OPTIMIZING MAINTENANCE AND REPLACEMENT ACTIVITIES FOR WATER DISTRIBUTION PIPELINES

By

Lindsay Michelle Jenkins

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Civil Engineering

December, 2014

Nashville, Tennessee

Approved:

Sanjiv Gokhale, Ph.D.

P.K. Basu, Ph.D.

Sankaran Mahadevan, Ph.D.

Mark McDonald, Ph.D.

Kenneth Pence, Ph.D.

Scott Potter

DEDICATION

"…to God, who by his mighty power at work within us is able to do far more than we would ever dare to ask or even dream of—infinitely beyond our highest prayers, desires, thoughts, or hopes". Ephesians 3:20

ACKNOWLEDGEMENTS

I would first like to thank my dissertation committee, without whom this work would not have been possible. Dr. Sanjiv Gokhale, when we first met, you explained that your name means eternal life. Thank you for living up to your name. I will be forever grateful to you for reviving my dreams of earning a PhD. Long after you have retired, your legacy will live on in the countless students you have guided and mentored. I am ever thankful to be one of them.

Dr. McDonald, thank you for not quitting on me, even upon departing Vanderbilt. It has been an honor and privilege to learn from you and work beside you at Lipscomb University. Dr. Mahadevan, thank you for stretching me academically. Dr. Basu, thanks for being an advocate for me in this long journey. Dr. Pence, thank you for your help and review of the material.

Mr. Potter, thanks to you and the staff at Metro Water Services for your countless assistance and checks to insure the relevance of my work. Additional thanks to Shanna Whitelaw, Rob Smith, Hal Balthrop, and Cyrus Toosi for their help and input.

Thanks to the staff at City of Brentwood, Todd Spangler and James McCord, and the City of Franklin, Mark Hilty. I hope you find this work beneficial as you provide quality service to the residents of Middle Tennessee.

To my colleagues at CH2M HILL, thank you for finding value in my work. I have been blessed by your constant support and encouragement. Working with each of you has greatly supplemented my knowledge and understanding of this problem and the industry as a whole. Thanks especially to Kevin Colvett. Your input significantly improved the quality and applicability of this research.

Thanks to the Department of Civil and Environmental Engineering for providing financial support during my tenure. Thanks especially to the administrative staff, who have always brightened my day. You are blessings. To the ladies past and present assigned to offices 261 and 262, thank you for being my friends and support system. Thanks to the Risk and Reliability Research group.

I would next like to thank the staff and students at the Raymond B. Jones College of Engineering, Lipscomb University. Without your encouragement, I would have never stayed in the Ph.D. program. Thanks for giving me the opportunity to teach, lead, and learn alongside you.

Thanks to my amazing family. Trey and Ashley, thank you supporting me and for always helping to turn my tears into laughter.

To my dad, when I asked if you had ever dreamed that you would have made it where you are today, you answered profoundly, “No one ever told me I couldn’t”. Thanks for showing me how to dream without limits. You have demonstrated how to live an impactful and adventurous life with hard work, courage, integrity, and sacrifice.

To my mom, thank you for teaching me to hope. You continuously reminded me of Romans 5, which explains that being patient through difficult times builds character. This character provides us with the strength to hope that God will use us and the challenging times in our lives to fulfill a greater plan. Without this hope, I would not be thankful for the struggles and work that brought me to this point, and joyfully anxious about the adventures to come as I try to live out God-inspired dreams.

Finally, to those who made this journey more difficult than it could have been, thanks for the opportunities to build patience, character, confidence, and most importantly, hope.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

Table Page

CHAPTER 1 - INTRODUCTION

1.1 Motivation

Over 54,000 drinking water systems in the United States serve over 320 million residents through 2 million miles (3.2 million km) of water distribution pipes. Many of these pipes were installed 50 to 100 years ago, with a design life of only 100 years. As more and more pipes are reaching the end of their useful lives, utilities and consumers must address the problems associated with pipe breaks. Pipe breaks not only cause disruption in water distribution service, they can wreak havoc on the surrounding environment and cause traffic delays and infrastructure damage. Many have tried to evaluate the total economic impact of pipe breaks that includes societal costs such as increased travel time due to lane closures and detours and increased carbon emissions associated with pipeline replacement [1]–[3]. This research presents improved models and frameworks to assist utilities in developing maintenance and pipe replacement programs to mitigate and reduce the risk of pipeline failures. The models presented limit the number of input parameters and are tested on utilities with varying size, uncertainty, and break rates.

Several studies have been released by the Environmental Protection Agency (EPA) and American Water Works Association that attempt to quantify the long-term needs associated with maintaining water distribution networks. Twenty-year financing needs estimates range from $280 billion [4] to $1 trillion [5] to maintain and replace assets in water distribution networks. These high estimates are due to prolonged deferred maintenance of ageing infrastructure. The current average annual pipeline replacement rate for utilities is just 1% of the total network [4].

Utility directors cite multiple pipeline breaks as a primary criterion for replacement [6]. Operations and maintenance (O&M) programs need to be adjusted to thoughtfully increase replacement rates to minimize the long-term economic consequences of deferred maintenance. Comprehensive asset management programs are needed in order to assist utilities in creating maintenance, rehabilitation and replacement (MR&R) spending plans that minimize the long-term costs to users to maintain a minimum level of service for water distribution networks. The need for asset management for water and sewer utilities was highlighted with the issuance of Government Accounting Standards Board (GASB) Statement 34. Issued in 2009, GASB 34 requires utilities to calculate and report the costs of maintaining and improving assets over a twenty-year period. To make these calculations, utility managers are given the option of using either a historical cost based depreciation model, or a comprehensive asset management program [7].

The Environmental Protection Agency (EPA) has outlined best practices for asset management. The five core framework should determine the current state of assets, level of service, critical assets, minimum life-cycle costs, and a long-term funding plan [8]. As part of asset management programs and in adherence to GASB 34 standards, much attention is paid to building and maintaining information systems with data describing the current state of assets. Many utilities have digitized as-built maps of their water and sewer networks in order to create Geographic Information System (GIS) models of the network, in order to keep network maps up to date. The network must be inventoried to account for network age and materials. The inventory data can be stored in the GIS database. Asset management teams must also maintain data describing segment inspections, and maintenance operations including the location, date, and actions taken.

The data available through asset management information systems can be utilized to develop decision support systems (DSS) for MR&R activities. These DSS can assist in the planning and/or design of pipeline rehabilitation and replacement activities for water and/or sewer networks. Though DSSs all differ in functionality, the most comprehensive decision support systems identify pipelines with the highest probabilities of failure, quantify the criticality of failure, and identify repair and rehabilitation strategies based on overall risk of pipe segments failing.

Recent surveys of utilities by Matthews et al. [9] and St. Clair and Sinha [10] revealed that few large utilities, serving hundreds of thousands of consumers, incorporate failure prediction models as a part of their DSS's. Though many statistical models have been developed and are summarized in Chapter 2, utilities are simply not embracing the new technologies and methods. The following paragraphs examine possible causes for the low adoption rate of failure prediction models among utilities.

First, utilities may not have the necessary data to develop a failure prediction model as specified by the literature. The statistical models are based on the assumptions that pipes of the same age, diameter, and material degrade at the same rate. The differences in failure times for homogenous pipe groups is due to outside random factors including pressure surges, loading above the pipe, acidity of the soil around the pipe, joint assembly, and many other "explanatory variables" or covariates in the stochastic models. As noted by Wood and Lence [11] instead of trying to calculate covariate value, surrogates are used to account for the factors contributing to pipe failure. For example, classifying a pipe as under a roadway accounts for factors due to increased loads on the pipe compared to a pipe that is under a sidewalk.

Many of the models reviewed in Chapter 2 include a significant number of these explanatory variables, requiring data that is not readily available for most utilities including direct condition assessment data, soil conditions around pipelines, pipe bedding depth and materials. Even the most basic information regarding pipe segment material, installation year, and length can be missing from an asset management database. Since most asset management database are only 10 to 20 years old, data about pipes installed and possibly repaired prior to the inception of the database can often be uncertain. Additionally, there are no standards for what information should be contained in an asset management database. The need for database standards for pipeline asset management has only recently been addressed [12].

The problem of missing or unknown information has been documented by several researchers [13]–[16]. Most approaches rely on excluding assets with missing data from the training data set, making an educated assumption, or assigning a median value to the unknown parameter based on the known data. None of these approaches take into account the potential model bias due to such assumptions. Not accounting for this bias leads to uncertainty and lack of confidence in failure prediction performance.

The lack of database and GIS model standards has also influenced how one of the most basic pipeline properties is stored. The pipe length stored in asset management databases is often the length resulting from digitized as built maps. This length does not represent the actual segment length in the field. Many models include pipe length as a covariate, yet the lack of standardized definition of pipe length can influence prediction results [17], [18]. The uncertainties associated with including pipe length in statistical models could negatively impact the transferability of the models.

In addition to missing data, medium sized and small utilities face even more challenges when adopting failure prediction modeling. Such utilities have sparse data with fewer than one hundred breaks per pipe material class. To develop reliable failure prediction models a statistically significant sample size of recorded pipe failures is needed, which some estimate to be at minimum five years of data [19]. Utilities could be hesitant invest the time and money to determine if they can develop statistically significant failure models given sparse data. Furthermore, medium and small utilities bring in less revenue that could be allocated for maintenance studies including failure prediction modeling.

Few case studies are available that address how to fully utilize failure prediction modeling in making asset management decisions. Without demonstrating how the failure predictions can be incorporated with risk analyses to make decisions, the benefit of using a failure prediction model over a Pareto analysis or multi-criteria decision based method for identifying pipes needing MR&R strategies in the presence of data uncertainties has not yet been demonstrated.

Though many researchers have applied optimization tools including linear programming and genetic algorithms to develop MR&R strategies [20]–[23], most have been tested on small networks or subsets of networks, with no investigation into the scalability of these models for larger utilities. None have been demonstrated using a model that accounts for substantial uncertainties. Utility decision makers are in need of case studies of easy to implement optimization routines that demonstrate how to incorporate failure likelihood and consequence analysis into MR&R planning. Additionally, most optimization/scheduling frameworks presented do not account for the spatial relationship of identified MR&R projects, which is crucial to decision makers planning inspection and rehabilitation activities in geographic subsets of the network.

Optimization results presented in literature are a listing of pipes of various lengths scattered throughout the utility. Many utilities do not have work crews that can be dedicated to replacing small amounts of pipe across the network. For larger utilities, this work must be undergo a procurement process and be contracted out. Economies of scale and constraints such as the ones listed dictate that most replacement projects consist of replacing at minimum a quarter of a mile (0.4 km) of linear feet of pipe. Optimization methodologies need to be refined to consider this constraint.

1.2 Research Goals

Based on the needs and motivation presented, the proposed research will accomplish the following goals:

1. Investigate the impact of uncertainty in failure prediction modeling
2. Examine the suitability of a model transfer to assist the development of failure prediction models for medium and small utilities with sparse data

3. Demonstrate how to integrate failure model results into MR&R planning

1.3 Research Objectives

The following objectives must be met in order to pursue the research goals:

Objective 1—Calibrate and validate a Weibull Hazard Rate Model (WHRM) that accounts for uncertainty in pipeline properties using binary variables and reduces the amount of explanatory variables needed by introducing a spatial failure clustering variable.

Objective 2—Develop an Excel-based genetic algorithm optimization tool to determine which assets should undergo replacement during a short-term planning horizon. This algorithm will minimize a risk-based penalty function calculated using survival model and criticality assessment results. This algorithm will also constrain projects to subsets of the network, that better reflect utility operations in replacement programs.

Objective 3—Calibrate and validate WHRMS for medium and small utilities and recalibrate parameters using model transfer techniques that explore the contribution of information from large utilities.

1.4 Organization of Dissertation

The dissertation is organized in the following manner:

***Chapter 2: Background* –** This chapter first discusses pipe failure modes, repair actions, and replacement activities. Next, a comprehensive review of statistical models for pipeline failure prediction is presented. Limitations of the models and contributions are discussed.

***Chapter 3: Utility Overview* –** The three utilities studied in this document are introduced. This chapter includes background information, material composition, and performance statistics for each network.

***Chapter 4: Comparison of Pipeline Failure Prediction Models for Water Distribution Networks with Uncertain and Limited Data* –** This chapter examines the changes in prediction performance when adding parameters to account for uncertainties and the spatial distribution of breaks to the widely used WHRM. Validation and prediction performance results from a case study of Utility A are presented.

***Chapter 5: Optimization of Maintenance and Replacement Activities for Water Distribution Pipes Using WHRM* –** This chapter provides the methodology for utilizing a validated WHRM to optimize maintenance and replacement activities for water distribution pipes based on failure risk and hydraulic reliability. Results of criticality and hydraulic reliability analyses are presented. An optimization routine is presented that prioritizes MR&R activities given budget constraints. This methodology is demonstrated on Utility A.

***Chapter 6: Investigating the Spatial Transferability of Pipeline Failure Prediction Models for Medium and Small Utilities* –** In this chapter, several model transfer techniques widely used in other applications are employed to examine if information from larger, neighboring utilities can be used to improve prediction performance for medium and small utilities with limited data. WHRMs using the model form described in the previous chapter are developed for Utilities A, B, and C. The results of three model transfer techniques from Utility A to Utilities B and C are presented. Recommendations for developing models to prioritize MR&R activities for medium and small utilities are presented.

Chapter 7: Framework for Prioritizing Pipe Maintenance and Replacement Activities for Small Utilities -- This chapter presents a framework for using clustering algorithms to identify high failure rate zones in medium and small distribution networks. Cluster analysis is utilized to examine potential root causes of failure and recommend MR&R activities. A criticality/consequence based prioritization method is introduced and MR&R activities are identified for Utility B.

Chapter 8: Conclusions and Future Recommendations – This chapter presents the conclusions from the research, addresses limitations of the work, and provides recommendations for future work in the field of study.

CHAPTER 2 – BACKGROUND

2.1 Introduction

This chapter presents the necessary background information to gather an understanding of water pipeline failure and MR&R activities. First, pipeline materials and failure modes are discussed. Typical pipeline repair operations are reviewed, and maintenance activities in the form of condition assessment technologies are introduced.

A review of models developed for prioritization of MR&R activities is also included. Though multiple model forms are discussed, the focus of the review is on statistical models, with specific attention paid to model training and validation. Case studies of pipe replacement optimization methodologies are presented along with a review of utility practice. The final section discusses limitations of the works presented and details how this work compliments the current body of knowledge.

2.2 Pipe Materials and Failure Modes

2.2.1 Ductile Pipe

Gray cast iron pipe is the most popular pipe material in the United States, comprising over 50% of the total US water main network [24]. Some of the earliest cast iron pipe in the U.S. was installed in the 19th Century and remained popular until the 1970's, when the popularity of ductile iron pipe grew. The root cause of failure in ductile pipes is excessive forces acting upon the pipe in the forms of internal pressure, bending, soil movement, and thermal expansion due to differences between the temperature of the water pipes and the surrounding soil or the pipes and joint mechanisms [25]. Table 2.1, adapted from [26] shows failure modes and mechanisms for ferrous and PVC pipes considered in this study.

Table 2.1 Failure Modes and Mechanisms

Failure Mode		Failure Mechanism	Material
Cracking	Circumferential	Bending moments applied to the pipe and soil movement which produce tensile forces on pipe	All
	Longitudinal	Internal water pressure, crushing and compressive forces acting on pipe	All
	Spiral	Pressure surges and/or combination of bending forces and internal pressure	All
	Mixed	Combination of stresses	All
	Ring	Axial tension, bending, traffic load, settlement, uplift, production, fatigue, residual stresses, temperature, and frost	PVC
	Axial	Internal pressure, bending, traffic load, production, residual stresses, and frost	PVC
	Irregular	Environmental such as chemical, UV, and stress cracking	PVC
Fracture	Circumferential	Bending moments applied to the pipe and soil movement which produce tensile forces on pipe	All
	Longitudinal	Internal water pressure, crushing and compressive forces acting on pipe	All
	Spiral	Pressure surges and/or combination of bending forces and internal pressure	All
	Mixed	Combination of stresses	All
Buckling	Axial	External pressure, axial compression, temperatures, fire, and interventions	PVC
	Transverse/ring	External pressure, axial compression, production, residual stresses, temperatures, fire, and interventions	PVC
	Non-symmetric	Longitudinal bending and brazier effect	PVC
	Longitudinal	Axial compression and thermal effects	PVC

The failure modes observed amongst ductile pipes are blowout holes, circumferential cracking, bell splitting, longitudinal cracking, bell shearing, and spiral cracking. Blowout holes are caused by corrosion pitting which causes wall thinning. Eventually, the pressurized water exceeds the strength of the thin pipe wall and a hole is formed. Circumferential cracking, the most common failure mode for pipes less than 14 inches, is caused by bending forces or tensile forces due to soil movement [25]. Another very common failure mode for small diameter cast iron pipes is bell splitting. Bell splitting is primarily found in pipes installed in the 1930's and 1940's when leadite, a sulphur-based jount-sealing compound was utilized to create joint seals at bells.. The failures occur due to the difference between the coefficient of thermal expansion in the joint-sealing compound and the metal in the pipes. In cold temperatures, the leadite expands differently than the cast iron pipe, causing splitting at the bell.

Large diameter ductile pipes are subject to longitudinal cracking [25]. These failures characterized by cracks that propagate along the length of the pipe wall are caused by increases in a combination of internal forces due to water pressure and external forces due to loading conditions above the pipe such as traffic. In some cases, longitudinal cracking can be seen on opposing sides of the pipe resulting in more expensive repairs [25].

Medium diameter ductile pipes are also subject to an additional type of pipe cracking called spiral cracking, in which the pipe crack propagates as a coil around the pipe. Again, this type of pipe failure is caused by bending and internal and external pipe forces [25].

2.2.2 PVC Pipe

The 1970's saw a transition to the use of Polyvinyl chloride (PVC) pipe because it was cheaper to buy, transport, and install than ductile pipes [27]. Another benefit of PVC is that it does not corrode like ferrous pipe. The material composition of PVC make it more brittle than ductile pipes under certain conditions, but also more prone to bending and flexure.   A study of PVC pipe used for gas distribution showed that manufacturing flaws and installation practices contributed the most of PVC pipe failures [28]. With respect to installation practices, pipes left out in the sun too long prior to bedding are subject to chemical breakdown, degrading the structural integrity of the pipe. Contractors must also be careful in bedding pipes to make sure that the pipe does not like on large, sharp rocks, which could eventually rupture the brittle pipe. An additional cause of pipe failure due to installation of PVC pipe is over insertion at the pipe joint, where the spigot joint is inserted too far into the bell, causing fracture [29].

The elastic properties of PVC also make it susceptible to rupture due to cyclical pressure loadings in the pipe, also known as “water hammer” [30]. Though most PVC pipe was installed in the 1970’s, researchers are still working to improve hydraulic calculations to quantify the impact of pressure hammer for design purposes [31].

There are many different mechanisms for similar failure modes. In order to truly capture pipe failure mechanisms, the long-term operational and environmental information regarding the specific pipe segment must be known. Information gathered during pipe repair can assist in determining failure mechanisms.

2.3 Pipe Repair

Each utility has operating procedures for addressing water main breaks. The following is a summary of the standards operating procedures for pipe replacement for Utility B in this study

[32]. At the instance of a break, a work order is placed for the repair work to be done. This work order is tracked using a spreadsheet, database, or more advanced computerized maintenances management software (CMMS).

After initializing the work order, the work crew is dispatched to repair the pipe. The crew gathers information about the pipe including material, size, upstream valve location, and downstream valve location. On occasion the appropriate valves are difficult to locate in the field. Valve closure requires precision in order to avoid pressure surges and losses in other parts of the network. The information about the network is often distributed to crews via paper maps, or sometimes digitally with mobile GIS maps.

Once information is gathered, the crew begins actions to repair the pipe. First, failure site needs to be excavated in order to expose the pipe. In some cases, this involves cutting into a major road, and diverting traffic. After the pipe is exposed, the break type should be identified and an appropriate repair method selected. Pinhole failures and short longitudinal cracks can often be fixed with a repair clamp, while other failure modes require replacement of the pipe section. When replacing a pipe section, water must be turned off and customers affected should be notified. Bypass pumping can be enabled when redundant lines are available to limit the customers affected by outages. The damaged section of the pipe is removed and a new pipe segment is installed. The line is sterilized and water is turned back on. The pipe then undergoes pressure testing and the water is sampled for chlorine and bacteria. Lastly, the trench is backfilled and pavement/surface restoration is performed.

The costs of pipe repair are dependent upon the location of the failure and the type of failure and material. For example, a pipe break under a major road has greater economic impact than one that occurs under a grassy surface. Pavement repair can add significant direct costs, and diverting traffic adds to societal costs of failures. A study of large utility main break repairs revealed that societal costs accounted for half of all costs of repair [2]. The authors state that for large diameter pipes, location and the amount of potential water losses are the driving factors of failure costs.

2.4 Pipe Maintenance

Maintenance activities can be performed to mitigate the risk of pipe failures. Specifically, condition assessment can be performed to try to assess the current state of the pipe and predict the future condition. A focus has been placed on non-destructive indirect condition assessment techniques, as they are less costly than alternative, destructive methods. Common indirect condition assessments include smart ball acoustic technologies, electromagnetic, and ground penetrating radar [12]. Grigg [33], asserts that the limitations of condition assessment technologies are economical, not technological. Due to the expense associated with these technologies, they are not used on the entire network and instead utilized on a site-specific basis. Additionally, more training and information is needed to help interpret condition assessment results to make meaningful asset management decisions, especially with respect to pipes that could result in catastrophic failure. A chart used by consultants at CH2M HILL to screen candidate condition technologies is shown in Table 2.2. Note that many other technologies exist, and this is just a summary of some of the most widely used.

