UNIVERSITY OF CINCINNATI

May 24 , 20 01

I, Trevor Ray Lindley ,
hereby submit this as part of the requirements for the degree of:

Master of Science

in:

Environmental Engineering

It is entitled:

A Framework to Protect Water Distribution Systems Against Potential Intrusions

Approved by:

Steven Buchberger
Jim Uber
Xinhao Wang
Daniel Schaefer

A Framework to Protect Water Distribution Systems Against Potential Intrusions

A thesis submitted to the

Division of Research and Advanced Studies
of the University of Cincinnati

in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

in the Department of Civil and Environmental Engineering
of the College of Engineering

2001

by

Trevor R. Lindley

B.S., Utah State University, 1996

Committee Chair: Dr. Steven G. Buchberger

ABSTRACT

A framework is developed to quantify the susceptibility of drinking water distribution systems to intrusion events. The framework integrates infrastructure information, hydraulic modeling, and demographic data. These elements are managed within a geographic information system (GIS). Using criteria that reflect system pressure, hydraulic intrusion pathways, and contaminant sources, the framework identifies locations within the distribution system susceptible to intrusion events. Locations found to be susceptible to intrusions are prioritized for attention based on proximity to sensitive populations, such as young children and the elderly.

The proposed method is demonstrated with a case study based on a real distribution system. The study area encompasses approximately 38 square miles, includes three service areas, contains over 280 miles of water main serving 18,900 connections with a total average demand of five to six million gallons per day. Susceptibility conditions exist at some locations throughout the system; however, only rarely do all three conditions coincide. Hence very few locations were deemed susceptible to intrusion events.

The framework may support capital improvement programs, operational decisions, and distribution system sampling designs. Methods such as this have been suggested as part of a larger distribution system management approach to improve water quality and at the same time reduce regulatory sampling requirements.

PREFACE

At some point prior to my entry into the field of Civil and Environmental engineering I began to develop an interest in this wonderful and amazing thing we call water. I can not say if the interest stemmed from being pulled behind the family ski boat across the flat glass surface of Flaming Gorge Reservoir or casting an elk-hair caddis to cutthroat trout on the upper reaches of the Logan River. It may have had its origins in time spent along the Gulf Coast of Texas seeing the need and subsequent impact of large petrochemical facilities. Even more distant memories exist of hunting with my Dad, camping with the Boy Scouts, and fishing at a small pond near my Grandma's house. It was likely a combination of multiple positive and negative experiences that instilled in me the value of nature and of water.

In 1993 the National Geographic Society published a special edition of their monthly periodical entitled "Water, The Power, Promise, and Turmoil of North America's Fresh Water." Review of the pictures and articles suggests not much has changed in the past eight years. Many societies are without clean water for drinking or bathing; many societies use the resource to excess without regard for future implications. One thing is certain, objective decision making related to the use of water (in all societies) is elusive. I hope to be able to use the knowledge and expertise I have gained to support decisions surrounding the beneficial use of this valuable resource.

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Steven Buchberger for all of his guidance throughout my time in Cincinnati. He allowed me the autonomy I needed while still keeping me pointed down the path I needed to be on. I would like to thank Dr. Jim Uber for discussions related to modeling and drinking water distribution systems.

I especially need to thank my family both extended, parents and siblings, and immediate here in Cincinnati. Emma Kate, our first daughter, will always be our Cincinnati baby. Tyson Ray, our oldest son, allowed his Dad to work on his “pipes, and tanks, and water” in between Lego building and baseball games (backyard and the Reds). The largest portion of thanks goes to my wife Tiffanie. She willingly left the steady and relatively simple life we had in Utah to embark on this big adventure called graduate school. She loves her children very much and instills in them a measure of kindness and an appreciation for nature I can only hope to have one day. I appreciate her love for us and her dedication to our family and my career.

I also need to thank the National Science Foundation and those involved with securing and managing my support funds, namely, Dr. Shafik Islam and other members of the Environmental Hydrology group at the University of Cincinnati.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF NOTATION

AWWARF	American Water Works Research Foundation
C	Discharge coefficient (gpm / $\sqrt{psi}$) associated with emitter function
EPS	Extended Period Simulation
HPC	Heterotrophic Plate Count
GIS	Geographic Information System
GPM	Gallons Per Minute
MGD	Million Gallons per Day
psi	pounds per square inch
p	Pressure (psi) associated with emitter function
q	Flowrate (gpm) associated with emitter function
USEPA	United States Environmental Protection Agency
US	United States

1.0 INTRODUCTION

Treatment of drinking water in the United States and many other countries can be viewed as a multiple barrier approach. These barriers often include protection of source waters, treatment by advanced processes, maintenance of a disinfectant residual, and distribution via a pressurized system. Collectively, these measures have been very effective in protecting public health and providing a safe supply of potable water. In recent years, the final barrier (i.e., the water distribution system) has come under increasing scrutiny.

Between 1971 and 1998 there were 619 reported waterborne disease outbreaks in the US (Craun and Calderon, in press). Of this total, 113 (18%) have been attributed to water quality problems in the distribution system. Owing to rigorous epidemiological reporting standards, it is suspected that the record of documented outbreaks represents a small fraction of the total waterborne illnesses (see Figure 1-1 Frost et al, 1996).

Payment et al, (1991) argue that many common gastrointestinal illnesses (not necessarily outbreak events) can be attributed to public water supplies that are in full compliance with North American regulatory requirements. These illnesses purportedly stem from treatment plant inefficiencies, pathogen regrowth, chemicals in the water, or breaks in the integrity of the distribution system. Additional work (Payment, 1997) suggests water flowing through the distribution system leads to a higher rate of illness than finished water collected at the treatment works—possibly incriminating the distribution system. Waterborne illness attributable to breakdowns in the distribution system should exhibit geographic clustering; however, sophisticated spatial and temporal analyses have not confirmed this behavior (Payment, 1998). These inconclusive and

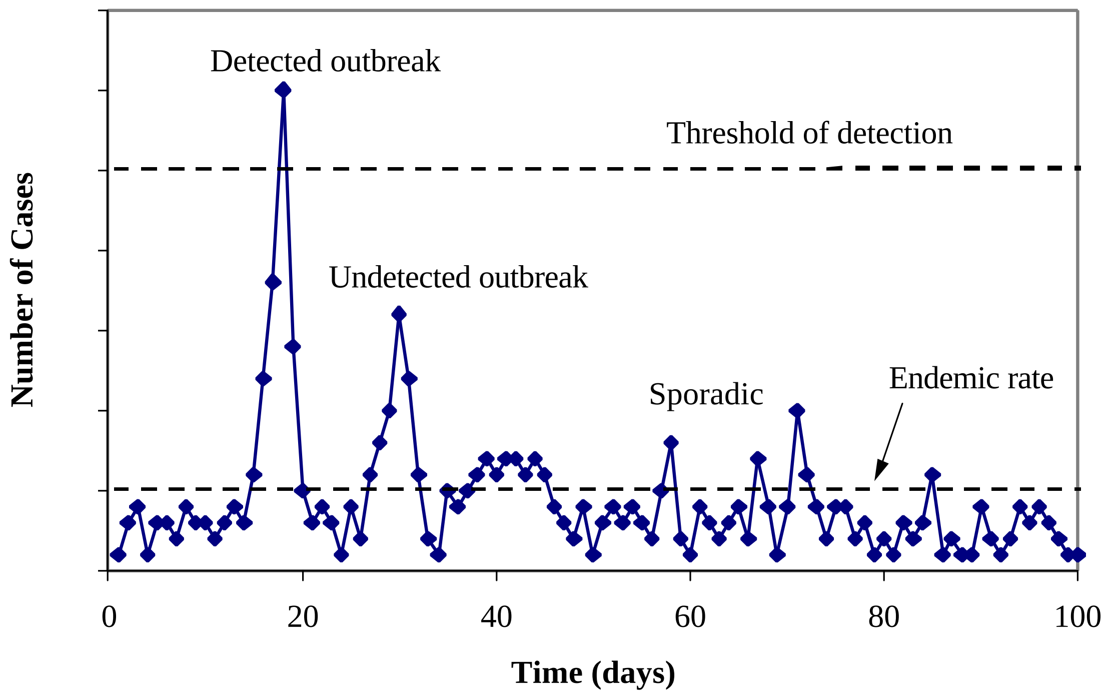


Figure 1-1. Hypothetical Incidence of Endemic vs. Epidemic Disease (adapted from Frost et al, 1996)

sometimes contradictory findings underscore the inherent difficulty in identifying and quantifying detrimental health effects attributable to the water distribution system. The growing challenges surrounding operation of water distribution systems have been accompanied by an explosion in the availability of utility data, dramatic advances in system hydraulic modeling, and a general industry shift towards best management practices (Brothers, 2000). Against this evolving landscape, this paper develops a holistic framework to help manage distribution system operations. This framework exploits tools that are currently moving out of engineering backrooms and into the forefront of informed utility decision making. These tools, namely distribution system hydraulic modeling and Geographic Information Systems (GIS), are ideally suited for addressing operational questions faced by distribution system managers. Using hydraulic modeling and GIS, this paper will demonstrate how to integrate multiple contributing risk factors in order to identify locations in the distribution system that may be susceptible to undesirable contaminant intrusion events.

1.2 Research Objectives

The research objectives are:

- Formulate a framework that first identifies locations within the distribution system that are susceptible to intrusions and second prioritizes management of these susceptible locations considering the potential influence on nearby sensitive population centers;
- Develop the framework considering data typically available to a water utility, and the tools of hydraulic modeling and GIS; and
- Apply the framework to an actual utility.

1.3 Thesis Organization

Section 2 introduces intrusion events, the water quality concern specifically addressed in the thesis. Section 3 discusses the susceptibly conditions that may result in an intrusion event. Section 4 introduces the framework that has been developed to identify locations susceptible to intrusion events. Section 5 shares the case study location and framework results. Section 6 provides a conclusion to both the method and the case study results. Section 7 contains the references cited in the body of the thesis. Additionally, numerous appendices provide more details relative to background information and implementation of the framework in the study area. Appendix A contains a complete literature review on the topic of distribution system water quality, sensitive populations, and the use of decision frameworks in infrastructure management. Appendix B discusses the details surrounding the pressure portion of the framework including the basis for the stochastic model inputs and Visual Basic code used to access the EPANET Toolkit. Appendix C provides numerous figures used in the pathway analysis; they summarize repair data for the case study area. Appendix D contains information describing the spatial GIS themes used in the analysis (i.e., metadata) [(1)].

[1]Sections 1.0 and 2.0-6.0 inclusive are to be submitted as a manuscript to the Journal of American Water Works Association.

2.0 INTRUSION EVENTS

Water quality concerns related to distribution system networks and water storage facilties include: source variability, cross-connections, leaky pipes, metal dissolution, pipe corrosion, disinfectant loss, undesirable reactions, bacterial regrowth, turbidity fluctuations, and improper construction, maintenance, or repair practices (AWWARF, 2001; Grayman, et al, 2000; Walski, 2000; Grayman and Kirmeyer, 2000). From this list, cross-connection contamination, improper construction, maintenance, or repairs, and contamination at leaky pipe joints (or at any location structurally deficient) are associated with contaminant intrusions. Intrusions are defined here as the unintended and uncontrolled introduction of an undesirable agent into the potable water distribution system. While all intrusions are potentially serious, those leading to waterborne illness are the primary concern.

An intrusion related waterborne illness or undesirable water quality excursion such as a high heterotrophic plate count (HPC) is a secondary manifestation of underlying factors that actually cause the quality upset. Underlying factors resulting in distribution system intrusions include low pressure, cross-connections, appurtenance malfunction, improper maintenance or repair, a contamination source, and structural integrity of the system (AWWARF, 2001; LeChevallier, 1999; Geldreich, 1996). In some instances the occurrence of just one of these factors may result in an intrusion event; in other instances, the joint occurrence of multiple factors is required for contamination to occur.

3.0 SUSCEPTIBLITY CONDITIONS

Three fundamental susceptibility conditions must be met for an intrusion event to occur: (1) adverse pressure gradient, (2) hydraulic pathway, and (3) contaminant source. This conceptual model of the conditions leading to distribution system intrusions is maintained throughout the development of the distribution system susceptibility framework. Table 3-1 shows the illness attributable to the distribution system and the susceptibility conditions that accompanied many of these events.

3.1 Adverse Pressure Gradient

System wide positive pressure is essential to maintain the quality and integrity of the distribution system (Geldreich, 1996). Locations experiencing low pressure are more susceptible to backflow at uncontrolled service connections or controlled service connections that may be failing (LeChevallier, 1999; Haas, 1999; Geldreich, 1996). Locations that experience extremely low pressure are at risk to intrusions not only at above-grade service connections but also at subsurface appurtenances and along buried sections of the pipe network. Maintenance of some minimum pressure (typically 20 - 35 psi) is a key component of utility operations and greatly reduces or even eliminates pressure as a susceptibly condition (AWWARF, 2001).

3.2 Hydraulic Pathway

The hydraulic pathway is the hydraulic connectivity route between the potable supply and the contaminant source. The pathways contributing to distribution system contamination include consumer service connections and structurally deficient locations (i.e., appurtenances, leaking pipes, uncovered tanks, etc). Obviously users are connected to the drinking water supply throughout the distribution system; thus, the first

Table 3-1. Deficiencies in Water Distribution Systems Resulting in a Documented Outbreak of Waterborne Illness in the United States from 1971 to 1998

			Susceptibility Condition(s) Met		
Cited Deficiency Causing Illness Outbreak [(1), (2)]	**Number of Outbreaks**	**Percent of Total**	**Adverse Pressure Gradient**	**Hydraulic Pathway**	**Contaminant Source**
Cross-Connection and Back-Siphonage	60	53.1	✓	✓	✓
Inadequate Separation of Water Main and Sewer	1	0.9	✓	✓	✓
Broken and Leaking Water Mains	10	8.8	✓	✓	✓
Contamination While in Storage	15	13.3		✓	✓
Contamination During Construction/Repair	6	5.3		✓	✓
Contamination of Household Plumbing [(3)]	8	7.1			✓
Metal Corrosion and Metal Leaching [(3)]	13	11.5			✓
Total	113 [(4)]	100			

[1] Adapted from Gunther and Craun (in-press).

[2] Both community and non-community systems.

[3] These categories arguably fit the definition of an intrusion, however they typically result from some contaminant already inside the pipe (or part of the pipe wall) and are not addressed in this work.

[4] Total waterborne outbreaks considering all causes from 1971 to 1998 = 619.

contamination pathway can never be eliminated. Connection risk is, however, actively managed and minimized through a utility's cross-connection control program. The second pathway exists at locations where the infrastructure may be structurally deficient. Escalating utility repair costs, leak detection reports, and unaccounted for water are all manifestations of structurally deficient infrastructure (USEPA, 2001; Haas, 1999). As with connections, the nature of distribution systems (i.e., expansive, buried, etc.) suggests the intrusion risk due to structural deficiencies will never be completely overcome; the risk due to structural deficiencies can only be managed through rigorous maintenance programs including inspection, repair, and replacement.

3.3 Contaminant Sources

Contaminant sources are reservoirs of biological or chemical contaminants connected to the potable water supply via one of the pathways mentioned above. These pathogen sources include consumer processes connected to the distribution network (i.e., service connections), or other environmental sources that transfer contaminants to the potable supply at a location that is structurally deficient. Examples of service connections managed in a cross-connection control program (due to their perceived risk as a contaminant source) include hospitals, mortuaries, dry cleaners, and industrial users. Residential service connections, except in the case of an onsite water supply, do not typically have cross-connection control. Reservoirs of environmental contaminants (generally external to the distribution system infrastructure) that may contaminate distribution systems include sanitary, storm, combined sewers, septic systems, waterbodies, and animals. Research addressing distribution system intrusions found pathogenic organisms (often associated with sanitary wastewaters) in soil and water samples external to buried drinking water distribution system infrastructure (AWWARF, 2001).

The susceptibly condition(s) leading to an intrusion event can be considered in the context of a Venn diagram as shown in Figure 3-1. As the number of susceptibly conditions increases at a given location in the distribution system, the potential for intrusion increases. Distribution system operating procedures are a series of management practices that seek to eliminate one or all of the susceptibly conditions thus minimizing the potential for contamination events. Eliminating all conditions requires continued vigilance. Additionally, even with rigorous management practices, other events can still initiate susceptible conditions.

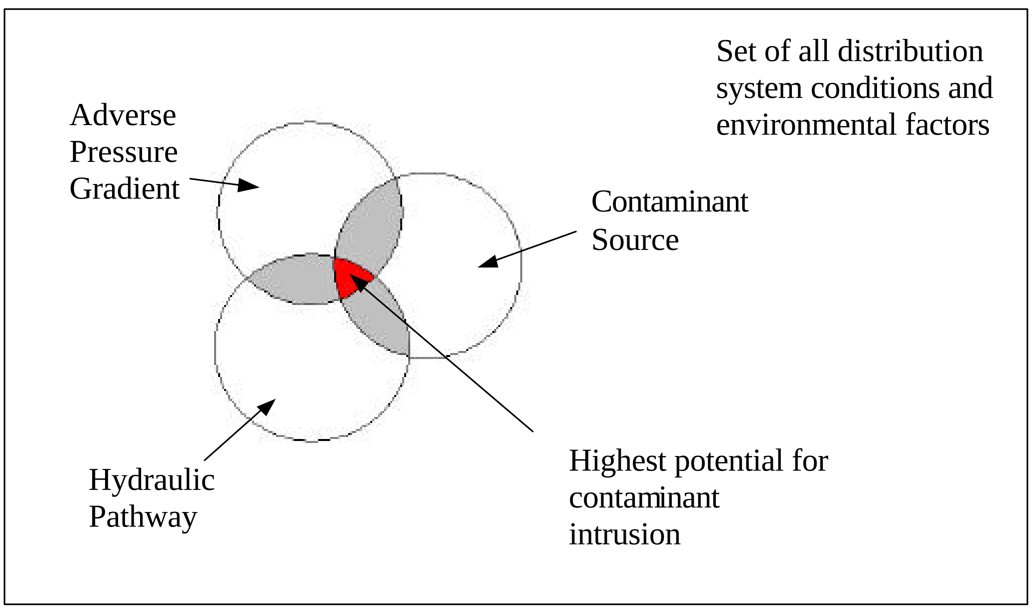


Figure 3-1. Susceptibility Conditions Contributing to Distribution System Intrusions

Table 3-2 lists some initiating events and notes which are explicitly addressed in this work. The extension of the management framework to all of the initiating events and subsequent susceptibility conditions is straightforward.

Table 3-2. Events That Initiate Susceptible Conditions

Susceptible Condition Initiated	**Initiation Event**	**Discussed Explicitly (1)**
Adverse Pressure Gradient	Extreme hydraulic stress (fire flow, pipe break, peak demand, hydrant flushing)	✓
	Extreme consumer pressure (residential on site water supply, high pressure industrial supply)	✓
Hydraulic Pathway	Cross-connection control failure	✓
	Infrastructure structural concerns (leaking pipes, valve malfunction, etc.)	✓
	Unsecured storage tank	NE
	Improper maintenance or repair practice	NE
Contaminant Source	Inadequate offset distance between distribution system and pathogen source	✓
	Residential/commercial connection outside utility cross-connection control program	✓
	Extreme environmental pathogen concentrations (i.e., sewer overflow near distribution system, etc.)	NE

[1] NE = not explicitly addressed in the case study.

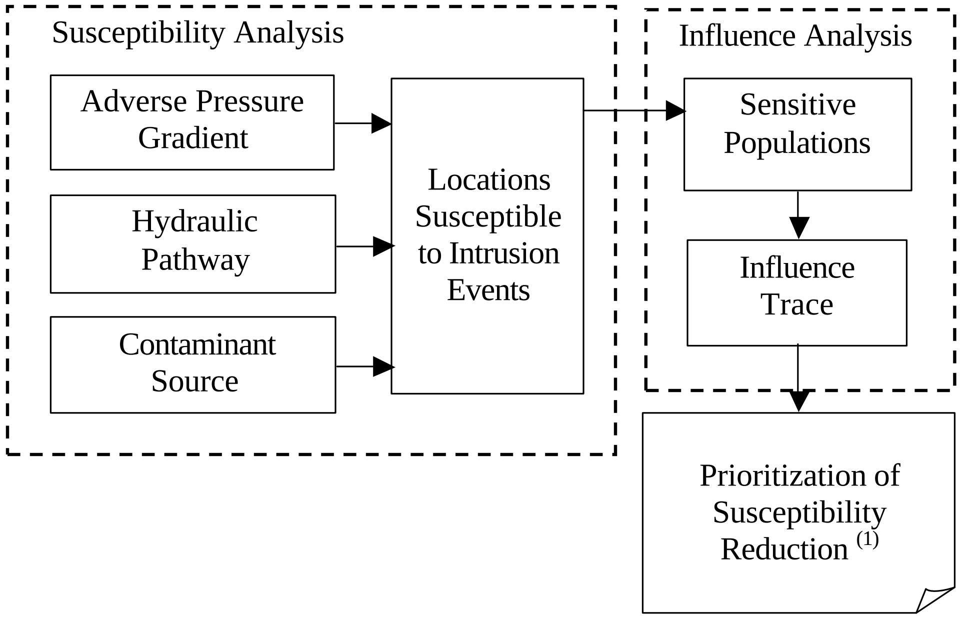
4.0 SUSCEPTIBLITY MANAGEMENT FRAMEWORK

The management framework is a method to synthesize susceptibility data and locate portions of the distribution system at risk to intrusion events. Actual implementation involves identifying and obtaining appropriate data sets, performing hydraulic simulations, estimating structural integrity, locating high risk connections or sensitive populations, and performing spatial data queries. The framework identifies locations susceptible to intrusion events by finding the joint spatial occurrence of susceptibility conditions. Once identified, the framework provides a mechanism to prioritize the susceptible locations considering the influence an intrusion event may have on hydraulically connected populations. Figure 4-1 shows how the management framework is built around the distribution system susceptibly conditions.

4.1 Susceptibly Conditions as Information Layers

Each susceptibility condition (i.e., pressure, pathway, source) is a spatial information layer. Locations susceptible to low pressure are identified through hydraulic model simulations. Locations with structural concerns are identified using leak inspection reports or repair data. In these cases, the pressure and repair data are primary data layers that when analyzed yield derived data layers of locations with pressure below some criteria or structurally unsound locations. Information acquired to support the framework will have spatial (where) and attribute (what) components. Layering data having both spatial and attribute components is a typical application of a digital Geographic Information System (GIS) (Clarke, 1999).

Barcellos (2000) used a GIS analysis to investigate health outcomes related to sanitation conditions for an urban area in Brazil. Each information source is a layer with



Sample Data Needs:
-Calibrated hydraulic model
-GIS data layers of water lines, sewer lines, repair events, service connections, other contaminant sources, sensitive populations, leak detection data

Sample Data Analyses:
-Extended period simulation hydraulic runs
-Spatial analysis in a GIS environment, geocoding, data queries

Figure 4-1. Susceptibility Management Framework

[1] Essentially a ranking of locations where steps could be taken to reduce intrusion susceptibility. Improving pressure, main repair, increased cross-connection control, etc. reduces intrusion susceptibility.

a distinct origin, purpose, and constructive characteristic that enable spatial operations and "population-at-risk" calculations in a GIS environment (Barcellos, 2000). Others have spatially correlated HPC excursions to local hydraulic conditions observed via hydraulic modeling (McMath and Casey, 2000).

Besides investigation of susceptibility layers and conditions, the proposed framework also includes an analysis associated with sensitive population groups. Once a set of susceptible locations is identified, the potential influence of an intrusion from this susceptible location on surrounding sensitive populations is determined. This influence is investigated considering the hydraulic connectivity between the potential intrusion location and the sensitive receptor(s). Modeling source water or contamination propagation is well documented (Maslia et al, 2000; Clark, 2000). The sensitive populations are included in the framework due to their perceived vulnerability to water quality upsets and the potential for future regulations addressing these populations (USEPA, 2000a; USEPA, 1999a). Table 4-1 shows how these sensitive populations are proportioned in the United States.

