turnover and implications for water age/quality
    b. Pump information
        i. Pump Curve characteristics
            (a) Impeller Characteristics (Diameter, type)
            (b) Pump Type
                (i) Constant speed vs. Variable Speed
                (ii) Turbine, Horizontal Split Case, End Suction, etc.

c. System Demand Patterns - note: these are statistical characterizations (i.e. all equivalent demand types (residential/commercial/etc) are assumed to have equal withdrawals rates during equivalent time periods; the actual demands are stochastic)
    i. Average Daily Demand
    ii. Peak Hourly Demand
    iii. Peak Instantaneous Demands
    iv. Transient System Fluctuations
        (a) Fire Flow
        (b) Line Break
        (c) Equipment/controls outage or malfunction
    v. Change in system characteristics over time (new development/industry, new booster station, close an old pumping station, etc.)

2. Water Quality
    a. Water characterization from plant
        i. Normal operation
        ii. Malfunction or atypical operation
    b. Pipe effects
        i. Biofilms and its impact on contaminant
            (a) Protective matrix
            (b) Potential nutrient source
            (c) Potential reproductive/breeding area
        ii. Pipe materials (ex. CIP/DIP, PVC (internal/external leaching), copper, etc.)
        iii. In-pipe water quality changes (pipe as a reactor/contact basin)

c. Chlorine residual maintained in water column

d. Transformations caused by chemical interactions between different water sources

    i. Multiple water supply sources/plants feeding the distribution system

    ii. Intersystem connections

e. Water age (related to hydraulic characteristics above)

f. Operational influences

    i. Line Maintenance

        (a) Flushing

        (b) Pigging

        (c) Repairs

    ii. Transient conditions

        (a) Hydraulic

        (b) Treatment fluctuations

    iii. Cross-connections

Appendix 3: Description of Computational Procedure Associated with Analysis of Anytown Water System

1. Develop and Calibrate a hydraulic model of the water system in question. This would include decisions regarding skeletonizing non-critical distribution elements (ex. pipe diameters below a given threshold, dead-end lines, etc.), verifying pump/system control sequences, and backchecking against actual operating records. The base computer model used should have a water quality feature available (ex. EPANET, WaterCAD, KYPIPE, etc.); General calibration procedures should be followed (see AWWA working papers).
2. Develop a Design Basis Threat (DBT) to describe the nature and extent of the hypothesized contamination event. This would include such factors as time of injection, duration of injection, volume/concentration of injection, identification and nature of the contaminant.
3. Convert the DBT into an injection event in the computer model. This includes modeling decisions such as time step selection, integration into diurnal variations, decay/growth parameters, pipe wall effects, etc.
4. One of the issues to be dealt with is 'what duration should be used when running the computer simulation?' The DBT should be test run at a number of locations to see how long the contaminant persists in the system. A decision criteria should be established such as calculated contaminant values equal to or less than a given threshold. This will establish the duration to be used as the time basis for the system definition runs.
5. Develop a notation convention to keep track of the computer runs and then move the injection event to every node in turn, run the network model and record the

results. The notation should delimit monitoring and injection nodes. WaterCAD has a Scenario/Alternative feature that is useful in the regard.

6. Take the results of all the runs and export them to Excel. Note: the number of runs will equal the number of nodes in your model if you moved the injection point to every node in turn. Because you are measuring the response of the injection at a specific node at every other node, this means that you will have N^2 node responses in one dimension. In the other dimension, the model will list the response by timestep. Several things must be kept in mind when performing this step. Excel is limited to 256 columns. Because the analysis will include N^2 node responses the data will need to be transposed in order to be imported into Excel. The number of timesteps required will generally be much less than 256 so the transposition should not be a problem. (All formulas and VBA code in the steps that follow are designed to work with time steps as column labels and nodes as row labels so, even if N^2 does not exceed 256, transpose the data.)
    a. Insert, Worksheet
    b. Return to original worksheet; select the full range of data, taking note of the size of the range as shown beside the cursor (or in the Name Box), ie. 46R x 256C; Copy
    c. Go to new, blank worksheet and select a range in reverse size, ie. 46C x 256R; Edit, Paste Special, Transpose, OK.
    d. Rename the new worksheet *Transposed CSV* and delete the original worksheet.
        i. Fractional timesteps will occur at transition points in the analysis (pump on/off, tank fill/drain, etc.) so if the duration is long or the system is unusually active, some data cleansing may be required.