Table 2.2: Condition Assessment Technology Screening Tool (Courtesy: CH2M HILL)

Tier	**1**	**1**	**1**	**1**	**1**	**1**
Technology	Pressure and Flow Monitoring	Soil Survey and Corrosion Analysis	Guided Wave	Infrared thermal	Ultrasonic	Acoustic Emissions
Pipe Service Condition	In Service, Pipe Full	In Service, Pipe Full	In Service, Pipe Full	In Service, Pipe Full	In Service, Pipe Full	In Service, Pipe Full
Pipe Material	Metal and Concrete	Metal and Concrete	All	All	Metal	Concrete (PCCP)
Variable Detected	Service Pressure and Flow Condition	Corrosion Potential	Remaining Wall Thickness	Leak Detection	Remaining Wall Thickness	Active Failure in Prestressing Wires
Technical Maturity	High	Mid	Mid	Low	Mid	Mid
Inspection Level	Survey Level	General Details	General Details	Survey Level	Specific Details	Survey Level
Cost Range	<$1/ft	<$1/ft	<$10/ft	<$5/ft	NA	<$20/ft
Pros	Establishes Criticality of Pipeline and Current Performance Requirement	Provide Detailed Knowledge of Risks from Corrosives Soils and Stray Currents	Can Inspect Segments of Pipeli	Remote Sensing Method - No Access Needed or operational Disruption	Gives Detailed information of Pipe Wall Thickness	Can Locate Specific Wire Breaks
Cons	Temporary Flow Measurement Can Be Access Issues	Not a Direct Inspection of Pipeline Conditon Establishes Risk Factors Only	Access Pits Needed; Follow up Inspection May be Needed at Specific Defects	Misinterpretation of Data - Voids or Leakage may be from Other Sources	Specific Only to the Location Tested	Long-Term Monitoring Strategy

With many condition assessment options available, researchers have developed decision support tools to evaluate the cost trade off and value of information provided by condition assessment technologies. Osman et al. [34] demonstrated the use of a partially observable Markov decision process and genetic algorithms to optimize the type of condition assessment technology and inspection interval for water distribution pipes in a Canadian utility based on value of information analysis. Kleiner [35] presented a decision support system for scheduling both inspection and renewal of large water distribution mains. This markov transition based decision tool considers the cost of inspection, the cost of preventative maintenance and the cost of replacement. This tool has only been demonstrated as proof of concept, and needs further refinement to be utilized by utilities.

With the options to replace or inspect assets, and the varying value of information gained from condition assessment technologies, a resource allocation problem is presented. Utilities need an assortment of models to prioritize MR&R activities, and decision support tools are needed to optimize assessment technology usage and replacement activities. Models to predict pipe failure rates and condition state are needed to help solve this resource allocation problem.

2.5 Water Pipeline Failure Prediction Models

Historical records for water main repair work orders can be synthesized with GIS data to develop failure prediction or survival models for prioritizing pipe replacement and maintenance activities. The following sections summarize pipeline failure/condition prediction models by model form. The model forms considered are deterministic, statistical and machine learning. The review of statistical and machine learning models includes subsections devoted to model calibration and model validation. A section devoted to utility practice describes the models currently being utilized by large utilities. Limitations and recommendations derived from the literature review are discussed at the end of this section.

2.5.1 Deterministic

Comprehensive reviews of deterministic models have been provided by St. Clair and Sinha [10], and Rajani and Kleiner [36]. The models are either mechanistic, empirical based or both. The long-term mechanical performance of the pipe is related to known parameters describing the pipe's physical characteristics, operational characteristics, and the environment around the pipe. The amount of data and computational effort required for deterministic models limits their application. Though they can provide a more accurate prediction of the long-term performance of a pipe, they are generally reserved for large-diameter pipes where more properties are known or can be gathered using direct condition assessment techniques [36]. Also, many of the physical models based on experimentation are site-specific and cannot be applied in other areas of the network [10]. The regional adoption of these models would not be appropriate. Moreover, these models are not appropriate for utilities that lack basic data describing the physical and environmental characteristics surrounding water pipes

2.5.2 Statistical

A review of statistical models presented in literature over the past ten years was written by St. Clair and Sinha [10] . The authors summarize model form and data requirements, but did not explicitly outline model training and validation techniques. The following section reviews many of the models reviewed by St. Clair and Sinha and includes more models which were not included and potential published after the review. This work differs from that of St. Clair and Sinha as the analysis of the models specifically considers how researchers evaluated and validated model performance.

Model Description: Le Gat and Eisenbeis [37] introduced a model for estimating the survival of pipelines using maintenance records. The model used is a parametric Weibull Proportional Hazard Model (WPHM). The WPHM is able to account for left truncated and right censored data, which is typical of most maintenance records for utilities. Failure times and explanatory variables are used to produce survival curves for pipe cohorts based on material. The model is demonstrated on two utilities in France, which vary in size, number of recorded failures, and data available describing pipe characteristics and environment.

Training and Validation: For each utility, one year of data was reserved for validation comparisons, and the remaining data was used to train and parameterize the models. Parameter significance was determined using p-tests. Two methods were used to evaluate the model results. The predicted failures and the observed failures for the validation data were presented. The alternative validation metric is a rank order chart. To perform this validation routine, first the pipes are divided in quantiles with respect to number of predicted failures. The total observed and predicted failures are summed for each quantile and compared graphically.

For the larger utility, the model over predicts the total number of failures for all pipe materials. The percent difference from observed versus predicted failures ranges from 8% for the pipe cohort with the most observed failures to 46% to the pipe cohort with only 5 recorded failures. The model for the smaller utility with an older accident database significantly over predicts failures up to 2 to 2.5 times the observed failures. The authors argue that though the models over predict failures, the rank order quantile charts show that the models fairly accurately identify the highest risk pipe groups, and can be used for project prioritization.

Model Description: Park and Loganathan [38] introduce a threshold break rate equation for determining the economically optimal time to replace a pipeline with respect to the number of observed failures and the cost of repairs, replacement, interest, and inflation. In a companion paper [39], the authors demonstrate how to optimize pipeline replacement by equating failure prediction models to the break rate threshold calculations. The failure models considered are linear and exponential break rate models with a Weibull-based Rate of Occurrence of Failure (ROCOF) model. ROCOF curves are commonly used for repairable systems and are based on counting functions that track the cumulative number of failures.

The optimum time for replacement is solved by setting the failure models equal to the break rate threshold equations. Examples are presented for three data sets using the Weibull-based ROCOF model. A methodology is presented for determining optimal replacement without relying on maintenance record databases.

Training and Validation: In a similar paper, validation of the example problems are presented by comparing empirical data with fitted Weibull ROCOF curve [40]. The authors do not provide information on model training.

Model Description: Pelletier et al. [41] demonstrate a survival model for use with utilities with limited break histories, with respect to overall network age. The time to first failure is fit to a

Weibull distribution, and the time between failures is modeled with an exponential distribution. The model is demonstrated on three utilities in Canada. The model is used to forecast future failures and investigate the impact of increasing replacement rates on long-term network reliability.

Training and Validation: The observed failures were plotted with against a simulated failure curve, and the R-squared values were computed. The R-squared value presented was low, less than 0.4, which the authors contribute to model simplicity, lacking explanatory variables for pipe failure, and the overall randomness of pipe ageing aging and failure. The authors argue that the low R-squared value is not indicative of the utility of the model, which can reconstruct pipe histories and capture the pipeline ageing trend.

Model Description: Vanrenterghem-Raven et al. [42] investigated the risk factors for pipe degradation in New York City using a Cox Proportional Hazard Model (CPHM) and a Weibull Hazard Rate Model. The model includes discrete, continuous, and categorical variables. The importance of each parameter is first determined by training a CPHM with one parameter at a time. Next, a CPHM is trained with the determined significant parameters. The hazard ratio for each parameter is calculated to identify the importance and interdependencies of the parameters.

A WPHM with parameters identified as significant from the test above is used to predict long term network performance and investigate the impact of new repair and replacement strategies. The percentage of breaks avoided is reported with respect to varying replacement rates.

Some of the data sources were uncertain, and some materials were assumed given installation year. The impacts of the assumptions were not considered.

Training and Validation: The maintenance database spans 20 years. 17 years of data are used to train the WPHM, and the remaining 3 years are used to validate the model. Two validation metrics are presented. First the observed breaks are compared to the predicted breaks, and an overestimate percentage is computed. The lowest percent difference reported is 11 percent while the highest is over 120%.

Model Description: Rogers and Grigg [15] introduce a failure modeling schema that differs based on the number of recorded pipe breaks. A power law Non Homogenous Poisson Process (NHPP) is used to model future failures for pipes that have experienced three or more breaks. A separate model is developed for each pipe. An NHPP model will not converge with fewer than three recorded pipe breaks, as the scale parameter in the denominator of the function has a tendency to reduce to zero, so an alternative prioritization method is introduced.

For pipes with fewer than three breaks, a multi criteria decision analysis (MCDA) model is used. The MCDA model uses a weighted scoring system to rank and prioritize pipes for replacement. Risk factors considered in the MCDA analysis include break rates, age, diameters, bedding type, and pressure.

Two case studies of model implementation were presented for Colorado Springs, with a population of 400,000 and 1800 miles (2,900 km) of pipe and Laramie, CO, with a population of

30,000 and 19 miles (30 km) of pipe. Data quality issues including missing installation dates, lack of soil and pressure data, and duplicate pipe identifiers were discussed.

Few pipes had three or more recorded failures, less than 0.04% for Colorado Springs and 2% for Laramie, limiting the failure prediction models. Additionally, only the pipes having experienced failures underwent multi-criteria decision analysis, which represents only 1% of Colorado Springs' network and 8% of Laramie's. The utility of the model for prioritizing projects where failures have yet to occur was not demonstrated.

Training and Validation: Information is not provided regarding the reservation of data for validation purposes. For pipes experiencing three or more failures and modeled using a NHPP, the R-squared statistic for each pipe is computed. For Colorado Springs, R-squared values ranged from 0.68 to 0.89 with a mean of 0.75, and for Laramie R-squared values varied from 0.66 to 0.94 with a mean of 0.78.

The MCDA technique is subjective and can only be validated using expert opinion and historical records. For the Laramie model, which lacked soil and pressure information, over 20% of the pipes undergoing MCDA received the same risk score. Such groupings make it difficult for the decision maker to prioritize and rank pipes and projects.

Model Description: Wood and Lence [11] introduce a model for use with small and medium utilities. Data mining is used to subgroup the asset database with respect first to material and then to pipe installation year. Another set of subgroups were formed first with respect to material and then diameter. For subgroups with two or more failures, time-linear and time-exponential deterministic statistical models were developed to estimate the cumulative annual failures of the subgroup.

This modeling procedure was investigated using data from Laity View, a geographical area which accounts for approximately 13% of Maple Ridge, British Columbia. For most every subgroup, the time-linear model predicted failures more accurately than the time-exponential model, with the exception being cast iron pipes. The authors also concluded that additional information describing the pipe environment such as soil conditions is beneficial but not necessary to develop a useful model.

Training and Validation: Twenty years of data was available for Maple Ridge. The authors reserved the most recent five years of data for validation purposes. The model was calibrated with the remaining fifteen years. The percent difference for the observed versus predicted failures for the validation period was presented. For the time-linear models, R-squared statistics were calculated. R-squared values ranged from 0.63 to 0.94 with means around 0.8.

Model Description: Alvisi and Franchini [43] compare the prediction performances of a Weibull-Exponential-Exponential (WEE) models and a Weibull Proportional Hazard Model (WPHM) using data from a utility in Italy. Ferrara, the utility observed, has less data available than the utilities used to introduce the original models. The WEE model selected does not include explanatory variables. Because of limited utility data, the only parameters considered for the WPHM are pipe length, age, and diameter. Separate models are created by stratifying data by material and installation period. Though both models produced acceptable results, researcher observed that the performance of the WEE model is influenced by stratification. Also noted, the

WEE model, lacking in explanatory variables, is less useful when planning rehabilitation activities over an extended horizon, as it does not account for possible changes in pipe properties, such as replacing smaller diameter pipe with larger diameter. Since the WPHM model includes covariates, the model can account for such changes.

Training and Validation: No information is provided as to if data was reserved for training or validation. The validation technique used was a comparison of both the observed total number of breaks and total number of broken pipes to the mean predicted breaks and broken pipes. Models were judged as acceptable when the observed value fell within one standard deviation of the mean predicted value.

Model Description: Carrión et al. [44] present a survival model that is based on a modified semiparametric Cox proportional hazard function that is better suited for managing data that is left truncated and right censored. A modified extended Nelson-Aalen estimator is used to formulate the hazard function. The Nelson-Aalen estimator was originally introduced to handle right censored data, replacing a non-parametric maximum likelihood estimator. The Nelson-Aalen estimator was later extended to also manage left-truncated data.

A case study of the survival function estimated using the extended Nelson-Aalen estimator was presented for a Spanish Mediterranean city with 330,000 residents. The pipe records extend five years with over 1,400 failures documented. Over 93% of the data is right censored. The parameters required in the model are the calculated left truncation time, the failure or censor time, and censor status. Non-parametric survival models are generates for all pipes, and then later stratified based on material, diameter, length, and traffic above the pipes. To evaluate the impact of parameters on survival, a semiparametric Cox Proportional Hazard model was implemented. The authors concluded that for the data presented, longer pipes with larger diameters under sidewalks were least likely to fail.

Training and Validation: No information is provided as to if data was reserved for training or validation. Parameter significance was determined using p-tests. Also a several checks were conducted on the training parameters, including Cox-Snell residuals to assess if the parameters vary over time and deviance residuals to identify outlier data. Comparisons of predicted compared to observed failures were not presented.

Model Description: Debón et al. [14] compare the results Cox proportional hazard model (CPHM), Weibull accelerated lifetime model and Generalized linear models (GLM) estimated for a medium-sized Spanish city. Models were developed for pipes installed after 1940. Models were not stratified by material cohort, as material is a parameter in the models.

Comparisons are first made by analyzing the regressions coefficients for the three models. The significance of the material parameter varies across all three models. Parameter significant was determined using p-values. Polyethylene pipes are only found significant in the GLM.

Next, the true and false positive rates for the CPHM and GLM in order to compare the hazard rate predictions for the models. ROC curves for both models are graphed using these values. The area under the ROC curves (AUC) were calculated to estimate the expected overall performance of the models. The results of the analysis show that the GLM is the better model as it tends to generate more true positives than the CPHM, and has a larger AUC value.

Training and Validation: No information is provided as to if data was reserved for training or validation. The validation metrics presented are the ROC curves and AUCs described above.

Model Description: Kleiner and Rajani [45] introduce a Zero-Inflated Poisson (ZIP) process to model pipe failure, which accounts for the problems exhibited from using a non-homogenous Poisson process with typical utility data. Because pipes do not experience breaks every year, and a small percentage of pipes actually fail, the counting process for break rates includes many zeros, which cannot be expressed by a non-homogenous Poisson process. The zero inflated Poisson process results in an additional regression parameter used in the model. The model also includes time-dependent covariates including rain deficit and freezing indices.

The model was demonstrated using data from a utility in Western Canada. Only 150mm diameter cast iron pipes installed between 1956 to 1960 were considered in the model. The model was successful in predicting the total number of breaks and breaks per year, but was unsuccessful in predicting number of breaks per pipe. The model still can be used to statistically rank pipes for rehabilitation and estimate the impact of explanatory variables on pipe failure.

In a later paper, Kleiner and Rajani [46] compare the NHPP to three other models to predict failures in individual pipes. Two data mining models are considered and two regression models. The first data mining model considered is an ordered lists model in which lists are created ranking pipes based on covariates such as number of previous observed failures, pipe length, and accident scatter. Weights are assigned to each covariate as described by the lists. An aggregation function is used to find a composite score for each individual pipe. The pipes are then ranked by composite scoring and then compared to a list containing pipes with the highest number of breaks. Next the best set of covariate weights are determined using a genetic algorithm in order to maximize the number of hits, or pipes contained in both the composite score ranking list and highest number of pipe breaks list.

The next model considered is a Naive Bayesian Classification (NBC) Model that partitions data into classes and determines the probability of a pipeline being in a certain class given a set of covariates. The likelihood ratios (LR) for the pipes is computed and the pipes are ranked by LR values. The highest ranking pipes are compared to the list of pipes experiencing the most breaks, and a genetic algorithm is used to find a set of class limits that maximizes the number of pipes or hits contained in both lists.

The third model considered is a logistic regression model which determines is a pipe is contained in the list of high failing pipes based on sets of independent covariates. The set of covariate coefficients is determined by using the maximum log-likelihood method.

Training and Validation: The models were trained on 40 years of data and validated with the most recent 5 years of data. Several validation metrics were presented. First, the cumulative number of breaks observed and predicted for the training and validation periods were presented. Next, a pipe-dimension and time-dimension coefficients of determination were introduced and computed. The coefficients are similar to R-squared statistics, but the observed versus predicted data is aggregated by pipe and year. A pipe-dimension coefficient of 0.43 was presented and a time-dimension coefficient of 0.61 was reported.

Model Description: Park et al. [47] investigate time-dependent parameters that influence pipeline failure, using a proportional hazards model. Survival models were developed for 150 mm cast-iron pipes in a study area in the U.S. The pipe size was chosen because it represents a majority of pipes within the network. The pipe inventory was divided into survival time groups (STGs) based on the number of previously observed pipe failures. Survival models were estimated for each STG.

To investigate the impacts of parameters on pipe failure rates, several studies were performed prior to estimating the final survival models. First, the time-dependency of each covariate considered in the models was examined using a scoring process based on Schoenfeld residuals.

Next, non-parametric baseline hazard functions were estimated for each STG. When the log-log transformed values of the baseline survival function were linear, a Weibull parametric model was assumed. LOESS regression models were fitted to STG baseline hazard models in which failure times did not fit a Weibull distribution. The survival functions were estimated by multiplying the resultant baseline hazard functions with the exponential covariate functions.

The estimated survival times for each STG shows a decrease in failure time as a pipe segment undergoes multiple failures. This pattern is in line with the common bathtub shaped curve of infrastructure degradation, when at the tail end of an asset's life, the failure probability with respect to time increases exponentially. The authors ascertain that the model presented can be used by decision makers allocate funds for maintenance, repair and replacement by knowing the conditions that increase failure probability, and knowing the break number at which a pipe enters the tail end of the bathtub curve, and future breaks occur more frequently.

Training and Validation: No information is provided as to if data was reserved for training or validation. In addition to the checks performed on the covariates considered in the models, deviance residuals were calculated to investigate the difference between the observed failures and the expected failure times. Two of the six models had deviance residuals exceeding an acceptable level.

Model Description: Malm et al. [48] use Herz and Weibull based survival models to investigate replacement rates for water pipes in a large utility in Sweden. The authors examine training survival models using pipe data for the entire age of the network, over 100 years, and training the models using replacement data, which only spans 14 years. An equation is presented that relates the replacement rate for future decades to the survival function values at a specific time. For the first scenario, residual pipe length, which is the percentage of the original pipe length for a certain decade that remains at a later time, was the input data for the models. For the second scenario, where more specific replacement data is known, the replacement rate for the decades prior to 1990 is extrapolated from the data spanning from 1991-2005. Survival curves are calculated using the extrapolated data.

The survival functions based on Alternative 1 tend to result in higher replacement rates for both failure models. The survival functions for Alternative 2, based on the extrapolated 15 year rehabilitation data, showed lower replacement rates. Alternative 2 assumes pipe replacement is due to degrading condition and not due to city planning activities that occurred as the network developed. The survival curves fit Alternative 2 data better than Alternative 1 data based on R-squared value, showing that survival curves fit data best when it is conditioned on condition based rehabilitation that does not account for city planning/development based rehabilitation decisions .

Training and Validation: No information is provided as to if data was reserved for training or validation. The validation metric used was computed the R-squared value for the survival curve plotted on a histogram chart of installation decade versus residual pipe length.

Model Description: Martins et al. [49] compare three stochastic models for predicting water pipeline failure. The models considered are a single-variate Poisson process, a Weibull accelerated lifetime model (WALM) and a Linear extracted yule process (LEYP). The prediction performance of each model was compared using data from a Portuguese water utility. The Poisson process model presented predicts the failure rate for all pipes within a categorical grouping based on material, diameter, and age. The failure rate is the number of failures divided by the sum of the product of pipe length and observed failure time. After predicting the length-dependent failure rate, maximum likelihood estimation is used to compute the expected number of failures for individual pipes within a time period.

The WALM is based on the WPHM presented by Le Gat and Eisenbeis [37], with some added improvements. First, a variable is introduced that accounts for the time between the start of the observation window or asset management database, and the last recorded failure. The survival function is altered to include this variable so that the time to failure follows the appropriate distribution.

The next improvement is considered is how to address time dependent variables such as pipe age. Instead of using age as a covariate, which the author suggests will not impact the distribution of failure time when a pipe does not fail, segmenting the data into pipe age groups is investigated. The results of implementing this method did not result in significantly better or worse prediction performance.