Table 4-1. Sensitive Subpopulations in the United States

Subpopulation [(1)]	**Number of Individuals**	**Percent of Population**
Pregnant Women	6,240,000	2.4%
Infants/Children (<10 years)	38,704,000	14.1%
Elderly (>65 years)	34,817,000	12.6%
Diabetic	15,700,000	5.8%
Liver Impairments	595,000	0.2%
Immunocompromised	400,500	0.2%
Total Estimate of Sensitive Subpopulations		**35.3%**

[1] USEPA (2000a).

Others have suggested using knowledge about sensitive population centers in drinking water distribution system management (Antoun et al, 1999). The use of potentially impacted populations adds another layer to help prioritize management issues in water distribution systems.

4.2 Implementation of the Framework

From Figure 4-1 it is clear that the hydraulic modeling used to generate pressure information and subsequent influence analysis is closely connected to other spatial information in the GIS and vice versa. Utilities will be advantaged to move beyond thinking of their hydraulic models and their GIS data as separate systems. Current modeling technologies allow for models to be built and running in time frames on the order days or weeks not months or years. Traditional skeleton models with a handful of modeled nodes connected by straight links (i.e., the links are not spatially correct except for length) will not easily support the type of analysis shown in the following case study or other analyses beyond traditional planning applications. Proper spatial representation of the modeled links allow for wider use of distribution system models in conjunction with other spatial data to solve a wide range of problems.

The framework does not implicate pipes that are routinely being contaminated. It merely synthesizes all of the known susceptibility conditions identifying those areas of the system relatively more susceptible to an intrusion. The data sources and steps shown in Figure 4-1 are not exhaustive; other data or analyses that support an assessment of distribution system susceptibly likely exist. A thorough assessment of system susceptibility and influence will have a majority of the components identified in the figure. At a minimum the susceptibility analysis must consider pressure as a risk factor.

5.0 CASE STUDY

The susceptibility management framework is tested on an actual utility. At the recommendation of utility personnel, this area was chosen due to its size, distribution of old and new lines, and relative compactness. The study area encompasses approximately 38 square miles, includes three service areas, with 280 miles of water mains (down to service connections), 18,900 connections, and an average demand of six MGD. Many of the data used in the analysis were extracted from a regional GIS; water main repair data and the hydraulic model (EPANET) were provided by the utility. Table 5-1 summarizes the data used in the case study data. The hydraulic model used for this analysis is a skeleton; it contains 100% of the pipe greater than 12 inches in diameter, 87% of the 12 inch pipe, 64% of the eight inch pipe, and 12% of the six inch pipe. The hydraulic model has been in use by the utility for planning purposes since 1995.

5.1 Susceptibility Conditions

5.1.1 Adverse Pressure Gradient Analysis

Extended period simulations (EPS) conducted under routine operating conditions (average demands) found no distribution system locations experiencing pressures less than 30 psi. (Certain locations do experience regular low pressure but these are at pump intakes or in the treatment works; these locations were not considered). To investigate the possibility of lower pressures during times of system stress, initiation events (Table 2) were investigated. Specifically, a high demand scenario equivalent to approximately two times the average demand (i.e., a summer month) was coupled with water main breaks to investigate possible pressure sensitive locations.

Table 5-1. Data Used to Implement Framework in Case Study Area

	Susceptibility Condition Or Influence Analysis	Data [1]	Source
Susceptibility	Adverse Pressure Analysis	-Calibrated hydraulic model -Repair data [2, 3, 4]	Utility hydraulic model
Susceptibility	Hydraulic Pathway Analysis	-Repair data -Water distribution system, slope and soil type, pitometer results -Service connection information [3] -Cross connection information	Utility maintenance data, regional GIS, utility pitometer surveys
Susceptibility	Contaminant Source Analysis	-Sanitary, storm, combined sewers, septic tanks -Service connection information	Regional GIS
Influence	Sensitive Populations	-Sensitive population centers; daycare centers, preschools, elementary schools, nursing homes included [3]	Health department, phone book
Influence	Trace	-Calibrated hydraulic model	Utility hydraulic model

[1]In general all data used have both spatial and attribute components. All layers had a common map projection and coordinate system to facilitate layering in a GIS environment.

[2] Utility repair data span 15 years, 1985-1999. Data set contains 748 total repairs, 520 repairs match lines still in use, 281 repairs match lines represented in the hydraulic model.

[3] Geocoded point data.

[4]All data points in the repair set are assumed to be associated with some sort of unscheduled leak or main break that required excavation and replacement. Scheduled installations, lining, replacement, bursting, etc. are not included in this set.

To couple the demands to pipe breaks, a stochastic modeling approach was used to estimate the intensity, duration, frequency, and location of low pressure. Stochastic modeling involves inputs of random variables rather than fixed values. The inputs take on a range of values based on a probability distribution; this approach captures the inherit variability of processes at work in water resource applications. Using the EPANET Toolkit and a programming algorithm, a long-term (one month) simulation was performed. Water mains were broken in a random fashion (with a historical basis) to induce sudden pressure stresses expected during an actual break event in a real distribution system. In this context, sudden is the comparison of an immediate break (lasting for a few hours) to the length of time during a one-month simulation; actual hydraulic transients were not simulated. Table 5-2 shows the random model inputs included in the pressure analysis.

Table 5-2. Stochastic Inputs to Model Water Main Breaks

Input	Probability Distribution	Typical Value	Basis
Break rate	Truncated normal	2-3 breaks/month	Utility repair history
Break attributes	Historical proportions	22% 6”, 61% 8”, 6% 12”, 10% >12”	Utility repair history
Break flowrates	Log-normal	150-300 gpm	Utility observations, fire flow available
Break duration	Function of flowrate	2-6 hours	Utility observations, crew response criteria
Break time	Uniform	Random	Engineering judgment

[1] Break rates were determined for each month, winter colder months tended to have higher break rates than warmer months; this trend been observed elsewhere (AWWARF, 1986). Typical value of 2 to 3 per month represents August conditions.

[2] Overall, six-inch diameter mains have the highest break rates, however, few six inch mains are in the hydraulic model. The percentages above specifically support the break algorithm (integrated within the hydraulic model) and therefore represent break proportions associated with modeled mains. Vintage was also an attribute used in determination of which mains to break (percentages not shown).

In a distribution system hydraulic model and in reality, pipe breaks are demands; demands can only be assigned to model nodes (junctions). To more equitably spread the break locations, analogous to actual system breaks, additional nodes (and links) were added to the skeleton hydraulic model as direct replacements to existing links (approximately 1152 new nodes in the study area). The connecting nodes between the new links are zero demand nodes serving as placeholders for a potential break location. These new nodes and links were created at the location of the actual pipe segments they represent in space. Adding the new nodes and model links where they actually occur in space moved the hydraulic model from a traditional skeleton, where only a hand full of nodes are modeled with straight links in between, to a skeleton whose modeled pipes actually have spatial representation. Figure 5-1 shows this comparison between the original skeleton and the newly created skeleton. Represented properly in space, the new nodes can inherit all the necessary spatial attributes needed to support pipe breaks in the pressure algorithm.

In the case study area, diameter and vintage are key components influencing break rates; therefore these new nodes inherit the diameter and vintage attributes of the pipes they represent. If bedding conditions significantly influence the break rates then these potential break nodes inherit a local soil type and a corresponding break rate for that soil type. This type of spatial attribute inheritance occurs only when locality is considered. Water mains (now represented by a set of nodes) are “broken” in a manner statistically indistinguishable from the actual repair record.

A.

B.

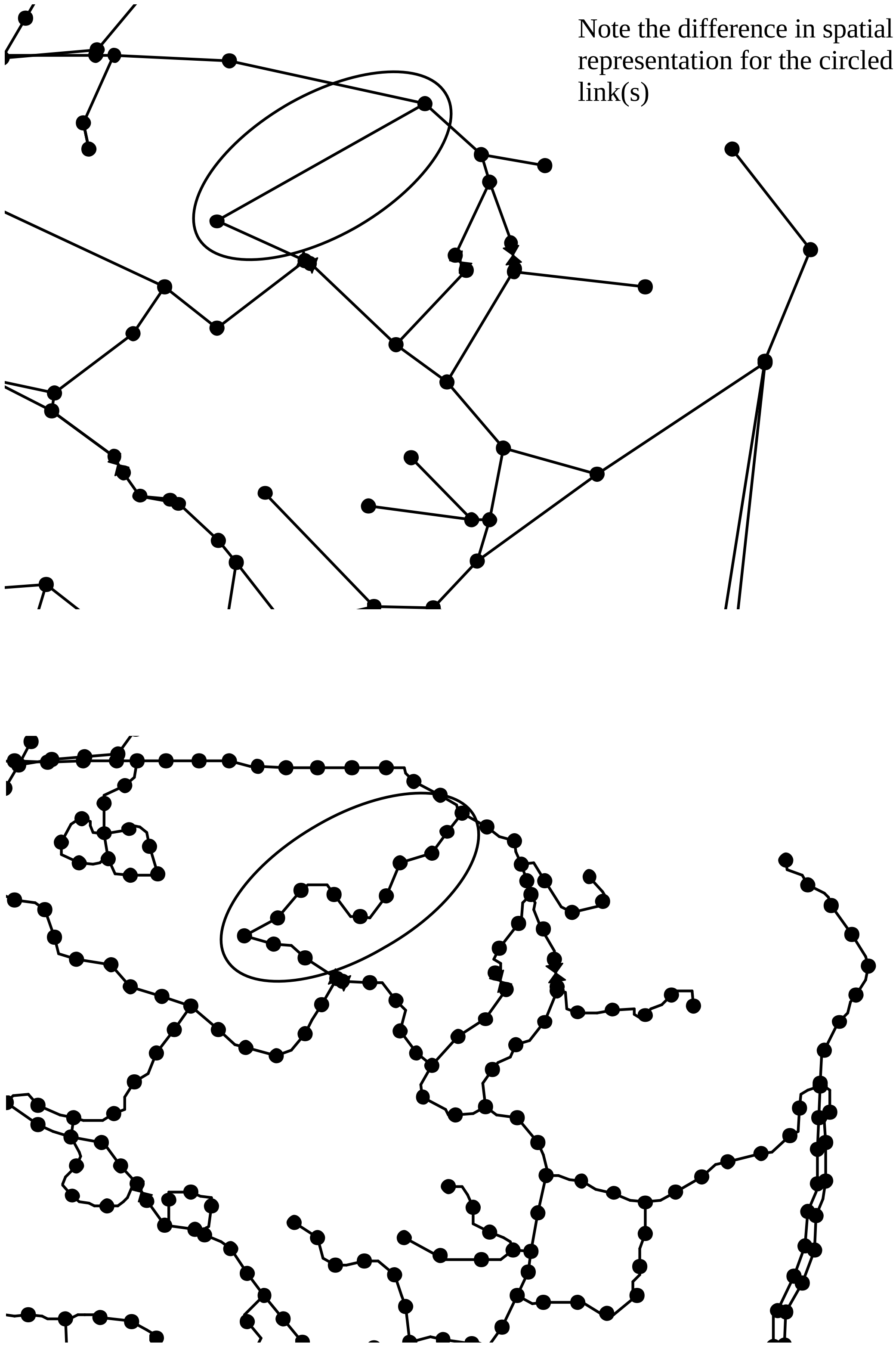


Figure 5-1. Original Skeleton (A) and New Skeleton (B)

A broken main is activated using a version of the emitter function in EPANET (USEPA, 2000b).

$$q = Cp^{0.5} \qquad (1)$$

where:

q = flow rate (gpm)

C = discharge coefficient (gpm / $\sqrt{psi}$)

p = pressure (psi)

In the algorithm developed for the pressure analysis, the discharge coefficient is a random variable whose range of outcomes has a basis in the flow rate of breaks observed by the utility and flow available for fire demands. Figure 5-2 shows log-normal discharge coefficient outcomes for simulated breaks of 6 inch, 8 inch, 12 inch, and 24 inch diameter pipe (100 breaks for each diameter). Figure 5-3 shows corresponding break flow rates at various pressures based on the median (measure of central tendency for a lognormal distribution) discharge coefficient value. From Figures 5-2 and 5-3 a “typical” break on a 6 inch line (log C at 50% = 1.46, C=29.1) would flow at about 200 gpm given a pressure of 50 psi. Smaller breaks and larger breaks are possible.

The single month simulation was repeated hundreds of times in the context of a Monte Carlo simulation, generating probability distributions for low pressures. Each one-month simulation is a historical realization of events, including system operations and main breaks that occur in the distribution system. Figure 5-4 shows locations expected to have a high frequency of break occurrences and locations expected to have a high frequency of low-pressure occurrences. Figure 5-4 reveals locations with higher

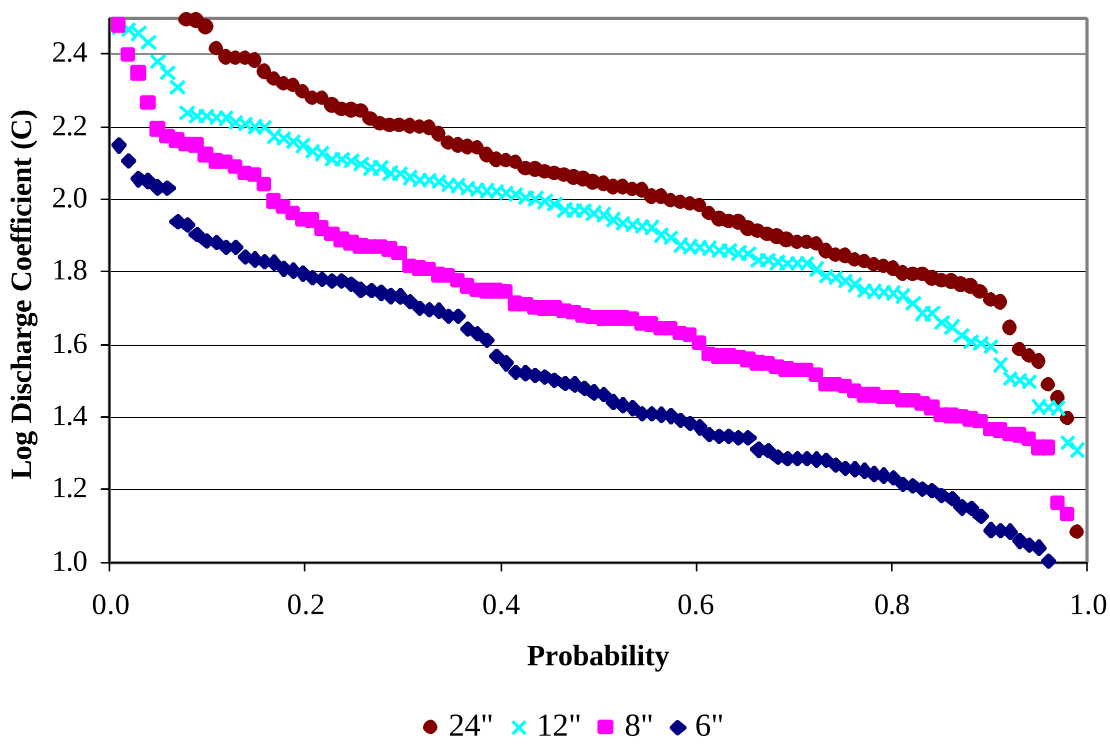


Figure 5-2. Log-Normal Probability Plots of The Discharge Coefficient (C) for 100 Simulated Breaks of Each Diameter (400 total breaks)

tendencies for low pressure are not correlated strongly in space to locations with high tendencies for breakage. In Figure 5-4, frequency is at a monthly time scale. The reported break locations experienced at least one break in five or more simulations (five out of 500) and the low pressure locations experienced pressures less than 20 psi at least once in five or more simulations (five out of 500).

Figure 5-5 summarizes the critical pressure data showing that over 12% of the sampled distribution system nodes (N=1524) experienced negative pressure between zero and five times during the Monte Carlo simulation. However, the frequency in Figure 5-5

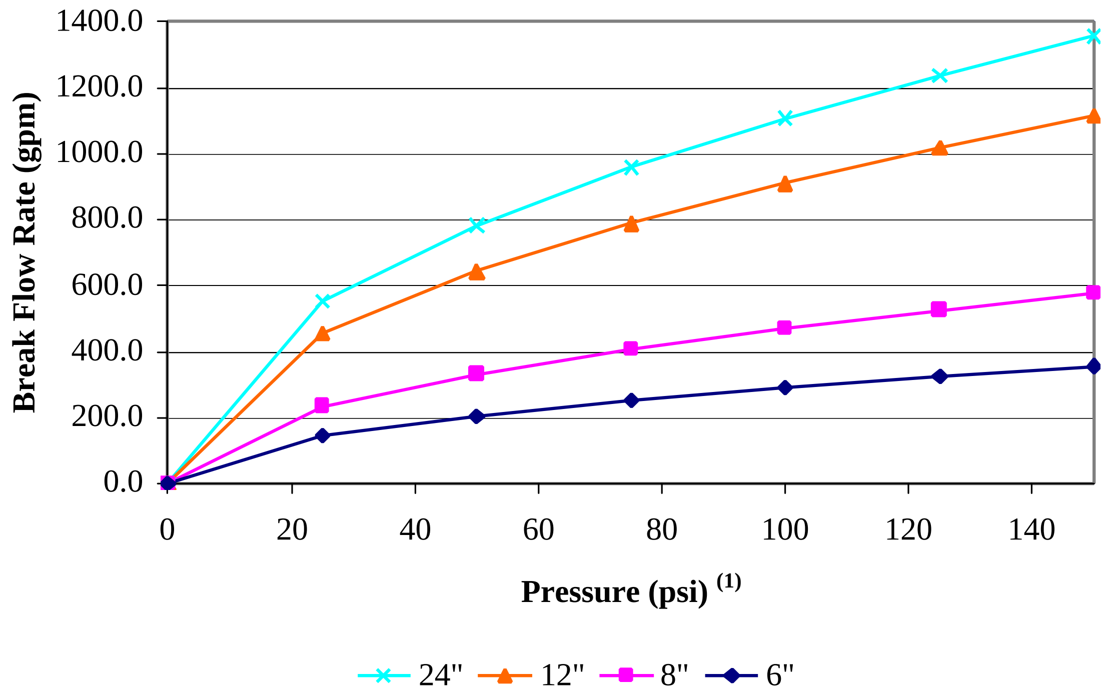


Figure 5-3. Typical Break Flow Rates Used to Initiate Pressure Susceptibility

[1] Curves generated from equation (1), C =50% value from Figure 5-2.

is at an hourly scale, the scale the pressures are sampled from the model. Five measured pressure occurrences at a given location must be compared to the number of times pressure was observed at that location. For hourly observations, over 31 days, simulated 500 times the result is 372,000 pressure observations at each node. Therefore a node that experiences five occurrences less than zero psi has approximately a 1 in 75,000 chance of experiencing negative pressure at any given hour. The location most sensitive to low pressure experienced 25 occurrences less than zero psi or about a 1 in 15,000 chance of negative pressure during any given hour. At least one negative pressure reading due to high demands and simulated breaks was observed 43 out of the 500 simulations

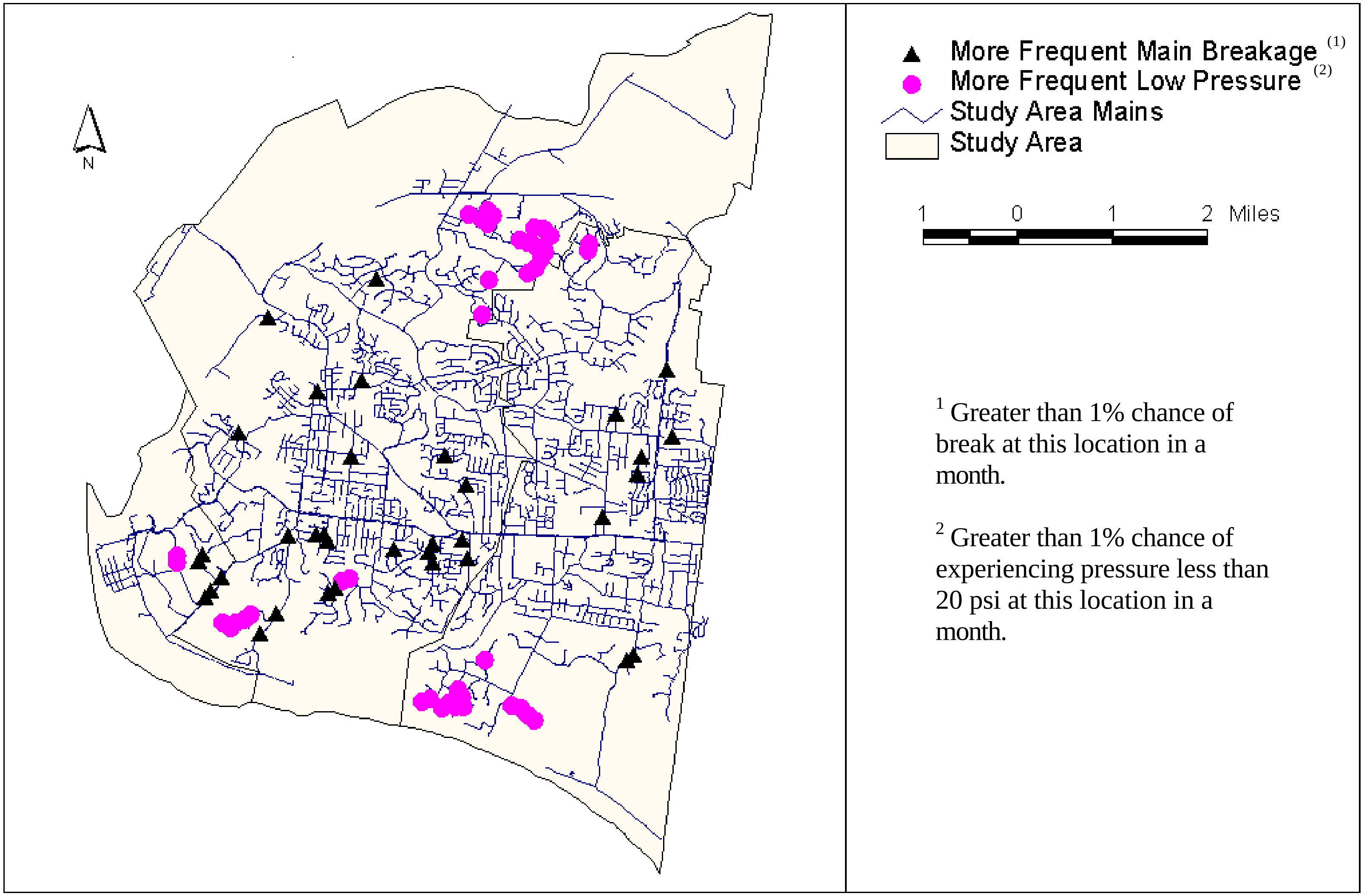


Figure 5-4. Locations With High Break Frequency and High Frequency of Low Pressure Events

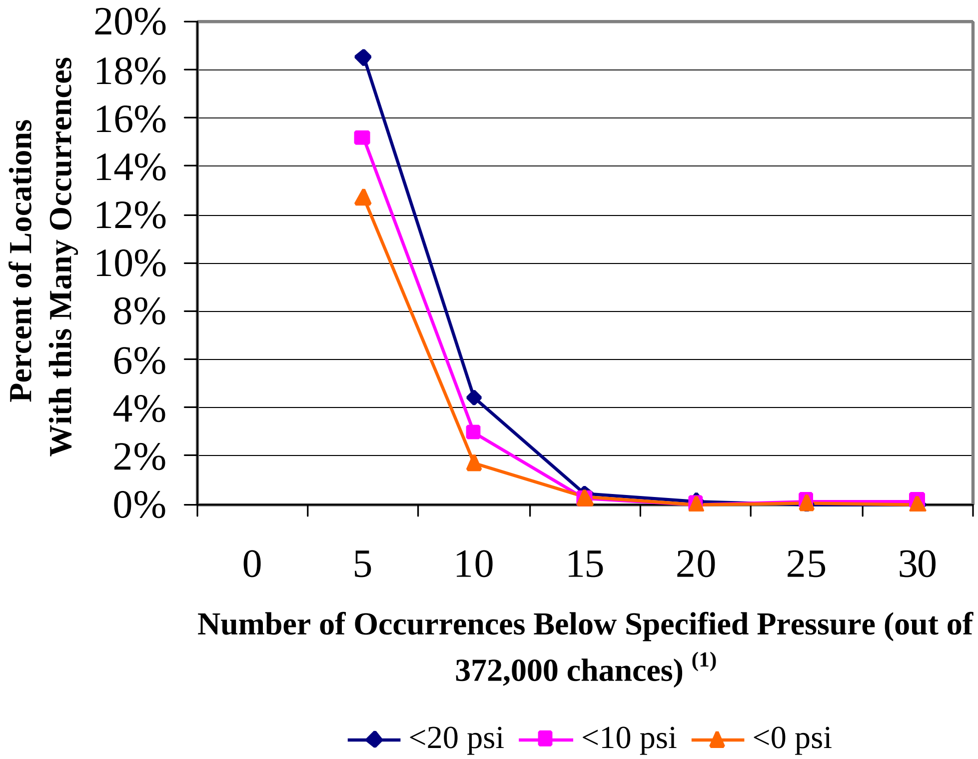


Figure 5-5. Pressure Summary From Hydraulic Model Run Under Summer Demand Scenario (10-12 MGD) and Random Main Breaks

[1] Five locations had more than 30 occurrences less than 20 psi, one location had more than 30 occurrences less than 10 psi (not shown).

representing an 8% chance of having the pressure less than zero somewhere in the study area during the month.