e. In a later step, you will need to filter for unique response and injection nodes.
    i. On the *Tranposed CSV* worksheet, type *rnode* in cell AA1; enter the formula =LEFT(A2,FIND(" ",A2,1)-1) in cell AA2; copy the formula in cell AA2 to the last row containing data; filter this new list for unique values by clicking Data, Filter, Advanced Filter… (if you get a warning message about the label, just click OK); in the Advanced Filter window, select Copy to another location; the List range is your list of values in the rnode column – in this example, cells $AA$1:$AA$257; in the Copy to box, type $AC$1; check the Unique records only box; click OK; this provides you with a list of response nodes that will be used as row labels.
    ii. On the *Tranposed CSV* worksheet, type *cnode* in cell AB1; enter the formula =MID(A2,LEN(AA2)+2,(FIND(" ",A2,LEN(AA2)+2))-(LEN(AA2)+2)) in cell AB2; copy the formula in cell AB2 to the last row containing data; filter this new list for unique values by clicking Data, Filter, Advanced Filter… (if you get a warning message about the label, just click OK); in the Advanced Filter window, select Copy to another location; the List range is your list of values in the cnode column – in this example, cells $AB$1:$AB$257; in the Copy to box, type $AD$1; check the Unique records only box; click OK; this provides you with a list of injection nodes that will be used as column labels; manually sort this list if necessary.

f. Also in a later step, you will need the time steps in rows as a list range for a combo box.
    i. On the *Transposed CSV* worksheet, select the time steps in row 1; Copy

ii. Go to cell AE2 on the same worksheet; Edit, Paste Special, Transpose, OK.; type the label *Time* in cell AE1.

7. Using the transposed data set, set a threshold as a basis for counting the values for each timestep from 0 to the end of the duration - always beginning with timestep 0. For example, at timestep 5, the count would be computed by counting the values from timestep 0 to timestep 5. Note: depending on how the network program is set up, the actual contamination will occur at a small fraction after point 0, say time 0.1, so the full initial timestep is used. This produces a slight error but it's certainly within reason given the overall uncertainty of the process. Also, if there is a preexisting level of the contaminant at a node, this step will be necessary to capture it's initial effect. It is suggested that only the even increment timesteps be used (typically this is the timestep used in the water system modeling).

   Two variable selections, five tables and a command button are required in order to summarize the data set. These are created on a new, separate worksheet and are comprised of both formulas and code.

   a. Insert, Worksheet; name it *Count Summary*.
   b. Name and create the first variable setting – a threshold variable by which data set values are determined to be less than or equal to, or greater than, the variable.
      i. In cell A1, paste this text: **ENTER *Threshold* VARIABLE (mg/L):** and format the row height at 21.75
      ii. In cell F1, paste this number: ***0***

c. Name and create the second variable setting – a combo box with time step selections.
    i. In cell L1, paste and right-align this text: Select Time Period:
    ii. Insert a Combo Box and size it to span row 1 and columns M & N
    iii. Right-click on the combo box and select properties; on the (Name) row, name the button *cmbCTime*; on the ForeColor row, drop down to select red from the Palette; on the Font row, click the ellipses and select Bold (Arial, 10 should already be set); on the LinkedCell row, type $N$1; on the ListFillRange row, paste this reference: 'Transposed CSV'!AE2:AE26; close Properties window; click the Design Mode button on the Controls toolbox to exit design mode
    iv. Select a value from the combo box (to place a value into cell N1 that will avoid formula errors in later steps).
d. Create the less than or equal to threshold variable table.
    i. In cell A2, paste this text: ***Less than or equal to variable:*** and format row height at 30
    ii. In cell A3, type *Node*
    iii. Starting in cell A4, paste the unique Response nodes from the Transposed CSV worksheet ('Transposed CSV'!AC2:AC17)
    iv. Starting in cell B3, transpose/paste the unique Injection nodes from the Transposed CSV worksheet ('Transposed CSV'!AD2:AD17)
    v. In cell B4, enter (Ctrl+Shift+Enter) this array formula:

    =COUNTIF(INDEX('Transposed CSV'!$B$1:OFFSET('Transposed CSV'!$B$1,256,$N$1-'Transposed CSV'!$B$1),MATCH(TEXT($A4&"

"&B$3&" Injection Concentration (mg/l)",0),'Transposed CSV'!$A$1:$A$257,0),0),"<= "&$F$1)

Copy the formula to the rest of the cells in the table

vi. Format data as Number with 0 (zero) decimals; format table as desired

e. Create the greater than threshold variable table.