The final improvement recommended is related to generated failure time distributions for subsequent failures. A binary covariate is used to address if a pipe has failed in the past or not. This method replaces using the number of previous failures or log of number of previous failures as a continuous variable. The author suggests that the recommended parameter will help insure that Monte Carlo simulations will not enter an infinite loop, as often the time to next failure decreases exponentially.

The LEYP model was originally introduced by Le Gat [50] is based on a pure birth Yule process, but adapted to better process failure data. The basic yule process assumes a Markov property, that there are at most one failure at a time, and the distribution of failures follows some geometric process. The LEYP builds upon the Yule process and is a special case of a Non Homogeneous Birthing Process (NHBP). The LEYP assumes the intensity function can vary with time, and the distribution of failures is a continuous extended Negative Binomial. Lastly, the intensity function can built on the Weibull power law, so that the function is based on pipe covariates. The parameters of the LEYP are estimated using the maximum log-likelihood method. The authors simplify the likelihood function to prevent computational problems associated with functions exceeded machine precision for high values of failure time.

The models were all tested using data from a Portuguese utility with over 360 km of pipe, primarily asbestos cement, high-density polyethylene (HDPE) and polyvinyl chloride (PVC). The material breakdown is worth noting considering HDPE and PVC pipes do not follow standard deterioration/corrosion processes like cast iron or ductile iron pipes. The failure rate for these materials is based less on simple decomposition rate and more so on random processes that contribute to failure such as poor construction, increased loadings, poor material quality, and pressure surges.

Using multiple comparisons described below, the WALM performed better than both the LEYP and the single-variate poisson process. The authors presents how to use the predictions from the three models to examine replacement rate scenarios for the network that minimize long-term costs of ownership due to deferred maintenance.

Training and Validation: Both temporal and random division methods of selecting training data were used. For the temporal method, the most recent three years of data are reserved for model validation. For the random division method, 50% of pipes are selected at random to be used for training, and the remaining 50% is used for validation.

Several comparisons are performed to validate the model. First, the rank order quantile method introduced by Le Gat and Eisenbeis [37] is used to compare observed and predicted failures for classes of pipe. Using this validation metric, the WALM performed better than the poisson process, while the LEYP over predicted failures in the highest risk quantile. This over estimation was only observed when temporal division rather than random division was used to select training points. Yet, the overall prediction accuracy of the WALM and LEYP are reduced when random division is used.

Next, a mean absolute error term was computed which normalizes the difference between predicted and observed failures based on the number of pipes evaluate in the material group. The WALM had the least error of the three models presented. Using the random division method, the absolute error term for all models increased.

In the related dissertation, ROC curves are also presented showing that the WALM out performs the other models. Also, charts showing the mean expected failures for the test sample of pipes which was divided into groups based on number of observed failures are presented. Using this validation metric, the LEYP performed better than the other models by predicting the most failures for pipes experiencing four or more failures. The LEYP model along with the other two models over predicted failures for pipes not experiencing failures. This was observed for both division methods for training data.

Model Description: Toumbou et al. [51] build upon the work of Pelletier et al. [41] and introduce a Weibull-Exponential-Exponential (WEE) pipeline failure model. The time from first failure to second failure is modeled using a Weibull distribution. The time from the second to third break is modeled using an exponential distribution, and the time to subsequent breaks is modeled using another exponential distribution. Unlike Pelletier et al. [41], the model presented includes pipe diameter and length as parameters in the survival model.

The model is demonstrated on a small city with less than 115 miles (185 km) of pipe. Three modeling scenarios are considered. First, the model is calibrated with covariates describing pipe diameter and length. Next, the model is calibrated without parameters. Lastly, pipes are group by diameter size, and models are calibrated for each sub group.

Analysis of the failure forecasts for all three models indicates that the inclusion of covariates has little impact on failure prediction for the duration of the training data. With respect to long-term forecasts, the model without covariates predicts more failures than the models based on diameter grouping and the model with covariates.

Training and Validation: No information is provided as to if data was reserved for training or validation. The validation metric presented is a chart showing the estimated failure curve and the observed cumulative number of annual failures. An R-squared statistic is not provided.

2.5.3 Spatial Models

Only in the most recent years have researchers started to focus utilizing the spatial distribution and clustering of failures as decision tools to pipe MR&R prioritization. These models are quicker and cheaper to implement, easy to interpret, and provide value for top level assessment of network risk. The following summarizes spatial-based models with a focus on clustering methods.

Oliveria et al. [52] investigate the density of breaks in water distribution networks. A spatial search algorithm is used to find subsets of the network where the break rate is higher. This comparison is made by computing a likelihood ratio, with the null hypothesis being that the break rate is homogenous across the network. A novel algorithm that relies on the *k*th nearest neighbor to define a rectangular search window, S, based on the point locations of breaks is used to detect potential areas where the break rate is elevated. The null hypothesis requires the break rate for the entire network to be computed. The underlying failure probability model is investigated, and null hypotheses are computed using a homogenous poisson process, a NHPP adjusted for age, and an NHPP adjusted for pipe diameter, since both age and diameter can influence break rate. The alternative hypothesis was computed by calculating separate break rates for all pipes within the search window and pipes outside of the search window. The new search algorithm for clustering was successful in identifying clusters of pipe failures. A brief section is included in which explanations for the increased break rates are investigated.

Oliveria et al.(b)[53] introduce an algorithm for determining clusters of pipeline failures within water distribution networks. The model is an extension of the very common DBSCAN algorithm [54], which is a clustering algorithm that relies on an input of a minimum number of points to define a cluster. Similar to the OPTICS [55] clustering algorithm, a second parameter describing the minimum distance between two failures is used to also define clusters. Using these two parameters, core points within potential clusters are identified. The OPTICS algorithm is improved in the method of selection of the next point to re-start the algorithm. Instead of restarting the cluster search algorithm at a random point, the algorithm is re-started at the closest point to the last analyzed point in order to reduce uncertainties and potential bias associated with the starting point of the cluster search.

The cluster search algorithm summarized above was demonstrated on a sample network, and the impact of explanatory variables on clusters is investigated. The statistics for pipe diameters and materials for clusters are compared. Additionally, the break rate for the clusters is compared to a global break rate. Finally, a variable called betweeness was computed for pipe clusters and groups based on explanatory variables. A high betweeness score indicates that when a pipe is disturbed, extra loads will placed on other pipe segments in close proximity to the pipe. The authors indicate that betweeness can be used to prioritize replacement. Also to help with prioritization, the break rate for clusters can be fit to stochastic models to predict future break rates that follow the spatial and temporal pattern of breaks. Additionally, the authors suggest that multivariate analysis can be used to determine if spatial clustering or high break rate can be used as a surrogate for explanatory variables used in models.

Bogárdi et al. [56] assigns network level failure probabilities using a space-time homogenous Poisson process. The ROCOF or break rate is calculated for the entire network. Next a grid is overlaid on the network, and the Poisson intensity is calculated and reported as the number of failures per cell. After verifying that the spatial distribution of breaks is homogenous, a space-time NHPP probabilistic model is used to generate potential failure patterns. First, using the computed ROCOF, a number of expected failures represented as points are uniformly distributed across the network. Next, using the spatial Poisson process model and intensity function, a random radius between the failure point and its nearest neighbor is generated. A failure location is recorded when this radius crosses a pipe. This process is repeated for all points in the network and is performed multiple times to generate possible failure patterns. The simulations are then used to compute expected repair costs over a time horizon.

The model was demonstrated on a subset of a German water network. The pipe grouping was homogenous, with similar pipe material, age, and diameter. As a result, adjustments for such explanatory variables were not made. Also, no justification is provided for the selection of the grid cell size. Checks were made however to compare the frequency at double the grid cell size to verify homogeneity.

In order to test for spatial homogeneity, several tests were performed. First, the failures per cell versus the frequency distribution and calculated Poisson probabilities are graphically compared. Next the failure distance distribution versus the relative frequency of failures and the calculated probability of failures are compared. Lastly the training set is segmented temporally, with three years of data being used for calibration and three for validation. The failures per cell versus relative frequency are compared.

Christodoulou et al. [57] expand on the DBSCAN algorithm to detect failure clusters by adding a parameter describing the time window before and after a break to account for the temporal variations associated with failure clustering. A failure point is considered to be in a cluster if it is within a threshold distance of a core point and the failure occurred within the time window specified with respect to the failure time of the core point.

Spatial-temporal clusters are computed for the network of Limassol in Cyprus, Greece. In a related work, the break rates for district metered areas in this utility were computed using a poisson model. The break rates differed across district meter areas. An extension of this work would be to calculate break rate for clusters, for use with rehabilitation planning.

Shi et al. [58] investigate network level break rate and break rates within clusters for the water distribution network in Hong Kong. First a grid was assigned to the network and the break rate for each cell was computed. Moren's I statistic is used to assess if the failure data is spatially clustered. The failure rates were assigned to the centroid points of the cells, and the Moren's I statistic for the break rate points was computed.

After determining cluster, regression models were developed for pipe failures with respect to diameter and age. Variation in failure rates due to pipe materials and temperature at time of failure was investigated.

The data set used in this study considers 80,000 failures, and is significantly larger than the data sets in the other models reviewed. The conclusions from the analysis show that both failures and failure rates calculated using the grid method are clustered. Pipe age is correlated to higher break rates, and break rates decrease exponentially as pipe diameter increases.

Information regarding the segmentation of data for calibration and validation purposes was not provided. Statistical metrics were used to determine clustering. R-squared values were used to address the goodness of fit of the regression models.

2.5.4 Machine Learning Models

Machine based models that are not constrained to a pre-determined model form could be a viable alternative for developing valid predictive condition assessment models. The machine learning based models developed for water pipes can be categorized as neural network based, fuzzy logic, polynomial regression, and Bayesian. A brief discussion of these model types and examples are included.

2.5.4.1 Neural Network

Though several artificial neural network models have been presented over the past decade [59]–[63] they can be difficult to implement on a broad scale. Several of these models predict condition rating rather than break rate or failure time. There is no standardized pipe condition rating model for water pipes; therefore, the models are difficult to compare and validate. Additionally, ANN models are criticized as being black-box and lack transparency in validation metrics. Lastly, ANN models require numerous parameters, which as previously mentioned, are often not available.

2.5.4.2 Fuzzy Logic

Fuzzy logic models provide a method for incorporating data uncertainties and expert knowledge into models prediction the condition state or break rate of assets. For uncertain pipe parameters, values are modeled as fuzzy membership functions. Typically, the development of these models requires several interviews or workshops with utility personnel to gather estimates of pipe data and performance to develop fuzzy rules that relate the parameters to pipe condition. Though several other fuzzy logic models have been introduced [45], [64]–[68], the most extensive and best tested model was developed by St. Clair and Sinha [69]. The authors developed a pipe condition rating prediction model that considers 27 pipe attributes. Laboratory experiments are used to validate the model results. The overall results of the model are very promising, yet the processes involved to develop and test it are extensive. This type of model might be best used to further classify pipes designated as having high failure likelihoods from statistical model results.

2.5.4.3 Polynomial Regression

Berardi et al. [13] introduce an Evolutionary Polynomial Regression (EPR) model for predicting the break rate of water distribution pipes. The EPR model uses genetic algorithms to find a model form, and the least squares approximation method to determine model parameter coefficients. The focus of the research is to develop statistically significant and parsimonious models that consider a limited amount of explanatory variables. The parameters considered in this work include pipe age, diameter, length, previous failures, and the number of properties served by the pipe under investigation.

The initial calibration results for the EPR model is a Pareto set of burst prediction models that can be used to evaluate the relationships between parsimony and model fit to the observed data. Model fit is evaluated using the Coefficient of Determination (CoD) metric that compares the predicted and observed data and also considers the sum of squared errors for the entire data set. As no information was provided as to a hold out sample for validation, it is assumed that the CoD was calculated for the entire data set.

Xu et al. [70] compare machine learning methods to estimate break rate prediction models using EPR and genetic algorithm based models. The pipe characteristics considered include pipe age, length, and diameter. The EPR and genetic algorithm models performed similarly, with reasonably good results for the training set, and poor results for the validation data set. The CoD metric was used to evaluate model performance. The CoD for the validation set for both the EPR and GA models was low, around 0.34. The underestimation rate was also high. The authors suggest that better leak detection efforts and maintenance activities in the region have improved system performance and decreased break rates in the most recent years.

Wang et. al [16] investigate using advanced regression models to estimate annual brake rate for individual pipes. The parameters considered include material, diameter, length, age, and break rate. Though the regression models fit the training data well, with R-squared values ranging from 0.7 to 0.8, they did not perform well at predicting failures on individual assets for the validation data. Misclassification rates for failed pipes ranged from 50% to 61%.

2.5.4.4 Bayesian

Watson et al. [71] introduced a Bayesian updating model to estimate the annual break rate for pipes. Heuristic knowledge was used to estimate the Bayesian prior. Markov Chain Monte Carlo analysis was used to update the posterior distribution. The model is tested on only two pipes and compared to a base Poisson distribution for break rate. The Bayesian model performed better than the base model.

Wang et al. [72] developed a Bayesian model for predicting the deterioration condition of pipes. This model includes an error term to account for uncertainties. Though the parameters do not include length, they do require information that is difficult to obtain including pipe coatings and trench depth. Overall, the model performs reasonably well. Results were compared to observations for twenty pipes, and R-squared values averaged around 0.7.

Expanding the use of Bayesian prediction models, Francis et al. [73] examined the use of Bayesian belief networks for predicting water main breaks. Interestingly, the parameters included in the data set do not include pipeline material and installation characteristics. The model performs poorly at predicting breaks. This concept is intriguing, and the model could possibly be improved by including material and operational properties of pipelines.

Li et al. [74] use Bayesian nonparametric learning methods to predict water pipe condition. Specifically, a hierarchical beta process (HBP) is used to predict pipe failures. The researchers demonstrate how models can be formulated using sparse data. The HBP model is compared to non-parametric Weibull and Cox survival models using data from a large utility and a suburban utility. Though the HBP does perform better than nonparametric Weibull and Cox models, the performance for all of the models is poor. The performance metric is an ROC curve of pipe length inspected versus failures detected. The AUC value for these curves is less than 0.61, which means there is little difference in sensitivity and specificity.

Scholten et al. [75] demonstrate how to use Bayesian methods to incorporate expert knowledge of systems with sparse data to improve service life modeling. The researchers introduce an expert elicitation methodology to estimate a prior distribution for a non-parametric Weibull-based survival model. The information from multiple experts is gathered to form an aggregated prior. The results of this work are promising, and could lead to a solution for developing informed pipe rehabilitation strategies for utilities with sparse data.

In subsequent work which examines creating rehabilitation strategies for small utilities with limited recorded pipe failures [76], the researchers note that the expert elicitation based aggregated prior methodology is significantly more complex than standard maximum likelihood estimation (MLE). Alternatively, the prior distribution for the small utility is estimated using data from three larger utilities in the same country. Parameter estimates for a Weibull-Exponential survival model for the three utilities are obtained using MLE. These estimates are combined to form a prior distribution. Markov Chain Monte Carlo (MCMC) sampling of the previously described prior, the conditional likelihood, and failure observations from the small utility is used to obtain the posterior distribution. The updated model is used to develop a rehabilitation framework, yet validation metrics of pipe break prediction performance for the small utility are not included.

2.7 Optimization of MR&R Activities

Fares and Zayed [48] developed a fuzzy logic model to predict the risk of pipe failure, which considers consequence of failure. The model results are used classify pipes according to the level of risk of failure. The authors then urge decision makers to map the pipes and use GIS to segment the network into regions where repair actions will be performed. Pipes with less failure consequences might be included in the replacement programs due to connectivity considerations and opportunity. This is the only model found in this review that makes references to the spatial relationship of pipe breaks, which is an important consideration in the planning process.

Alvisi and Franchini [77] demonstrate the use of a multi-objective genetic algorithm to minimize water losses and breakage repair costs. The failure model used a non-parametric WPHM. No validation metrics are presented. The authors integrate hydraulic modeling in the analysis to optimization routine. A case study is presented on a small utility in northeast Italy. The authors use post-processing to identify zones where the majority of leak detection or replacement activities should be performed.

Kleiner and Rajini [78] present a Markov-transition based framework for prioritizing inspection and replacement activities of large water pipes. The framework considers the total cost of a pipe including failures, inspection activities, and replacement. A proof of concept application was presented, but the researchers state that more work is needed to make the tool widely useable for utilities.

The genetic programing based optimization model developed by Xu et al. [79] uses a data mining prediction model to determine the optimal time to replace an asset with respect to minimizing the annual cost of pipe replacement activities and break repairs. A case study is presented using a large subset of data from a very large utility and studies are performed examining the impact of discount factors and break rate prediction models on the optimization function. Pareto front charts, run time, and pipes identified for replacement were not provided.

Dandy and Englehart [23] also use a genetic algorithm technique to prioritize pipe replacement with respect to minimize the cost of breaks and costs of replacement. A polynomial regression model is used to estimate failure potential. The impacts of breaks are evaluated considering societal costs and hydraulic impact. Repair costs multipliers based on failure location are used to estimate indirect failure costs. With respect to hydraulic impact, velocity and pressure is calculated at each node after the GA model run. This allows the decision maker to include upsizing of the pipe. Pressure and velocity constraints are applied to the GA through the use of penalty costs. The framework is demonstrated on a subset of a large utility, consisting of less than 500 pipes. Analysis of the results includes an evaluation of the impact of changing pipe size. The authors include information on run time and cost savings associated with replacement scenarios.

Giustolisi and Berardi [80] demonstrate a sorting based multi-objective optimization algorithm to prioritize pipe replace. In a previous work, [81], the authors demonstrate the use of a multi-objective genetic algorithm (OPTIMOGA) to identify pipes for replacement and/or upsizing that minimize the financial impacts of pipe failures. This study improves upon the OPTIMOGA model by demonstrating how to prioritize pipes post OPTIMOGA simulations. The authors note issues with OPTIMOGA algorithm including the lack of reproducibility of results, i.e. multiple runs do not identify the same pipes for replacement. This suggest that more optimal solutions can be found.

Post-processing is used to further refine replacement selection and scheduling decisions. Pipes are prioritized for replacement based on the number of times they appear in a solution on the

Pareto front. This methodology was evaluated on a subset of a network consisting of less than 2,000 pipes. The authors provide model run-time for OPTIMOGA routines. The post-processing sorting prioritization resulted in more optimal solutions than the OPTIMOGA routines.

2.8 Utility Practice

Given the models presented, several utility surveys have been performed to investigate current practices and adoption of pipeline failure models. Both surveys by Matthews et al. [9] and St. Clair and Sinha [10] show that utilities are not taking advantage of the new statistical models for pipeline failure being introduced. While some utilities have adopted long-term economic forecast models, less than half of the utilities surveyed are using statistical models for short-term investment planning. The failure prediction models utilized most utilized by utilities are Weibull based or LEYP.

Though pipeline breaks are a commonly cited reason for pipe replacement, other common reasons are related to hydraulic concerns including low flow and the need to change pipe size [6]. Many of the utilities surveyed also use hydraulic models to evaluate network performance and plan for water line extensions and upgrades. An all-encompassing decision support system for MR&R improvements would include at minimum hydraulic criticality analysis results from a hydraulic model. Ideally, hydraulic upgrades could be coordinated with failure mitigation activities.

2.9 Discussion

Though many models have been presented in literature, the utility adoption rate for the models is low. The primary reasons for the poor adoption rates can be contributed to the amount of data needed, the data preparation required, the computational effort required, and the lack of perceived confidence in the models. Clarity in model training and validation metrics is needed to convey confidence in model predictive performance and suitability for using in scheduling replacement or rehabilitation projects.

With respect to model parameters, one of the most commonly used, yet difficult to define parameters is material length. Used in the majority of the statistical and machine learning models, the definition of pipe length is vague. Though seemingly trivial, this definition is actually quite important. The assumption would be that pipe length reflects the joint to joint or valve to valve length of pipe currently in the ground. The problem associated with this assumption is that GIS models were not developed considering the in-situ length of pipes. Often, the pipe length is the result of digitization methods in which the cartographer lifts his/her pen or mouse and completes a line segment. The problems associated with variable pipe length definitions have been reported by [17], [18], who propose extensive pre or post GIS processing to try to aggregate pipes to city block level. These processing efforts are tedious and have not been studied in detail as to how they impact model performance and accuracy.

Another one of the most significant and commonly used parameters is a parameter describing the evidence of previous breaks. The problem associated with this parameter is rooted in the definition of pipe length. For digitized pipes with varying pipe lengths, the parameter may not be unbiased in its representation. Longer pipes will be more likely to experience more failures. When databases can range with pipes lengths modeled from 1 to 2,000 feet, this parameter can be biased. Though important, a more appropriate, non-biased parameter to consider is the spatial clustering of pipe breaks.

Only recently have researchers applied spatial clustering techniques have been implemented to evaluated water distribution system performance. For binning type clustering applications, determining the appropriate size of the bin can be problematic. The problem with these models is that few incorporate statistical methods to predict future breaks. The frameworks reviewed that do incorporate predictions with the spatial analysis do not consider training and validation of the statistical models.

As an alternative, the clustering of breaks can be considered as a parameter in statistical models. Only one model framework reviewed considers clustering of breaks [46]. Though clustering was not determined to be a significant parameter, this could be due to the clustering algorithm utilized. There is a need to examine alternative methods for spatially identifying areas of increased break rate and incorporating those observations into a condition assessment model.