This result does not include other low-pressure initiation events such as maintenance flushing, fire flows, or hydraulic transients; transients cannot be addressed with a model such as EPANET (AWWARF, 2001). Table 5-3 shows a correlation matrix between

Table 5-3. Correlation Matrix Between Frequency of Low Pressure and Other Model Parameters (Pearson Coefficient, n=225)

	Number of Occurrences < 0 psi at a Node [(1)]	Elevation of the Node	Minimum Pressure at Node Under Normal Conditions	Water Age at the Node at the Time Minimum Normal Pressure Was Observed
Number of Occurrences < 0 psi at a Node	1.0	-	-	-
Elevation of the Node	0.435	1.0	-	-
Minimum Pressure at Node Under Normal Conditions [(2)]	-0.215	-0.669	1.0	-
Water Age at the Node at the Time Minimum Normal Pressure Was Observed	-0.051	-0.092	0.193	1.0

[1] n=225 is the number of pressure monitoring locations (model nodes) with at least one pressure observation < 0 psi during the Monte Carlo simulation. The correlated variables are collected at these 225 locations. The correlation is for a location and compares the number of hourly observations less than zero to the value of the parameter being correlated to.

[2] Normal conditions imply average demands and routine operations.

three parameters and the number of times certain locations experience pressures less than 0 psi. The table portrays how these locations of critical pressure are generated by the inherent complexity of the system and the random nature of a perturbation such as a water main break. Looking simply at the high points in the system, locations with normally lower pressures, or areas with older water will not always yield locations susceptible to low pressure and subsequent water quality upsets. Older water, not associated with low pressure, is noted in this context because some have suggested water age may be a generic indicator of water quality. Water age maybe a valid indicator for parameters such as taste, odor, or disinfectant residual but from this result (and intuition) it is likely not a valid parameter to assess the potential for intrusion related water quality concerns.

5.1.2 Hydraulic Pathway Analysis

The hydraulic contamination pathways investigated in the case study include service connections (controlled and uncontrolled) and structurally unsound locations along the pipe network. All service connections (industrial, commercial, residential) were assigned to a modeled node (where the pressure is known). This allows quantification of the number of service connections associated with locations of low pressure. Figure 5-6 shows the distribution for the number of service connections assigned to each pressure monitoring location. On average (arithmetic) each pressure monitoring location in the hydraulic model serves 13 to 14 connections.

Estimating the structural integrity of buried infrastructure in the study area to identify contamination pathways is less straightforward. Much of the structural analysis shown here is based on guidance given in an AWWARF research report (1986) using available GIS layers. Break variables diameter, vintage, material, bedding slope, and bedding soil type were investigated using 15 years (1985-1999) of utility repair data. This analysis was done within the GIS. Geocoded break data (from utility repair histories) were matched to water mains to summarize diameter, vintage, and material break trends. The

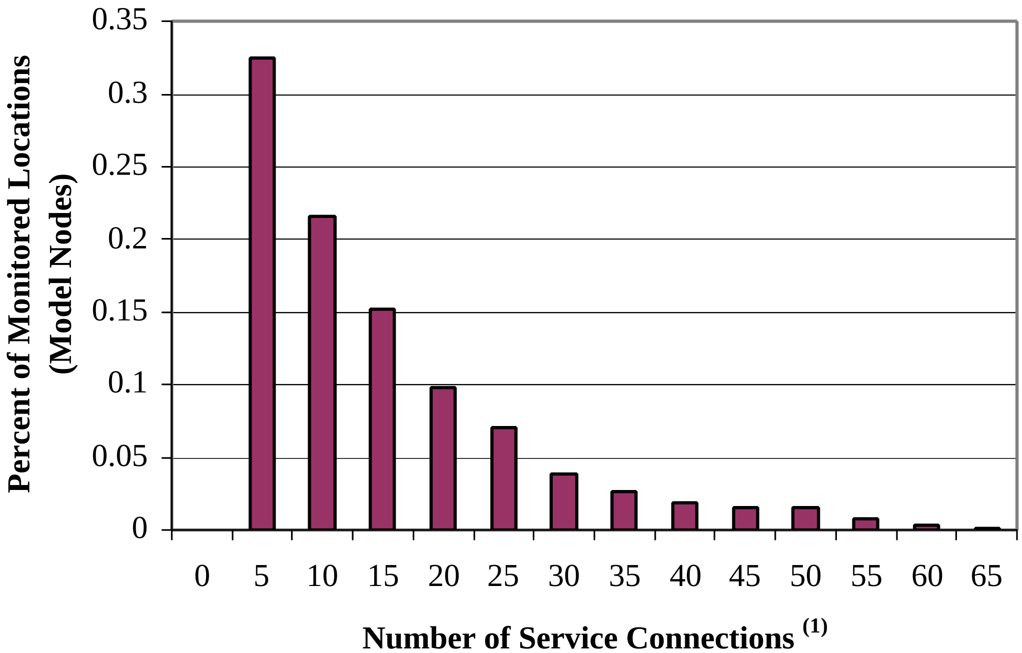


Figure 5-6. Distribution of Service Connections As Assigned to Monitored (Modeled) Nodes

[1] Assignment of service connections (i.e., users) to a modeled node was done in the GIS strictly based on the straight-line distance between the modeled nodes and the users location.

break locations were also summarized against a slope layer and soil type layer. All results were normalized against the length of pipe in the respective category in existence for each year of the break record (e.g. six inch breaks in 1985/miles of six inch pipe existing in 1985). This activity resulted in a break rate measured in terms of breaks/mile for each year and category. These yearly break rates were then investigated over the 15 year data set to investigate any apparent correlations or underlying structure of the data.

The results of this analysis suggest for this study area, diameter, material, and vintage play important roles in pipe breaks. However, those vintages with poor break histories were all of the same pipe material thus the same information is available from two key variables, diameter and vintage (in general, over 90% of the breaks were in cast iron pipe all installed prior to 1975). The structural findings for this study area are in line with observations by the utility relative to water mains with less than favorable break histories. Six-inch diameter pipes installed between 1940 and 1970 have break rates much higher than the nearest category investigated.

5.1.3 Contamination Source Analysis

The contamination sources investigated include sanitary sewer lines, septic systems, and known high risk service connections. For the subsurface contaminant sources, the offset distance between the potable water mains and the sources were analyzed. Figures 5-7 and 5-8 show the results of this analysis for sanitary sewer lines and septic systems.

A layer of service connection information and parcels with zoning information was used to investigate high-risk service connections as a source of contamination. The utility maintains a rigorous cross-connection control program including process risk assessment at the time of connection and yearly inspections of control devices. Data collected by the utility at the time of connection (i.e., information about a customer's process) are not accessible in a digital format. Therefore, explicit identification of connections considered high risk was done by filtering on the connection's branch size in the customer information attributes. This resulted in identification of 161 locations (with service branches greater than 4 inches in diameter). These locations were then screened manually (generally by name) to determine if the facilities could be considered high risk. 28 facilities were eventually identified as high-risk connections. They are generally

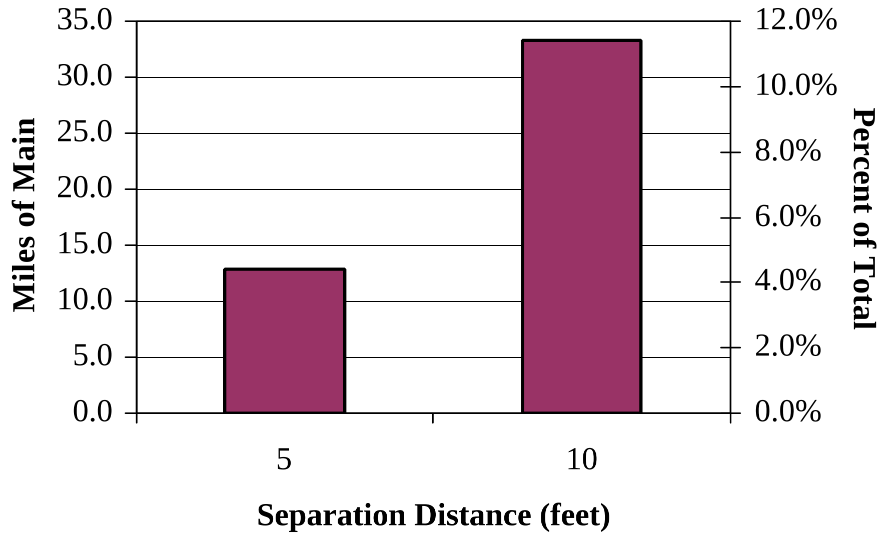


Figure 5-7. Miles of Water Main Within Given Distance of Sanitary Sewer Lines

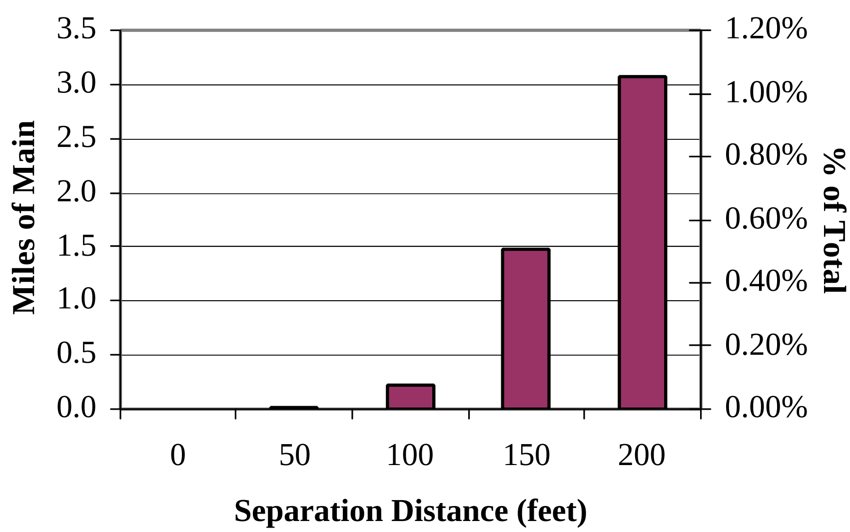


Figure 5-8. Miles of Water Main Within Given Distance of Septic Tanks

industrial or commercial facilities with one or more, large branch connections (> 4 inch) and contain some type of internal process deemed a concern (e.g. hospitals, mortuaries, chemical process locations). All of these are in the utilities cross-connection control program.

5.2 Combined Susceptibility Results

At this point, data associated with the three susceptibility conditions namely adverse pressure gradient, hydraulic pathway, and contaminant source have been investigated. These results were combined yielding a set of locations deemed more susceptible to contamination by external intrusion. As depicted in Table 3-1 and Figure 3-1, the joint set of one, two, or three susceptibly conditions yields different potential intrusion mechanisms.

5.2.1 Backflow Susceptibility

Figure 5-9 shows the number of service connections served by each node experiencing pressures less than 20 psi during the critical pressure simulations (high demands coupled with main breaks). This plot identifies locations experiencing the joint susceptibility conditions of adverse pressure gradient and hydraulic pathway via a direct service connection. A data point plotting far to the right on the x-axis and near the bottom has a higher frequency of low pressure but its influence sphere includes fewer service connections. Points high on the y-axis and to the left serve many connections but have a lower frequency of low pressure observations. Points to the right and upper parts of the graph would be deemed most susceptible to backflow during low pressure at a service connection where flow control is absent of failing. The result suggests no location, for this study area under these conditions, has both a high probability of low

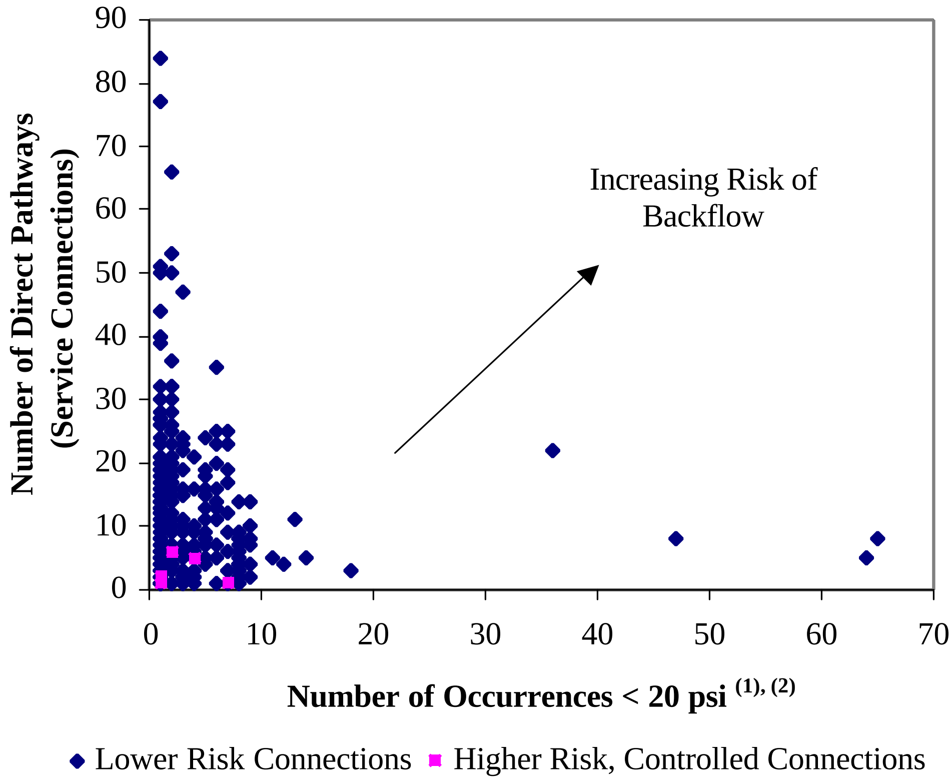


Figure 5-9. Susceptibility Plot for Backflow Events at Controlled and Failing or Uncontrolled Service Connections

[1] Each data point represents a node that experienced at least one pressure observation less than 20 psi during a simulated main break.

[2] Series labeled "Lower Risk Connections" represent modeled nodes (where pressure is known) that are serving generally residential or lower risk commercial connections. The other series labeled as higher risk serves some connections (all controlled) felt to have internal processes that could serve as a contamination source as described in the text.

pressure and a high number of service connections, although some locations have one or the other. The joint occurrence of adverse gradient, hydraulic pathway (via a controlled or uncontrolled connection) and contaminant source (i.e., explicitly identified high-risk connection), exists at nine locations in the study area (see Figure 5-9). The locations farther to the right and higher in Figure 5-9 (12 locations) and the locations with at least one low-pressure observance serving a high-risk connection will be saved as locations more susceptible to back flow (n=21 total locations).

5.2.2 Subsurface Intrusion Susceptibility

Additional layering of susceptibly conditions provides subsurface intrusion information. Table 5-4 shows this result.

Table 5-4. Susceptibility Results for Intrusion Events at Structurally Deficient Locations

Susceptible Condition	**Criteria**	**Length in category, and percent of system total (1)**
Adverse Pressure Gradient	At least one occurrence < 0 psi (2)	32.7, 11.6%
Hydraulic Pathway	6" diameter mains installed between 1940 and 1970 (3)	35.0, 12.4%
Contaminant Source	Mains within 10 feet of sewers and 200 feet of septic tanks	35.4, 12.5%
	Joint set of Pathway and Source	**4.4, 1.6%**
	Joint set of Pressure, Pathway, Source	**0.6, 0.02%**

[1] Approximately 282 total miles of water main.

[2] At least one occurrence out of 372,000 hourly observations at each modeled node where the pressure is known.

[3] Utility repair data suggest six inch pipe in these vintages have excessive break rates. This knowledge is now assumed as a predictor for pipes with structural concerns and prone to contaminant entry during depressurization.

As with Figure 5-9, the results for the study area suggest only a handful of locations have the three susceptibility conditions in common in regards to subsurface contamination via structurally unsound water mains. The lack of having all susceptibility conditions in common minimizes the intrusion possibility.

Figure 5-10 shows the study area with the resultant locations having susceptibility conditions in common. In Figure 5-10, the resultant layer scripted with a two represents locations (water mains) with both a structural pathway condition and a contaminant source condition in common. These locations are not a concern during normal or even stressed operation due to a pressure surface resistant to extreme fluctuations. However, they could become a concern during actual repair events. Improper disinfection practices during repair at these locations may result in pathogenic contamination due to the relative proximity of the pathogen source. Increased application of existing disinfection practices or construction methods maybe warranted at these locations.

The final group of susceptible locations (21 potential backflow and 0.6 miles of subsurface intrusion) can now be further analyzed by investigating the influence these locations have on sensitive populations.

5.3 Influence Analysis

5.3.1 Sensitive Populations

Population centers identified in the study area include day care centers, preschools, elementary schools, adult day care centers, retirement communities, and nursing homes. The locations were added to the GIS via geocoding against a matchable street layer (Clark, 1999). These facilities were also contacted and questioned relative to the

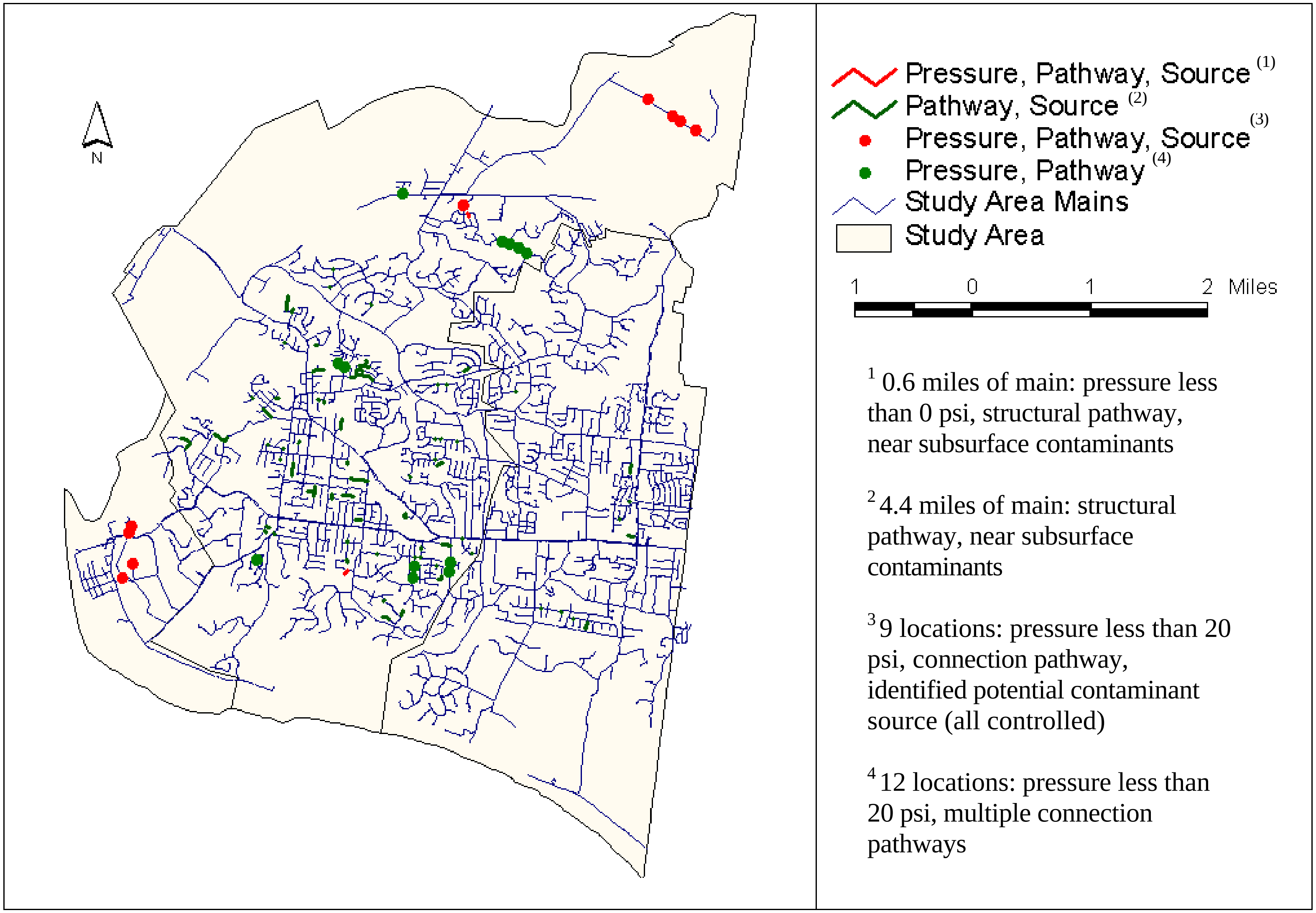


Figure 5-10. Locations with Combined Susceptibility (see Table 5-4 and Figure 5-9 for Criteria Discussion)

Table 5-5. Study Area Summary of Sensitive Populations

Group	Number	Total Number of Occupants in Group [(1)]	Average time at location (years) [(1)]
Preschools/Day Care Centers	10	789	14
Elementary Schools	8	5398	34.5
Elder Care Centers	8	1040	12
Totals	26	7227	21.5

[1] Based on simple telephone survey.

numbers of inhabitants at the location on a typical day and the length of time the facility has been located at its present address. Table 5-5 summarizes the results of the questions posed to the population centers. The total estimate of the sensitive inhabitants (7,227) represents approximately 10% of the total study area population. However, the number of sensitive service connections, 26, only represents 0.14% of the total connection count. This observation alone suggests targeting these populations as receptors allows a utility to address a large percentage of the population more susceptible to water quality upsets while managing fewer locations.

5.3.2 Influence Trace

The susceptible locations are now combined with the sensitive populations to complete the framework by investigating the hydraulic connectivity between a potential source and the sensitive population. In the event the susceptibility analysis yields multiple locations of concern, (perhaps too many to address at once), this final analysis serves as a prioritization mechanism. Connectivity was investigated using the trace function in EPANET. The susceptible locations (i.e., model nodes representing susceptible locations) were set as source nodes and the number of receptors hydraulically connected to the source was totaled. Of the 23 locations set as source nodes (21 locations associated with the connection pathway and 2 locations representing the 0.6 miles of main susceptible to subsurface intrusion) two were hydraulically connected to sensitive populations. These two locations are potential backflow locations meeting pathway (via connections) and pressure conditions. Figure 5-11 identifies shows the susceptible locations hydraulically connected to the sensitive populations.