i. In cell A21, paste this text: ***Greater than variable:*** and format the row height at 30

ii. Copy table 1 A3:Q19 (includes row and column labels) and paste to cell A22

iii. In cell B23, enter (Ctrl+Shift+Enter) this array formula:

=IF(COUNTIF(INDEX('Transposed CSV'!$B$1:OFFSET('Transposed CSV'!$B$1,256,$N$1-'Transposed CSV'!$B$1),MATCH(TEXT($A23&" "&B$3&" Injection Concentration (mg/l)",0),'Transposed CSV'!$A$1:$A$257,0),0),"> "&$F$1)=0,"",COUNTIF(INDEX('Transposed CSV'!$B$1:OFFSET('Transposed CSV'!$B$1,256,$N$1-'Transposed CSV'!$B$1),MATCH(TEXT($A23&" "&B$3&" Injection Concentration (mg/l)",0),'Transposed CSV'!$A$1:$A$257,0),0),"> "&$F$1))

Copy the formula to the rest of the cells in table 2.

iv. Name table 2's data cells B23:Q38 (excludes row and column labels) *countgreat* (type countgreat in the Name Box and press Enter)

v. In cell R22, paste **COUNT**; in cell S22, paste *Ave*; in cell T22, **Total**

vi. In cell R23, enter this formula and copy it down the rows:

=COUNT(B23:Q23)

vii. In cell S23, enter this formula and copy it down the rows: =AVERAGE(B23:Q23)

viii. In cell T23, enter this formula and copy it down the rows: =R23*S23

ix. Format data as Number with 0 (zero) decimals; format columns as desired

f. Create a third table that will sort the values of table 2. In a later step, code will be written in VBA to re-create table 3 but data must first be visually placed on the worksheet in order to name ranges that VBA will be programmed to refer to.

i. In cell A40, paste this text: ***Sorted values:*** and format the row height at 30

ii. Select cells A22:T38 - the entire table 2 including row and column headings, as well as the three additional count and average columns, but not the cell containing the table heading *Greater than variable:*; Copy

iii. Paste the *values* of table 2 to cell A41 (Edit, Paste Special, Values, OK)

iv. With the pasted values still highlighted (A41:T57), type *countsorttab* in the Name Box; press Enter

v. Select the data in the Count column of table 3 R42:R57 (excludes the column heading, Count) and name the range *countcv*

vi. Select the data in the Total column of table 3 T42:T57 (excludes the column heading, Total) and name the range *countcav*

vii. Select the Injection node column headings in the top row of table 3 B41:Q41 (excludes column A and the 3 Count and Average columns) and name the range *countcn*

viii. Select the Response node row headings in column A of table 3 A42:A57 (excludes the row containing the Injection node column headings) and name the range *countrn*

ix. Select the first row of *data* in table 3 B42:Q42 (excludes row and column node headings and the 3 count and average columns) and name the range *counter*

x. Select all of the *data* in table 3 B42:Q57 (excludes row and column node headings and the 3 count and average columns) and name the range *countar*

g. Create a fourth table that will extract sorted data from table 3. In a later step, code will be written in VBA to re-create table 4 but data must first be visually placed on the worksheet in order to name ranges that VBA will be programmed to refer to.

i. In cell A59, paste this text: ***Summary of filtered values:*** and format row height at 30

ii. In cell A60, paste *Col Node*, in cell B60, paste *Mon Node*, in cell C60, paste **COUNT**, in cell D60, paste **Total**

iii. Copy the row labels that identify the Response nodes from table 3 (A42:A57) to cells A61:B61

iv. Name the Injection nodes in column A of table 4, A61:A76 (excludes headings), *countcnn*

v. Name the Response nodes in column B of table 4, B61:B76 (excludes heading), *countrnn* (note that the values are not accurate now but will be later)

vi. Name the blank cells in column C of table 4, C61:C76 (excludes heading), *countcvn*

vii. Name the blank cells in column D of table 4, D61:D76 (excludes heading), *countcavn*

viii. Name the entire table 4, A60:D76 (excludes the table heading, Summary of filtered values:), *countfiltval*

ix. Format table as desired

h. Create a fifth table that will sort the values of table 4. In a later step, code will be written in VBA to re-create table 5 but data must first be visually placed on the worksheet in order to name ranges that VBA will be programmed to refer to.