Regarding model form, utility directors and asset managers still want to have ownership and input on prioritization. Some of the models presented would require significant training to understand the background of the model and computations required to train the model. Such models would have to be calibrated by very specialized consultants. Furthermore, the current state of the literature does not show dramatic improvements in condition prediction using machine learning models.

With respect to implementation, few case studies have been presented using data from medium and small utilities with fewer than 100 failures in material classes. Researchers instead have made suggestions rooted in model transfer theory to train both statistical and machine learning based models by utilizing data from larger utilities with more extensive databases. Only some of these suggested have been thoroughly tested using common validation metrics. More research is needed to examine the suitability and effectiveness of these methods.

Finally, though extensive work has been performed on the modeling side, the amount of research detailing how to incorporate models into decision making is scarce. Only several optimization frameworks have been presented, and even fewer have been demonstrated on entire networks, which is extremely important when examining scalability of models. The majority of these frameworks use genetic algorithms to minimize or maximize a penalty or benefit function.

Genetic algorithms, (GA's) are used to solve multi-objective, non-linear optimization problems by defining the problem using a fitness functions. Solutions of the fitness or objective function yield a set of Pareto-optimal solutions. Solutions that result in improvements of the fitness function are found by modifying a genetic representation, often called a chromosome, of the solution space. Some chromosomes improve the objective function, and these chromosomes are allowed to reproduce or mutate in order to search for even better solutions. This process is repeated until a convergence criterion is met, which is often simply the number of simulations or chromosomes created, but could also be related to some improvement in the fitness function.

Though GAs can be power decision making and design tools when considering multiple objective problems, they do have several documented limitations. The major limitation, noted by Deb et al. [82] is related to the computational complexity of the algorithm. The expense of the algorithm is directly related to the number of objectives and the number of decision variables. Increasing the number of decision variables can result in an exponential increase in the computation time for each iteration.

The other common limitation associated with the GA is the problem of premature convergence. When a search space is narrowed rapidly by a chromosome that causes a marked improvement in the fitness function, this chromosome can reproduce rapidly, narrowing the decision space [83]. This causes the GA to converge at a local minimum/maximum rather than the best solution or global minimum/maximum.

These issues related to computational effort and premature convergence have not been fully explored in the literature. With respect to computational effort, the largest case study presented in literature consisted of 2,000 pipe segments, while Utility A, considered in this paper, consists of over 20,000 segments. An order of magnitude increase in the pipes results in an order of magnitude increase in the decision variables which has an exponential impact on the search space. Increasing the search space increases model run time to reach convergence.

On the topic of convergence, the stopping criteria for the optimization routines presented in literature is subject to scrutiny, especially given the lack of replicability of results. Though establishing a threshold number of simulations is a valid stopping criterion [84], better stopping methods exist, including stopping when a threshold change is no longer observed in the objective function value. Though potentially increasing the run time, such a stopping criterion could result in more optimal solutions.

Finally, lacking in the replacement prioritization models is a focus on how to utilize optimization results to develop capital projects. To model real world decision making, an emphasis needs to be placed on the spatial relationship and connectivity of pipes identified for replacement. This spatial relationship can incorporated into the optimization routine through spatial binning of the distribution network, decreasing the amount of effort required by the decision makers to spatially evaluate optimal pipe replacement strategies.

2.10 Contribution to Literature

The proposed research will contribute to the body of knowledge in several ways. First, the pipeline replacement prioritization model presented in this paper will attempt to account for uncertainties in input parameters. Of note is that pipe length will not be included as a model parameter. The researchers argue that the parameter does not help describe physical pipe properties that might contribute to pipe failure, as digitization efforts result in pipe lengths that do not represent joint to joint lengths in the field. Additionally, this research investigates the contribution of incorporating the spatial distribution of breaks into statistical models. The researcher hypothesizes that the break rate distribution variable serves as a surrogate for a multitude of other model parameters, which could result in more parsimonious models.

Secondly, an optimization routine that utilizes the WHRM results to prioritize pipe replacement and minimize the consequences of failure while considering the spatial dependency of pipe replacement projects is demonstrated. This routine is used to investigate how the proposed model improvements impact decision making by comparing prediction results to a base model which includes the commonly used and available WHRM model parameters.

Though many papers demonstrate the use of models for prioritizing pipe replacement and maintenance activities, few case studies are available showing researchers and practitioners how to incorporate these models into risk-based planning frameworks. This paper presents a case study of utilizing the validated WHRM to minimize the consequences of failures with specific considerations to hydraulic reliability and the spatial relationship of pipe MR&R activities. Top down and bottom-up approaches to asset management within the utility are described. An investigation of using binning methods to prioritize pipe replacement in specific regions is performed. Binning methods are commonly used in GIS environments as a pre-processing technique to display and visualize density for large data sets by replacing multiple observations with a single value in a bin [85]. The most common form of binning, the grid method, is utilized in this study. Additional research is performed to address convergence and scalability of models for large data sets.

Another major contribution to the literature is the investigation of using WHRMs to prioritize pipe replacement for medium and small utilities, which has not been studied in detail. The suitability of model transfer techniques is also investigated. The investigations are performed using data from one large and two medium sized utilities in the Southeast U.S. A model transfer technique widely used in transportation problems, yet never before applied to the transfer of WHRM parameters for pipe failure prediction is evaluated, along with the other model transfer techniques suggested in the literature.

Finally, a risk-based framework for prioritizing MR&R activities for small and medium utilities is presented. The methodology is rooted in cluster analysis using popular cluster algorithm, DBSCAN. Hydraulic information and heuristic knowledge is also incorporated. This framework is demonstrated in a case study with data from a medium utility in the Southeast US.

CHAPTER III – UTILITY INFORMATION

3.1 Introduction

Maintenance records and GIS models were collected from three neighboring utilities in middle TN. Shown in FIGURE 3.1, Utility A is much larger than the other suburban utilities. The work presented in the remainder of this dissertation is novel in that no case studies or example problems have been presented using a consortium of data from neighboring utilities of different sizes. This regional focus reduces some potential variations in data such as environmental differences, material sourcing, and construction practices, allowing for a less biased comparison of model performance for the varying utilities.

The following sections describe in detail the data provided by the three utilities. First a brief history of the utility is provided. The frequency of pipe materials and sizes are reported in Figure 3.2 and Figure 3.3. Details regarding data storage, and the processing steps required to identify failures is described in section.

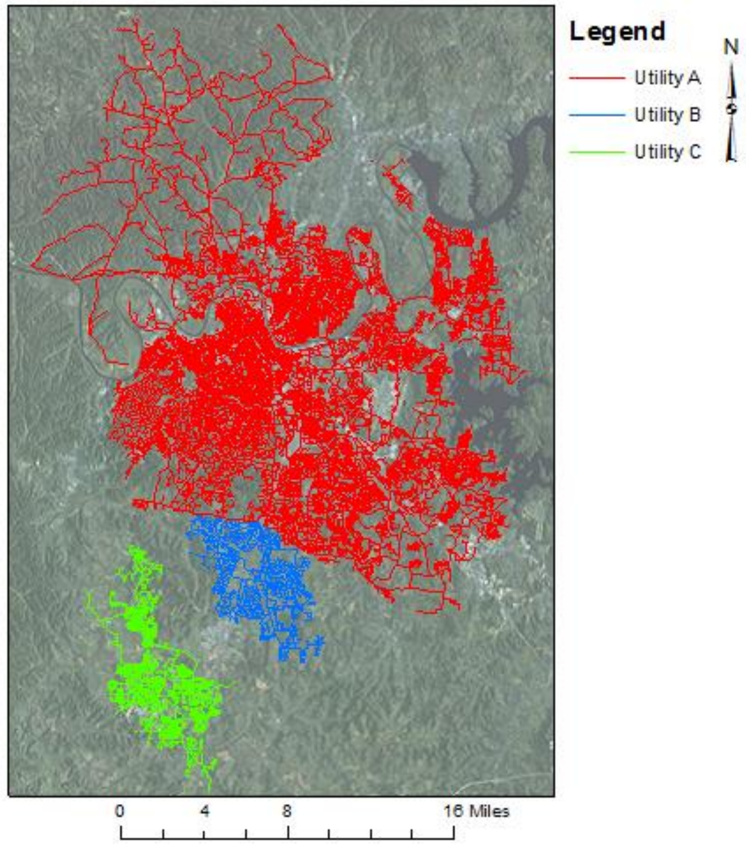


FIGURE 3.1: REGIONAL UTILITY MAP

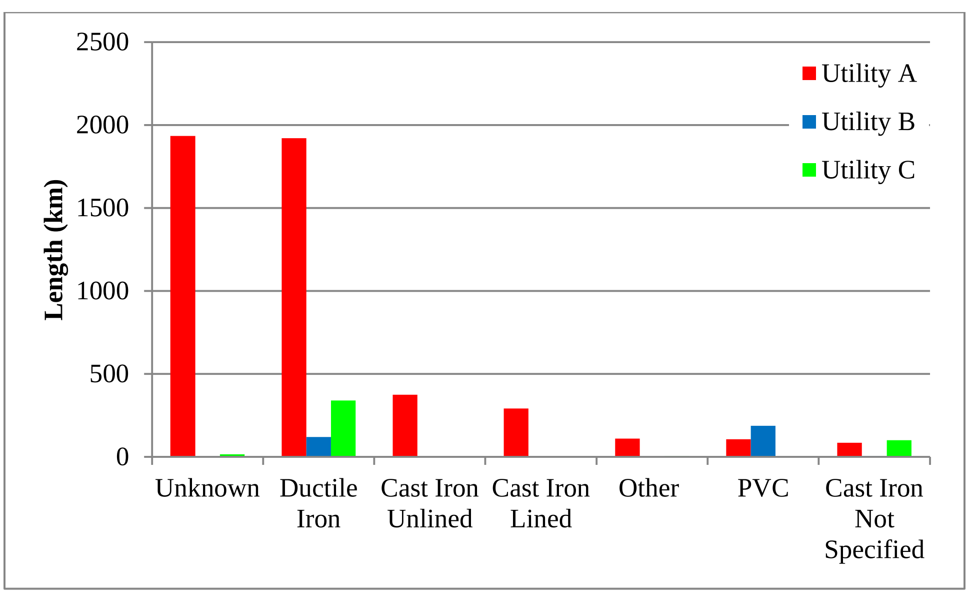


FIGURE 3.2: MATERIALS BY LENGTH

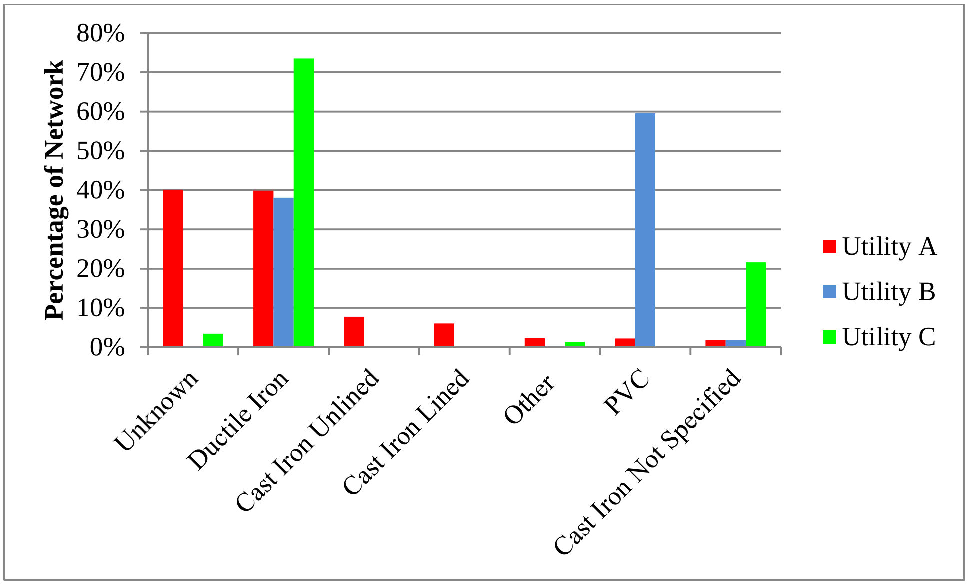


FIGURE 3.3: NETWORK MATERIAL COMPOSITION

3.2 Utility A

Utility A is a large utility providing water services to over 600,000 residents. Originally established in the 1800's, the city utility was annexed with the surrounding county utilities in the 1960's. Thus, the utility inherited assets not originally owned and operated by the city.

The utility is primarily composed of cast iron and ductile iron pipes. The utility maintains a computerized maintenance management system (CMMS) software, which is used to track work orders, which interfaces with a Geographic Information Systems (GIS) model. The CMMS tracks work orders with respect to GIS pipeline segment ID. The work orders related to breaks were extracted from the CMMS database, and the failure times with respect to the start of the asset management database were computed.

Data quality concerns for Utility A are primarily related to unknown material and installation date. Approximately 40% of the materials and installation dates are unknown. Using expert knowledge of the network, missing materials and install dates were imputed, and binary categorical parameters were introduced to account for the assumed material and installation dates. This process is described in further detail in Chapter 4.

3.3 Utility B

Utility B is within a suburb of Utility A, and was founded in the 1960's. The utility's age is represented in the material makeup of the network, which is primarily PVC and DI pipe. Utility B maintains a spreadsheet of failures that includes details including failure time, pipe material, diameter, and the GPS location of breaks. Independent from the maintenance database is a GIS model that describes pipe material, diameter, and installation date. The material fields were well populated, but in some cases the installation date for pipes was estimated by utility personnel. For these cases, an assumption covariate was assigned.

To identify failed pipe segments, buffers were placed around the GPS location of failure points. The buffers were intersected and joined to the closest pipe segment. In some cases, the pipe material listed in the failure record did not match any neighboring pipes in the GIS model. It was assumed in these instances that the failure record accurately described the pipe material, and the material was updated in the GIS model. The assumptions were noted, but the assumed material parameter was not utilized for these instances. The inclusion of the parameter could introduce potential bias into the model, as only pipes that failed would have a positive value for this parameter.

3.4 Utility C

Utility C is also within the suburban area of Utility A, yet is much older than Utility B. Within the past several decades, urban sprawl has resulted in large growth in the area. The network now has a mix of very old cast iron pipe and much newer DI pipe.

Similar to Utility B, work orders related to pipe breaks are stored in a spreadsheet, and a GIS model is available which describes pipe locations and attributes. Using a process similar to Utility A, expert elicitation was used to fill in data gaps related to missing materials and installation dates, and categorical variables were used to account for these assumptions.

Utility C has the most uncertainty with respect to failure locations and information describing pipe failures. The failure records provided addresses or intersection locations, and many records did not describe the pipe material. To identify failed pipes, the failure locations were geocoded and a shapefile of failure points was created. Using the same process described for Utility B, a buffer was created around the failure points, and the closest pipe segment to the failure point was identified.

3.5 Conclusions

Though the utilities presented are geographically similar, they are still very different with respect to size, material composition, and data uncertainty. Additional differences are evident in utility performance. Table 3.2 shows failure statistics with respect to annual break rate.

Table 3.2: Utility Break Rates

Utility	**A**	**B**	**C**
Pipe Length (mi)	2,996	208	506
Area (sq. mi)	526	34.7	30.1
DI Break Rate (brks/1000ft/yr)	0.008	0.006	0.007
CU Break Rate (brks/1000ft/yr)	0.021	--	0.03
Database Duration (yrs)	10.5	8.5	7.25

These differences should be kept in consideration as model prediction performance comparisons are made. This research highlights the importance of testing models on utilities of varying size and composition.

CHAPTER IV– COMPARISON OF WATER PIPELINE FAILURE PREDICTION MODELS FOR NETWORKS WITH UNCERTAIN AND LIMITED DATA

4.1 Introduction

Though over 50 stochastic and probabilistic models for pipeline failure prediction have been presented over the past twenty years, most utilities are not implementing these technologies for rehabilitation and replacement planning. These models, reviewed by Rajani and Kleiner [36], [86], St. Clair and Sinha [10], and Yamijala et al. [87] are powerful tools in capital improvement planning. Recent surveys of large and very large utilities show that few are using failure prediction models, especially advanced models including artificial neural network and fuzzy logic, also known are “gray box” or “black box” models [9], [10]. The Environmental Protection Agency (EPA) suggests that the majority of the financial burden associated with water distribution needs in the next 20 years will come from medium and large sized utilities serving over 3,300 persons [4]. These utilities could be greatly assisted by reliable pipeline performance prediction models, which can help utilities analyze and reduce long term costs through proactive asset management while maintaining or improving current levels of service [10].

The lack of utilization of these models can be attributed to the data needs required and the complexity of the models. In their review of failure prediction models and utility practices, St. Clair and Sinha [10] describe that most models presented in the past decade are too sophisticated to be put into practice by utilities and require extensive amounts of pipeline data. This data can be extremely difficult to acquire for an entire network. Even basic pipeline property data such as material and installation date are commonly missing from asset management databases, as noted in literature[13]–[16]. The treatment of uncertain data in failure modeling could introduce bias in the model predictions. This potential bias has yet to be investigated.

Most recent models require more parameters than just physical pipeline properties. These parameters serve as surrogates for data that explains variations in break rates and failure modes for otherwise homogenous groups of pipe [11]. Utilities might not have data nor the resources to collect data that is required in many models presented. The more computationally intensive, data mining models such as artificial neural networks and evolutionary polynomial regression models require quite a significant amount of additional parameters in order to collect enough explanatory variables to train and validate a model.

Examining the break rate within clusters of failures could lead to a surrogate parameter for data such as soil conditions and traffic levels that are difficult to collect and have inherent uncertainty. By collapsing several parameters into one, the modeling effort is decreased, saving utilities time and resources. Additionally, by reducing the number of parameters, the chances of over-fitting a model are decreased, which is of concern for medium utilities with limited numbers of recorded failures.

4.2 Objectives

The goals of this paper are to investigate the impacts of uncertain data on model performance and to explore the suitability of using break rate as a surrogate for sparse data in the context of a survival model. These goals are achieved by introducing two improvements to the Weibull Proportional Hazards Model (WPHM) and comparing those to a base model. The base model is typical of other WPHMs introduced, and includes pipe installation date and diameter as explanatory variables. This model is stratified by pipe material. The first improvement to the WPHM is to introduce two categorical binary variables to account for expert elicitation of pipe material and installation date.

The second improvement builds upon the first and adds a localized distributed break rate as a covariate. The break rate is obtained by developing a kriging model of break rate at failure areas in the network. Kriging models are commonly used in spatial statistics to interpolate values based on observations or training points using weighted spatial covariance values. Within the past few decades, kriging models have been applied in several ways in the water/sewer/pipeline industry to solve estimate groundwater levels [88] to model contaminant concentration in groundwater [89] to predict the impact of earthquakes on water transmission pipeline breaks [90], and to estimate water flow through pipelines [91]. The distributed break rate covariate obtained from the kriging model is used to account for unknown parameters that cause increases in internal and external loadings resulting in elevated break rates. This model is also compared to the base WPHM model.

This chapter is organized in five sections: (1) a literature review of Weibull-based failure prediction models, and efforts investigating the spatial relationship of pipeline failure is presented; (2) Utility A is introduced, the parameters to be included in the models are described, and data quality concerns are addressed; (3) the methodology section introduces a method for calculating and interpolating break rate to be used as a data surrogate, and also described the methodology for estimating failures; (4) model calibration and prediction results are presented, and comparisons of the models are made; and (5) the conclusions are discussed.

4.3 Literature Review

One of the first statistical models used to predict pipeline failures in water distribution networks was developed by Andreou [92], which applies a proportional hazard model developed by Cox [93], to examine the influence of pipeline covariates on failure times.

Le Gat and Eisenbeis [37] introduced a model that utilized limited duration maintenance records to forecast failures in pipes. The model is a Weibull Proportional Hazard Rate Model (WPHM) that relates pipeline properties to the time to failure.

The Weibull hazard rate model is written as:

$$h(t_o) = \lambda p(\lambda t)^{p-1} \quad [4.1]$$

Where t is the failure time, λ is the Weibull intercept and p is the Weibull scale parameter. For a function with covariates, the Weibull hazard model is

$$h(t, \boldsymbol{\beta}, \boldsymbol{z}) = \lambda p(\lambda t)^{p-1}\exp(\boldsymbol{z}'\boldsymbol{\beta}) \quad [4.2]$$

Where $\boldsymbol{z}'$ is a vector containing explanatory variables also known as covariates that influence pipeline survival and $\boldsymbol{\beta}$ is vector of regression coefficients corresponding to the covariates.

The Weibull distribution can also be parameterized as an Accelerated Failure Time (AFT) model. An AFT model can be used to evaluate the impact of covariates on pipeline survival, whereas the WPHM is used to determine the impact of covariates on the hazard rate. A linear function describes the relationship between the covariates and log of the scale parameters and time [94].

$$\ln T = \alpha + \boldsymbol{z}' \boldsymbol{\beta}^{*} + \sigma \boldsymbol{W} \tag{4.3}$$

Where T is failure time, $\alpha = -ln\lambda$, $\sigma = 1/p$ and is the scale parameter, and $\boldsymbol{\beta}^{*} = -\sigma\boldsymbol{\beta}$. $\boldsymbol{W}$ is a vector of errors, each with an extreme value distribution. $\boldsymbol{W}$ can be rewritten given a specific failure as

$$w(t) = \frac{\ln t - \boldsymbol{z}' \boldsymbol{\beta}^{*}}{\sigma} \tag{4.4}$$

Assuming an extreme value distribution for the survival function as recommended for censored data [95] the survival function can be rewritten in terms of w as:

$$S(w) = \exp[-\exp(w)] \tag{4.5}$$

Substituting $w(t)$into the survival function yields:

$$S(t, \boldsymbol{\beta}^{*}, \boldsymbol{z}) = \exp\left[-\exp\left(\frac{\ln t - \boldsymbol{z}' \boldsymbol{\beta}^{*}}{\sigma}\right)\right] = \exp\left[-t^{1/\sigma} \exp\left(\frac{-\boldsymbol{z}' \boldsymbol{\beta}^{*}}{\sigma}\right)\right] \tag{4.6}$$

The WPHM model is valuable for ranking groups of pipe with respect to probability of failure. The model introduced by LeGat and Eisenbeis was the first to explore survival theory with short maintenance records, spanning less than 20 years and was also the first to account for left truncation and right censoring of failure records that is typical of failure data stored in relatively new digital asset management databases. The model allows decision makers to analyze the impact of properties describing the pipeline and surrounding environment on the survival of a pipe.