5.3.3 Intrusion Modeling

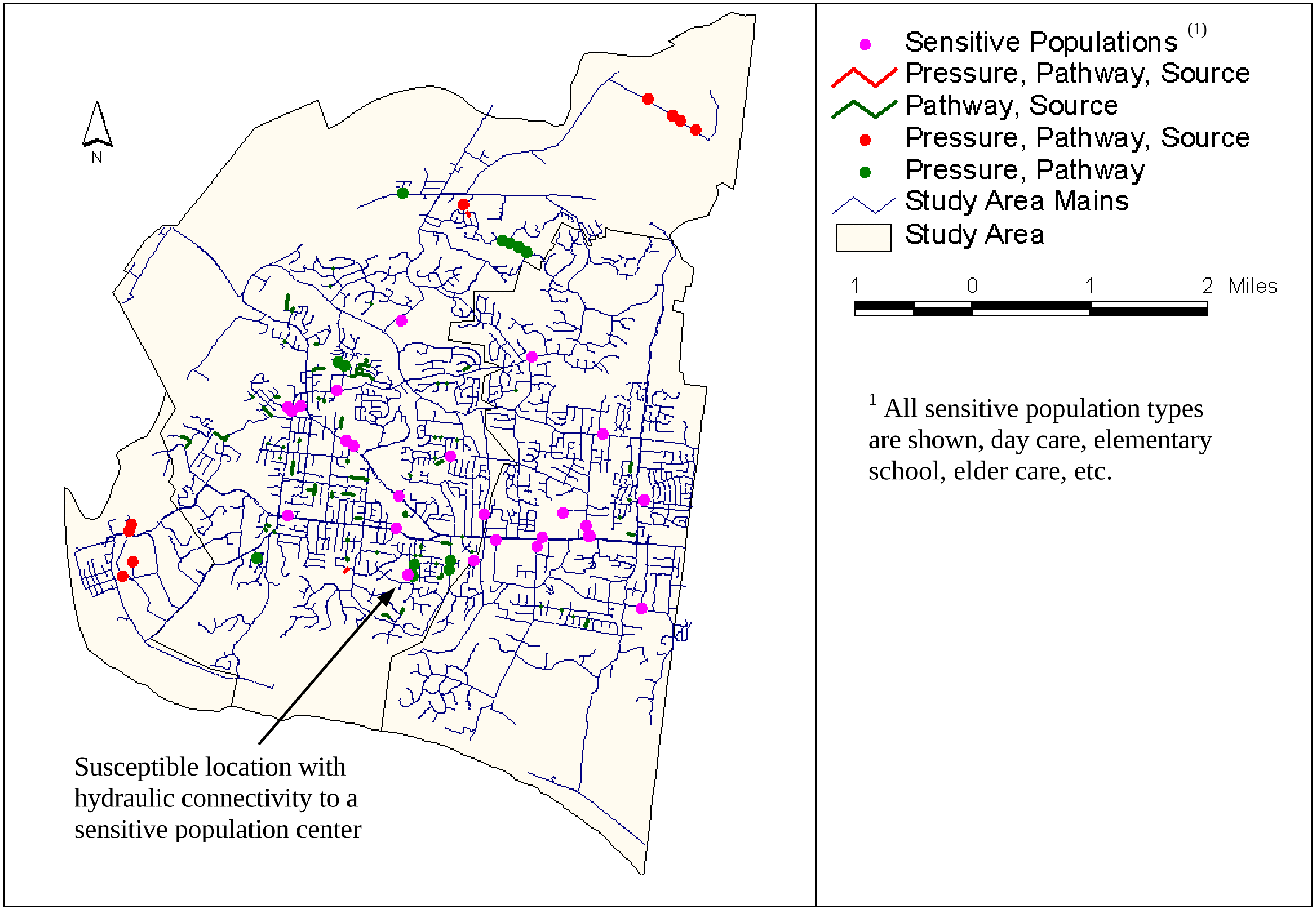
Besides using the trace function one potential source node at a time, the critical pressure algorithm developed to simulate breaks was also used to simulate intrusion events into the distribution system. During the simulation, an intrusion was assumed for any location that experienced pressures less than 1 psi. In the interval of low pressure, the intrusion was simulated via placement of a random mass inflow into the line at this location. The mass inflow was shut off when the pressure exceeded 1 psi. This analysis implies that a hydraulic pathway and contaminant source exist at all locations throughout the distribution system and only a low-pressure situation is required to initiate a contamination event. The random mass inflow at low-pressure locations results in pulses of mass at other locations connected to the source. The area under the curve for each of these pulses was calculated. The results were ranked and plotted on normal probability paper resulting in Figure 5-12. Figure 5-12 can be repeated for every location receiving mass from nodes where intrusions have occurred.

5.4 Discussion

The reliability of a looped system is reiterated in these results. The majority of the locations experiencing rare pressure problems are located on the system periphery where the network is more branched than in the central part of the study area. Even though the central part of the study area has mains with a predicted high chance of failure (Figure 5-4), connected mains are buffered against these initiation events due to the looping of the network. This resistance to pressure fluctuation eliminates a susceptibility condition and makes contamination events practically impossible.

In the case study, susceptible conditions exist at some locations in the system but few locations have all conditions in common at the same time. Additionally, the sensitive populations are for the most part located in the central regions of the area, where the pressure buffering due to the looped network is most profound. This spatial "serendipity" means those locations with the highest potential for intrusion events have little affect (considering hydraulic connectivity as a metric) on the populations of concern.

The susceptibility investigation for this study area suggests the system is possibly more at risk to backflow via the service connection pathway than due to a subsurface intrusion along a structurally deficient water main. This result is likely true of most systems noting the distribution system deficiencies resulting in illness outbreaks (see Table 3-1).



[1] All sensitive population types are shown, day care, elementary school, elder care, etc.

Figure 5-11. Susceptible Locations and Sensitive Populations

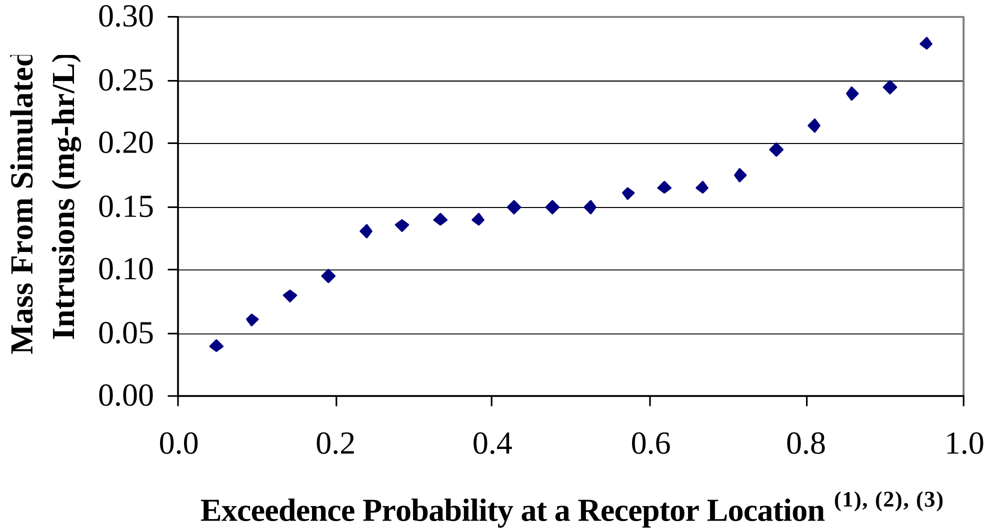


Figure 5-12. Normal Probability Plots of Area Under the Mass Curve for Any Location Influenced by A Node Susceptible to Intrusions

[1] Each data point represents one mass event passing the node being observed. The figure above applies to a single observation node. The location being influenced in this figure arises from random intrusions at low-pressure locations during simulated pipe breaks and high demands. The results shown in Figure 5-12 are for a location in the northern portion of the study area.

[2] Plotting positions on the figure are identified by ranking each mass value (1 to n), the corresponding probability is determined from $p = m/(n+1)$ where p is the probability, m is the rank, and n is the number of values in the set. The exceedence probability is calculated as 1-p.

[3] The units on the y-axis arise from calculation of the area under the curve for each mass pulse. The hydraulic model reports a concentration passing a location in mg/L. This is then multiplied by the time the concentration is observed (hours). Multiplication of this value (mg-hr/L) by the flowrate (L/hr) of water moving with the mass pulse will yield the mass in milligrams of intrusion at an influence location.

6.0 CONCLUSION

Waterborne illness attributable to distribution system intrusions is a secondary manifestation of three underlying susceptibility conditions: adverse pressure gradient, hydraulic pathway, and contaminant source. A hydraulic model and spatial information related to specific intrusion risk factors were integrated within a GIS to identify areas of the distribution system susceptible to intrusion events. Once identified, they were further prioritized considering how they influence (hydraulic connectivity) local sensitive populations.

The distribution system is the final barrier to waterborne illness. However, certain initiation events (i.e., pipe breaks coupled with large demands), which may be extreme but not improbable, can result in the occurrence of multiple susceptible conditions at a single location. The presence of the susceptible conditions does not mean intrusion is imminent; it simply means the necessary conditions exist for an intrusion to occur. Identification of these critical conditions has been suggested as part of a comprehensive operating plan or distribution system sanitary survey (AWWARF, 2001). The results of the framework, may support utility capital improvement plans, infrastructure maintenance, improved cross-connection control, accreditation procedures, and provide a basis for regulatory sampling designs. This kind of approach may actually improve water quality and protect public health more than extensive sampling and monitoring efforts (Allen et al, 2000).

7.0 REFERENCES

Allen, M.J., Clancy, J.L., and Rice, E.W., 2000. The Plain, Hard Truth About Pathogen Monitoring. *Jour. AWWA* 92:9:64.

Antoun, E.N., Dyksen, J.E., and Hiltebrand, D.J., et al, 1999. Unidirectional Flushing: a Powerful Tool. *Jour. AWWA 9*1:7:62.

AWWA Roundtable discussion, 1999. The disinfectant residual dilemma. *Jour. AWWA* 91:1: 24.

AWWARF (AWWA Research Foundation), 1986. *Water Main Evaluation for Rehabilitation/Replacement.* AWWA, Denver, CO.

AWWARF (AWWA Research Foundation), 2001. *Pathogen Intrusion Into the Distribution System.* AWWA, Denver, CO.

Barcellos, C., 2000. Health Risk Analysis of the Rio de Janeiro Water Supply Using Geographic Information Systems. *Interdisciplinary Perspectives on Drinking Water Risk Assessment and Management* (Proceedings of the Santiago (Chile) Symposium, September 1998) *IAHS*, 260:95

Brothers, K.J., 2000. Water Leakage and Sustainable Supply-Truth or Consequences? *Jour. AWWA*, 93:4:150

Clark, K., 1999. *Getting Started With Geographic Information Systems.* Prentice Hall, Upper Saddle River, NJ.

Clark, R., 2000. Water Quality Modeling-Case Studies, in L.W. Mays, ed., *Water Distribution Systems Handbook*, McGraw-Hill, New York, N.Y., chapter 13.

Crabtree, K.D., Gerba, C.P., Rose, J.B., and Haas, C.N., 1997. Waterborne Adenovirus: A Risk Assessment. *Water Science and Technology*, 35:11-12:1.

Craun, G.F. and Calderon R.L. Waterborne Outbreaks Caused by Distribution System Deficiencies in the United States. In press.

Frost, F.J., Craun, G.F., and Calderon, R.L., 1996. Waterborne Disease Surveillance. *Jour. AWWA*, 88:9:66.

Fenner, R.A., and Sweeting, L., 1999. A Decision Support Model for the Rehabilitation of Non-Critical Sewers." *Water Science and Technology*, 39:9:193.

Geldreich, E.E., 1996. *Microbial Quality of Water Supply in Distribution Systems.* CRC Lewis Publishers, Boca Raton, FL.

Grayman, W.M., Rossman, L.A., and Geldreich, E.E., 2000. Water Quality, in L.W. Mays, ed., *Water Distribution Systems Handbook,* McGraw-Hill, New York, N.Y., chapter 9.

Grayman, W.M. and Kirmeyer, G.J., 2000. Quality of Water in Storage, in L.W. Mays, ed., *Water Distribution Systems Handbook*, McGraw-Hill, New York, N.Y., chapter 9.

Griffis, S.M. and Ivey, R.J., 1998. Aging Sewers Represent Opportunity for Industry." *Environmental Technology,* March/April.

Haas, C.N., 1999. Benefits of Using a Disinfectant Residual. *Jour. AWWA,* 91:1:65.

Habibian, A., 1992. Developing and Utilizing Data Bases for Water Main Rehabilitation." *Jour. AWWA,* 83:7:75.

Hadzilacos, T., Kalles, D., Preston, N., Melbourne, P., Camarinopoulos L., Eimermacher, M., Kallindromitis, V., Frondistou-Yannas, S., and Saegrov, S., 2000. "UtilNets: a Water Mains Rehabilitation Decision-Support System." *Computers, Environment and Urban Systems,* 24:215.

ILSI Risk Science Institute Pathogen Risk Assessment Working Group, 1996. A Conceptual Framework to Assess the Risks of Human Disease Following Exposure to Pathogens. *Risk Analysis,* 16:6:841.

Kessler, A., Ostfeld, A., and Sinai, G., 1998. "Detecting Accidental Contaminations in Municipal Water Networks." *Jour. Water Resources Planning and Management,* 124:4:192.

Kiene, L., Lu, W., and Levi, Y., 1998. Relative Importance of the Phenomena Responsible for Chlorine Decay in Drinking Water Distribution Systems. *Water Science and Technology,* 38:6:219.

Kleiner, Y., and Rajani, B., 1999. Using Limited Data to Assess Future Needs. *Jour. AWWA* 91:7:47.

LeChevallier, M.W., 1999. The Case for Maintaining a Disinfectant Residual. *Jour. AWWA,* 91:1:86.

Lee, B.H., Deininger, R.A., and Clark, R.M., 1991. Locating Monitoring Stations in Water Distribution Systems. *Jour. AWWA,* 82:7:60.

Maslia, M.L., Sautner, J.B., Aral, M.M., Reyes, J.J., Abraham, J.E., and Williams, R.C., 2000. A Protocol for Using Water-Distribution System Modeling as a Tool to Assist Epidemiological Investigations. *Jour. Water Resources Planning and Management*, 126:4:180.

McMath, S.M., and Casey, R.J., 2000. The Use of Network Modelling Techniques to Resolve Water Quality Issues in Distribution Systems. Proceedings AWWA WQTC, Nov. 2000, Salt Lake City, Ut.

Morris, R.D. and Levin, R., 1994. Estimating the Incidence of Waterborne Infectious Disease Related to Drinking Water in the United States. *Assessing and Managing Health Risks from Drinking Water Contamination* (Proceedings of Rome (Italy) Symposium, 1994) *IAHS*, 233:75.

New York Department of Environmental Protection, 2000, www.ci.nyc.us/html/dep/html/wdrap.html (accessed 6/2000)

Nuckols, J.R., Stallones, L., and Reif, J., 1994. Evaluation of the Use of a Geographic Information System in Drinking Water Epidemiology." *Assessing and Managing Health Risks from Drinking Water Contamination* (Proceedings of Rome (Italy) Symposium, 1994) *IAHS*, 233:111.

Payment, P., Richardson, L., Siemiatycki, J., Dewar, R., Edwardes, M., and Franco, E. 1991. A Randomized Trial to Evaluate the Risk of Gastrointestinal Disease Due to the Consumption of Drinking Water Meeting Currently Accepted Microbiological Standards. *Amer J Public Health,* 81:703.

Payment, P., Siemiatycki, J., Richardson, L., Renaud, G., Franco, E., and Prevost, M., 1997. A Prospective Epidemiological Study of Gastrointestinal Health Effects Due to The Consumption of Drinking Water. *Int. J. Environ. Health Res*., 7:5.

Payment, P., 1998. Distribution System Impacts on Microbial Diseases. *Water Supply* 16:3/4:113.

Rossman, L.A., 2000. Computer Models/EPANET, in L.W. Mays, ed., *Water Distribution Systems Handbook,* McGraw-Hill, New York, N.Y., chapter 12.

Shaw, S.E. and Regli S., 1999. US Regulations on Residual Disinfection. *Jour. AWWA*, 91:1:75.

Schwartz, J., Levin, R., and Goldstein R., 2000. Drinking Water Turbidity and Gastrointestinal Illness in the Elderly of Philadelphia. *British Jour. of Epidemiology and Community Health*, 54:1:45.

Teunis, P.F.M., Medema, G.J., Kruidenier, L., and Havelarr, A.H., 1997. Assessment of the Risk of Infection by Cryptosporidium or Giardia in Drinking Water from a Surface Water Source. *Water Research,* 31:6:1333.

Trussel, R.R., 1999. Safeguarding Distribution System Integrity. *Jour. AWWA*, 91:1:46.

USEPA, 2001. *Drinking Water Infrastructure Needs Survey Second Report to Congress*. Office of Water, EPA 816-R-01-004.

USEPA, 2000a. *Report to Congress EPA Studies on Sensitive Subpopulations and Drinking Water Contaminants*. Office of Water, EPA 815-R-00-015.

USEPA, 2000b. *EPANET 2 Users Manual.* Office of Research and Development. EPA 600-R-00-057.

USEPA, 1999a. *25 Years of the Safe Drinking Water Act: Protecting our Health from Source to Tap*. Office of Water, EPA 810-K-99-004.

USEPA, 1999b. *Guidance for People with Severely Weakened Immune Systems*. Office of Water, EPA 816-F-99-005.

USEPA, 1997. *Drinking Water Infrastructure Needs Survey First Report to Congress*. Office of Water, EPA 812-R-97-001.

Walski, T.M., 2000. Water Quality Aspects of Construction and Operation, in L.W. Mays, ed., *Water Distribution Systems Handbook,* McGraw-Hill, New York, N.Y., chapter 8.

Ysusi, M.A., 2000. System Design: An Overview, in L.W. Mays, ed., *Water Distribution Systems Handbook*, McGraw-Hill, New York, N.Y., chapter 3.

APPENDIX A. LITERATURE REVIEW

The following sections summarize literature of related topics. The topic areas are divided into three sections (1) distribution systems, (2) population demographics and waterborne exposures, and (3) decision support systems in infrastructure. The references are shown in Section 7.0 in the body of the thesis.

A.1 Distribution Systems

Grayman *et al*. (2000) provided the following possible water "deterioration" outcomes observed in a drinking water distribution system: contamination via cross-connections or from leaky pipe joints, corrosion of iron pipes and dissolution of lead and copper from pipe walls, loss of disinfectant residual in storage facilities with long residence times (to include dead-end mains), reactions of disinfectants with organic and inorganic compounds resulting in taste and odor problems, bacterial regrowth and harboring of opportunistic pathogens, increased turbidity caused by particulate resuspension, and formation of disinfectant by-products some of which are suspected carcinogens. Principle causative factors underlie these deterioration outcomes, namely, the quality of the treated water supplied to the distribution system, material and condition of the transmission facilities (pipes, valves, storage tanks, etc.) and the amount of time the water is kept in the system. Independent of water quality issues, distribution system managers also seek system reliability. The design engineer or system operator, concerned with reliability, desires a larger system (in terms of pipe size) with more storage. Larger systems provide a safety factor and allow demands to be met during peak periods or in the case of fire flows. However, system redundancies and safety factors fundamental to reliability are counter to the needs for delivery of water that is relatively unchanged from when it left the treatment plant. An optimal combination of system operation and maintenance involves "trade-offs between cost, hydraulic reliability, and risk of impaired water quality" (Grayman *et al.*, 2000). Trussel (1999) cites Geldreich (1996) and makes the case for maintaining a residual in the distribution system to overcome system contamination that may occur as a result of uncovered reservoirs, cross-connections, construction related contamination, line repairs, inadequate separation between water and sewer lines, poor system flushing, inadequate system pressure, and old corroded water lines. He notes inadequate pressure and back-siphonage are by far the most common sources of contamination worldwide. Kiene *et al.* (1998) provide information on the relative importance of factors influencing chlorine decay in pipe networks. Lee *et al*. (1991) have proposed a method for locating monitoring stations based on the idea of coverage, essentially carefully selecting sample nodes considering flow direction in the pipe network. Kessler et al. (1998) apply an all-shortest-paths algorithm to an example water distribution system in order to identify the nodes to be monitored for accidental contamination events. They assume the probability of contamination at all locations in the system is the same but note this assumption may need further investigation. Antoun *et al*. (1999) explain a specific technique for distribution system maintenance known as unidirectional flushing. They recommend unidirectional flushing as part of a maintenance program that includes many things one of which they call "enhanced monitoring." They say enhanced monitoring may "…include such measures as revising the total coliform sampling plan to better reflect spatial and population distributions and take into account sensitive populations such as hospitals and day care centers." Shaw and Regli (1999) report of waterborne disease outbreaks (epidemics) occurring between 1971 and 1994, 30 percent of the 272 outbreaks reviewed have been attributed to problems with the distribution system. One study has suggested that the incidence of gastroenteritis (endemic) attributable to tap water may be as high as 14-40 percent, part of which may be due to the distribution system (Shaw and Regli, 1999). At a recent roundtable related to the subject of disinfectant residuals, a participant noted current research indicates contamination may occur in the distribution system on a daily or even hourly basis (Journal of the American Water Works Association, January, 1999).

Underlying water quality concerns and reliability is the overall age and condition of distribution systems in the United States. A recent survey of water utilities in the U.S. showed on average 58% of a given distribution system is over 20 years old. Additionally considering only the two largest population categories (systems surveyed serving over 50,000 people) the average age of the oldest section of pipe is 79 years with pipes over 100 years old reported as still in service. The same survey also found the average amount of unaccounted for water to be between 15 and 20 percent of the total volume put into the distribution system (Haas, 1999). The 1997 Infrastructure Needs Survey (USEPA, 1997) estimated the cost to upgrade transmission and distribution systems in the U.S. to be 77.2 billion dollars. This was based on a 20 year planning horizon to meet water quality standards. Storage needs were an additional 12.1 billion dollars. These costs represent 56% and 9% respectively of the total drinking water infrastructure cost estimates given in the 1997 report (including distribution, storage, treatment, source water protection, etc.).

A.2 Population Demographics and Waterborne Exposure

Payment *et al*. (1991) has investigated an increased incidence of gastrointestinal (GI) illness and respiratory symptoms attributable to drinking water. Schwartz *et al.* (2000) suggest correlation between water treatment plant effluent turbidity and hospital admissions of elderly patients with GI illness. The EPA in conjunction with the Center for Disease Control (CDC) has published guidelines to consider for immunocompromised individuals who may be concerned over water born illness, particularly cryptosporidium (USEPA, 1999b). The city of New York has developed a comprehensive risk assessment program as part of its Filtration Avoidance Determination. The program (1) seeks rates of giardiasis and cryptosporidiosis along with demographic and risk factor information on case patients, (2) provides a system to track diarrheal illness and (3) determines contribution of tap water to GI illness. The departments of health and environmental protection jointly administer the program (New York City Department of Environmental Protection, 2000). Teunis *et al*. (1997) provide probability distributions of human infection by cryptosporidium or giardia for drinking water using surface water as a source. They provide dose and probability relations. Crabtree *et al*. 1997 perform a risk assessment due to human exposure to adenovirus in both drinking water and recreational water. They modify the assessment for the elderly versus the general population. Morris and Levin (1994), using available data, estimated the incidence of mild to severe waterborne infectious disease in the United States. Recently, stakeholders concerned with the future of drinking water in the country held a series of meetings known as the SDWA 25 Futures Forum (USEPA, 2000a). One topic of the forum was the issue of vulnerable subpopulations, or those people more susceptible to contaminants that may be present in a public water supply. The Futures Forum cited the SDWA, which identifies infants, children, pregnant woman, elderly, and individuals with a history of serious illness as vulnerable subpopulations. To that list, the forum added all women of childbearing age, seriously immunocompromised persons including transplant recipients, people with AIDS, cancer patients treated with immunosuppressive drugs, the frail elderly, and people with poor nutrition. Future drinking water regulations may address "… increased percentage of an aging population that may require special health considerations…" (USEPA, 1999a).