i. In cell G59, paste this text: ***Nodes not covered by first level node:***

ii. Go to the named range countfiltval (drop down the name box and click on the name); Copy

iii. Click on cell G60, Paste

iv. With the pasted cells still selected (G60:J76), name the new range *countfinval*

i. Add a command button control to the worksheet. VBA code will be assigned to this button in a later step.

i. Insert a command button and size it to span G1:I1 (open the Control Toolbox, if necessary, to find the command button control)

ii. Right-click on the command button and select properties; on the (Name) row, name the button *cmdRank*; on the Caption row, type the caption View Ranked Data; on the Font row, click the ellipses, select Arial, Bold,

8; close Properties window; exit design mode (first button on control toolbar)

8. Write procedures to reproduce tables 3, 4 and 5.
   a. In Visual Basic, insert a module and name it Count (View, Properties, Name). Following the Option Explicit statement, paste each of the following six procedures.
   b. This first procedure automates pasting of the values in table 2 to table 3 and renames table 3.

```
Sub CountPTV()

    ' This first procedure automates pasting of the values
    ' in table 2 (Greater than variable) to table 3 (Sorted values)
    ' and renames table 3
      Worksheets("Count Summary").Range("countgreat").Select
      Selection.CurrentRegion.Select
      Selection.Copy
      Range("countsorttab").Select
      Selection.CurrentRegion.Select
      Selection.PasteSpecial Paste:=xlPasteValues
      Range("countsorttab").Cells(0, 1).Select
      ActiveCell.Value = "Sorted values:"

End Sub
```

   c. The second procedure sorts table 3 first by Count, then by Total, both descending.

```
Sub CountSNT()

    ' The second procedure sorts table 3 (Sorted values)
    ' first by Count, then by Total, both descending
    Range("countsorttab").Sort Key1:=Range("countcv").Cells(0),
    Order1:=xlDescending, Key2:= _
    Range("countcav").Cells(0), Order2:=xlDescending, Header:=xlYes

End Sub
```

d. The third procedure searches for columns with no values in the first row of sorted values in table 3.

```
Sub CountRCN()

    ' The third procedure searches for columns with no values
    ' in the first row of sorted values in table 3
    Dim countcn, counter, countcnn As Range
    Dim n As Integer

    ' Clear existing values in column A of table 4
    Worksheets("Count Summary").Range("countcnn").ClearContents
    Worksheets("Count Summary").Activate

    Set countcn = Range("countcn")   ' Refer to the first row in table 3 that
    contains the column (injection) node labels
    Set counter = Range("counter")   ' Refer to the second row in table 3 to
    determine if that row contains any empty cells
```

```
Set countcnn = Range("countcnn") ' Refer to the first column in table 4
where the column labels that match empty cells will be returned

' Loop through cells in table 3
   Do

   ' Count the number of columns in the countcn range and
   ' assign the number to the variable n
   For n = 1 To countcn.Columns.Count

      ' Use the variable n as the row or column identifier
      ' where applicable, ie. in the counter row, we
      ' use 1 as the row and n to cycle through the columns
      ' If a non-numeric cell is found in counter, the code
      ' returns the label from countcn to the appropriate row
      ' in table 4.  Otherwise, the row label in table 4 will
      ' have a value of "-"
      If IsNumeric(counter.Cells(1, n)) = False Then
         countcnn.Cells(n, 1) = countcn.Cells(1, n)
      Else
         countcnn.Cells(n, 1) = "-"
      End If
   Next n
   Exit Do
   Loop Until n = countcn.Columns.Count
```

```
End Sub
```