3.3.1 Weibull Based Failure Models

Implementations, variations, and suggested improvements of Weibull based break prediction rate or failure prediction models have been documented in literature [14], [41]–[43], [47]–[49], [51]. The models vary in methodology for predicting successive failures on the same pipe segments, and the types and forms of model parameters. The parameters required for stochastic models including the Weibull-based models referenced above are shown in FIGURE 4.1. In addition to basic pipeline data, such as material and installation date, some models require parameters that serve as surrogates to account for variations in break rates and failure modes for otherwise homogenous groups of pipe. For example the pipe bedding and backfill material is a surrogate for increased external loading by construction practices and structural resistance of a pipe that could elevate local break rates compared to the average network break rate.

Parameters	Author(s) 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Material	X	X		X	X	x	X	X	X	X	X				X	
Diameter	X		X	X	X	X	X	X	X	X	X		X		X	X
Age	X	X		X	X	X	X	X	X	X	X		X	X	X	X
Length	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X
Pressure	X				X			X		X		X				
Soil Type	X	X			X			X			X					
Pipe Protection								X			X					
Land Use		X		X	X			X				X				
Traffic	X			X				X	X	X						
Under Water				X												
Infrastructure Above Pipe				X	X			X		X						
Gravity/Operational	X															
Assembly Type	X											X				
Properties/Customers Supplied						X						X				
Flow								X								
Depth							X	X								
Pump Failures											X			X		
Freezing Index											X			X		
Rain Deficit											X			X		
Recency														X		
Clustering														X		

Legend: 1. LeGat & Eisenbeis, 2000; 2. Mailhot, et al, 2000; 3. Park & Loganathan, 2002; 4. Vanrenterghem-Raven, 2007; 5.Rogers & Grigg, 2009; 6.Berardi et al., 2008; 7. Wang, et al, 2009; 8. Wood & Lence, 2009; 9. Carrión, et al, 2010; 10. Debón, et al., 2010; 11. Kleiner & Rajani, 2010; 12. S. Park et al., 2011; 13. Xu, et al, 2011; 14. Kleiner & Rajani, 2012; 15. Martins et al., 2013; 16. Toumbou et al., 2013

FIGURE 4.1: MODEL PARAMETERS

While these models have been successful in prioritizing pipes for replacement, they have not been widely adopted by local utilities for the purposes of capital improvement planning. Many of these models are too sophisticated to calibrated by utilities in house [10], requiring the additional expensive of hiring consultants to develop and update the model(s). Additionally, these models require parameters that are often not available and difficult to acquire. Even basic pipeline property data can be missing from asset management databases. Researchers have noted problems with missing and unreliable pipeline parameters and have proposed methods for addressing data gap problems including assigning the average installation date by material and diameter [13], excluding pipes with property uncertainty from the model [14], or leaving the missing attribute blank [15]. The potential bias due to uncertainties associated with educated assumptions of pipe material and age, which could be significant in networks with large data gaps. The impact of these assumptions on model prediction performance has yet to be investigated.

3.3.2 Spatial Models: Clustering Analysis

As an alternative to collecting missing pipeline data and vast amounts of surrogate data, spatial models are now being used to investigate clustering of accidents, to determine candidates for pipe replacement and evaluate parameters that influence clustering. Models for spatially correlated survival data have been successfully used to detect and analyze patterns in health and ecological related survival data [96]–[99]. Spatial models commonly used in other fields are now being applied to pipeline failure analysis. Clusters of pipeline failures are being detected by using distance and/or time based search algorithms or by finding areas of increases pipe break rate intensity.

Oliveria et a. [53] introduce a new search window algorithm to estimate pipeline break within clusters of failures. Hypothesis testing is used to compare the network level break rate to the break rate within clusters as defined by the search window area. The authors investigate pipe break density using an expanded Ordered Points to Identify Clustering Structure (OPTICS) algorithm, which is based on the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) cluster search algorithm, which uses a minimum number of points and minimum threshold distance to define clusters. The impact of explanatory variables within clusters is investigated. Christodoulou, et al,[57] further expand the DBSCAN algorithm to consider the time window before and after a break to account for temporal variations in spatial clustering of pipeline failures.

Though these new search algorithms to identify spatial and temporal patterns of breaks are beneficial in detecting current high risk areas, they cannot be used to predict the future state of the network. Additionally, clusters of failures in dense areas of the network may not be of as high of concentration of breaks in the less dense area, as utility directors cite multiple pipeline breaks as a primary criterion for replacement [6].

3.3.3 Spatial Models: Break Rate Analysis

Instead of evaluating clusters, other researchers are examining areas of elevated break rate. Shi et al. [58] investigate variations of break rates within a network due to clustering of accidents by overlaying a grid on the network to calculate a relative break rate for the water distribution network in Hong Kong. Bogárdi [56] investigates spatial homogeneity of pipe failures by computing break rate to be used as the intensity term in a Poisson distribution. A grid is overlaid on the network, and the average number of failures per grid is calculated and used as the intensity value in a Poisson probability distribution describing pipeline failures. The observed frequency of failures per cell is compared to the theoretical distribution of failures per cell. Using the intensity function, a space-time Poisson process is used to simulate potential failure patterns in a water distribution network.

The grid method for calculating break rate can be problematic with respect to determining the appropriate cell size for the grid [100]. The method is somewhat arbitrary and can lead to artificially increased break rate where there is little pipe within the cell. Bogárdi's model is one of the first to use the current break rate to estimate the future state of the network. Though successful, this model does not take into account known parameters that influence break rate such as pipe size and installation period.

3.3.4 Clustering in Statistical Models

Oliveira, et al, [53] hypothesize that clusters could be used as inputs in multi-variate replacement planning models. This hypothesis has yet to be tested fully. Only Kleiner and Rajani [46] have incorporated clustering into stochastic models to predict the likelihood of individual pipe failure. Failure clusters were determined using the K-means search algorithm and a binary covariate for clustering was introduced. Pipes within a cluster were assigned a value of one, and pipes outside of clusters were assigned values of zero. Clustering was considered to be a statistically significant covariate when included with other pipeline attributes in a non-homogenous Poisson process model.

The models used by Kleiner and Rajani also include a number of other explanatory variables such as rain and freezing indices. These parameters are time dependent and difficult to process, requiring the database manager to distribute values from measurement locations and assign them to specific pipes. In the case study presented of a Canadian utility, rain and freezing indices are not found to be statistically significant with respect to model calibration [46]. Clustering is determined using a k-means algorithm which does not take into account the density of the network where clustering is observed and as a result may not adequately describe areas of increased break rate.

The model improvement recommended in this paper is to include a localized break rate as a covariate in the WPHM, which is not calculated using arbitrary mesh sizes. The following sections of the paper describe the utility data used to test the models, and the methods for adding the assumption and break rate covariates.

4.4 Utility Data

Data has been provided by Utility A, located in Southeast United States. This Utility and the corresponding data are described in Chapter 3. In inspecting the data for the utility, several problems arise. First, the material for over 40% of pipe segments is unknown. Corresponding with the unknown material, the installation date for over 47% of pipe segments is missing. A possible explanation for this high percentage of unknowns is that Utility A inherited several smaller utilities through annexation and county-city consolidation over 50 years ago. The data for these inherited networks could have been lost.

Though researchers have demonstrated ways of determining materials and installation periods, none have addressed the potential bias these uncertainties can have in model prediction performance. Without addressing such bias, asset managers might lack confidence in the models to correctly predict failures to prioritize pipes for rehabilitation or replacement.

The next data quality concern is related to the recorded pipe lengths in the GIS model for Utility A. Due to segmentation of the network in the process of digitizing maps, the length of pipes in the GIS model for Utility A does not correspond to the in situ length of pipe in the field. A method for correcting inaccurate pipe lengths due to segmentation has been described by [17].

Though pipe length is used as a covariate or explanatory variable in many statistical models, the uncertainty associated with attempting to correct the segment lengths makes it impractical to consider length in the model.

4.5 Addressing Unknown Material and Installation Dates

Data imputing methods, which assume complete randomness of missing data, allow for the missing data to be discarded before training, a learning algorithm capable of handling missing data to be specified, or missing variables to be estimated [101]. The most common method for imputing is assigning a median or mean value of the known values for the parameter [101]. Discarding records with unknown pipe parameters similar to Røstum [102] and [17] would result in a large data loss and is not recommended when the amount of missing data is large [101]. A learning algorithm is not applicable with the model form, so the remaining alternative is imputing. Filling in data gaps in asset management through educated assumptions of pipe age and material based on the vicinity of known pipes and knowledge of development and pipe installation decades is a method that has been utilized by Rogers and Grigg [15]. Since the correlation between urbanization and pipe material is strong [41] imputing missing pipe materials and installation dates through expert elicitation based on known utility practices during urbanization periods is a viable process, yet it does introduce bias. The method of accounting for this bias through a categorical parameter [101] was employed.

The expert opinions of utility personnel with extensive network knowledge were used to estimate material. Within the GIS model, areas of unknown pipe were outlined. Knowing the history of the network and development in the area, and being able to see the material of surrounding pipes, utility personnel estimated pipe material in subdivisions. Given the material, the age of the pipe was assumed based on known practices. For example, the transition from cast-iron pipe to ductile iron pipe for Utility A occurred around the early to mid 1990's. Therefore, pipe segments estimated to be ductile iron pipe were assigned an installation date of 1992. Also, pipe segments known to be ductile iron (DI), yet were only missing an installation date, were assigned an assumed installation date of 1992.

In order to account for the bias associated with these assumptions, two assumption parameters are introduced to the model. The parameters *AssumedMaterial* and *AssumedDate* are binary. Shown in Table 4.1, when the material for a pipe segment is assumed, the *AssumedMaterial* parameter value is true, and assigned a value of 1. When the material is known, the parameter value is 0.

Table 4.1: Example of Records with Assumption Covariates

COMPKEY	Install Date	Material	Updated Material	Updated Install Date	Assumed Material	Assumed Date
739973	1996	DI	DI	1996	0	0
811219	UNK	UNKNOWN	DI	1992	1	1

Though the assumptions were not verified and in many cases could not be verified due to missing or damaged as-built construction documents, the impact of these assumptions was investigated. Preliminary modeling efforts not reported in detail in this paper investigated modeling unknown materials as one large cohort of unknowns which resulted in extreme over predictions of failure. The cohort of known materials modeled did not perform substantially differently than the models presented later in this paper, with respect to comparisons of predicted and observed failures.

4.6 Spatial Variation of Break Rates

The spatial variation of breaks is analyzed by computing and comparing local break rates for subsets of the network. This break rate parameter considers the number of breaks over the duration of the asset management database. Though researchers have used grids to calculate break rate, the sometimes arbitrary selection of an appropriate cell size can be difficult (Shi et al,, 2013). Alternatively, a localized break rate can be calculated within Thiessen polygons calculated around failure points. Thiessen polygons are commonly used in hydraulic analysis of water distribution networks to define service areas around GPS locations of meters and to distribute demand across nodes within the service areas. A similar method can be used to distribute break rate across a failure region.

The boundaries of a Thiessen polygon are the bisectors of the lines from a core point and the points surround it shown in FIGURE 4.2 The break rate as defined as the number of breaks per 1,000 feet (305 m) of pipe per year is calculated for each Thiessen area and assigned to the failure point(s) within the Thiessen polygon.

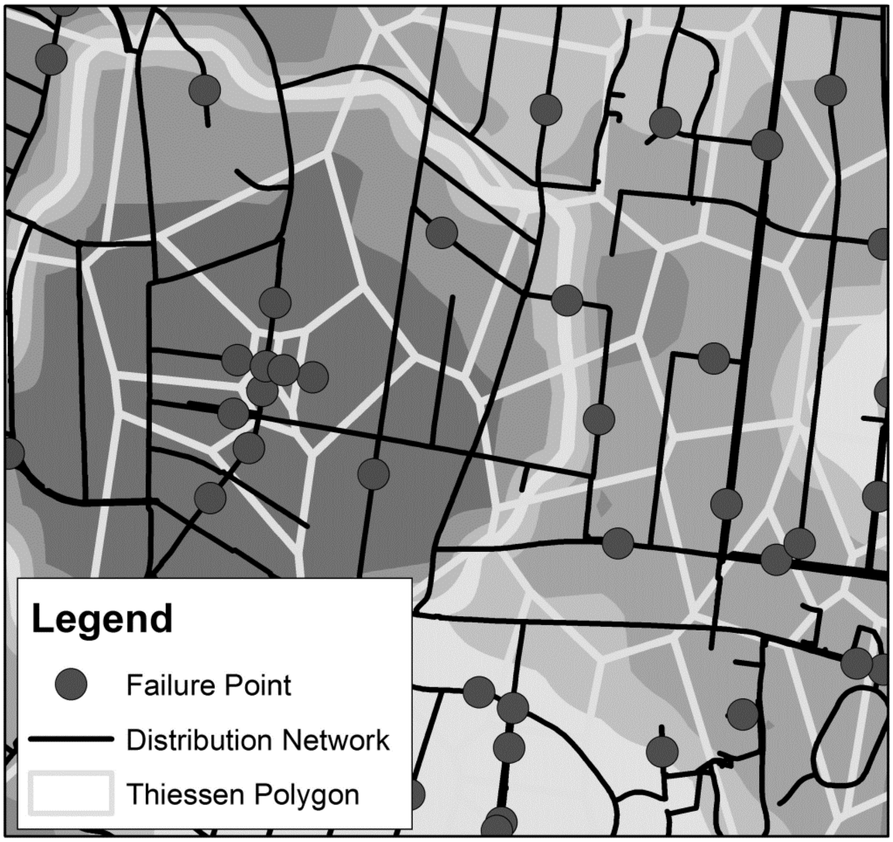


FIGURE 4.2: BREAK RATE DISTRIBUTION

To distribute the break rate across the network, a stationary Kriging model is used to interpolate the expected break rate between failure points. Kriging models are used to approximate functions through interpolation and are defined by a mean function and covariance function. The failure points are assumed to comprise a random field of Gaussian random variables. This random field is assumed to be stationary, meaning that patterns do not change over time. A non-stationary model would not be appropriate due to the brevity of the failure database.

The Kriging model inputs, in this case x-coordinate and y-coordinate of failures, are assumed to correspond to set of random functions indexed by observations, or in this case, calculated break rates at failure locations. The response function of the GP model is described by the following equation:.

$$\hat{Z}(s_o) = \sum_{i=1}^{N} \lambda_i \, Z(s_i) \tag{4.7}$$

Where λ_i = an unknown weight for the measured value at the *i*th location.

N = the number of measured values

$\boldsymbol{Z}(\boldsymbol{s_i})$ = the measured value at the *i*th location.

It is assumed that the break rate is a spatially autocorrelated process with independent random errors described by a mean and error function below:

$$Z_t(s) = \mu(s) + \varepsilon_t(s) \tag{4.8}$$

Where $\mu(s)$ is a unknown, deterministic mean value and $\varepsilon_t(s)$ is a function that accounts for random measurement and model fitting errors. For more information on the spatial Kriging model, refer to Appendix A.

The output of the GP model allows for an estimated mean break rate and variance at every coordinate within the utility district. Since the break rate varies spatially, and pipe segments extend across space, the average of the expected break rate and variance across each pipe segment is computed and assigned to the pipe segments. The average estimated break rate is included as a covariate in the WPHM. Background information kriging model within ArcGIS 10 and results including raster images of the estimated break rate are included in Appendix A.

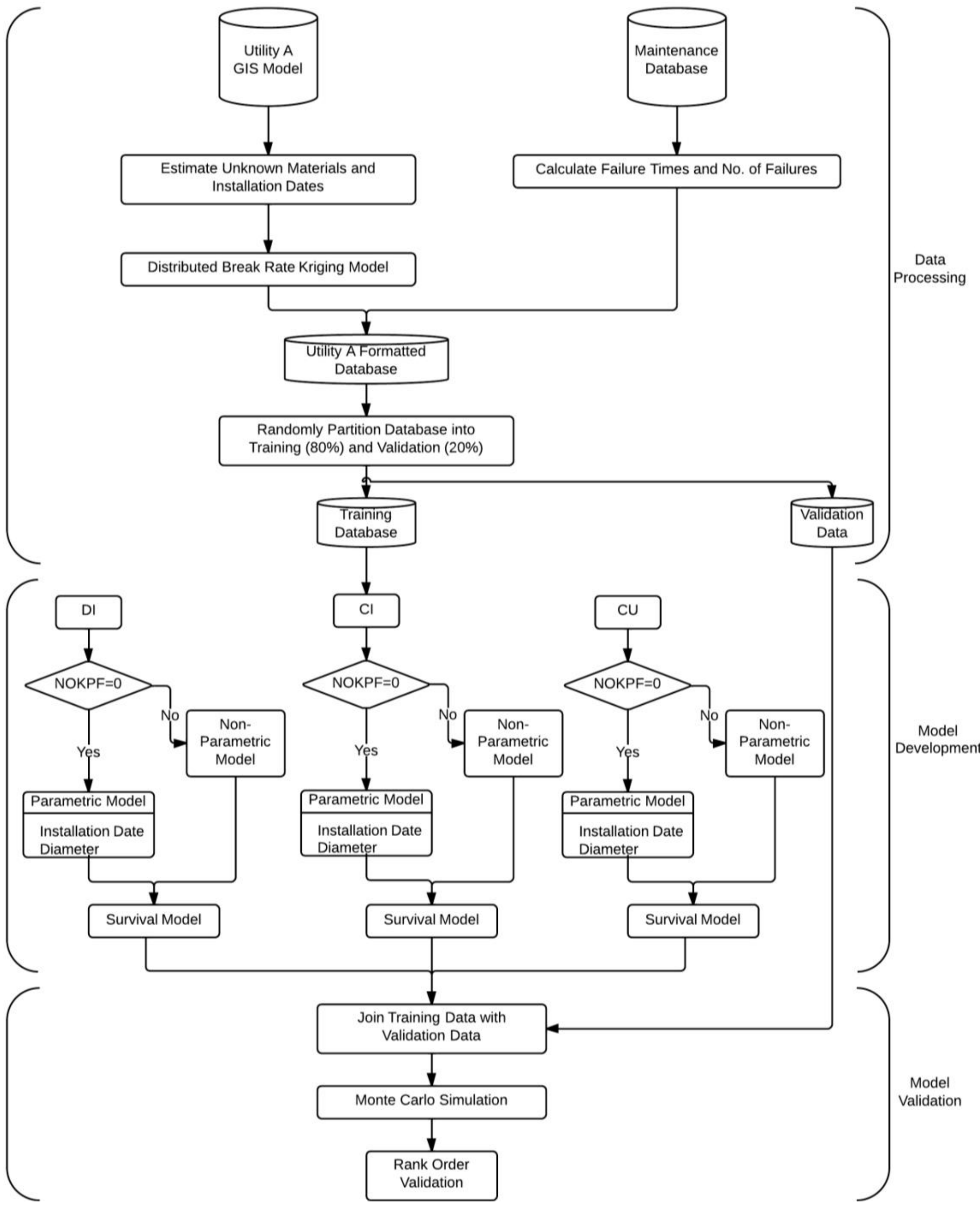


FIGURE 4.3: METHODOLOGY FLOW CHART

4.7 Methodology

The models for ductile iron (DI), lined cast iron (CL), and unlined cast iron (CU) pipe cohorts will take the form of the Weibull Hazard Rate Model described by Le Gat and Eisenbeis [37] as shown in Equation 4.6. Shown in the Model Development section of the flow chart in Figure 4.3, separate models will be calibrated for each material cohort of pipe. Separate model forms will be used to model the first failures expected for the pipe segment, in which the Number of Known Previous Failures (NOKPF), a term introduced by LeGat and Eisenbeis [37] and used by Røstum [102] and Kleiner and Rajani [46], is zero. Since repaired pipelines either perform "good as new", "good as old" "worse than new" or "worse than old" [103] a separate model is needed to account for pipe behavior after repair. Since few pipes have multiple recorded failures, in order to avoid over-fitting the models, a WPHM with no explanatory variables is used to predict future failures where NOKPF is greater than zero. The parameters for the models will vary, as only significant parameters as determined by p-test statistics will be included in the models. These parameters could differ based on material.

To test the models' prediction accuracy, 20% of failure records and pipes are randomly selected to be used as a hold out sample for model validation. This allows for a good representative sample to be used for testing, while retaining valuable data needed for model calibration, especially the failure records for subsequent failures on pipes. The regression parameters are tested for significance, and coefficients of parameters are estimated using the remaining 80% of failure records and pipe inventory. The Maximum Log Likelihood method is used to estimate the WPHM coefficients.