EPANET, a public domain drinking water distribution system model, has been used to investigate waterborne disease outbreaks or exposure to a given source water. In these cases, EPANET is a tool to estimate the possibility or likely amount of water from a contaminated source that reached the population. In one instance the EPANET results were overlain with a map of CDC confirmed cases of salmonellosis, showing correlation between confirmed cases and contaminated source water (Clark, 2000). EPANET has been used to assess potential exposure at a Superfund site in New Jersey (Maslia *et al*, 2000).   Nuckols *et al.* (1994) have proposed the use of EPANET in conjunction with reproductive outcome data (birth weight, fertility, gestational age, etc.) within a GIS to investigate spatial correlations between the birth outcomes and potential trihalomethane (THM) exposure.   To couple contamination and exposure specifically related to biological agents, a working groups has proposed a conceptual risk assessment framework (ILSI Risk Science Institute, 1996).

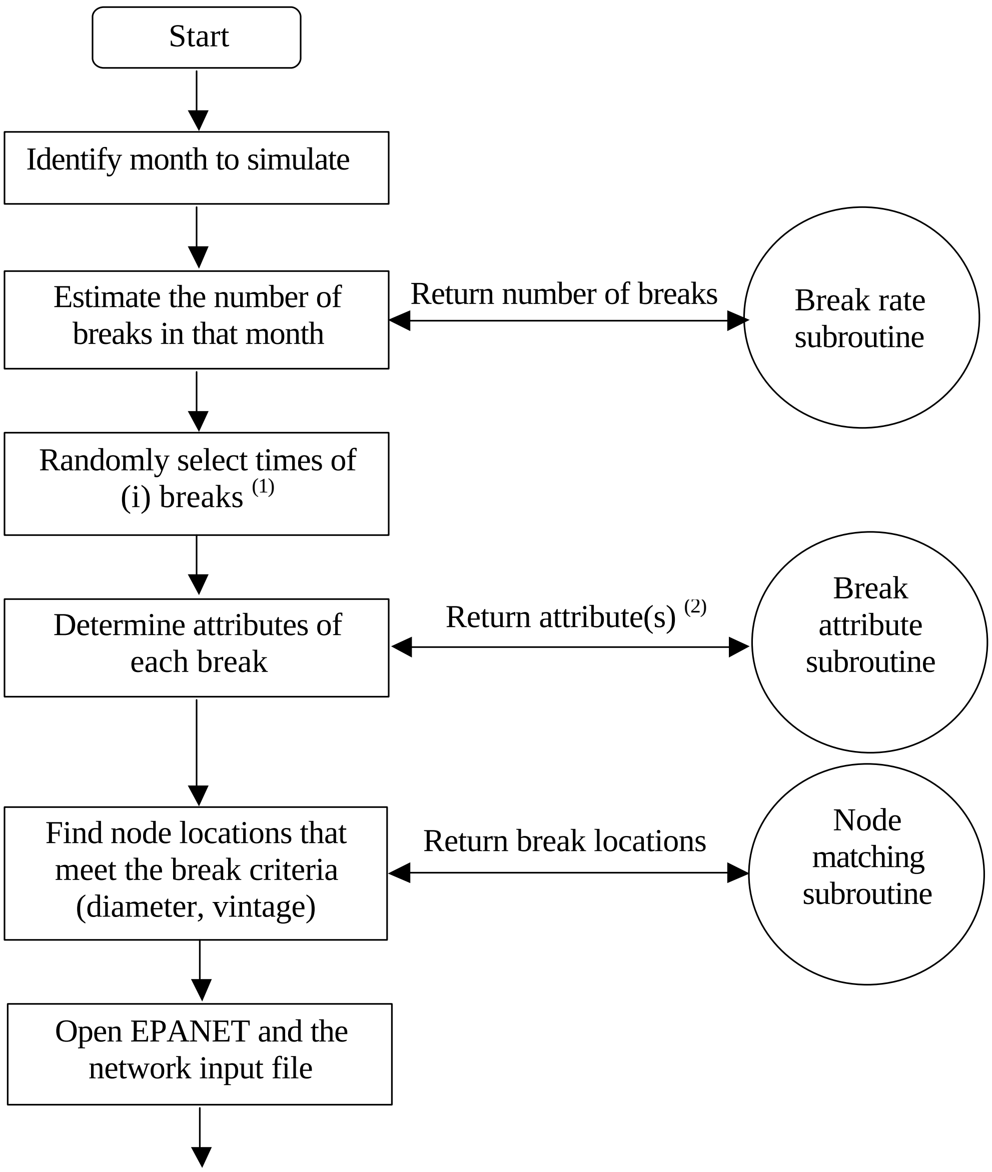
A.3 Decision Support Systems in Infrastructure

Habibian (1992) suggested the use of a computer database for managing information related to maintenance of a water distribution network. The European Union has recently funded development of a comprehensive drinking water distribution system decision support database. The database estimates the probability of main failure given such information as age, pipe material, surrounding soil type, and loadings (deterministic and probabilistic algorithms are applied). Once a failure is predicted the model also considers replacement cost and the risk to surrounding areas (damage assessment) or potential impact to downstream consumers (e.g. hospitals). Using all of this information, priorities are developed for pipeline replacement or rehabilitation (Hadzilacos *et al*., 2000). The sewer industry has used geographic principles and risk as a decision tool to replace sewers. In one case (Griffis and Ivey, 1998) the cost of a sewer failure in petrochemical process sewers was evaluated. The cost was not just the cost to repair but also other costs such as plant downtime in the area affected by a potential failure and environmental liability. Fenner and Sweeting (1999) present a statistical method for rehabilitation of "non-critical" sewers and propose a risk scoring system. They develop a risk plot of consequence of failure versus likelihood of failure. The data points were developed from sewer pipes grouped by spatial grid squares.

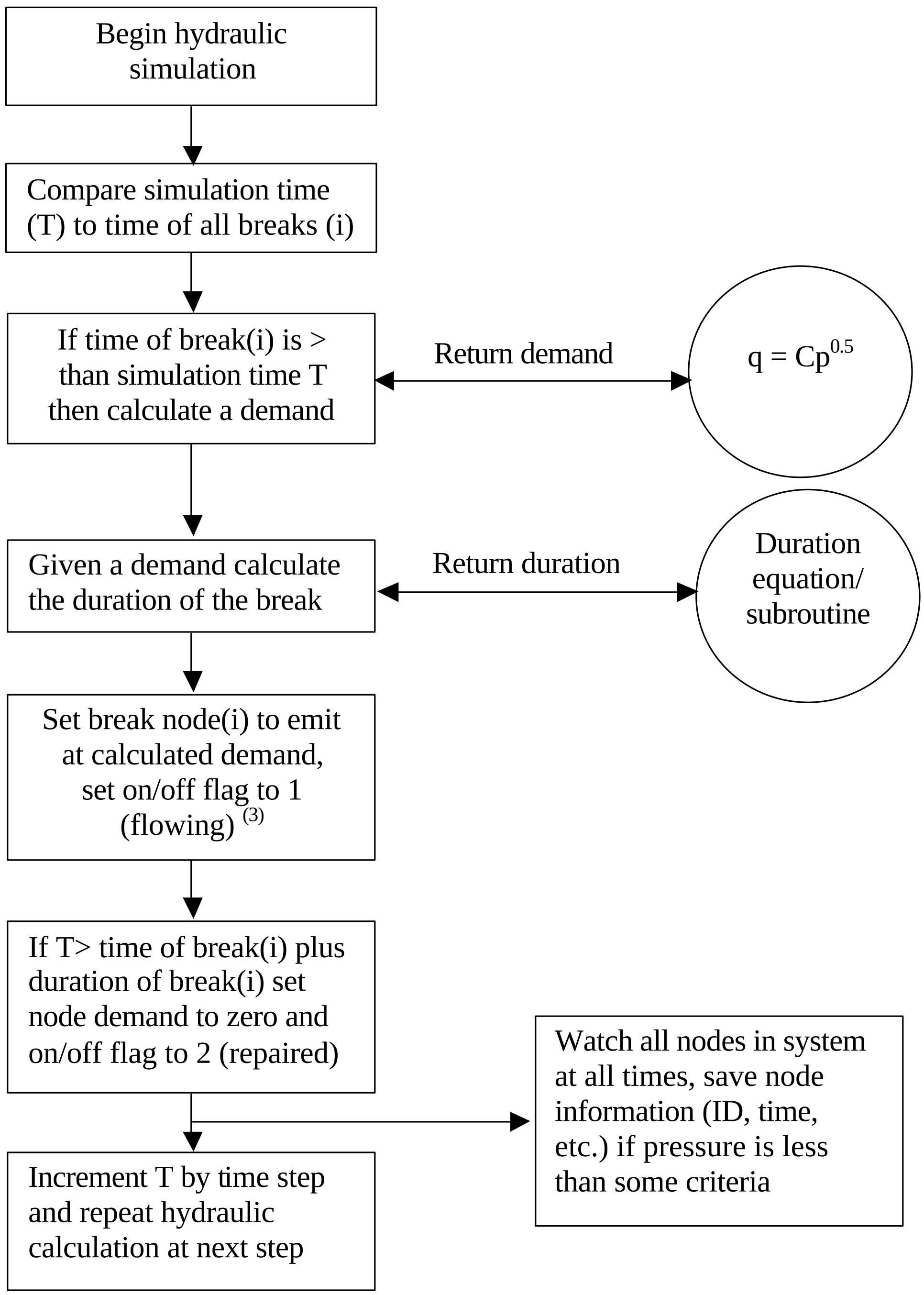
APPENDIX B. ADVERSE PRESSURE GRADIENT ANALYSIS

B.1 Main Break Algorithm Flow Chart

The following flow chart shows how the main break algorithm steps from estimating the number of breaks, through the breaking process, to the pressure analysis.



Continue next page



[1] Letter "i" is an index for tracking the number of breaks (i.e., break1, break2, etc.).

[2] Diameter, vintage, discharge coefficient.

[3] While an emitter is on, the demand is recalculated at each hydraulic time step, the duration is based on the initial demand.

B.2 Visual Basic Computer Code

Visual Basic allows coding to be done in forms or modules; the code associated with the forms or modules is reusable in other portions of the code. The algorithm implemented to simulate random main breaks contains six distinct pieces of code. They area summarized below with the detailed code following:

- EPANET2.BAS: Code provided with the EPANET Toolkit. This code declares all global variables and functions that are called within the EPANET dll computational engine. The EPANET developer wrote this piece of code.
- Subprocedures.bas: Code written as part of algorithm to estimate break attributes, read input file for potential break locations, generate zero mean, unit variance normal distribution, search for break locations that match criteria. Routines in this module are called by other pieces of code prior to running EPANET.
- Outputprocedures.bas: Code written to output a test file for number of breaks and break attributes.
- frmSiteMap: Piece of code to navigate screens at program startup (not integral to analysis).
- frmStatisticaInputs: Code written to test break rates and break attributes. Used in braek algorithm.
- frmCriticalPressure: Code that estimates number of breaks, calls for break attributes, starts EPANET, randomly breaks water mains, and watches pressure surface. Results are written to a file.
- frmTraceAnalysis: Identical to frmCritcalPressure with addition of random mass input when pressure drops below some determined value.