e. The fourth procedure returns the corresponding response node, Count and Total values from the first maximum value rows in table 3 to the rows containing values in table 4.

```
Sub CountRRN()

    ' The fourth procedure looks at the columns in table 3
    ' that match the column (injection) node labels returned to table 4
    ' to determine the first maximum value within table 3's column
    ' and return the respective row (response) node label, Count value,
    ' and Total value to table 4.
    Dim myrange, countcn, countrn, countar, countcv, countcav, countcnn,
    countrnn, countcvn, countcavn As Range
    Dim maxnum, m, n, r As Integer

    ' Clear columns 2-4 in table 4
    Worksheets("Count
    Summary").Range("countrnn,countcvn,countcavn").ClearContents
    Worksheets("Count Summary").Activate

    Set countcn = Range("countcn")  ' Refer to column (injection) node labels
    in table 3
    Set countrn = Range("countrn")  ' Refer to row (response) node labels in
    table 3
    Set countar = Range("countar")  ' Refer to data in table 3
```

```
Set countcv = Range("countcv")  ' Refer to Count calculations in table 3
Set countcav = Range("countcav")    ' Refer to Total calculations in table 3
Set countcnn = Range("countcnn")    ' Refer to Col Node column in table
4
Set countrnn = Range("countrnn")    ' Refer to Mon Node column in table
4
Set countcvn = Range("countcvn")    ' Refer to Count column in table 4
Set countcavn = Range("countcavn")  ' Refer to Total column in table 4

' Loop through cells in table 3
Do

   ' Count the number of columns in the countcn range and
   ' assign the number to the variable n
   For n = 1 To countcn.Columns.Count

   ' Cycle through each row (using the variable n)
   ' in column A of table 4
   ' to determine if the row contains a value
   ' in order to proceed with code
   If countcnn.Cells(n, 1) Like "J*" = True Then

   ' Count the number of rows in the countrn range and
   ' assign the number to the variable m
   ' Use the value assigned to m as the highest value for r
```

```
    m = countrn.Rows.Count
    For r = 1 To m

    ' Once a value has been identified in column A of table 4
    ' the variables r, n & m serve to select data rows 1 to m in table 3
    ' in the column that matches the value identified
    ' and assigns that selected range to the variable myrange
    Set myrange = countar.Range(Cells(r, n), Cells(m, n))

    ' Using the myrange variable, maximum value is located
    ' From the row containing the first value equal to the maximum,
    ' the row (response) node label from table 3 is returned to the Mon
Node column in table 4
    '   on the same row as the matching Col Node value
    ' the Count value from table 3 is returned to the Count column in table 4
    '   on the same row as the matching Col Node value
    ' the Total value from table 3 is returned to the Total column in table 4
    '   on the same row as the matching Col Node value
    maxnum = Application.Max(myrange)
       If countar.Cells(r, n) = maxnum Then
          countrnn.Cells(n, 1) = countrn.Cells(r, 1)
          countcvn.Cells(n, 1) = countcv.Cells(r, 1)
          countcavn.Cells(n, 1) = countcav.Cells(r, 1)
          Exit For
       End If
```

```
            Next r
            End If
            Next n
        Exit Do
        Loop Until n = countcn.Columns.Count
    End Sub
```

f. The fifth procedure copies table 4 to table 5 and renames table 5.

```
    Sub CountPFV()
        ' The fifth procedure copies table 4 to table 5 and renames table 5
            Worksheets("Count Summary").Range("countfiltval").Select
            Selection.CurrentRegion.Select
            Selection.Copy
            Range("countfinval").Select
            Selection.CurrentRegion.Select
            Selection.PasteSpecial Paste:=xlPasteValues
            Range("countfinval").Cells(0, 1).Select
            ActiveCell.Value = "Nodes not covered by first level node:"
    End Sub
```

g. The last procedure sorts table 5

```
    Sub CountSFT()
        ' The last procedure sorts table 5 by Count column (descending), then by
        Total (descending).
```

```
        Range("countfinval").Sort Key1:=Range("countfinval").Offset(0, 2),
    Order1:=xlDescending, Key2:= _
        Range("countfinval").Offset(0, 3), Order2:=xlDescending,
    Header:=xlYes

End Sub
```

9. Assign code to the command button control.
   a. In Visual Basic's Project Explorer window, double-click Sheet1 (Count Summary) and paste the following procedure into the blank Code sheet, after Option Explicit. This code calls the six previously written procedures, in their correct order, when the user clicks the View Ranked Data command button.

```
Private Sub cmdRank_Click()

    Call CountPTV
    Call CountSNT
    Call CountRCN
    Call CountRRN
    Call CountPFV
    Call CountSFT
    Worksheets("Count Summary").Range("countfinval").Select

End Sub
```

Click File, Close and Return to Microsoft Excel. Save the spreadsheet.

About the Author

James R. Chastain, Jr. attended the University of Florida where he received his Bachelor of Science in Civil Engineering with honors in 1971 and a Master of Engineering in Environmental Engineering in 1972. He became a registered Professional Engineer in the state of Florida in 1976 and has been a practicing engineer since then. He became president of Chastain-Skillman, Inc., a consulting engineering and environmental science firm, in 1982.

Mr. Chastain received a Master's of Public Health from the University of South Florida's College of Public Health in 1993. After receiving his MPH, Mr. Chastain became interested in projects that combined his engineering training with public health issues. He subsequently entered the Ph.D. program at the University of South Florida where he focused on research that sought to protect public water systems.