In considering the impact of new covariates on prediction performance, a base model is calibrated which considers the most commonly used Weibull model parameters that were available for the network as summarized in Figure 4.1. The base model for each material cohort is then re-calibrated to consider the base model parameters and the *AssumedMaterial* and *AssumedDate* categorical assumption variables previously introduced. The final model calibrations consider the base model parameters, the assumption covariates, and the distributed break rate parameter described in the previous section.

Each model depicted in Figure 4.3 will be used to simulate failures using the Monte Carlo simulations. Monte Carlo Simulations are widely used to evaluate deterministic functions that are dependent upon input variables that can be represented as distributions that account for parameter uncertainty [104]. The Monte Carlo Simulation technique generates random samples of each input using a pseudo-random number generator, to generate an output distribution of the deterministic function results. This distribution can be used for assessing the predictive performance of the calibrated model.

With respect to evaluating the comparative performance of the models presented, the survival model is solved for time, shown in Equation 4.9. In context of Monte Carlo Sampling, time now serves as the deterministic function that is dependent upon the survival probability and regression parameter estimates.

$$t = \left(\ln\left(\frac{1}{S}\right) \exp\left(\frac{\alpha + \boldsymbol{z}'\boldsymbol{\beta}^{*}}{\sigma}\right) \right)^{\sigma} \qquad [4.9]$$

To execute the Monte Carlo Simulation, a uniform random number, u, is sampled between 0 and 1 to represent the survival probability, S. Using Equation 4.9, the corresponding failure time is computed. The survival time is compared to an established time horizon. A pipeline failure is indicated when the survival time is less than the established time horizon, which in this study is the duration of the asset management database, which is used to compare predicted failures to observed failures. In simulating future failures, the time horizon will extend beyond the duration of the asset management database. The distribution of survival times and as a result failures or non-failures can then be used to estimate the expected failures during the time interval and assign confidence intervals to these estimates.

The routine described will be used to simulate the number of expected annual failures for the duration of the observation period for the asset management database for Utility A. The routine will be repeated one thousand times to generate an appropriate number of points for consideration.

The Monte Carlo Simulation to generate expected failures assumes that the failure times for pipe segments follow the Weibull distribution previously specified. Any underlying randomness not captured by the Weibull distribution will not be represented in the Monte Carlo Simulation. Monte Carlo Simulation analyses are also highly dependent upon the number of simulations, so care must be taken to insure the convergence of the simulation number [104].

The results of the simulations vary, and a range of potential failure histories are produced. For model comparison purposes, the averages of the simulation results are reported. The decision maker needs to keep in mind that the probability always exists that the expected failures could be higher or lower than the averages reported. Rather than assuming that the Monte Carlo Simulations accurately predict exact facility failure times, the model is used to identify the pipes with the highest simulated risk of failure over a time horizon.

As a surface level evaluation of prediction performance, the total average predicted failures for each model are compared to the observed failures. To validate the models' ability to identify high risk pipes, a validation metric that examines the failure predictions for quantiles is used. This validation metric introduced by Le Gat and Eisenbeis [37] also used by Røstum [102] Martins et al. [49] evaluates the prediction performance at the pipe or level. Pipes are gathered into cohort quantiles based on predicted failures, which are correlated to the number of observed breaks. The cumulative total of observed and predicted breaks for the cohorts are compared.

4.8 Results

Models were developed for the three materials that incorporated estimates for both material and installation date. Table 4.2 and Table 4.3 show the calibration results for lined CL, CU, and DI pipes calibrated without the assumption parameters and with the assumption parameters. Table 4.4 shows the results for model with assumption parameters and the estimated break rate covariate.

Table 4.2: Base Model Parameters

Model Parameter	CL		CU		DI	
	NOKPF=0	NOKPF>0	NOKPF=0	NOKPF>0	NOKPF=0	NOKPF>0
Intercept	4.65358	0.82463	4.42538	0.81347	-45.3759	0.66264
Scale	1.02810	0.97515	0.97866	1.04012	0.91560	1.02660
Parameters						
Pipe Diameter	0.16934	--	0.16981	--	0.16615	--
Installation Date	--	--	--	--	0.02535	--

Table 4.3: Model Parameters with Assumption Covariates

Model Parameter	CL		CU		DI	
	NOKPF=0	NOKPF>0	NOKPF=0	NOKPF>0	NOKPF=0	NOKPF>0
Intercept	125.0983	0.82463	26.55634	0.813478	-103.45	0.662648
Scale	1.023578	0.975154	0.970991	1.040124	0.913859	1.026608
Parameters						
Pipe Diameter	0.177721	--	0.138171	--	0.129147	--
Installation Date	-0.06152	--	-0.01154	--	0.054973	--
Assumed Installation Date	-1.92709	--	-1.23252	--	-1.32742	--
Assumed Material	2.315415	--	1.930053	--	--	--

Table 4.4: Model Parameters with Assumption and Break Rate Covariates

Model Parameter	CL		CU		DI	
	NOKPF=0	NOKPF>0	NOKPF=0	NOKPF>1	NOKPF=0	NOKPF>0
Intercept	112.6289	0.82463	27.34865	0.813478	-94.1319	0.662648
Scale	1.021532	0.975154	0.968981	1.040124	0.91303	1.026608
Parameters						
Pipe Diameter	0.187801	--	0.145772	--	0.119572	--
Installation Date	-0.05481	--	-0.01169	--	0.050541	--
Assumed Installation Date	2.050982	--	-1.23687	--	-1.17224	--
Assumed Material	-1.68658	--	1.934948	--	-0.31484	--
Average Break Rate	-0.00423	--	-0.00305	--	-0.00276	--

Evaluating the impact of covariates on the survival function, for each material, the average break rate is significant given p-test results and influences the survival function in the same fashion. The negative value for pipe break rate indicates that pipes with higher break rates have greater failure probabilities. Also for each material, pipe diameter is a significant covariate and acts in the same way for all materials. The positive value indicates that pipes with larger diameters are less prone to fail, because as the pipe size increases, the failure time also increases. The positive coefficient value for pipe diameter aligns with the convention that smaller pipes are more prone to fail, as they have thinner wall thickness, and are less resilient to external and internal forces and corrosion.

For CL and CU, the assumed installation date parameter is negative, meaning that older pipes have decreased survival probabilities. This is not the case for DI, but this could be attributed to the vast number of pipes with an assumed installation date of 1992, which is at the high end of the range of installation periods for DI.

Additionally concerning DI, the intercept is negative for the first cohort model (NOKPF=0), indicating that the failure increases at a lower rate than the other pipe cohorts. DI pipe is the newest material in the network, and now comprises approximately 40% of Utility A's network. The failure rate for DI is less than the other pipe materials, and as a result has more right censored pipes. Additionally, the last two years of failure records showed a marked decrease in failures for DI. Since DI has only recently been installed, it is possible that a majority of the DI pipes installed have yet to reach wear out stages that occur at the end of a pipe's useful life. It is possible that the DI pipes are experiencing normal wear and the model is not capturing the end of life break rate, associated with the right hand side of the "bathtub curve" describing infrastructure failures..

Using the model calibration results obtained above, Monte Carlo Simulation routines were used to simulate failures for the duration of the asset management database. For each model, 1,000 simulations were performed, and the average of the predicted failures rounded to the nearest integer are reported in Table 4.5.

Table 4.5: Prediction Performance

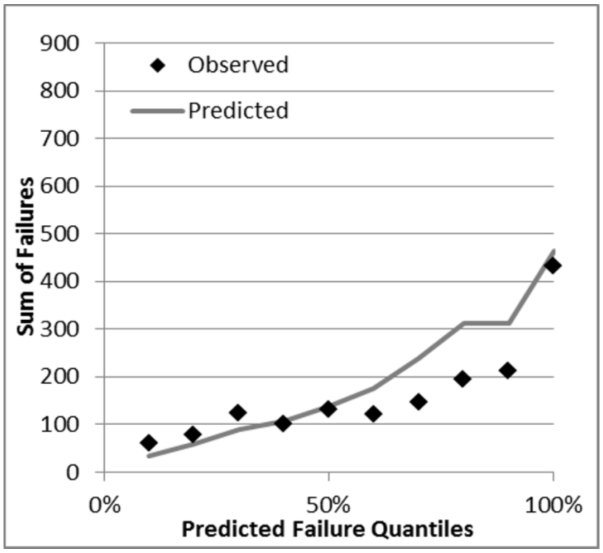
Material	Observed Failures	Predicted Failures		
		Base Model	Base Model And Assumption Covariates	Base Model With Assumption and Break Rate Covariates
CL	733	594	594	584
CU	953	761	778	755
DIP	668	553	555	554

Model Comparison

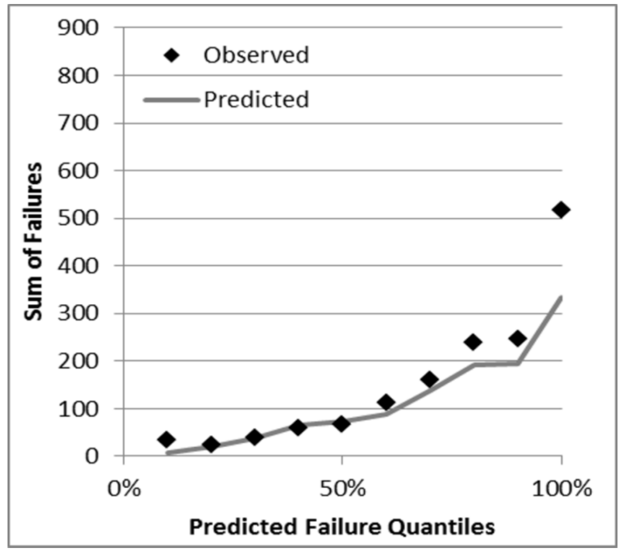
In order to evaluate the prediction capabilities of the models, the rank order validation metric is utilized. Pipelines are sorted according by the average number of predicted breaks. Quantiles for the set of predictions are determined. Pipes are grouped by quantile and the sum of the predicted breaks for each quantile is compared to the sum of observed breaks for each group. Since the data is heavily censored, we expect to see a large jump between the 90% and 100%

quantiles, which represents the pipelines with the highest risk pipe segments. An aggregate comparison is made in Figure 4.4, comparing predictions for all three pipe material cohorts.

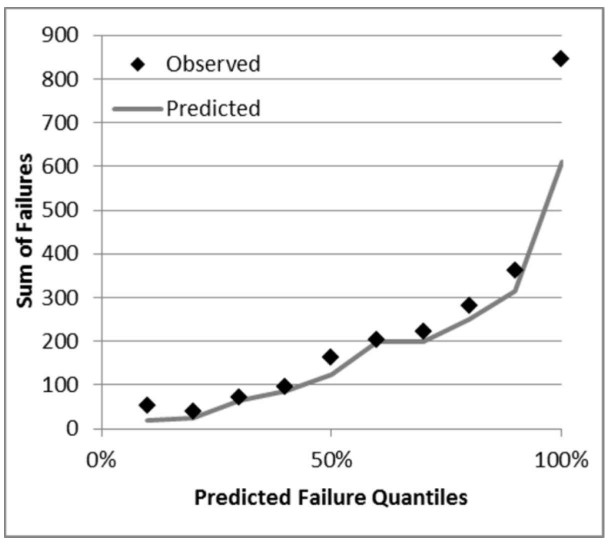
Though the models without assumption parameters perform well on a network level by satisfactorily predicting the number of observed failures, the models do not accurately predict failures for the highest risk cohorts. Quantiles are more readily observed when adding the assumption variables, and a marked increase is shown in the model when the break rate covariate is added.



(a) Base Model



(b) Model with Assumption Covaraiates



(c) Model with Assumption and Break Rate Covariates

FIGURE 4.4: VALIDATION RESULTS

Conclusion

This paper has addressed the problems of uncertain and limited data that attributes to the low adoption rate of pipeline failure prediction models. It has solved these problems by introducing two improvements to the Weibull Hazard Rate Model (WPHM). Upon application of these improvements, it was found that utilities with uncertainties about basic pipeline properties can incorporate expert elicited assumptions of these properties to develop improved models based on comparisons of observed data and Monte Carlo Simulation results in which the models calibrated with 80% of the recorded data were used to simulate failures for all of the facilities within the network. When these assumptions are acknowledged, in the form of categorical binary covariates, the ability to detect the highest risk pipes increases, as evidenced by the commonly utilized quantile rank validation routine. This finding supports the idea that not accounting for pipeline property assumptions introduces bias into failure prediction models.

It was also discovered that a marked increase in the prediction performance for the pipes with the highest likelihood of failure exhibited in the eighth through tenth deciles occurs when the spatially distributed break rate is added to the model that includes the categorical assumption variables. This increase is realized when compared to both a base WPHM and the WPHM with the added assumption covariates. The value of this contribution is that an acceptable failure prediction model was developed with a limited number of explanatory variables. The number of explanatory variables required was reduced by examining break rate on a local level using Thiessen polygons to subset the network, and distributing the break rate spatially using a kriging model. This single break rate covariate can be a surrogate for many other explanatory variables, which will limit the amount of data collection required for utilities to develop failure prediction models. Additionally, by reducing the amount of explanatory variables required, the chance of over fitting a model decreases, which is beneficial to medium sized utilities with smaller numbers of recorded failures which can be used to train models.

The limitations of these model improvements are that they have not been tested on networks varying in uncertainty, size, and data availability. Future work is needed to investigate the impact of the assumption covariates on networks with lesser and greater pipeline property uncertainty than Utility A. These limitations are investigated in Chapter 6, where the model is demonstrated on medium sized utilities.

Additionally, the break rate covariate model should be compared to a model that does include a number of additional explanatory variables including bedding type, infrastructure above the pipe, cathodic protection of the pipe, soil type, and ground water levels. Most of these parameters were not available for Utility A. Finally, after more breaks have been observed, temporal validation metrics should be used to compare predicted breaks to observed breaks for specific time intervals.

Though every effort should be made to maintain an accurate and reliable inventory database, the results of this study are encouraging for utilities wishing to develop risk-based capital improvement plans, but do not have exhaustive asset management databases and have uncertainties regarding network properties. Failure predictions can be synthesized with criticality studies and hydraulic models to predict future level of service, and prioritize pipeline rehabilitation and replacement projects in order to minimize the long-term cost of network ownership.

CHAPTER V – RISK-BASED OPTIMIZATION OF MR&R ACTIVITIES

5.1 Introduction

Compared to the amount of papers demonstrating failure prediction models for water pipes, there are few resources available describing how to utilize the predictions to prioritize pipe replacement and maintenance activities. For utilities to adopt strategic MR&R plans, there needs to be more real examples of risk-based asset management frameworks. The following chapter demonstrates a risk-based prioritization routine for pipe MR&R activities using the survival models presented in the previous chapter. A case study of this framework is presented using data from Utility A. The first section of this chapter provides background information describing previous works to prioritize water pipe replacement. Section 5.3 describes the methodology for quantifying the consequence of pipe failures in terms operational, environmental, economic, and hydraulic impacts.

The next sections focus on a genetic algorithm based approach to determining the MR&R actions and areas of the network to undertake such actions that minimize the short-term impacts of pipe failures, given budget constraints. Geographic binning of the network is utilized to constrain replacement work to specific areas of the network, helping the asset manager develop replacement projects.

The greedy-heuristic genetic algorithm is demonstrated using data from Utility A, which is much larger than utilities presented in case studies reviewed in Chapter 2. With 250,000 assets, a pipe replacement GA model like those presented in literature would have a search space of $2^{250,000}$

which is a very large problem and cannot be solved in a reasonable amount of time using a standard PC. In investigating the use of GA-based pipe replacement algorithms, a single year analysis required over 12 hours to run and converge.

As an alternative to using an algorithm to make binary decisions on which individual assets to replace, a novel method is introduced that is used to decide which part of the network to perform work. This reduces the search space by several orders of magnitude. Cost of repair versus replacement ratio analysis is used to determine which areas of the network to do work in and how much to invest in those areas to mitigate the impacts of failure.

5.2 Background

In order to thoughtfully develop maintenance and rehabilitation strategies, utilities are adopting risk-based approaches to pipeline prioritization. To assess the risk of infrastructure failure, the consequences associated with failure must be quantified, often using costing models. The most robust of these models consider the not just the costs of replacement, but the operational and social costs of failure. Many of these costs of failure tools are utilized in rehabilitation and design optimization studies and primarily consider the direct costs of replacement [3].

Other models attempt to quantify the societal costs of failures, or the indirect costs not paid by utilities. These can include costs associated with traffic due to detours, energy loss, customer outage costs, and flood restoration. The Grand Central model (GCM) [1] is a spreadsheet tool capable of analyzing both the direct and social costs of failures for smaller diameter pipes. As a follow-up study, Gaewski and Blaha [2] updated the Excel-based tool to analyze the total costs of large diameter pipes. This tool considers both the direct and social costs of pipe failures. In their assessment, the authors determine that the highest societal costs due to failures are associated with costs due to flood restoration and traffic disruptions which comprised over 80% of the social costs for the failures studied. Flooding and traffic costs are directly related to land use and urban density.

Piratla and Ariaratnam [3] simplify a societal costs model by adding a location multiplier to the direct costs of repair. This multiplier is based on the locality of the break including industrial and environmentally sensitive areas, which accounts for the built environment around the buried infrastructure. In assessing the total costs due to failure, the researchers also introduce an equation for calculating the costs of lost water, assuming an average break diameter and velocity.

The loss of water due to breaks also results in increase energy costs due to increased pressure requirements as a result of water losses. Though these costs are significant, accurate quantification of these costs depend on network topography and the spatial distribution of breaks [105]. Cabrera et al. [106] include a parameter developed by Colombo and Karney [105] in a simplified equation to calculate the energy losses due to leakage and breaks that is based on volume of water lost, the average pressure, and pumping efficiency.

The decrease in pressure and/or supply associated with leaks and breaks can also cause Level of Service failures, such as minimum fire flow protection, which could have a negative financial impact on the utility. Cabrera et al. [106] estimate a penalty for such costs and demonstrates that including these occasional costs can alter the optimum cost-effective time for pipe replacement.

In the PARMS-PRIORITY decision support system, Moglia et al. [107] introduce a probabilistic approach for calculating costs due to service disruptions. This approach considers the costs associated with commercial customers experiencing repeated interruptions will expect to be compensated for their losses. Dandy and Engelhardt [108] consider two forms of service loss: local interruption which accounts for service loss to customers and global interruption which takes into effect reduced pressure due to the break. The local interruption factor is simplified and assumes an average number of customers per land usage. The global interruption requires running simulations of valve closures for each pipe segment in the network and determining areas of critically low pressure and the associated customers. The result is a Total Expected Number of Customers Impacted by pipe failure, which is used in a rehabilitation optimization routine.

To fully quantify the impacts of failures, both the direct costs and the indirect costs paid by society need to be quantified. Though very robust and user friendly models like the Grand Central model (GCM) and the companion to the GCM for large diameter pipes have been developed, more can be done to validate these model results with case studies of pipe failures.

5.3 Consequence Analysis

The consequence analysis attempts to assess the environmental, operational, economic, and hydraulic consequences of pipe failures. The hydraulic consequence study was performed by an outside consultant. This analysis was performed using a hydraulic model with pipe failure simulations. The hydraulic consequence study considered the impact of pipe failures on customers with respect to supply and minimum pressures for fire flow protection. The results of the hydraulic consequence study are the most critical pipes in the network with respect to water losses and outages. Failures along these lines are considered to be catastrophic because if they fail, there are no redundant pipe lines to help supply the demand. These lines should drive some capital project decisions including the projects to add redundant lines. Utility A is already developing capital projects to mitigate the impacts of failure along these lines by adding hydraulically redundant pipelines. As a result, these pipelines were excluded from the optimization analysis to develop routine asset management programs. Instead, a ranking of these pipes based on failure probability was provided to Utility A, to use in prioritizing capital projects for the catastrophic lines.

For the remaining pipe segments, a consequence analysis was performed based on the work of [17] which is a weighted approach to assessing risk with respect to the environmental, economic, and operational impacts of failure. Tables 5.1 to 5.3 demonstrate show the weighted matrices, data, and geoprocessing steps taken to perform the analysis. Weights were determined with input from asset managers from Utility A.