Lines preceded by an apostrophe are comment lines.

```
        'EPANET2.BAS
'
'Declarations of functions in the EPANET PROGRAMMERs TOOLKIT
'(EPANET2.DLL)

'Last updated on 10/25/00

' These are codes used by the DLL functions
Global Const EN_ELEVATION = 0     ' Node parameters
Global Const EN_BASEDEMAND = 1
Global Const EN_PATTERN = 2
Global Const EN_EMITTER = 3
Global Const EN_INITQUAL = 4
Global Const EN_SOURCEQUAL = 5
Global Const EN_SOURCEPAT = 6
Global Const EN_SOURCETYPE = 7
Global Const EN_TANKLEVEL = 8
Global Const EN_DEMAND = 9
Global Const EN_HEAD = 10
Global Const EN_PRESSURE = 11
Global Const EN_QUALITY = 12
Global Const EN_SOURCEMASS = 13

Global Const EN_DIAMETER = 0      ' Link parameters
Global Const EN_LENGTH = 1
Global Const EN_ROUGHNESS = 2
Global Const EN_MINORLOSS = 3
Global Const EN_INITSTATUS = 4
Global Const EN_INITSETTING = 5
Global Const EN_KBULK = 6
Global Const EN_KWALL = 7
Global Const EN_FLOW = 8
Global Const EN_VELOCITY = 9
Global Const EN_HEADLOSS = 10
Global Const EN_STATUS = 11
Global Const EN_SETTING = 12
Global Const EN_ENERGY = 13

Global Const EN_DURATION = 0      ' Time parameters
Global Const EN_HYDSTEP = 1
Global Const EN_QUALSTEP = 2
Global Const EN_PATTERNSTEP = 3
```

```
Global Const EN_PATTERNSTART = 4
Global Const EN_REPORTSTEP = 5
Global Const EN_REPORTSTART = 6
Global Const EN_RULESTEP = 7
Global Const EN_STATISTIC = 8
Global Const EN_PERIODS = 9

Global Const EN_NODECOUNT = 0     'Component counts
Global Const EN_TANKCOUNT = 1
Global Const EN_LINKCOUNT = 2
Global Const EN_PATCOUNT = 3
Global Const EN_CURVECOUNT = 4
Global Const EN_CONTROLCOUNT = 5

Global Const EN_JUNCTION = 0      ' Node types
Global Const EN_RESERVOIR = 1
Global Const EN_TANK = 2

Global Const EN_CVPIPE = 0        ' Link types
Global Const EN_PIPE = 1
Global Const EN_PUMP = 2
Global Const EN_PRV = 3
Global Const EN_PSV = 4
Global Const EN_PBV = 5
Global Const EN_FCV = 6
Global Const EN_TCV = 7
Global Const EN_GPV = 8

Global Const EN_NONE = 0          ' Quality analysis types
Global Const EN_CHEM = 1
Global Const EN_AGE = 2
Global Const EN_TRACE = 3

Global Const EN_CONCEN = 0        ' Source quality types
Global Const EN_MASS = 1
Global Const EN_SETPOINT = 2
Global Const EN_FLOWPACED = 3

Global Const EN_CFS = 0           ' Flow units types
Global Const EN_GPM = 1
Global Const EN_MGD = 2
Global Const EN_IMGD = 3
Global Const EN_AFD = 4
Global Const EN_LPS = 5
Global Const EN_LPM = 6
Global Const EN_MLD = 7
```

```
Global Const EN_CMH = 8
Global Const EN_CMD = 9

Global Const EN_TRIALS = 0       ' Misc. options
Global Const EN_ACCURACY = 1
Global Const EN_TOLERANCE = 2
Global Const EN_EMITEXPON = 3
Global Const EN_DEMANDMULT = 4

Global Const EN_LOWLEVEL = 0     ' Control types
Global Const EN_HILEVEL = 1
Global Const EN_TIMER = 2
Global Const EN_TIMEOFDAY = 3

Global Const EN_AVERAGE = 1      'Time statistic types
Global Const EN_MINIMUM = 2
Global Const EN_MAXIMUM = 3
Global Const EN_RANGE = 4

Global Const EN_NOSAVE = 0       ' Save-results-to-file flag
Global Const EN_SAVE = 1
Global Const EN_INITFLOW = 10    ' Re-initialize flow flag

'These are the external functions that comprise the DLL

 Declare Function ENepanet Lib "epanet2.dll" (ByVal F1 As String, ByVal F2 As String,
ByVal F3 As String, ByVal F4 As Any) As Long
 Declare Function ENopen Lib "epanet2.dll" (ByVal F1 As String, ByVal F2 As String,
ByVal F3 As String) As Long
 Declare Function ENsaveinpfile Lib "epanet2.dll" (ByVal F As String) As Long
 Declare Function ENclose Lib "epanet2.dll" () As Long

 Declare Function ENsolveH Lib "epanet2.dll" () As Long
 Declare Function ENsaveH Lib "epanet2.dll" () As Long
 Declare Function ENopenH Lib "epanet2.dll" () As Long
 Declare Function ENinitH Lib "epanet2.dll" (ByVal SaveFlag As Long) As Long
 Declare Function ENrunH Lib "epanet2.dll" (T As Long) As Long
 Declare Function ENnextH Lib "epanet2.dll" (Tstep As Long) As Long
 Declare Function ENcloseH Lib "epanet2.dll" () As Long
 Declare Function ENsavehydfile Lib "epanet2.dll" (ByVal F As String) As Long
 Declare Function ENusehydfile Lib "epanet2.dll" (ByVal F As String) As Long

 Declare Function ENsolveQ Lib "epanet2.dll" () As Long
 Declare Function ENopenQ Lib "epanet2.dll" () As Long
 Declare Function ENinitQ Lib "epanet2.dll" (ByVal SaveFlag As Long) As Long
 Declare Function ENrunQ Lib "epanet2.dll" (T As Long) As Long
```

```
 Declare Function ENnextQ Lib "epanet2.dll" (Tstep As Long) As Long
 Declare Function ENstepQ Lib "epanet2.dll" (Tleft As Long) As Long
 Declare Function ENcloseQ Lib "epanet2.dll" () As Long

 Declare Function ENwriteline Lib "epanet2.dll" (ByVal S As String) As Long
 Declare Function ENreport Lib "epanet2.dll" () As Long
 Declare Function ENresetreport Lib "epanet2.dll" () As Long
 Declare Function ENsetreport Lib "epanet2.dll" (ByVal S As String) As Long

 Declare Function ENgetcontrol Lib "epanet2.dll" (ByVal Cindex As Long, Ctype As
Long, Lindex As Long, Setting As Single, Nindex As Long, Level As Single) As Long
 Declare Function ENgetcount Lib "epanet2.dll" (ByVal Code As Long, Value As Long)
As Long
 Declare Function ENgetoption Lib "epanet2.dll" (ByVal Code As Long, Value As
Single) As Long
 Declare Function ENgettimeparam Lib "epanet2.dll" (ByVal Code As Long, Value As
Long) As Long
 Declare Function ENgetflowunits Lib "epanet2.dll" (Code As Long) As Long
 Declare Function ENgetpatternindex Lib "epanet2.dll" (ByVal ID As String, Index As
Long) As Long
 Declare Function ENgetpatternid Lib "epanet2.dll" (ByVal Index As Long, ByVal ID As
String) As Long
 Declare Function ENgetpatternlen Lib "epanet2.dll" (ByVal Index As Long, L As Long)
As Long
 Declare Function ENgetpatternvalue Lib "epanet2.dll" (ByVal Index As Long, ByVal
Period As Long, Value As Single) As Long
 Declare Function ENgetqualtype Lib "epanet2.dll" (QualCode As Long, TraceNode As
Long) As Long
 Declare Function ENgeterror Lib "epanet2.dll" (ByVal ErrCode As Long, ByVal
ErrMsg As String, ByVal N As Long)

 Declare Function ENgetnodeindex Lib "epanet2.dll" (ByVal ID As String, Index As
Long) As Long
 Declare Function ENgetnodeid Lib "epanet2.dll" (ByVal Index As Long, ByVal ID As
String) As Long
 Declare Function ENgetnodetype Lib "epanet2.dll" (ByVal Index As Long, Code As
Long) As Long
 Declare Function ENgetnodevalue Lib "epanet2.dll" (ByVal Index As Long, ByVal
Code As Long, Value As Single) As Long

 Declare Function ENgetlinkindex Lib "epanet2.dll" (ByVal ID As String, Index As
Long) As Long
 Declare Function ENgetlinkid Lib "epanet2.dll" (ByVal Index As Long, ByVal ID As
String) As Long
 Declare Function ENgetlinktype Lib "epanet2.dll" (ByVal Index As Long, Code As
Long) As Long
```

```
 Declare Function ENgetlinknodes Lib "epanet2.dll" (ByVal Index As Long, Node1 As
Long, Node2 As Long) As Long
 Declare Function ENgetlinkvalue Lib "epanet2.dll" (ByVal Index As Long, ByVal Code
As Long, Value As Single) As Long

 Declare Function ENgetversion Lib "epanet2.dll" (Value As Long) As Long

 Declare Function ENsetcontrol Lib "epanet2.dll" (ByVal Cindex As Long, ByVal Ctype
As Long, ByVal Lindex As Long, ByVal Setting As Single, ByVal Nindex As Long,
ByVal Level As Single) As Long
 Declare Function ENsetnodevalue Lib "epanet2.dll" (ByVal Index As Long, ByVal
Code As Long, ByVal Value As Single) As Long
 Declare Function ENsetlinkvalue Lib "epanet2.dll" (ByVal Index As Long, ByVal Code
As Long, ByVal Value As Single) As Long
 Declare Function ENsetpattern Lib "epanet2.dll" (ByVal Index As Long, F As Any,
ByVal N As Long) As Long
 Declare Function ENsetpatternvalue Lib "epanet2.dll" (ByVal Index As Long, ByVal
Period As Long, ByVal Value As Single) As Long
 Declare Function ENsettimeparam Lib "epanet2.dll" (ByVal Code As Long, ByVal
Value As Long) As Long
 Declare Function ENsetoption Lib "epanet2.dll" (ByVal Code As Long, ByVal Value As
Single) As Long
 Declare Function ENsetstatusreport Lib "epanet2.dll" (ByVal Code As Long) As Long
 Declare Function ENsetqualtype Lib "epanet2.dll" (ByVal QualCode As Long, ByVal
ChemName As String, ByVal ChemUnits As String, ByVal TraceNode As String) As
Long

'*****************************************************************
'**************    Module SubProcedures    **********************
'Loads all forms, reads in attributes from potential breaknodes,
'contains general normal distribution, has search routine to match
'potential break nodes to desired break attributes, determines
'diameter and vintage characteristics when a break is needed
'***********Developed by TREVOR LINDLEY; Jan-May, 2001***********
'*****************************************************************
Option Explicit   'general declaration and variable declaration
Option Base 0

Public ZeroMean As Single
Public sum As Single
Public BreakRate As Single
Public VintageCode As Integer
Public Diameter As Integer
Public RandomSeed As Single
Public PotBreakNodes(1154, 3) As String
Public MatchedNodes(1154) As String
```

```
Public temp As String
Public mFileSysObj As New FileSystemObject
Public mFile1 As File
Public mFile2 As File
Public mFile3 As File
Public mTxtstream1 As TextStream
Public mTxtstream2 As TextStream
Public mTxtstream3 As TextStream
Public counter As Integer
Public selectindex As Integer

'******Sub procedure to load the form at program startup*********
Public Sub Main()
  'loads project forms, shows starting form
  Load frmHelp
  Load frmSiteMap
  Load frmStatisticalInputs
  frmSiteMap.Show
End Sub

'**Sub procedure to sample historical distribution for break diameter****
Public Sub FindDiameter()
Call Randomize
RandomSeed = Rnd()  'Provides random number between 0 and 1

 'VB Switch function provides If/Then/Else logic to map random
 'numbers to a pipe diameter, percentages from historic repair data
 Diameter = Switch(RandomSeed >= 0.943, "24", _
                RandomSeed >= 0.915, "16", _
                RandomSeed >= 0.851, "12", _
                RandomSeed >= 0.836, "10", _
                RandomSeed >= 0.224, "8", _
                RandomSeed >= 0, "6")
End Sub

'*********Sub procedure to sample distribution for vintage*******
Public Sub FindVintage()
'Once a diameter is established, the following finds a
'vintage given the break diameter

  If Diameter = 6 Then          'given a 6" break, vintage
    Call Randomize             'distribution follows as:
    RandomSeed = Rnd()
    VintageCode = Switch(RandomSeed >= 0.984, "10", _
                RandomSeed >= 0.937, "8", _
                RandomSeed >= 0.921, "7", _
```

```
                  RandomSeed >= 0.571, "6", _
                  RandomSeed >= 0.397, "5", _
                  RandomSeed >= 0.27, "4", _
                  RandomSeed >= 0.048, "3", _
                  RandomSeed >= 0, "1")

  ElseIf Diameter = 8 Then      'given a 8" break, vintage
    Call Randomize              'distribution follows as:
    RandomSeed = Rnd()
    VintageCode = Switch(RandomSeed >= 0.988, "10", _
                  RandomSeed >= 0.971, "9", _
                  RandomSeed >= 0.901, "8", _
                  RandomSeed >= 0.57, "7", _
                  RandomSeed >= 0.076, "6", _
                  RandomSeed >= 0.07, "5", _
                  RandomSeed >= 0.006, "3", _
                  RandomSeed >= 0, "1")

  ElseIf Diameter = 10 Then     'given a 10" break, vintage
     VintageCode = 7            'distribution equals 7:

  ElseIf Diameter = 12 Then     'given a 12" break, vintage
    Call Randomize              'distribution follows as:
    RandomSeed = Rnd()
    VintageCode = Switch(RandomSeed >= 0.722, "10", _
                  RandomSeed >= 0.611, "9", _
                  RandomSeed >= 0.5, "7", _
                  RandomSeed >= 0.444, "6", _
                  RandomSeed >= 0.222, "4", _
                  RandomSeed >= 0, "1")

  ElseIf Diameter = 16 Then     'given a 16" break, vintage
    Call Randomize              'distribution follows as:
    RandomSeed = Rnd()
    VintageCode = Switch(RandomSeed >= 0.625, "7", _
                  RandomSeed >= 0, "3")

  ElseIf Diameter = 24 Then     'given a 24" break, vintage
    Call Randomize              'distribution follows as:
    RandomSeed = Rnd()
     VintageCode = Switch(RandomSeed >= 0.938, "6", _
                  RandomSeed >= 0, "5")
  End If

End Sub
```

```
'****Sub procedure to build a zero mean unit variance
'*****normal distribution****
Public Sub BuildNormal()
Dim i As Integer

sum = 0
For i = 1 To 12
  Call Randomize
  RandomSeed = Rnd()
  sum = sum + RandomSeed
Next i
  ZeroMean = sum - 6
End Sub


'****Sub procedure that loads a set of potential break locations
'****(non-demand nodes along the pipe).  The procedure is called to
'****match these potential break nodes to the sampled diameters and
'****vintage.
Public Sub SearchNodes(Diameter As Integer, VintageCode As Integer)
Dim i As Integer

'Reads three files to get break node information (from input file)
Set mFile1 = mFileSysObj.GetFile("h:\thesis\PBNID.txt")
Set mTxtstream1 = mFile1.OpenAsTextStream(ForReading)
Set mFile2 = mFileSysObj.GetFile("h:\thesis\PBNDiameter.txt")
Set mTxtstream2 = mFile2.OpenAsTextStream(ForReading)
Set mFile3 = mFileSysObj.GetFile("h:\thesis\PBNVintage.txt")
Set mTxtstream3 = mFile3.OpenAsTextStream(ForReading)

'Loads possible break locations (nodes in EPANET input file) into a
'matrix (table)
For i = 1 To 1153
  PotBreakNodes(i, 1) = mTxtstream1.ReadLine   'Vector of IDs
  PotBreakNodes(i, 2) = mTxtstream2.ReadLine   'Vector of Diameters
  PotBreakNodes(i, 3) = mTxtstream3.ReadLine   'Vector of Vintage
Next i

'Matches sampled results (diameter and vintage) to available
'break locations and counts the number of matches
counter = 0
For i = 1 To 1153
  If PotBreakNodes(i, 2) = Diameter _
    And PotBreakNodes(i, 3) = VintageCode Then
    MatchedNodes(i) = PotBreakNodes(i, 1)
      If IsNumeric(MatchedNodes(i)) Then
```

```
        Else
        counter = counter + 1
        'temp = MatchedNodes(i)
        End If
    End If
Next i

'Extracts potential break ID given a single match
If counter = 1 Then
    For i = 1 To 1153
      If PotBreakNodes(i, 2) = Diameter _
      And PotBreakNodes(i, 3) = VintageCode Then
      MatchedNodes(i) = PotBreakNodes(i, 1)
      temp = MatchedNodes(i)
      End If
     Next i
End If

'Extracts potential break ID given more than one match
If counter > 1 Then
    Dim tempArray() As String
    Dim Index As Integer
    Index = 1

    Call Randomize
      selectindex = Int(counter * Rnd()) + 1

    For i = 1 To 1153
      'Index = i

      If PotBreakNodes(i, 2) = Diameter _
      And PotBreakNodes(i, 3) = VintageCode Then
      MatchedNodes(i) = PotBreakNodes(i, 1) 'Assigns Node ID

        'Index = 1
        ReDim tempArray(1154, 2)
        If IsNumeric(MatchedNodes(i)) Then
          Else
          tempArray(i, 1) = Index
          tempArray(i, 2) = MatchedNodes(i)
          Index = Index + 1

            If tempArray(i, 1) = selectindex Then
              temp = tempArray(i, 2)
            End If
        End If
```

```
    End If
  Next i
End If
End Sub


'************************************************************
'**************Submodule Outputprocedures********************
'writes results of testing break rates and attributes to file
'***********Developed by TREVOR LINDLEY; Jan-May, 2001********
'************************************************************

Option Explicit 'General Declaration
Public mFileSysObj As New FileSystemObject
Public mFile As File
Public mTxtstream As TextStream

Public Sub OutputAttributes()
  'Sets up file system for writing results
  Call mFileSysObj.CreateTextFile("d:\attributes.dat")
  Set mFile = mFileSysObj.GetFile("d:\attributes.dat")
  Set mTxtstream = mFile.OpenAsTextStream(ForWriting)
End Sub

Public Sub OutputBreaks()
  'Sets up file system for writing results
  Call mFileSysObj.CreateTextFile("d:\breaks.dat")
  Set mFile = mFileSysObj.GetFile("d:\breaks.dat")
  Set mTxtstream = mFile.OpenAsTextStream(ForWriting)
End Sub

'***************************************************************
'**************    Form Code SiteMap    *************************
'Shows desired forms upon button selection***********************
'***********Developed by TREVOR LINDLEY; Jan-May, 2001***********
'***************************************************************

Option Explicit     'general declaration

Private Sub cmdCriticalPressure_Click()
'show the form with model inputs
  frmCriticalPressure.Show
End Sub

Private Sub cmdHelp_Click()
'show the form with model inputs
```

```
  frmHelp.Show
End Sub

Private Sub cmdTestInputs_Click()
  'show the form with model inputs
  frmStatisticalInputs.Show
End Sub

Private Sub cmdTraceAnalysis_Click()
'show the form with model inputs
  frmTraceAnalysis.Show
End Sub

'*****************************************************************
'**************    Form Statistical Inputs   **********************
'Trial form to test break rates and attributes, estimates number
'of breaks given a month, determines break attributes, writes
'500 breaks to a file.
'***********Developed by TREVOR LINDLEY; Jan-May, 2001***********
'*****************************************************************

Option Explicit     'General Declaration
Option Base 0

Dim Coefficient As Double
Dim NumberofBreaks As Integer
Dim dv(500, 3) As Single
Dim breaks(500, 1) As Integer
Dim StDev As Single
Dim Mean As Single
Dim i As Long
Public mFileSysObj As New FileSystemObject
Public mFile1 As File
Public mFile2 As File
Public mFile3 As File
Public mTxtstream1 As TextStream
Public mTxtstream2 As TextStream
Public mTxtstream3 As TextStream
Dim BreakC() As Single
Dim j As Integer
Dim TimeofBreak As Long
Dim status As Integer
Dim pressure As Single
Dim N As Integer
Dim T As Long
Dim k As Integer
```

```
Dim Tstep As Long
Dim Simulation As Integer
Dim P() As Single
Dim Q() As Single
Dim NodeCount As Long
Dim step As Long
Dim demand As Single
Dim iterations As Integer
Dim simlength As Long
Dim BreakID() As String
Dim BreakDuration() As Single
Dim BreakIndex() As Long
Dim BreakFlag() As Integer
Dim BreakTime() As Long
Dim results() As Single
Dim NodeID As String
Dim BaseDemand As Single
Dim Base As Single
Dim quality As Single
'****Event procedure to sample diameter and vintage 500 times****
'****and then write results to an output file to verify statistics****
Private Sub cmdOutputAttributes_Click()
  Call OutputAttributes
    'Finds 500 diameters and vintages
    For i = 1 To 500
      Call cmdBreakAttributes_Click
      'saves results in array dv
      dv(i, 1) = Diameter
      dv(i, 2) = VintageCode
      dv(i, 3) = Coefficient
      'writes results to file
      Call mTxtstream.WriteLine(dv(i, 1) & " " & dv(i, 2) & " " & dv(i, 3))
    Next i
    mTxtstream.Close
  txtOutputAttributes.Text = mFile
End Sub

'****Event procedure to get a break diameter, vintage, and ****
'****emitter coefficient, will be called by the final model simulation****
Private Sub cmdBreakAttributes_Click()
Dim fireflow As Single
Dim CV As Integer
Dim lambda As Single
Dim sigmasquared As Single
Dim fireC As Single
```

```
CV = 1   'Set coefficent of variation to 1

        'Calls subroutines held in SubProcedures.bas
        Call FindDiameter
        Call FindVintage

        'Sets St. deviation and mean for determination of the
        'emitter coefficient
        fireflow = Switch(Diameter = 6, 1000#, _
                    Diameter = 8, 1500#, _
                    Diameter = 10, 2000#, _
                    Diameter = 12, 2500#, _
                    Diameter = 16, 3000#, _
                    Diameter = 24, 3500#)

        'converts normal moments to log-based
        fireC = fireflow / 20 ^ 0.5

        Mean = fireC / (1 + 2 * CV)

        StDev = CV * Mean

        sigmasquared = Log(1 + CV ^ 2)

        lambda = Log(Mean) - Log(1 + CV ^ 2) ^ 0.5


   Call BuildNormal     'Calls zero mean unit variance distribution
   Coefficient = ZeroMean * sigmasquared + lambda
   Coefficient = Exp(Coefficient)
   txtCoeff.Text = Coefficient

   txtDiameter.Text = Diameter 'writes single output to field on form
   txtVintage.Text = VintageCode
   txtGamma.Text = 0.5
  End Sub

'****Event procedure to sample break rate distribution; ****
'****will be called by the final model simulation****
Public Sub cmdBreakHistory_Click()
StDev = Switch(cboMonth.Text = "January", 0.028, _
                    cboMonth.Text = "February", 0.009, _
                    cboMonth.Text = "March", 0.006, _
                    cboMonth.Text = "April", 0.003, _
                    cboMonth.Text = "May", 0.008, _
                    cboMonth.Text = "June", 0.01, _
```

```
                    cboMonth.Text = "July", 0.009, _
                    cboMonth.Text = "August", 0.0156, _
                    cboMonth.Text = "September", 0.012, _
                    cboMonth.Text = "October", 0.011, _
                    cboMonth.Text = "November", 0.011, _
                    cboMonth.Text = "December", 0.016)

Mean = Switch(cboMonth.Text = "January", 0.036, _
                    cboMonth.Text = "February", 0.013, _
                    cboMonth.Text = "March", 0.011, _
                    cboMonth.Text = "April", 0.003, _
                    cboMonth.Text = "May", 0.011, _
                    cboMonth.Text = "June", 0.015, _
                    cboMonth.Text = "July", 0.018, _
                    cboMonth.Text = "August", 0.0135, _
                    cboMonth.Text = "September", 0.017, _
                    cboMonth.Text = "October", 0.017, _
                    cboMonth.Text = "November", 0.017, _
                    cboMonth.Text = "December", 0.021)

  Call BuildNormal    'Calls zero mean unit variance distribution
  BreakRate = ZeroMean * StDev + Mean
  NumberofBreaks = BreakRate * 150
  If NumberofBreaks < 0 Then
   NumberofBreaks = 0
  End If
  txtBreaks.Text = NumberofBreaks
End Sub

'****Event procedure to sample break rate distribution 500 times****
'****and then write results to an output file to verify statistics****
Private Sub cmdOutputBreaks_Click()
Call OutputBreaks

'User chooses a month and then mean and deviation is then chosen
'(from historical break data)
StDev = Switch(cboMonth.Text = "January", 0.028, _
                    cboMonth.Text = "February", 0.009, _
                    cboMonth.Text = "March", 0.006, _
                    cboMonth.Text = "April", 0.003, _
                    cboMonth.Text = "May", 0.008, _
                    cboMonth.Text = "June", 0.01, _
                    cboMonth.Text = "July", 0.009, _
                    cboMonth.Text = "August", 0.0156, _
                    cboMonth.Text = "September", 0.012, _
                    cboMonth.Text = "October", 0.011, _
```

```
                cboMonth.Text = "November", 0.011, _
                cboMonth.Text = "December", 0.016)

Mean = Switch(cboMonth.Text = "January", 0.036, _
                cboMonth.Text = "February", 0.013, _
                cboMonth.Text = "March", 0.011, _
                cboMonth.Text = "April", 0.003, _
                cboMonth.Text = "May", 0.011, _
                cboMonth.Text = "June", 0.015, _
                cboMonth.Text = "July", 0.018, _
                cboMonth.Text = "August", 0.0135, _
                cboMonth.Text = "September", 0.017, _
                cboMonth.Text = "October", 0.017, _
                cboMonth.Text = "November", 0.017, _
                cboMonth.Text = "December", 0.021)

NumberofBreaks = -1
For i = 1 To 500

  'Calls zero mean distribution to get a single random number
  'in this case result will have negative numbers which can be
  'considered zero breaks
  Call BuildNormal
    BreakRate = ZeroMean * StDev + Mean
    NumberofBreaks = BreakRate * 150
    breaks(i, 1) = NumberofBreaks

    'writes results to file
    Call mTxtstream.WriteLine(breaks(i, 1))
Next i
     txtOutputBreaks.Text = mFile
End Sub


'****************************************************************
'**************    Form TraceAnalysis   **********************
'Estimates number of breaks, calls for break attributes, calls
'for potential break locations matching attributes, randomly selects
'times for breaks, starts EPANET, turns on breaks at appropriate
'times, watches pressure, drops tracer in line if pressure is below
'some criteria.
'***********Developed by TREVOR LINDLEY; Jan-May, 2001***********
'****************************************************************


Option Explicit    'General Declaration
```

```
Option Base 0

Dim Coefficient As Double
Dim NumberofBreaks As Integer
Dim dv(500, 2) As Integer
Dim breaks(500, 1) As Integer
Dim StDev As Single
Dim Mean As Single
Dim i As Long
Public mFileSysObj As New FileSystemObject
Public mFile1 As File
Public mFile2 As File
Public mFile3 As File
Public mTxtstream1 As TextStream
Public mTxtstream2 As TextStream
Public mTxtstream3 As TextStream
Dim BreakC() As Single
Dim j As Integer
Dim TimeofBreak As Long
Dim status As Integer
Dim pressure As Single
Dim N As Integer
Dim T As Long
Dim k As Integer
Dim Tstep As Long
Dim Simulation As Integer
Dim P() As Single
Dim Q() As Single
Dim NodeCount As Long
Dim step As Long
Dim demand As Single
Dim iterations As Integer
Dim simlength As Long
Dim BreakID() As String
Dim BreakDuration() As Single
Dim BreakIndex() As Long
Dim BreakFlag() As Integer
Dim BreakTime() As Long
Dim results() As Single
Dim NodeID As String
Dim BaseDemand As Single
Dim Base As Single
Dim quality As Single
Dim test As Integer
Dim duration As Single
Dim demandscenario As String
```

```
'****Event procedure to get a break diameter, vintage, and ****
'****emitter coefficient, will be called by the final model simulation****
Private Sub cmdBreakAttributes_Click()
Dim fireflow As Single
Dim CV As Integer
Dim lambda As Single
Dim sigmasquared As Single
Dim fireC As Single

CV = 1
        'Calls subroutines held in SubProcedures.bas
        Call FindDiameter
        Call FindVintage

        'Sets St. deviation and mean for determination of the
        'emitter coefficient
        fireflow = Switch(Diameter = 6, 500#, _
                    Diameter = 8, 1000#, _
                    Diameter = 10, 1500#, _
                    Diameter = 12, 2000#, _
                    Diameter = 16, 2500#, _
                    Diameter = 24, 3000#)

        fireC = fireflow / 20 ^ 0.5

        Mean = fireC / (1 + 2 * CV)

        StDev = CV * Mean

        sigmasquared = Log(1 + CV ^ 2)

        lambda = Log(Mean) - Log(1 + CV ^ 2) ^ 0.5

   Call BuildNormal     'Calls zero mean unit variance distribution
   Coefficient = ZeroMean * sigmasquared + lambda
   Coefficient = Exp(Coefficient)

End Sub

'****Event procedure to sample break rate distribution; ****
'****will be called by the final model simulation****
Public Sub cmdBreakHistory_Click()
StDev = Switch(cboMonth.Text = "January", 0.028, _
                cboMonth.Text = "February", 0.009, _
                cboMonth.Text = "March", 0.006, _
```

```
                cboMonth.Text = "April", 0.003, _
                cboMonth.Text = "May", 0.008, _
                cboMonth.Text = "June", 0.01, _
                cboMonth.Text = "July", 0.009, _
                cboMonth.Text = "August", 0.0156, _
                cboMonth.Text = "September", 0.012, _
                cboMonth.Text = "October", 0.011, _
                cboMonth.Text = "November", 0.011, _
                cboMonth.Text = "December", 0.016)

Mean = Switch(cboMonth.Text = "January", 0.036, _
                cboMonth.Text = "February", 0.013, _
                cboMonth.Text = "March", 0.011, _
                cboMonth.Text = "April", 0.003, _
                cboMonth.Text = "May", 0.011, _
                cboMonth.Text = "June", 0.015, _
                cboMonth.Text = "July", 0.018, _
                cboMonth.Text = "August", 0.0135, _
                cboMonth.Text = "September", 0.017, _
                cboMonth.Text = "October", 0.017, _
                cboMonth.Text = "November", 0.017, _
                cboMonth.Text = "December", 0.021)


  Call BuildNormal     'Calls zero mean unit variance distribution
  BreakRate = ZeroMean * StDev + Mean
  NumberofBreaks = BreakRate * 150
  If NumberofBreaks < 0 Then
   NumberofBreaks = 0
  End If
End Sub

'****Event procedure to perform a EPANET stochastic simulation of
'****breaks.  The event will call subs to sample diameter, vintage,
'****coefficient, and duration.
Private Sub cmdRunTraceAnalysis_Click()

'Starts a stochastic simulation
NodeID = "               "
N = cboSimulations.Text
For Simulation = 1 To N

'Calls a procedure to estimate the number of breaks in a
'user specified month
  Call cmdBreakHistory_Click
    ReDim BreakID(NumberofBreaks)
```

```
    ReDim BreakC(NumberofBreaks)
    ReDim BreakDuration(NumberofBreaks)
    ReDim BreakTime(NumberofBreaks)
    ReDim BreakFlag(NumberofBreaks)
    ReDim BreakIndex(NumberofBreaks)

'Given the number of breaks, the time of the breaks within a month are
'randomly selected and the break attributes (diameter, vintage, emitter C)
'are assigned to each break
  For i = 1 To NumberofBreaks
    Call Randomize
    TimeofBreak = 172800 + (2592000 * Rnd()) 'Time of break estimate (in seconds)
    Call cmdBreakAttributes_Click 'Returns Diameter, Vintagecode, Coeff
    Call SearchNodes(Diameter, VintageCode)
                'Matches available break locations to requested
                'break attributes

    BreakID(i) = temp             'Vector of node IDs
    BreakID(i) = Trim$(BreakID(i)) 'Trim the Node IDs (remove trailing spaces)
    BreakTime(i) = TimeofBreak    'Vector of Break times
    BreakC(i) = Coefficient       'Vector of discharge coefficients
    BreakDuration(i) = 0          'Vector of durations (set here to 0)
    BreakFlag(i) = 0              'Vector of on/off flags (set here to 0)
    Open "h:\thesis\breakattributes15.csv" For Append As #1
      Print #1, _
      Simulation, ","; BreakID(i), ","; BreakTime(i), ","; _
      BreakC(i), ","; Diameter, ","; VintageCode
      Close #1
  Next i

 'Opens EPANET
  status = ENopen("h:\thesis\version2\maxmonth.inp",
"h:\thesis\version2\maxmonth.rpt", "")
  ENgetcount EN_NODECOUNT, NodeCount
  NodeCount = Int(NodeCount)

 For i = 1 To NodeCount
    ENgetnodevalue i, EN_BASEDEMAND, BaseDemand
    If BaseDemand > 0 Then
    ENsetnodevalue i, EN_SOURCETYPE, EN_MASS
    End If
  Next i

  'Assigns nodal index to vector BreakIndex based on BreakID
  For i = 1 To (NumberofBreaks)
    status = ENgetnodeindex(BreakID(i), BreakIndex(i))
```

```
Next i

'Opens and initializes hydraulic solver
status = ENopenH
status = ENinitH(1)

Do

'Starts hydraulic simulation
  status = ENrunH(T)

    'Checks simulation time (T) against a prior determined break time
    For i = 1 To NumberofBreaks
      If T >= BreakTime(i) And BreakFlag(i) = 0 Then

        'If time criteria met, gets pressure at aprior break node
        'with a lower constraint
        ENgetnodevalue BreakIndex(i), EN_PRESSURE, pressure

        'Calculates demand at break node and duration with an
        'upper and lower duration constraint, flag is set to
        '1 indicating break has occurred; demand calculation here is
        'once to get a duration
        demand = (BreakC(i) / 694.44) * pressure ^ 0.5
        BreakDuration(i) = 3600 * (50 * Exp(-0.02 * demand * 694.44))
          If BreakDuration(i) < 7200 Then
            Call Randomize
            duration = Rnd()
            If duration < (1 / 3) Then
               BreakDuration(i) = 10800 '(3 hr interval to flow for 2)
            ElseIf duration > (2 / 3) Then
               BreakDuration(i) = 18000 '(5 hr interval to flow for 4)
            Else
               BreakDuration(i) = 14400 '(4 hr interval to flow for 3)
            End If
          End If
        BreakFlag(i) = 1

        'Demand set back to zero at end of break duration
      ElseIf T >= (BreakTime(i) + BreakDuration(i)) _
                    And BreakFlag(i) = 1 Then
        demand = 0
        ENsetnodevalue BreakIndex(i), EN_BASEDEMAND, demand
        BreakFlag(i) = 2

        'During break demand is calculated to establish all
```

```
        'demands during the break
      ElseIf BreakFlag(i) = 1 Then
        ENgetnodevalue BreakIndex(i), EN_PRESSURE, pressure
            If pressure < 0 Then
               pressure = 5
            End If

            If pressure > 175 Then
               pressure = 175
            End If

        demand = BreakC(i) / 694.44 * pressure ^ 0.5
        ENsetnodevalue BreakIndex(i), EN_BASEDEMAND, demand

        'Demand results written to a file
        Open "h:\thesis\flowattributes15.csv" For Append As #2
               Print #2, _
               Simulation, ","; T, ","; BreakID(i), ","; _
               demand, ","; BreakDuration(i)
               Close #2

      End If   'Ends first if statement that began time check
       'Ends for loop that began check of break times
     Next i

        'During break times only (flag=1); all nodes with demand>0
        'are watched and those with low pressure are saved to a file
        For k = 1 To NodeCount
          ENgetnodevalue k, EN_BASEDEMAND, BaseDemand
             If BaseDemand > 0 Then
               NodeID = "              "
               ENgetnodeid k, NodeID
               NodeID = Trim$(NodeID)
               ENgetnodevalue k, EN_PRESSURE, pressure
                 If pressure < 15 Then
                   Open "h:\thesis\hresults15.csv" For Append As #3
                   Print #3, _
                   Simulation, ","; T, ","; NodeID, ","; _
                   pressure
                   Close #3
                 End If
             End If
        Next k

 'Begin next time T
 status = ENnextH(Tstep)
```

```
  Loop Until (Tstep = 0)

  ENcloseH

'************** Begin WQ simulation (intrusion trace)*********************

T = 0
Tstep = 0
i = 0
Dim unitmass As Single
Dim mass As Single
status = ENopenQ
status = ENinitQ(0)
ENgetcount EN_NODECOUNT, NodeCount
NodeCount = Int(NodeCount)

Do
 status = ENrunQ(T)

 'Following searches nodes for low pressure and drops tracer in line
 'or removes intrusion if pressure is high enough
  For i = 1 To NodeCount
    ENgetnodevalue i, EN_BASEDEMAND, BaseDemand

      'We only want to observe system nodes not zero demand
      'nodes around treatment/storage works
      If BaseDemand > 0 Then
        ENgetnodevalue i, EN_PRESSURE, pressure

          'Drop tracer in line
          If pressure <= 1 Then
            Call Randomize
            unitmass = Rnd()
            mass = unitmass * 1000
            ENsetnodevalue i, EN_SOURCEQUAL, mass
            NodeID = "              "
            ENgetnodeid i, NodeID
            NodeID = Trim$(NodeID)
            Open "h:\thesis\massin15.csv" For Append As #4
            Print #4, _
            Simulation, ","; T, ","; NodeID, ","; mass
            Close #4

          'Turn tracer off
          ElseIf pressure > 1 Then
            ENsetnodevalue i, EN_SOURCEQUAL, 0
```

```
          End If

        'Watch quality at all demand nodes, write results to a file
        ENgetnodevalue i, EN_QUALITY, quality
          If quality > 0.05 Then
            NodeID = "               "
            ENgetnodeid i, NodeID
            NodeID = Trim$(NodeID)
            Open "h:\thesis\qresults15.csv" For Append As #5
            Print #5, _
            Simulation, ","; T, ","; NodeID, ","; quality, ","; pressure
            Close #5
          End If
      End If
  Next i

 'Begin next time T
 status = ENnextQ(Tstep)
 Loop Until (Tstep = 0)
 status = ENcloseQ
 ENclose

Print Simulation

'Begin next simulation (new month)
Next Simulation
Print "Finish"

End Sub
```

The form code that runs only a hydraulic simulation (breaks and pressure observations) is identical to the code for TraceAnalysis, down to where the water quality simulation starts. This code is not repeated here.