Table 5.1: Operational Impact Matrix

Index Weight	OPERATIONAL							
	0.33							
Category Weight	Customer Impact						Traffic Impact	
	0.7						0.3	
Variable	Critical Customer		Material		Pipe Size		Road Type	
Weight	0.7		0.1		0.2		1	
Data Source	Critical Customer Data File		WM_Database		WM_Database		Streets	
Data Field	Address		MATERIAL		PIPE_SIZE		Class	
Processing	Geocode addresses then spatial intersect with customer points		None		None		Spatial intersect with buffer around streets. Note: AADT for spot locations also available	
Valid Entries	Value	Score	Value	Score	Value	Score	Value	Score
	Intersects With Critical Customer	100	OTH	1	0.5-2.0	1	0-1	100
	Doesn't Intersect	0	PVC	5	2.25-6	5	2	80
			CI, CL, COPP,CU, DIP, STEE	10	6-20	10	3	40
			AC, CONC	100	24-36	50	4	20
					42-60	100	5+	1

Table 5.2: Economic Impact Matrix

Index Weight	ECONOMIC					
	0.33					
Category Weight	REPAIR COSTS					
	1					
Variable	Land Use		Material		Pipe Size	
Weight	0.4		0.2		0.4	
Data Source	Property_March22		WM_Database		WM_Database	
Data Field	LAND_USE		MATERIAL		PIPE_SIZE	
Processing	Intersect pipeline buffers with property polygons. Intersect Census populatio n densities with land use polygons.		None		None	
Notes	Land use categories must be aggregated into valid entry categories.					
Valid Entries	Value	Score	Value	Score	Value	Score
	Other	1	OTH	1	0.5-2.0	1
	Park	5	PVC	5	2.25-6	5
	Residential	10	CI, CL, COPP,CU, DIP, STEE	10	6-20	10
	Commerical, Industrial	25	AC, CONC	100	24-36	50
	High Density	100			42-60	100

Table 5.3: Environmental Impact Matrix

Index Weight	ENVIRONMENTAL					
	0.33					
Category Weight	ENVIRONMENTAL IMPACT					
	1					
Variable	Critical Habitat		Wetlands		Pipe Size	
Weight	0.5		0.3		0.2	
Data Source	US FWS Crit Habitat		US FWS Wetlands		WM_Database	
Data Field	CRIT_HAB Poly		CONUS_Poly		PIPE_SIZE	
Processing	Intersect pipeline buffers with critical habitat polygons		Intersect pipeline buffers with wetland polygons		Only used if a pipeline intersects environmental area	
Valid Entries	Value	Score	Value	Score	Value	Score
	Intersects	100	Intersects	100	0.5-2.0	1
	Doesn't Intersect	0	Doesn't Intersect	0	2.25-6	5
					6-20	10
					24-36	50
					42-60	100

The total weighted consequence score is the sum of the weighted impact matrices. This score was computed for all pipe segments. Breaks in scores were assigned based on shape length to classify the pipes as high, medium, and low consequence. These scores are utilized as multipliers to account for the social costs of failures in the pipe replacement optimization routine described in the next section. It is important to note that the index weights can be modified to reflect the priorities and potential costs differences between operational, economic and environmental failure impacts. The asset manager should review and validate these failure consequence results and make any changes as deemed necessary. Specifically, one should compare case studies of previous pipe failures within the network and the associated societal costs to the results of the consequence analysis multiplier used in the optimization methodology.

5.4 Optimization Methodology

The most commonly used optimization methodology presented in the literature is genetic algorithms. These frameworks are discussed in Chapter 2.7. The optimization methodology proposed improves upon previous optimization frameworks by investigating the spatial binning to constrain pipe replacement projects to subset(s) of the network. The goal of the optimization routine is to help the decision maker develop pipe replacement projects that minimize the risk of pipe failures within a three year planning period by maximizing a costs ratio that compares the repair costs to replacement costs. Risk is quantified using a penalty function that considers both the probability and consequences of failure expressed in monetary terms. The methodology for the optimization routine is described below.

5.4.1 Probability of Failure

The probability of pipeline failure is calculated using the WPHM described in Chapter 4. Since the equations presented in Chapter 4 are survival functions, the calculation of the cumulative probability of failure is:

$$P_f(x) = 1 - S(x) \tag{5.1}$$

Where $P_f(x)$ is the probability of pipe failure, $S(x)$ is the survival function and x is time post the end of the asset management database duration defined in equation 5.2.

$$x = x_0 + t \tag{5.2}$$

Where x_0 is the duration of the asset management database used to calibrate survival model parameters, and t is the time in years post update of the database in which the failure probability is to be calculated.

5.4.2 Consequence of Failure

The cost of failure C_f, shown in Equation 5.3 considers both the probability of failure as well as the consequence, in order to account for the social costs of pipe failures. The criticality scores described in the previous section are used as multipliers to the base cost of a point repair.

$$C_f(x) = \left(P_f(x)C_{PR} \times CF\right) \qquad [5.3]$$

Where C_{PR} is the cost of a point repair at the present year provided in CF is a consequence factor multiplier. Values for these variables are shown in Appendix B.

5.4.3 Cost of Replacement

The cost of pipe replacement is a function of pipe diameter and pipe length shown in Equation 5.4. A lookup table included in Appendix B shows the unit costs of pipe replacement given diameter per linear foot. It is assumed that regardless of the original pipe material, the replaced pipe will be ductile iron.

$$C_R(x) = (RC \times L) \qquad [5.4]$$

Where RC is the replacement cost per unit length and L is the recorded shapelength of the asset.

5.4.4 Decision Variables

The following subsections describe the decision variables for the optimization function. Two decision variables are considered in this algorithm. First, the decision maker needs to decide where geographically to perform work. This decision is based on a grid that bins the network. Next, one must decide how much to invest in a bin in order to maximize the total costs ratio for the entire region. The following subsections describe the decision variables $z(x)$.

5.4.4.1 Binning of Assets and Projects

A binning method is used in order to subdivide the network into project. Using the fishnet the feature in ArcGIS, a grid is overlaid on the network.

The decision variable related to these binned zones is binary variable that controls whether replacement work is performed in a tract shown in Equation 5.5.

$$z_c(x) = \begin{Bmatrix} 1\ if\ work\ is\ \ to\ be\ performed\ in\ zone\ c\ in\ year\ x \\ 0\ if\ no\ work\ is\ to\ be\ done\ in\ zone\ c\ in\ yearr\ x \end{Bmatrix} \qquad [5.5]$$

For each year within the planning horizon, the decision is made to perform or not perform work. The number of zones in which work can be done in a given year is constrained. In addition, the structure of the algorithm is such that work can only be done in a zone once within the planning horizon.

5.4.2 Costs Ratio Function

The costs ratio for an individual asset, k, is defined as the penalty cost of failure at the end of year x, divided by the cost of replacement.

$$(CR)_k(x) = \frac{\left(C_f\right)_k(x)}{(C_R)_k(x)} \qquad [5.6]$$

The assets are ranked by benefit cost ratio and the cumulative benefit cost ratio $(CCR)(x)_N$ after replacing each asset is determined, where N is the CR rank order of the asset.

$$(CCR)(x)_N = \begin{cases} (CR)_k(x)\,, & for\ N = 1 \\ (CCR)_{N-1} + (CR)_k(x), & for\ N > 1 \end{cases} \qquad [5.7]$$

The cumulative cost $C_{cm}(x)_N$ is then calculated, which considers the financial impact of replacing assets according to B/C ratio ranking.

$$C_{cm}(x)_N = \begin{cases} (C_R)_k(x)\,, & for\ N = 1 \\ C_{cm}(x)_{N-1} + (C_R)_k(x), & for\ N > 1 \end{cases} \qquad [5.8]$$

Using the rankings, a cost benefit curve is plotted for each zone in each year. This curve is fit to a logarithmic function to develop the CR ratio cost functions, $\left(CCR(cap)\right)_z$, where the costs ratio is a function of the capital spent in a zone. As shown in Figure 5.1, there is a point at which investing more capital in a zone has a minimal impact on the CR value. The goal of the decision maker is to invest enough capital, or replace enough pipes to maximize the CR value, while keeping the capital investment to a minimum.

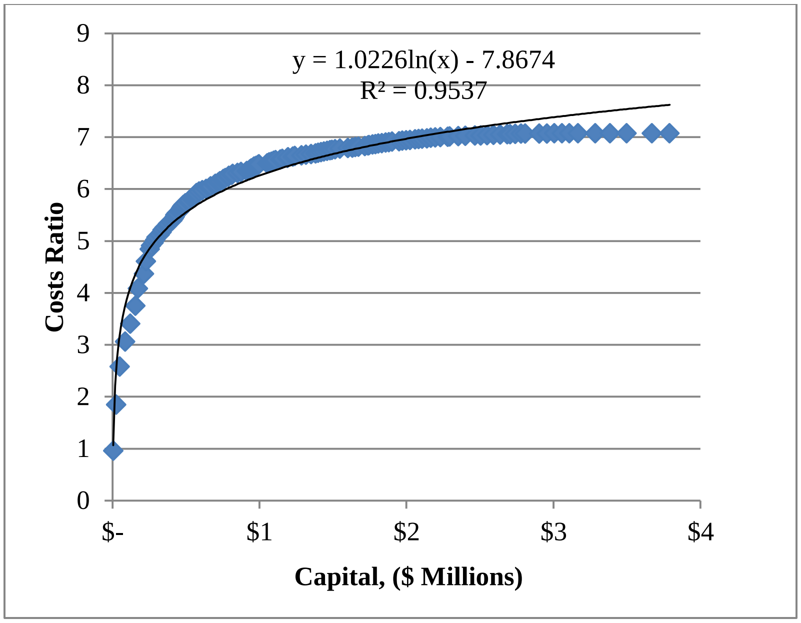


FIGURE 5.1: CUMULATIVE COSTS RATIO CURVE

5.4.4 Optimization Problem Formulation

$$\max_{z_{ij}c_{ij}} \sum_i \sum_j z_{ij}\, \mathrm{CCR}_{ij}(z_{ij}c_{ij}) \qquad [5.9]$$

s.t.

$$c_{ij}^{min} \le c_{ij} \le c_{ij}^{max} \quad \forall\ i \in I, j \in J \qquad [5.10]$$

$$z_j^{min} \le \sum_i z_{ij} \le z_j^{max} \quad \forall\ j \in J \qquad [5.11]$$

$$B_j^{min} \le \sum_i z_{ij}c_{ij} \le B_j^{max} \quad \forall\ j \in J \qquad [5.12]$$

$$z_{ij} \in \{0,1\} \qquad [5.13]$$

Where:

I = set of zones

J = number of years in planning horizon

z_{ij}= binary decision to work in zone *i* in year *j*

c_{ij} = dollars to be spent in zone *i*

$z_{ij}c_{ij}$= actual expenditure in zone *z* in year *j*

B_j^{min}=minimum capital expenditure for total network in year *j*

B_j^{max}=maximum capital expenditure for total network in year *j*

This optimization problem is solved using the OptQuest Engine available in the Excel add-on tool, Evolver[109]. The OptQuest algorithm incorporates Tabu search, scatter search and integer programming to solve linear and non-linear problems. Tabu search allows for change in the search direction, which is not typical of genetic algorithms [84]. The engine remembers solutions that solved the constraints, and recombines them to search for new solutions. As a result, this solving algorithm is less likely to get stuck in local minimum/maximums and more likely to find a global solution.

The OptQuest Engine allows for a convergence stopping criterion, which is a number of runs without a certain percentage of improvement in the solution. Though it is less likely to produce sub-optimal results, and replicability rate of results is high, the sensitivity of the model still needs to be examined. To further insure a more optimal solution, the engine is run 10 times for each evaluation. The best result from all runs is selected as the optimal solution

5.5 Case Study and Analysis

The following sections describe a case study of this framework evaluated on assets within a pressure zone of Utility A. Pressure zones are separated areas in the network with common elevation. Pipe networks within these areas are constructed to maintain pressure targets and operate independently from networks in other pressure zones [110]. For these reasons, the operational characteristics of networks can vary dramatically across zones. In considering maintenance and replacement programs to mitigate the impacts of future failures, one should consider projects isolated to specific pressure zones.

The pressure zone selected for consideration in this case study contains a large number of break rate distribution hot spots. Break rate distribution or cluster analysis can be used to identify priority zones for considerations in developing maintenance and replacement programs. The zone considered in this case study recently experienced a massive line break and outage and contains a concentration of critical customers. A main break occurred on a catastrophic line which distributes water in the pressure zone from a pumping station. The consequence of this break included temporary water outages for over 15,000 residents, closure of local businesses including a large shopping mall, and several residents being without water for a week. The mitigation of future failures in this zone is important to the public's perception of level of service.

5.5.1 Identifying Work Zones

Using the Fishnet tool in ArcGIS, a grid is overlaid on the network. The intersect tool is used to assign a bin ID to each asset in the network. The bin size for the case study is 5,000 ft by 5,000 ft (1.5 km by 1.5 km).

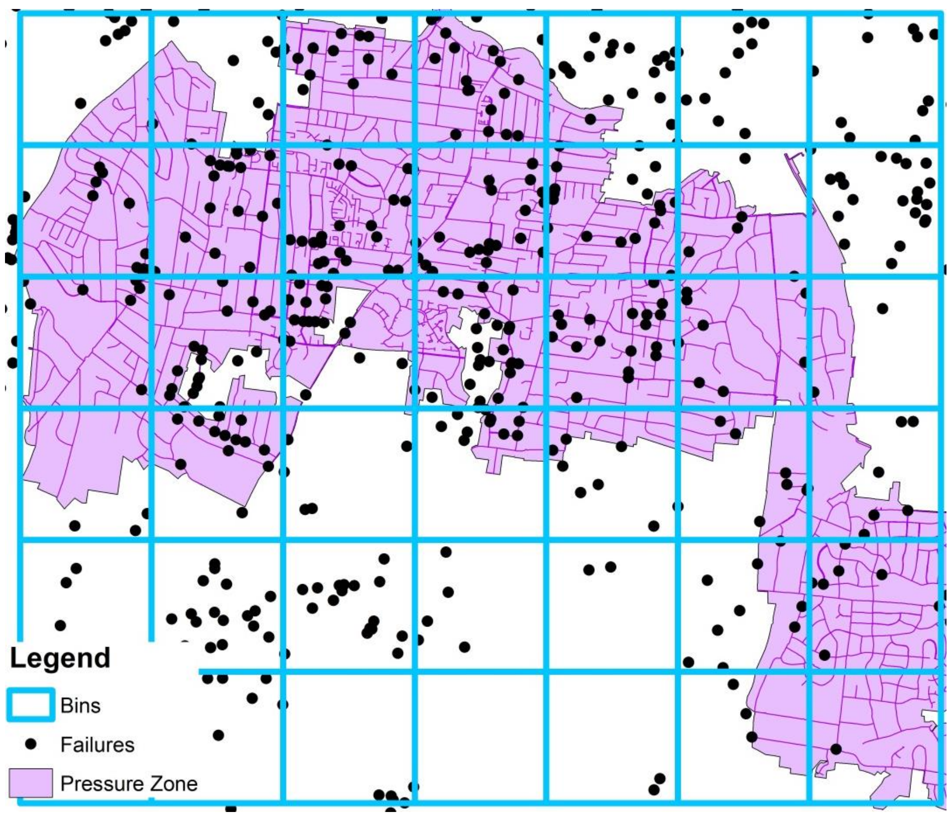


FIGURE 5.2: PRESSURE ZONE AND BINS

For each year in the three year evaluation period, the constraints in Table 5.1 are used. It is assumed that once work has been done in the zone, no more work will be done in that area for the remainder of the planning horizon. As a result, the decision variables are further condensed as the evaluation progresses and zones are no longer included in the analysis.

Table 5.1: Subset Constraints

Constraint	**Value**
B_j^{min}	\$500,000
B_j^{max}	\$1,000,000
z_j^{min}	1
z_j^{max}	2
c_{ij}^{min}	\$100,000
c_{ij}^{max}	\$1,000,000

The zones selected for prioritization efforts from the best solution and the percent annual budget allocated in each zone are shown in Figure 5.3, which includes the zones and pipes overlaid on the break rate distribution map. The cumulative cost benefit ratio for a $3 million investment is 25.16. The optimization was performed with convergence stopping criterion of 20,000 simulations with less than 0.01 percent improvement. To insure the most optimal solution was selected, the optimization routine was run ten times. The run time for this study was approximately 20 minutes per run on a PC with a 2.80 GHz processor with 8 GB RAM.

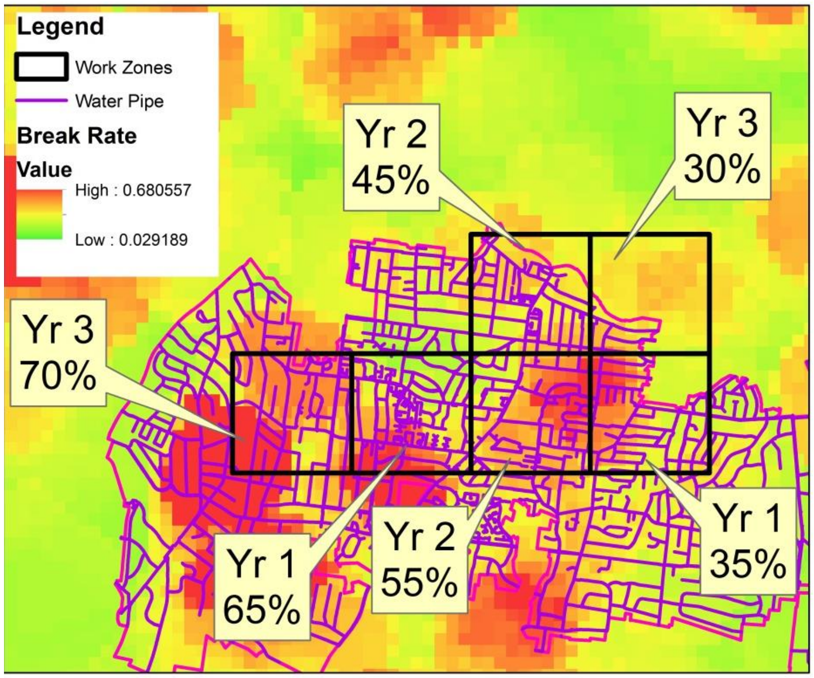


FIGURE 5.3: OPTIMIZATION RESULTS

The zones selected for prioritization efforts are shown in Figure 5.3. To demonstrate how this method can be used to prioritize assets on a street level, the pipes with the pipes with the highest CR values for the right two southwest zones are shown in Figure 5.4.

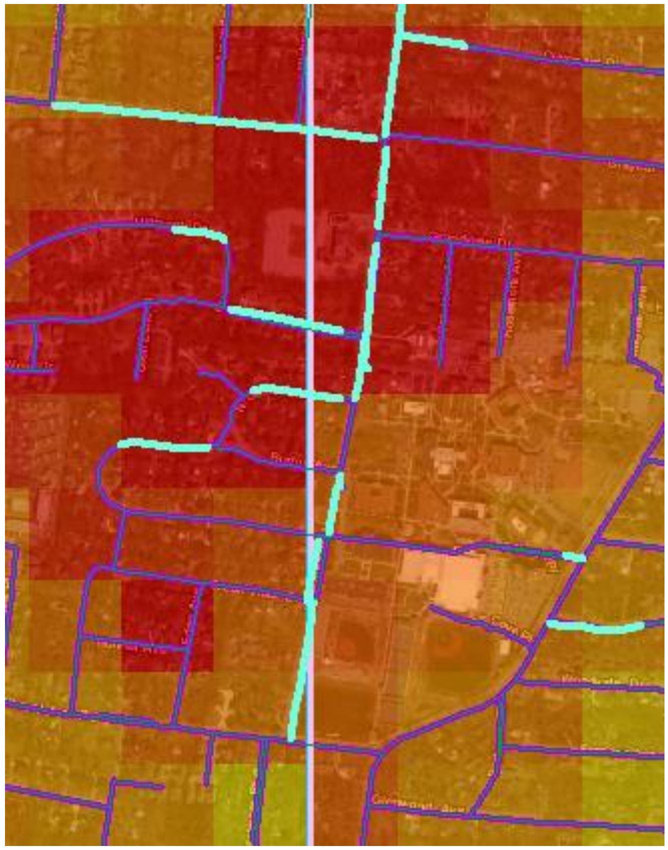


FIGURE 5.4: HIGH PRIORITY PIPES

The white line segmenting the picture is a zone division. Aerial photograph has been turned on in the background to allow for reference scale visualization. Once plotting the assets, the decision maker can start evaluating the connectivity of high priority pipes and start designing a replacement project. One can see the decision maker still has trade-off decisions to make in prioritizing pipes for condition assessment and/or replacement (see for example the two southwest pipe segments), but decisions can now be made on the street/neighborhood level, which is representative of such projects. Additionally, expert opinion may dictate the re-ordering of prioritization zones. For example, the zones shown above were scheduled for replacement activities in subsequent years. Though some optimality might be lost, obviously replacement work would be done in these zones at the same time. This highlights again that not this framework, or any framework, can replace expert opinion and analysis. The framework presented is a tool that can be used to help decision makers narrow down alternatives and make well-informed asset management decisions.

5.5.2 Impact of Binning

To fully evaluate the benefits and trade-offs of the framework presented, a separate optimization routine which does not consider binning of assets was performed in order to examine the effects of binning on asset prioritization and the maximum cost/benefit ratio. For simple comparison purposes, this algorithm was considered run for year one of the planning horizon and compared to the binning optimization results for year one. The CR curve as a function of capital for all assets in year one is shown in Figure5.5.