B.3 Basis for Stochastic Main Break Inputs

The following provides rationale for random variables used to support the stochastic main break algorithm.

B.3.1 Number of Breaks in a Month

The number of breaks estimated within the algorithm has a basis in the utility provided historical repair data. It is a maintenance event or repair database; there is no clear indication whether these repair events were in response to small leaks or large breaks. The definition of leak and break varies by utility. Generally a leak is something they schedule into their maintenance routines. A break is a high priority leak that must be addressed immediately. Discussions with the utility suggest the repair data provided can be considered as repairs associated with "main breaks" without much error. They are all repair events that required subsurface excavation. The utility provided data for 748 events over a 15-year period (1985-1999). From these 748 events, 520 events match water mains still in use in the system. Of this 520, 281 events match mains represented in the model. All repair data are used to support the structural assessment as part of the pathway analysis but only those repair data matching modeled mains supported the main break algorithm within the hydraulic model.

To simulate the number of breaks in a given month, the monthly repair data (related to modeled lines) are normalized against the length of line in existence each year of repair history (i.e., 1985, 1986, etc). This results in a break rate measured as breaks/mile for the given year. Figure B-1 shows the yearly break rates for the repairs associated with modeled lines (n=281 repair events).

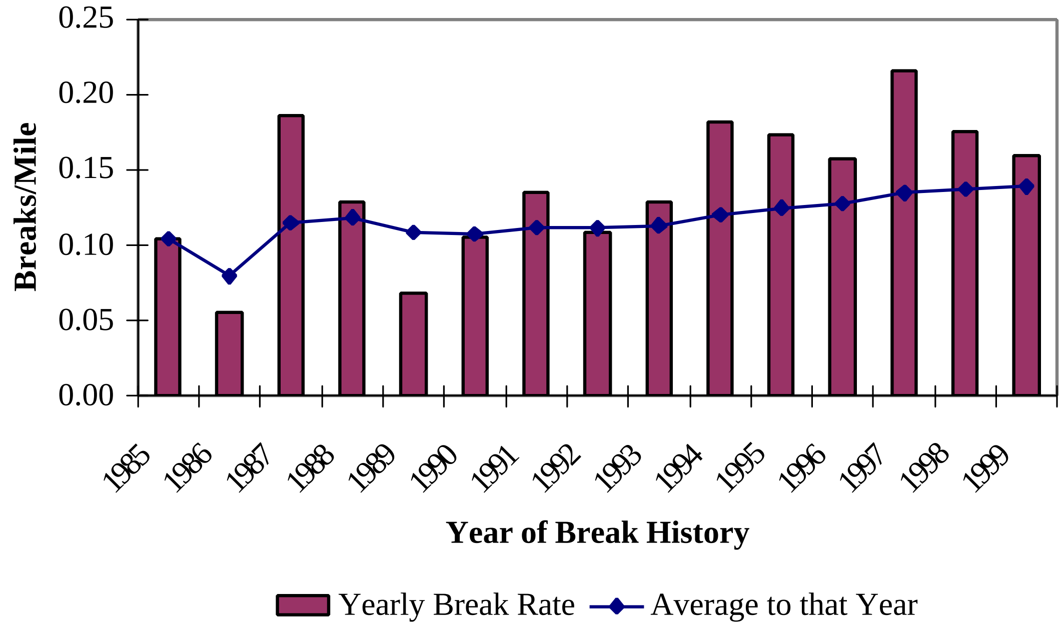


Figure B-1. Yearly Break Rates Associated with Water Mains Represented in the Model

The break rates shown in Figure B-1 are lower than the overall system break rate that includes all the data for all utility mains. Figure B-2 shows the averages plus or minus one deviation by month (n=15 for each month, i.e., 15 years of repair data). The averages and deviations for each month (figure B-2) are the two moments used to build the random variable for the number of breaks. The actual number of breaks was estimated by sampling an assumed normal distribution of break rates, (to return a random rate); the rate was then multiplied by the current length of mains in the model (150 miles). Negative breaks are assumed to be zero and thus the resulting distribution is truncated normal.

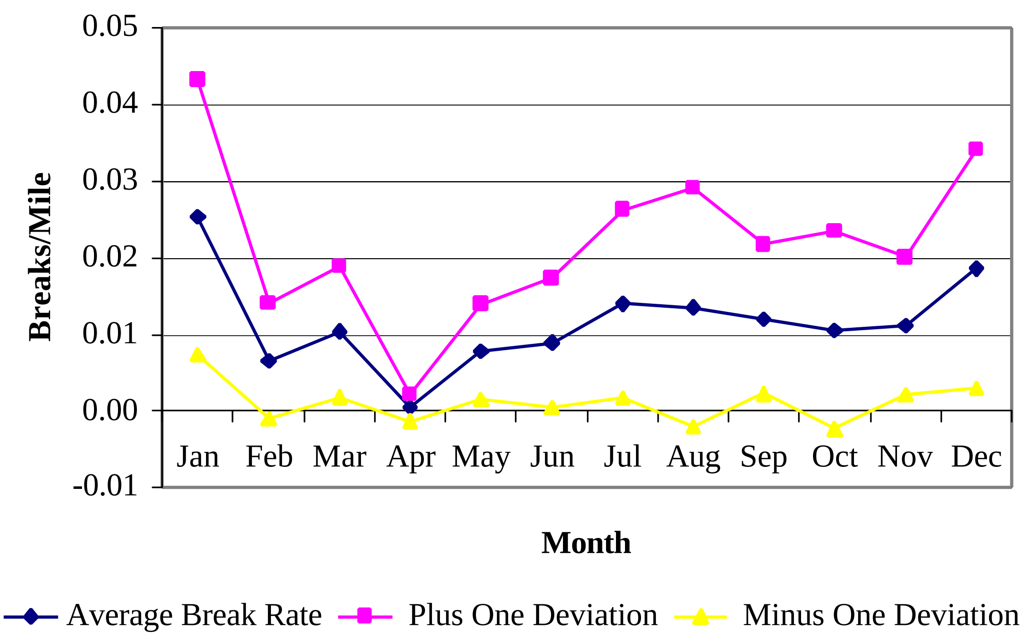


Figure B-2. Average Break Rates for Modeled Mains by Month of Repair

B.3.2 Basis for Break Attributes

The basis for the nature of the breaks (which pipes break) is also based on the maintenance data. The geocoded maintenance data (point events) were matched to nearby pipes or other spatial data within the GIS. From this matching exercise break patterns were investigated. Breakage related to diameter, material, vintage, slope (where break occurred), and soil type factors were investigated. These factors are typically cited as influential in potable water main breaks. The results (See Appendix 3) show diameter, vintage, and pipe material type influence break characteristics in the study area. However, pipe material and vintage tend to give similar information. This is primarily due to the existence of two predominant pipe types, gray cast iron and ductile iron. Most of the pipe placed prior to 1975 is cast iron and most of the pipe since 1975 is ductile iron. All vintages with high break rates are in the cast iron group. Noting this, diameter and then vintage were used as selection criteria for which pipes to break. Table B-1 shows the percentages of breaks by diameter and vintage for repairs on modeled mains. The proportions shown in Table B-1 are coded into the algorithm, once a break is specified the selection of diameter and vintage is weighed according to the table.

B.3.3 Basis for Selecting Actual Break Locations

Once a pipe diameter and vintage were selected (for example 12-inch 1965 water main) then these results were matched to those nodes in EPANET where a break can be modeled (only nodes can emit water in EPANET). This was done as follows. Multiple nodes were added to a pipe length at approximately 500 foot intervals. These new nodes become placeholders for a potential break location. The new nodes have associated with them a diameter and vintage for that piece of pipe they represent. Given a break diameter and vintage all matching nodes were found; if more than one match exists then the break location was selected completely at random from the set of matching nodes.

B.3.4 Basis for Calculating Demands (Flowrate) at a Break

In reality a broken mains flowrate is dependent on pressure. A logical way to handle this in EPANET is with the emitter property of a node (see Section 5.1.1). Assuming the emitter equation can be used to model a pipe break then there is a range of C and pressure exponent values that fit any main break if flowrate and pressure at the break location are known. The pressure exponent and discharge coefficient likely vary based on size of pipe, type of break (i.e., longitudinal, circumferential, etc.) and burial conditions. For the purpose of the main break algorithm, the pressure exponent was fixed at

Table B-1. Historical Distribution of Repairs By Diameter and Vintage for Modeled Water Mains

Diameter (inches)	% of Breaks This Diameter	Vintage [1]	% of Breaks This Vintage Given Diameter
6	22.4%	Pre-1910	4.8%
		1920s	22.2%
		1930s	12.7%
		1940s	17.4%
		1950s	35.0%
		1960s	1.6%
		1970s	4.7%
		1990s	1.6%
8	61.2%	Pre-1910	0.6%
		1920s	6.4%
		1940s	6.0%
		1950s	49.4%
		1960s	33.1%
		1970s	7.0%
		1980s	1.7%
		1990s	1.2%
10	1.5%	1960s	100%
12	6.4%	Pre-1910	22.2%
		1930s	22.2%
		1950s	5.6%
		1960s	11.1%
		1980s	11.1%
		1990s	27.8%
16	2.8%	1920s	62.5%
		1960s	37.5%
24	5.7%	1940s	93.8%
		1950s	6.2%

[1] Vintages not shown for a given diameter had no reported repairs.

0.5 and C was assumed to be a random variable with first and second moments that are a function of pipe diameter only. The moments were determined considering the following as a guide.

Referencing typical fire flow requirements (noting distribution systems are designed to meet these requirements) provides the following.

Table B-2. Typical Distribution System Design Fire Flow Requirements

Service area (1)	**Range required (gpm)**	**Typical pipe diameters serving these areas (2)**
Single family residential	500-2000	6, 8
Multi family residential	1500-3000	6, 8, 12
Commercial	2500-5000	12-16
Industrial	3500-10,000	12-24
Central business district	2500-15,000	>12

[1] Table after Ysusi (2000).

[2] The first two columns are in the reference the third column was added based on engineering judgment. These rates must typically be delivered with pressures not less than 20 psi.

For the main break algorithm it was assumed the values in Table B-2 are comparable to large but not improbable breaks for the respective diameters; i.e., a six inch line is designed to flow between 500 and 2000 gpm at 20 psi, therefore, breaks of this magnitude are possible. Assuming a fire flow rate is a large break, the average or median value is for a typical break is somewhat less. At the same pressure changes in the discharge coefficient will increase or decrease the flow rate. Therefore, at 20 psi it was assumed a C value yielding flowrates in accordance with Table B-2 is two standard deviations higher than a C value yielding "typical" flow rates associated with a break. For example, based on the emitter equation, a fire flow from a hydrant attached to a six-inch pipe is modeled as:

$$500 \text{ gpm} = (C)\,(20 \text{ psi})^{0.5}$$

resulting in a C value of 111.8 gpm/psi$^{0.5}$. Now to get the distribution of C it was assumed this value is equal to the mean plus two standard deviations. Furthermore, assuming a fixed coefficient of variation first and second moments are specified for each diameter. To avoid the possibility of a negative discharge coefficient associated with one side of the normal distribution, these normal moments were converted to log-normal moments. See figures 5-2 and 5-3 in the text for typical coefficients and expected flow rates.

B.3.5 Basis for Calculating Duration of Break

Intuition and discussion with utility personnel suggest the duration of the break is correlated to the size of the break. Extreme breaks require immediate attention (within a few hours) whereas barely perceptible leaks may be scheduled for repair as crews become available. There likely exists more than one function to describe the correlation between main break volume and response time. Leaks are likely responded to at some lower rate based on scheduling and availability; breaks are responded to based on proximity of crews to the break site or difficulty of repair. A single function with some response criteria was used to model the duration of the breaks.

The utility is required to restore service as soon as physically possible and generally shuts off significant main breaks in two to four hours. For the purpose of this work, a flowrate of approximately 125 gpm or less was assumed to have a response time (leak duration) described by:

$$\text{Duration} = 50e^{(-0.02(\text{Demand}))}$$

This suggests long leak times are possible for low flowrate breaks/leaks (maximum leak time is 50 hours). Leaks greater 125 gpm were responded to at a random rate of either two, three, or four hours.

APPENDIX C. RESULTS FROM PATHWAY ANALYSIS

The pathway analysis focuses on two main mechanisms: (1) direct service connections (2) structurally deficient locations determined from utility repair data. Figures and Tables related to these two pathways follow. The figures mostly summarize infrastructure data supporting assessment of locations with structural concerns.

C.1 General Infrastructure Information

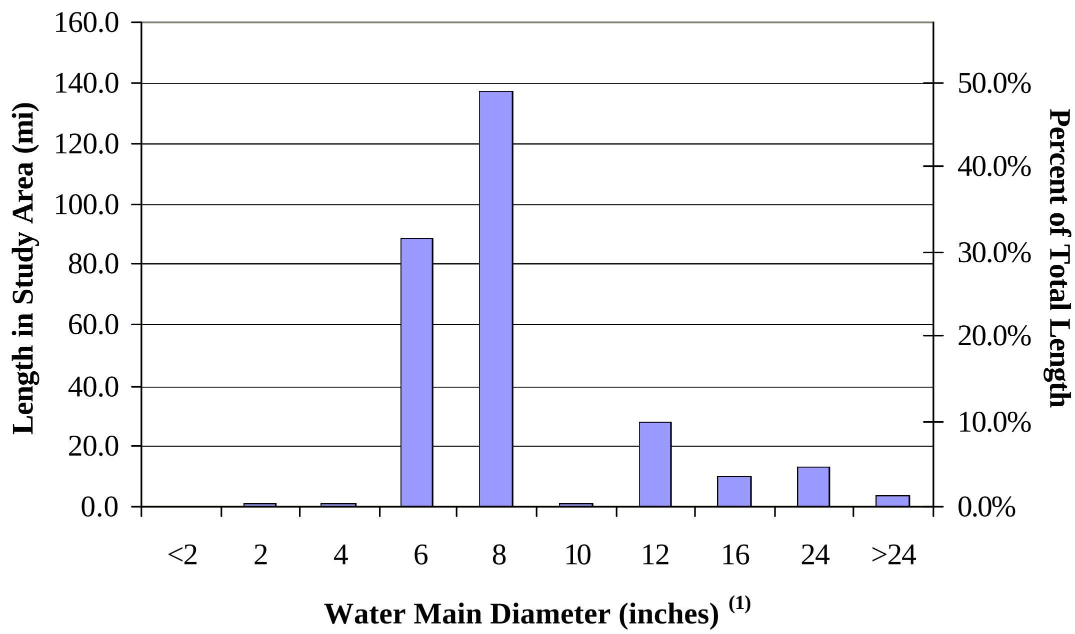


Figure C-1. Length Summary By Diameter

[1]For all water mains, all diameters, material, vintage, etc., extracted from regional GIS data and utility provided attribute data.

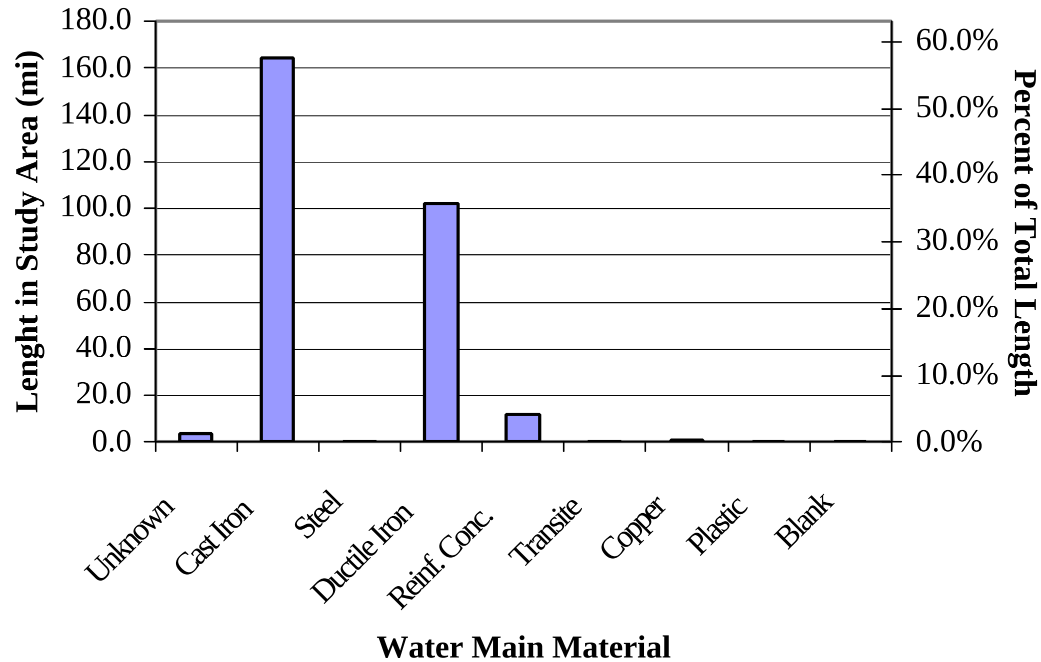


Figure C-2. Length Summary By Material

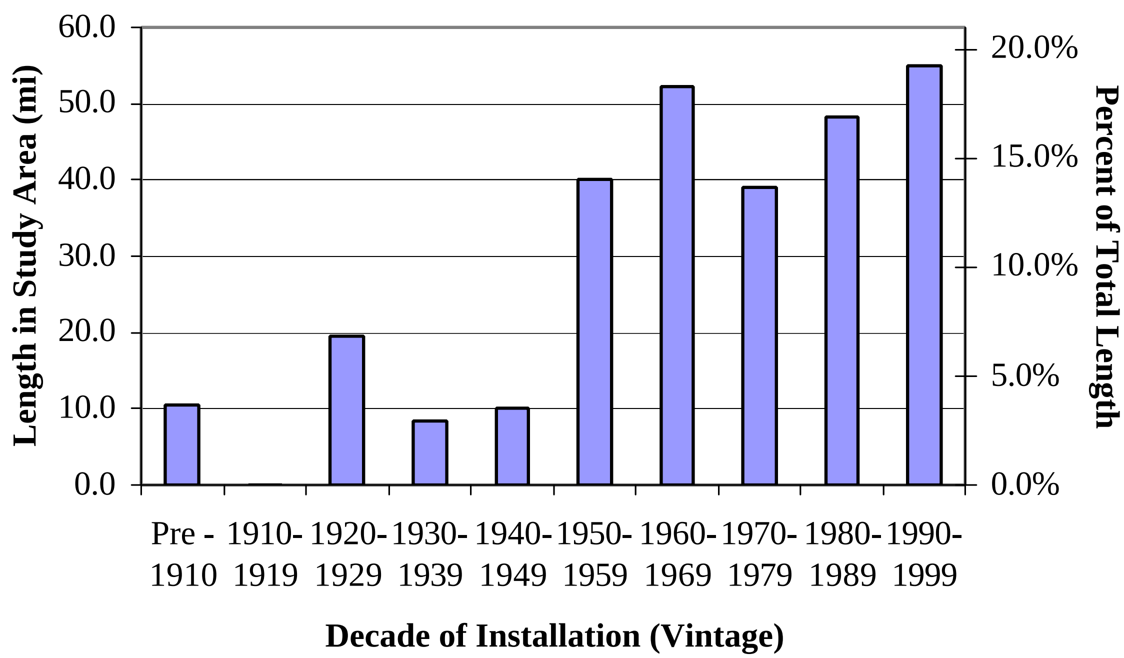


Figure C-3. Length Summary By Decade of Installation

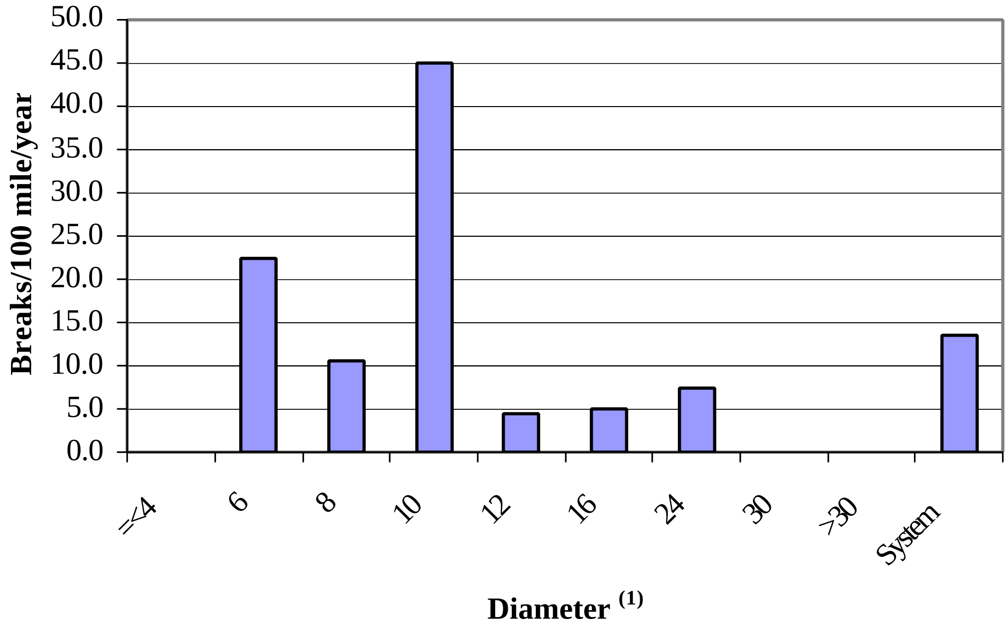


Figure C-4. Break Rate by Diameter

[1]This figure is developed using all repair data that matched pipes still in use in the system (520 repair events). For each year of repair history (i.e., 1985, 1986, etc.) the length of line in each diameter category was used to normalize the number of breaks. This activity resulted in a break rate in terms of breaks/mile for that diameter in that year. Averaging fifteen years of data and multiplying by 100 results in the values in Figure C-4 above. The 10-inch diameter pipe has an excessive rate; however this results from only 4 breaks over 0.6 miles of pipe. Most of the breaks occur in the six and eight inch categories.

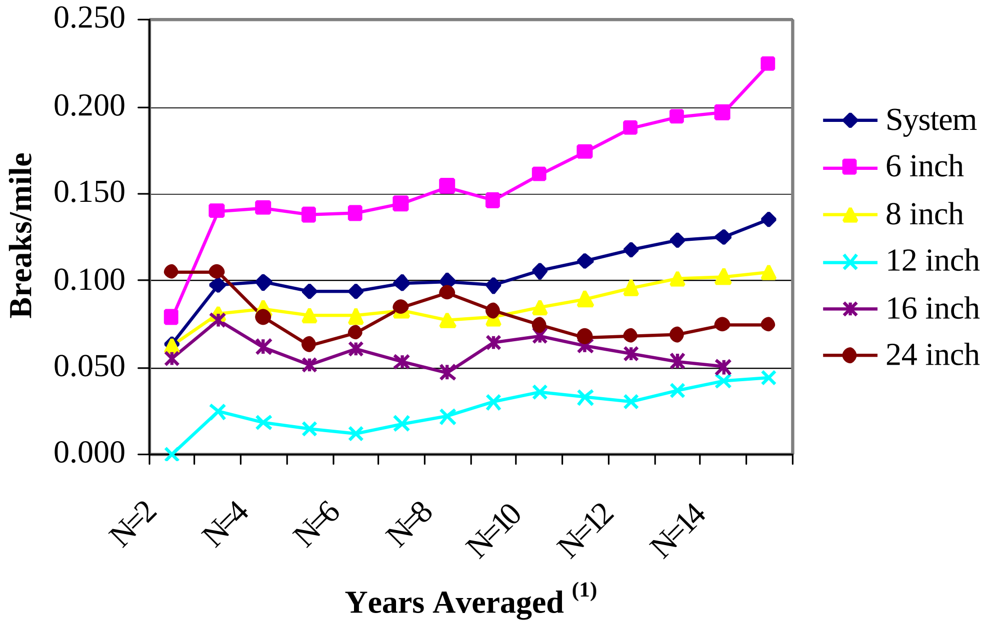


Figure C-5. Break Rate Stationarity by Diameter

[1]This figure is developed with the same results shown in Figure C-4. In this figure the break rates for each successive year (1985, 1986, etc.,) are averaged and averaged again (N=1, 2, 3, etc.) as a way to investigate the stationarity of the break rates. Data series that level off suggest steady break rates.

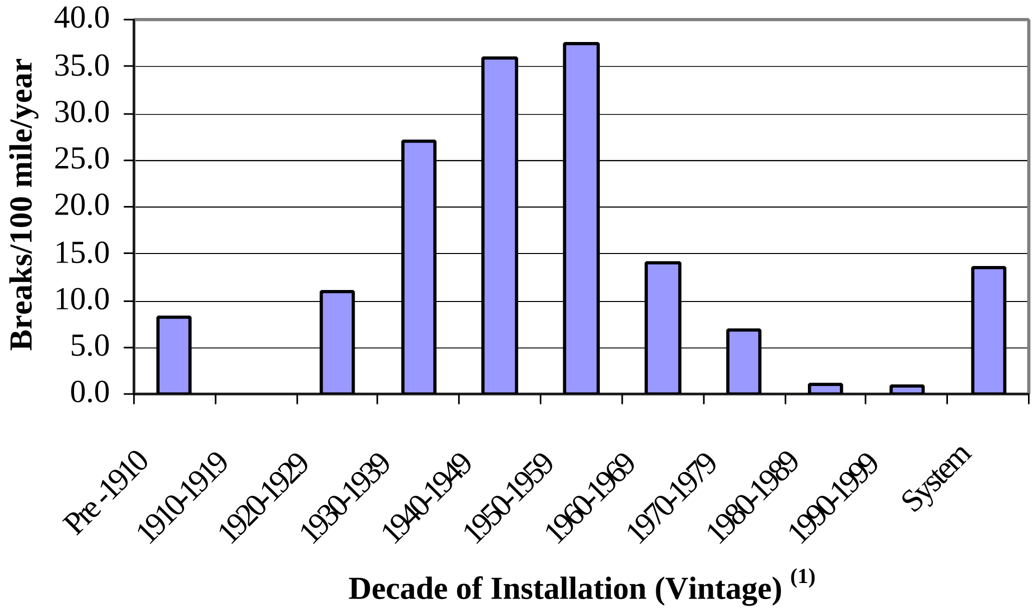


Figure C-6. Break Rate by Vintage

[1]This figure is developed using all repair data that matched pipes still in use in the system (520 repair events). For each year of repair history (i.e., 1985, 1986, etc.) the length of line in each vintage category was used to normalize the number of breaks. This activity resulted in a break rate in terms of breaks/mile for that diameter in that year. Averaging fifteen years of data results in the values in Figure C-6 above.

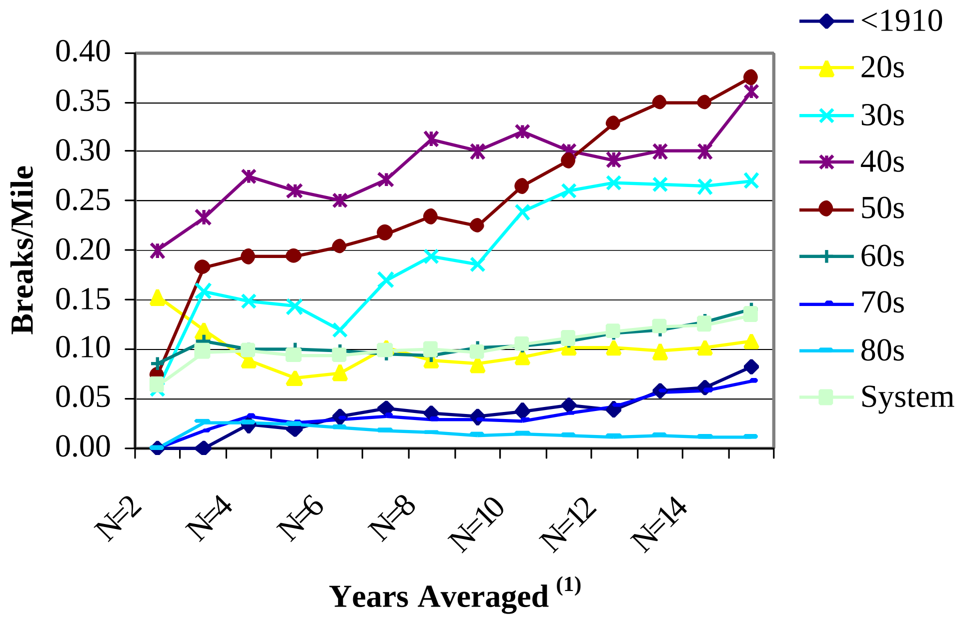


Figure C-7. Break Rate Stationarity by Vintage

[1]This figure is developed with the same results shown in Figure C-6. In this figure the break rates for each successive year (1985, 1986, etc.,) are averaged and averaged again (N=1, 2, 3, etc.) as a way to investigate the stationarity of the break rates. Data series that level off suggest steady break rates for the vintages shown.

APPENDIX D. GIS METADATA

A number of data sources were used to implement the management framework described in the text. The following is brief description of data generalities followed by specific tables that contain more unique details.

Most of the spatial data used in this work originated in a regional GIS system. These data were acquired by accessing a secure server and transferring the data in shape file format to a disk for analysis at another location. The spatial analysis was done using ArcView GIS 3.1; the hydraulic modeling was carried out in EPANET 2.0 using both the graphical user interface and the Toolkit. The EPANET Toolkit functions were accessed via a Visual Basic programming routine. Many of the intermediate data were managed using MS Access.   The spatial dat have the following projection characteristics:

Projection: Stateplane (Ohio South)

Units: Feet

Zone: 5001

Datum: NAD83

Spheroid: GRS1980

The following terms require further definition.

Clipping: Spatial technique available in ArcView that allows extraction of a subset of entities based on a geographic boundary.

Data Type: In this context, data type references a geographic entity or feature indicating whether the feature is a point, line, or area (polygon). Within ArcView this is often referred to as a feature's shape.

Geocoding: Spatial operation that uses a feature's address and layer of street information to locate the feature within the GIS.

Metadata: Information about the source, quality, and nature of the data; essentially data about the data. The metadata in this appendix briefly describe the data or refer to how the data were used or developed as part of this project. Extensive descriptions of the original sources (i.e., digitized from 1988 blue line drawings, aerial photography, etc.) are not provided here.

Shape Files: ArcView file format for storing location, attribute(s), and shape (i.e., point, line, area) of a geographic features.

Theme: ArcView's term for a spatial information layer.

View: Collection of themes within ArcView.

Table D-1. File Description and Metadata for Data Themes (Layers) Contained in the Pressure View

Theme [1]	**Shape File**	**Data Type**	**Metadata**
Monitored Locations	monitored locations .shp	point	Locations where the pressure is known from hydraulic simulations conducted in EPANET. Some of these locations (372) were nodes in the original utility provided model; an eight-digit ID number specified by the utility represents these. The other locations (1152) were added along modeled links in support of this work; they can be identified by their distinct PBN (potential break node) prefix. The X, Y, Z, coordinates for the original set (372) were provided as part of the model input file and added to the GIS as an event theme. The new nodes were created inside of EPANET resulting in an assignment of X and Y relative to the existing nodes. The new nodes inherited elevation (Z) from point layer of elevations via a nearest spatial query.
Point Elevations	point_ elev.shp	point	Layer of point data from regional GIS; each point has an associated elevation value.
More Frequent Main Breakage	break frequency .shp	point	From stochastic main break model. These nodes broke at least five times out of 500 total simulations. Output from simulation as a .csv file, saved in MS Access, summarized via a Pivot Table query and imported to the GIS. Data are results of runs conducted on 4/14/2001.
More Frequent Low Pressure	lowp_ frequency .shp	point	From stochastic main break model. These are locations with >1% chance of experiencing pressure <20 psi during 500 simulations. Each simulation is counted once if the event occurs. Therefore, these points have a 1% chance of pressure < 20 psi on a monthly time basis. (During a single simulation a location may experience low pressure due to a break for multiple hours, this is counted as one time, i.e. it happened this month). Data are from runs conducted on 4/14/2001.
Matched and Modeled Breaks	matched modeled .shp	point	Repair events that occurred on water mains still in existence in the study area and represented in the hydraulic model. These repairs are within 100 feet of a main in the model and have reported repair diameter matching the modeled main diameter. They also have a repair date after the main's installation date. In places where multiple lines existed in close proximity, some events were hand matched to nearby mains so the break attributes (diameter, vintage, etc.) could be known with surety. The field named "metadata" in the theme's attribute table is populated when a hand match occurred.

Table D-1. File Description and Metadata for Data Themes (Layers) Contained in the Pressure View (continued)

Theme	**Shape File**	**Data Type**	**Metadata**
Breaks on Modeled Lines	modeled breaks.shp	point	Repair events that are within 100 feet of a main represented in the hydraulic model.
Study Area Mains	tlmains .shp	line	Drinking water mains including attributes such as vintage, material, and diameter. Provided by the utility.
Study Area	study_ areas.shp	poly-gon	This theme originates from regional GIS related to utility service areas. The three areas analyzed as part of this work were then selected and made into a new shape file.

[1] Themes in this view were utilized in support of the adverse pressure gradient analysis described in the body of the thesis.

Table D-2. File Description and Metadata for Data Themes (Layers) Contained in the Pathway View

Theme [1]	Shape File	Data Type	Metadata
Structural Pathway	structural pathway.shp	line	All water mains > 10 ft in length and 6-inch in diameter and installed between 1940-1970 [2].
Study Area Mains	tlmains.shp	line	Drinking water mains including attributes such as vintage, material, and diameter. Provided by the utility.
Monitored Locations	monitored locations.shp	point	Locations where the pressure is known from hydraulic simulations conducted in EPANET. Some of these locations (372) were nodes in the original utility provided model; an eight-digit ID number specified by the utility represents these. The other locations (1152) were added along modeled links in support of this work; they can be identified by their distinct PBN (potential break node) prefix. The X, Y, Z, coordinates for the original set (372) were provided as part of the model input file and added to the GIS as an event theme. The new nodes were created inside of EPANET resulting in an assignment of X and Y relative to the existing nodes. The new nodes inherited elevation (Z) from point layer of elevations via a nearest spatial query.
Matched Repair Events	matched_ breaks.shp	point	Subset of theme All Repair Events. The total number of breaks in this table (520) is less than the total number of breaks reported. This is because in some situations a main has been replaced recently. For example, a repair has occurred and is reported in 1987, the nearest main to the break is reported as being installed in 1996. This would suggest the older main that was repaired in 1987 has since been replaced. Thus it is unknown what vintage of main was in place during the 1987 repair event. To determine which mains have been replaced, the fields for year of repair and year installed are compared. Year installed must be prior to year repaired for the match of maintenance to main to be accepted. Also fields for diameter repaired and the diameter of the main matched to the repair must be the same.
Study Area Connections	clipped_ addresses.shp	point	Locations served by the water utility (residences, businesses, etc.). From regional GIS, clipped to contain only those connections in the study area.

Table D-2. File Description and Metadata for Data Themes (Layers) Contained in the Pathway View (continued)

Theme	**Shape File**	**Data Type**	**Metadata**
Study Area Pitometer Districts	clippedpito.shp	polygon	Shape file of pitometer districts provided by the utility.
Study Area Soil Type	Clipped_ soils.shp	polygon	This theme originated from the regional GIS theme soils.shp. It was then clipped with the geoprocessing tool to only contain soils in the study area. The attribute table was then reviewed carefully after the clipping. This was done because during the clipping process some shapes on the borders are split but the areas are not recalculated. For example a large area that is a single shape may span the border multiple times and result in three or four slivers of area remaining. In the table each of these slivers has the same area as the original shape. The user will note some of these slivers that do not contain mains have been deleted.
All Repair Events	Mtce.shp	point	This theme originates from a repair shape file mtce.shp provided by the utility. It contains repair information for water mains in the study area from 1985 to 1999. The locations are geocoded; i.e., crew reports repair by address where work is conducted. The breaks were assigned to the nearest main using a spatial join. After the join it was noted the reported repair diameter did not always match the nearest main. This discrepancy is due to (1) multiple nearby mains did not allow the nearest query to match the correct main to the correct event (i.e., found the wrong main) (2) main has been replaced.

Table D-2. File Description and Metadata for Data Themes (Layers) Contained in the Pathway View (continued)

Theme	Shape File	Data Type	Metadata
Study Area Slope	Clipped_ slope.shp	polygon	This theme originates from the CAGIS theme slope.shp. It was then clipped with the geoprocessing tool to only contain slopes in the study area. It has six slope categories, (<10, 10-15%, 15-20%, 20-25%, 25-30%). After clipping, the attribute table was then reviewed carefully after the clipping. This was done because during the clipping process some shapes on the borders are split but the areas are not recalculated. For example a large area that is a single shape may span the border multiple times and result in three or four slivers of area remaining. In the table each of these slivers has the same area as the original shape. The user will note some of these slivers that do not contain main have been deleted.
Study Area	study_ areas.shp	polygon	This theme originates from regional GIS related to utility service areas. The three areas analyzed as part of this work were then selected and made into a new shape file.

[1] Layers in this view were utilized in support of the hydraulic pathway analysis described in the body of the thesis.

[2] Those lines < ten feet are generally small service connections. Approximately three miles of the entire system is made up of GIS records < 10 feet long. They were filtered out here as a way to identify sections of main that could logically be addressed through some type of maintenance activity.

Table D-3. File Description and Metadata for Data Themes (Layers) Contained in the Pathogen View

Theme [(1)]	**Shape File**	**Data Type**	**Metadata**
Monitored Locations	monitored locations.shp	point	Locations where the pressure is known from hydraulic simulations conducted in EPANET. Some of these locations (372) were nodes in the original utility provided model; an eight-digit ID number specified by the utility represents these. The other locations (1152) were added along modeled links in support of this work; they can be identified by their distinct PBN (potential break node) prefix. The X, Y, Z, coordinates for the original set (372) were provided as part of the model input file and added to the GIS as an event theme. The new nodes were created inside of EPANET resulting in an assignment of X and Y relative to the existing nodes. The new nodes inherited elevation (Z) from point layer of elevations via a nearest spatial query.
Controlled High Risk Connections	Highrisk-connections .shp	point	Connections greater than 4 inches screened by visual inspection of their name.
Subsurface Pathogen Risk	Sewerseptic .shp	line	Water mains having their centers in the 10 ft MSD buffer and the 200 ft septic buffer [(2)].
Study Area Connections	clipped_ addresses.shp	point	Locations served by the water utility (residences, businesses, etc.). From regional GIS, clipped to contain only those connections in the study area.
Study Area Mains	tlmains.shp	line	Drinking water mains including attributes such as vintage, material, and diameter. Provided by the utility.
Sanitary/Combined Sewer	clipped_msd sewer.shp	line	Sanitary and combined sewer lines from regional GIS, clipped to manage only those in the study area.
Study Area Septic Tanks	Clipped_septic .shp	point	Originates from regional GIS file septic.shp. It was then clipped to find only those septic tanks in the study area.
50' Septic	50_ft_ septic.shp	polygon	This theme is a result of a buffer query performed on the theme "Clipped_septic.shp." The buffer was set to fifty feet; i.e., the total diameter of any feature is 100 feet [(3)].

Table D-3. File Description and Metadata for Data Themes (Layers) Contained in the Pathogen View (continued)

Theme	Shape File	Data Type	Metadata
100' Septic	100_foot_ septic.shp	polygon	This theme is a result of a buffer query performed on the theme "Clipped_septic.shp." The buffer was set to 100 feet; i.e., the total diameter of any feature is 200 feet.
150' Septic	150_foot_ septic.shp	polygon	This theme is a result of a buffer query performed on the theme "Clipped_septic.shp." The buffer was set to 150 feet; i.e., the total diameter of any feature is 300 feet.
200' Septic	200_foot_ septic.shp	polygon	This theme is a result of a buffer query performed on the theme "Clipped_septic.shp." The buffer was set to 200 feet; i.e., the total diameter of any feature is 400 feet.
5' MSD Sewer	5_ft.shp	polygon	This theme is a result of a buffer query performed on the theme "Clipped_msdsewer.shp." The buffer was set to five feet; i.e. the total width of any feature is ten feet.
10' MSD Sewer	10_foot_ buffer.shp	polygon	This theme is a result of a buffer query performed on the theme "Clipped_msdsewer.shp." The buffer was set to ten feet; i.e. the total width of any feature is 20 feet.
Study Area	study_ areas.shp	polygon	This theme originates from regional GIS related to utility service areas. The three areas analyzed as part of this work were then selected and made into a new shape.

[1] Layers in this view were utilized in support of the contamination source analysis described in the body of the thesis.

[2] Options for finding one feature relative to another (i.e., proximity of sewer lines to water mains) include intersection, separation distance, the presence of one feature completely within another feature, or having the center of one feature within another. Iterations between these options suggested the "having centers in" approach returned results most representative of the objective of this proximity analysis.

[3] A buffer analysis builds an area of influence around a selected feature.

Table D-4. File Description and Metadata for Data Theme (Layers) Contained in the Population View

Theme [1]	**Shape File**	**Data Type**	**Metadata**
Monitored Locations	monitored locations.shp	point	Locations where the pressure is known from hydraulic simulations conducted in EPANET. Some of these locations (372) were nodes in the original utility provided model; an eight-digit ID number specified by the utility represents these. The other locations (1152) were added along modeled links in support of this work; they can be identified by their distinct PBN (potential break node) prefix. The X, Y, Z, coordinates for the original set (372) were provided as part of the model input file and added to the GIS as an event theme. The new nodes were created inside of EPANET resulting in an assignment of X and Y relative to the existing nodes. The new nodes inherited elevation (Z) from point layer of elevations via a nearest spatial query.
Sensitive Centers	sensitive_ pops.shp	point	Sensitive population groups were identified in a local phone directory and geocoded against area_streets.shp to locate in the GIS.
Study Area Streets	area_ streets.shp	line	Layer of streets from the regional GIS, clipped to manage only those in the study area.
Study Area	study_ areas.shp	polygon	This theme originates from regional GIS related to utility service areas. The three areas analyzed as part of this work were then selected and made into a new shape
Study Area Mains	tlmains.shp	line	Drinking water mains including attributes such as vintage, material, and diameter. Provided by the utility.

[1] Layers in this view were utilized in support of influence analysis described in the body of the thesis.

Table D-5. File Description and Metadata for Data Themes (Layers) Contained in the Combined Susceptibility View

Theme	Shape File	Data Type	Metadata
Sensitive Populations	sensitive_ pops.shp	point	Sensitive population groups were identified in a local phone directory and geocoded against area_streets.shp to locate in the GIS.
Pressure, Pathway, Source	pressurepath-waysource.shp	line	Mains with at least one occurrence < 0 psi (4/14/2001 data), within 10 feet of sewers or 200 feet of septic tanks and six-inches in diameter, and installed from 1940 to 1970 and greater than 50 feet.
Pathway, Source	pathwayandpat hogen50.shp	line	Mains within 10 feet of sewers or 200 feet of septic tanks and six-inches in diameter and installed from 1940 to 1970 and greater than 50 feet. Represents a contamination risk during repair at these locations.
Pressure, Pathway, Source	backflowrisk2	point	Locations with at least once occurrence < 20 psi and serve a high risk connections based on review of service connection data. Pressure data are from 4/14/2001 simulations.
Pressure, Pathway	backflowrisk .shp	point	These locations have occurrence(s) < 20 psi (some high) and/or serve many connections. These locations are data points plotted higher and farther to the right on Figure 5-9.
Study Area Mains	tlmains.shp	line	Drinking water mains including attributes such as vintage, material, and diameter. Provided by the utility.
Structural Pathway	structural pathway.shp	line	All water mains > 10 ft in length and 6-inch in diameter and installed between 1940-1970.
Subsurface Pathogen Risk	sewerseptic .shp	line	Water mains having their centers in the 10 ft MSD buffer and the 200 ft septic buffer.
Mains < 0 psi	Mainslessthan0 psi.shp	line	Mains connected to monitored locations experiencing at least one occurrence less than 0 psi. These were selected manually considering the topology between the mains and the modeled nodes.

Table D -5. File Description and Metadata for Data Themes (Layers) Contained in the Combined Susceptibility View (continued)

Theme	**Shape File**	**Data Type**	**Metadata**
Locations < 0 psi	monitored _ locations.shp	point	Layer of monitored nodes that had at least one occurrence less than 0 psi (from 4/14/2001 data). The locations with the occurrence live in a data table that is joined to the attribute table for Monitored Locations. This gives a True/False field that can be filtered to find those below zero psi.
Monitored Locations	monitored locations.shp	point	Locations where the pressure is known from hydraulic simulations conducted in EPANET. Some of these locations (372) were nodes in the original utility provided model; an eight-digit ID number specified by the utility represents these. The other locations (1152) were added along modeled links in support of this work; they can be identified by their distinct PBN (potential break node) prefix. The X, Y, Z, coordinates for the original set (372) were provided as part of the model input file and added to the GIS as an event theme. The new nodes were created inside of EPANET resulting in an assignment of X and Y relative to the existing nodes. The new nodes inherited elevation (Z) from point layer of elevations via a nearest spatial query.
Pathway and Pathogen	pathwayand-pathogen.shp	line	Mains within 10 feet of sewers or 200 feet of septic tanks and six-inches in diameter and installed from 1940 to 1970. Represents a contamination risk during repair at these locations.
More Frequent Low Pressure	lowp_ frequency .shp	point	From stochastic main break model. These are locations with >1% chance of experiencing pressure <20 psi during 500 simulations. Each simulation is counted once if the event occurs. Therefore, these points have a 1% chance of pressure < 20 psi on a monthly time basis. (During a single simulation a location may experience low pressure due to a break for multiple hours, this is counted one time, i.e. it happened this month). Data are from runs conducted on 4/14/2001.
Study Area	study_ areas.shp	polygon	This theme originates from regional GIS related to utility service areas. The three areas analyzed as part of this work were then selected and made into a new shape file.