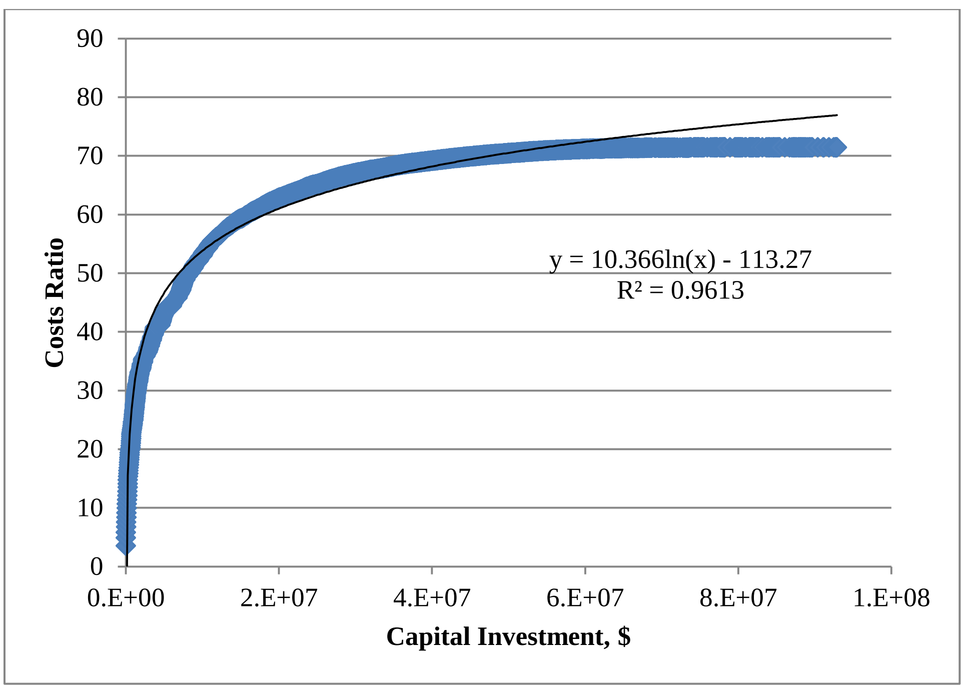


FIGURE 5.5: CR CURVE FOR ALL ASSETS IN YEAR 1

No changes were made to the budgetary constraints shown in Table 5.2. The differences in B/C ratios are noted in Table 5.3.

Table 5.2: Alternative Framework Constraints

Constraint	Value
B_{min}	\$100,000
B_{max}	\$1,000,000

Table 5.3: B/C Comparison

Framework	B/C Ratio
Binning	25.16
Individual Assets	37.9

Figure 5.6 shows assets selected for replacement or condition assessment using the binning method versus evaluating the costs ratio at the individual asset level.

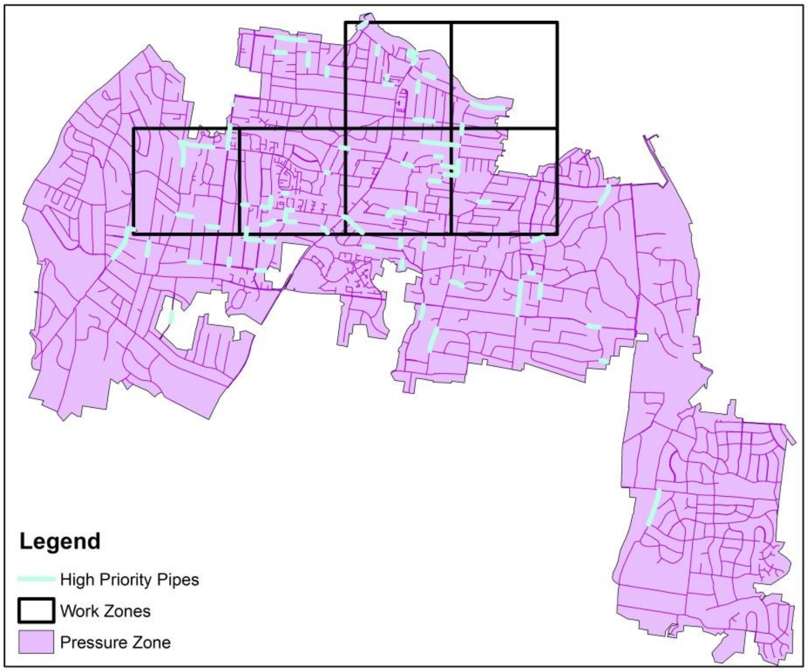


FIGURE 5.6 INDIVIDUAL ASSETS COMPARED TO ZONES

Though there is agreement between the high priority pipes selected from the alternative framework and the work zones, once can see that there is still a large percentage of pipes that fall outside of these zones, making project selection difficult. Though optimality suffers, the binning method is a better approach for prioritization.

5.6 Conclusions

This chapter presented a risk-based approach for prioritizing assets considering both the likelihood failure, quantified using survival models presented in Chapter 4, and consequence of failure using an expert opinion based weighted methodology. An optimization routine was presented to assist decision makers in selecting regions to undergo replacement programs. This framework differs from those presented in literature as it helps prioritize pipes considering the actual operations of a utility, and the size of replacement projects which typically include replacement of 0.25 mile (0.4 km) to 1 mile (1.6 km) of pipe. The sensitivity of bin size and shape on the optimization res

This methodology was compared to a framework that did not consider restricting work to subsets of the network. The comparison shows that the costs ratio is higher when considering prioritization of zones rather than individual assets, which is to be expected. Also expected was the wide spatial distribution of priority assets across the network. Utility operations dictate that the most optimal solutions through replacement will not be made due to the large spatial distribution of assets. The framework decision shows how to maximize costs ratios while considering how a utility actually designs, procures, and prioritizes pipe replacement or condition assessment projects.

CHAPTER VI – INVESTIGATION OF TRANSFER TECHNIQUES FOR WATER PIPE FAILURE PREDICTION MODELS FOR MEDIUM AND SMALL UTILITIES

6.1 Introduction

As water distribution pipes reach the end of their useful lives, the long-term costs of maintaining and operating water distribution networks are increasing. A recent study by the Environmental Protection Agency (EPA) suggests that the 20-year cost of maintaining water distribution assets will exceed $384 Billion [4]. Approximately 65% of the funding needs are expected to come from medium and small sized utilities serving less than 100,000 persons.

Though many statistical tools have been introduced to assist in the long-term planning and budget forecasting for water distribution linear assets, most of these tools have been demonstrated using data from large utilities with sophisticated asset management databases and maintenance inventories. Developing failure prediction models for medium and small utilities can be challenging due to limited number of recorded failures, which are needed to train statistical models. Also, medium and small utilities typically have less sophisticated asset management databases with fewer fields describing the physical properties of the pipe segment and the environmental conditions around the pipe.

Hypothesized solutions for the problems associated with limited data associated with medium and small utilities are rooted in model transfer theory. These methods assume that the parameters from a model trained on extensive data can be transferred to the smaller model, in order to improve prediction performance; information gained from a larger utility is useful to the smaller utility. These potential solutions have not been thoroughly examined. Investigating the potential spatial transferability of models could lead to a solution for the problems associated with training survival models for medium and small utilities with limited data sets.

This paper explores if model transfer can solve the problems associated with limited data for small and medium utilities using the Weibull Proportional Hazard Model as a base failure prediction model for one large utility and two medium sized, suburban utilities in the southeast United States. The three utilities are described in Chapter 3. Three model transfer techniques are evaluated for improvement in prediction performance using standard validation metrics used for pipe replacement prioritization models presented in literature.

The paper is organized as follows: A review of prioritization models and the development of models for medium and small utilities is introduced. The data for the utilities is described. The methodology for the model calibration and transfer are detailed. Model calibration and validation results are discussed. Conclusions and recommendations are made.

6.2 Background

Though numerous pipeline failure prediction models have been introduced over the past few decades, few have been demonstrated on medium and small utilities. Wood and Lence [11] investigate using time-linear and time-exponential models on various pipe groupings to estimate a cumulative number of breaks for small utilities. Though the models did reveal the most significant pipe parameters when considering break potential, model results were mixed, with some pipe groupings having high percent errors in prediction performance. Additionally, the methodology was demonstrated on only one utility.

Toumbou et al. [51] used a Weibull-Exponential-Exponential (WEE) to model pipe breaks for a small Canadian utility. Parameterized and non-parameterized forms of the WEE model accurately predicted the long-term break trend. Information regarding a holdout sample for model validation was not given. Additionally, the database for the utility spanned over twenty years with approximately one thousand failure records, which is comparable in record size to data sets from large utilities.

Xu et al. [70] demonstrate the use of evolutionary polynomial regression (EPR) to predict failures in a utility with brief data. Though the data set is described as brief in terms of observation period, the number of breaks is significant, ranging from 100 to 250 breaks per data set.

More sophisticated, Bayesian updating models have recently been applied to water pipeline failure prediction for a small utility in Switzerland [76]. Describing the current state of knowledge of network performance, the estimation of the prior parameter distributions for a Weibull model is obtained by estimating model parameters for three larger utilities, and then aggregating the three parameter distributions. Posterior parameter estimations for all utilities, including the small utility, are obtained through Markov Chain Monte Carlo (MCMC) sampling using the aggregated parameters from the large utilities and the data from the small utility.

Uncertainty is reduced by incorporating the data from the small utility. The authors provide guidance for validating the model, yet do not provide validation results.

Model transfer has been proposed as a potential solution to training valid models for small utilities. Common model transfer techniques include a naïve or direct transfer model, joint estimation model, and a combined transfer estimator model. Some of these methodologies have been applied or indirectly suggested by researchers for predicting water main failures in medium and small utilities with limited data, yet have not been fully validated using metrics common to the pipeline replacement prioritization models presented in literature.

The naïve transfer model assumes a direct transfer of estimators from the large utility to the small utility. Martins et al. [49] suggest using this method of model transfer for small utilities, but does not demonstrate its application. Instead, the author examines transferability across a singular network by training a model on a random selection of 50% of the data from a large utility, and validating the model using the remaining 50% of data. The validation shows that transferability of the model across the same network is viable, but does not definitively answer if the transferability of models across utilities is a valid approach.

In order to retain properties associated with the smaller network, joint estimation can be used to estimate model parameters. The data from multiple utilities is combined, and model parameters are estimated using data from all utilities. Renaud et al. [111] investigated joint estimation models when developing the prioritization software, SIROCO. The SIROCO model was developed using data from various utilities across France. The combined or amalgamated models were significant, but did not necessarily improve base model performance because the only utilities with the sufficient amount of records required for the Weibull based model were a medium and large utility. The researchers conclude that small utilities with less than 200 pipes would benefit from the amalgamated database approach, and would not be able to develop significant models otherwise. The joint context estimation methodology assumes the spatial homogeneity of factors influencing pipe breakage rates. Kleiner and Rajani [46] explain that non-pipe intrinsic factors such as climate can impact breakage rates in otherwise homogenous groups of pipe. The combination of these non-intrinsic factors accelerate or decelerate failure in otherwise homogenous groups of pipe. The assumptions that the groupings of these parameters have the same impact on network as the other may not be appropriate. For example, not all the data needed to capture such non-intrinsic parameters could be included in the model.

Similar to the naïve transfer model application demonstrated by Martins et al. [49], Savic et al. [112] demonstrate the applicability of joint estimation by examining model performance at individual zones across a network. Using data aggregated from over forty water quality zones within a utility, evolutionary polynomial regression (EPR) is used to formulate break rate prediction equations. The prediction performance in each water quality zones within the distribution network is evaluated using the coefficient of determination metric. The EPR model was valid, and the model is spatially transferable across regions in the same network, providing a proof of concept of joint context estimation, though not explicitly described. The authors determined that the EPR model successfully provides insight to the underlying physics of a failure model, and produces parsimonious models that are less likely to be fit to data noise.

Again, joint context estimation of the model was not explicitly described, nor was joint context estimation demonstrated with this model using data from more than one utility.

A third model transfer technique that has not been applied to pipeline survival models is the combined transfer estimation (CTE) model proposed by Ben-akiva and Bolduc [113]. CTE is a less sophisticated Bayesian model that uses weighted averaging to account for bias associated with transferring parameters from the large utility, known as the transfer region, to the small utility, known as the application region. This method has been demonstrated with much success for transportation planning problems [114], [115]. CTE method also allows for the asset manager to select which parameters from the large utility to transfer to the small utility, and which parameters to excuse. For example, a physical property such as pipe diameter might be transferable, while parameters describing the clustering of accidents are likely network specific and cannot be transferred across time and space.

Though model transfer techniques have been theorized and demonstrated in limited capacity, they need to be tested on more medium and small utilities. Additionally, the potential improvements in model transfer gained by the CTE method needs to be explored. Lastly, more thorough validation techniques need to be utilized to understand prediction performance at the asset level.

6.2 Model Parameters

The average break rate serves as a surrogate for many other parameters including soil conditions, pressure changes, and traffic levels above the pipe. Other model parameters considered include pipe diameter and installation date. Two binary variables, *Assumed Install* and *Assumed Mat* are used to account for assumptions made with respect to pipe installation date and material. These variables are described in Chapter 4.

6.3 Methodology

First, survival models are estimated for the data from each utility. These models are stratified by material and the number of observed failures. When the number of previous known failures (NOPKF) is zero, a parametric survival model is calibrated. When NOPKF is greater than zero, a non-parametric survival model is calibrated.

Monte Carlo simulation is used to estimate predicted failures within a specified observation period, using the estimated survival functions. The prediction performance of the models is evaluated using two validation metrics. A baseline comparison between the average expected number of failures and observed failures within a time window is made. Then the decile ranking of the predicted failures versus the observed failures is analyzed. Model calibration and validation are described in detail in this section. Models transfer techniques are employed, and prediction performance is compared using the validation metrics described.

6.3.1 Model Calibration

Data is analyzed from Utilities A, B, and C, previously described. Base statistics for all three networks are calculated. Parameter estimates for a base model are estimated for all networks using a randomly selected 80% of data, with 20% reserved as a holdout sample for validation.

The base model selected is a parametric Weibull Hazard Rate Model (WHRM) introduced for predicting pipe failures described in detail in Chapter 4. Due to the limited number of recorded multiple failures, a non-parametric WHRM is used to model subsequent failures. Model parameter estimates are obtained through the maximum likelihood estimation method.

The survival models for each utility are stratified by material. The parameters included for each of these models include diameter, installation date, assumed installation date, assumed material, and local break rate. Parameter estimates are obtained using maximum log likelihood estimation, and only significant parameters with p-values less than 0.05 are retained.

Using the parameter estimates, several methods for model transferability are tested. First, the naïve transfer method is evaluated by using the parameter estimates from Network A to estimate failures in Networks B and C. Let the subscript *i* refer to Utility A and subscript *j* refer to the suburban utilities. The naïve transfer method assumes that the Equation 6.1 survival function regression parameters estimated for Utility A are directly transferable to Utilities B and C.

$$\boldsymbol{\beta_j} = \boldsymbol{\beta_i} \qquad [6.3]$$

$$\sigma_j = \sigma_i \qquad [6.4]$$

Where β_i is a vector of regression parameters from Utility A, $\beta_i$is a vector of Weibull regression parameters for the suburban utilities, σ_i is the Weibull scale parameter for Utility A, and σ_j is the Weibull scale parameter for the transfer region(s).

Next, joint context estimation for transferability of model parameters is explored. The data from all three utilities is aggregated. The explanatory variables and failure times for the three utilities are aggregated as shown below:

$$\boldsymbol{z} = \begin{bmatrix} \boldsymbol{z_i} \\ \boldsymbol{z_j} \end{bmatrix} \qquad [6.5]$$

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{T_i} \\ \boldsymbol{T_j} \end{bmatrix} \qquad [6.6]$$

Where $\boldsymbol{z}$ is a vector of explanatory variables and $\boldsymbol{T}$ is a vector of failure times. Survival function parameters are then estimated using these vectors.

Lastly, the feasibility of CTE method is explored. The CTE methodology introduced by Ben-akiva and Bolduc [113] investigates the variance, or transfer bias, between parameter estimates for the application region and the transfer region by evaluating the Mean Square Error of the combined estimator. A generalization of the Bayesian Updating method [115], this method allows for greater contribution from the estimation region when the transfer bias is low, and a decreased contribution when the transfer bias is high. This transfer method could be particularly beneficial in the development of pipeline survival models as operational, environmental and maintenances differences vary across utilities, causing increases or decreases in model performance, which would impact regression parameters for survival models. The combined transfer regression parameter is computed as follows:

$$\boldsymbol{\theta}_{CTE} = \left((\boldsymbol{\Sigma}_i^{-1} + \Delta\Delta^{\mathrm{T}})^{-1} + \boldsymbol{\Sigma}_j^{-1}\right)^{-1} + \left((\boldsymbol{\Sigma}_i^{-1} + \Delta\Delta^{\mathrm{T}})^{-1} + \boldsymbol{\theta}_i + \boldsymbol{\Sigma}_j^{-1}\theta_j\right) \qquad [6.7]$$

Where:

$\boldsymbol{\theta}$ is a vector of Weibull regression parameters, $\begin{bmatrix}\boldsymbol{\beta}\\ \sigma\end{bmatrix}$

$\boldsymbol{\theta}_{CTE}$ is the transferred parameters of the suburban utility

$\boldsymbol{\theta}_i$ is the estimated parameters of Utility A, the estimation region

$\boldsymbol{\theta}_j$ is the estimated parameters of the suburban utility, the application region

$\boldsymbol{\Sigma}_i$ is the covariance matrix of the regression parameters for the estimation region

$\boldsymbol{\Sigma}_j$ is the covariance matrix of the regression parameters for the application region

Δ is the transfer bias, $(\boldsymbol{\theta}_i - \boldsymbol{\theta}_j)$

$\Delta^{\mathrm{T}}$is the transposed transfer bias matrix

It should be noted that though the CTE method is dependent upon weighted combined transfer estimator covariates, the weights selected are solely dependent upon the differences in the regression parameters. The differences due to the number of training points in the transfer and estimation regions is not explicitly accounted for in the CTE weighting method, or any of the transfer methods presented above.

Model Validation

Monte Carlo simulation is used to simulate failures from the survival function, using the estimated parameters. A survival probability is selected at random, and the corresponding survival time is computed. If the survival time is less than the duration of the asset management database, or time horizon under consideration, then a failure is recorded. Over 1000 simulations are performed, and the average number of predicted failures is reported.

Two validation metrics commonly presented in literature are utilized. The first is a baseline comparison of the observed failures within the observation period, and the average expected failures simulated from the survival models using Monte Carlo simulation.

The second validation metric is rank order validation introduced and utilized in Chapter 4.

6.4 Results

6.4.1 Base Model

Separate models were calibrated for all materials in each Utility. The calibration and validation results for these models are shown in Tables 6.1 and 6.2. The tables comparing the predicted failures to observed failures show reasonable agreement, suggesting that the models for all three utilities are statistically valid, yet the rank order validation graphs show otherwise.

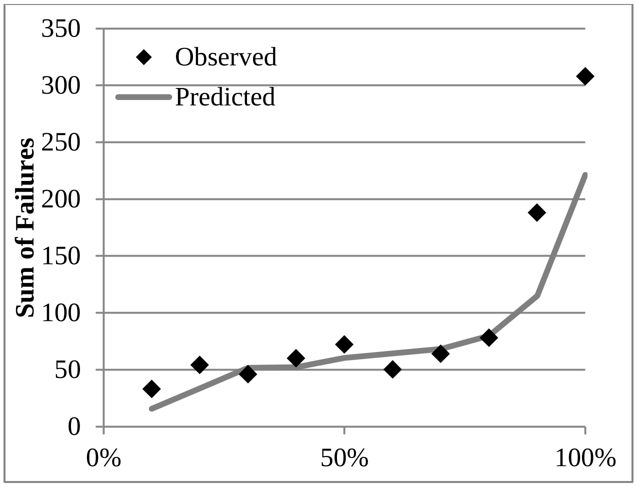
Table 6.1 Model Calibration Results

Utility	B		C			
Material	DI		CU		DI	
Model	NOKPF=0	NOKPF>0	NOKPF=0	NOKPF>0	NOKPF=0	NOKPF>0
Intercept	6.739164	6.109009	4.0704	0.2783	7.036589	0.25432
Scale	0.925587	0.926909	1.147384	1.3545	1.190347	1.15607
Parameters						
Diameter	--	--	--	--	-0.09805	--
Installation Date	--	--	--	--	--	--
Assumed Installation Date	--	--	--	--	--	--
Assumed Material	--	--	--	--	27.817	--
Average Break Rate	-0.00843	--	--	--	--	--

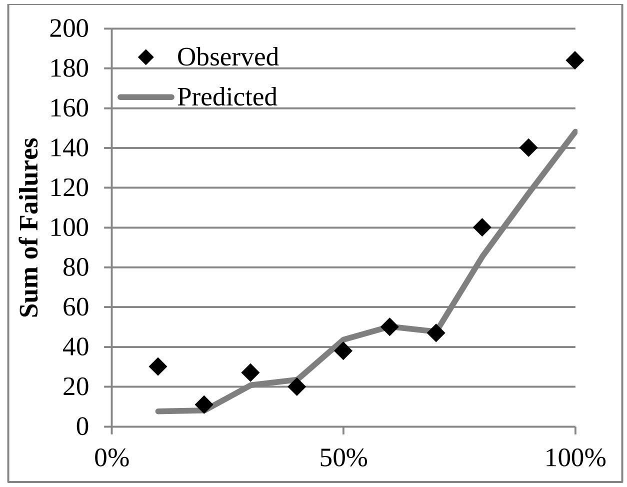
Table 6.2 Prediction Performance Results

	Predicted Failures				
Model	Base Model	Transfer	Amalgamated Model	Joint Context Transfer	**Observed Failures**
Utility A, DI	554	--	--	--	668
Utility A, CU	755	--	--	--	953
Utility B, DI	17	16	10	16	20
Utility C, DI	47	37	28	47	54
Utility C, CU	65	53	42	65	71

The model validation graphs for the base models for Utility A, shown in Figure3, suggest that the base models are valid, as there is reasonable agreement between predicted and observed deciles and a marked jump in the eighth through tenth deciles, representing pipe cohorts with the greatest probability of failure. One should note that the prediction performance for CU pipe was better than DI, with better realization of the highest risk pipes with respect to likelihood of failure.



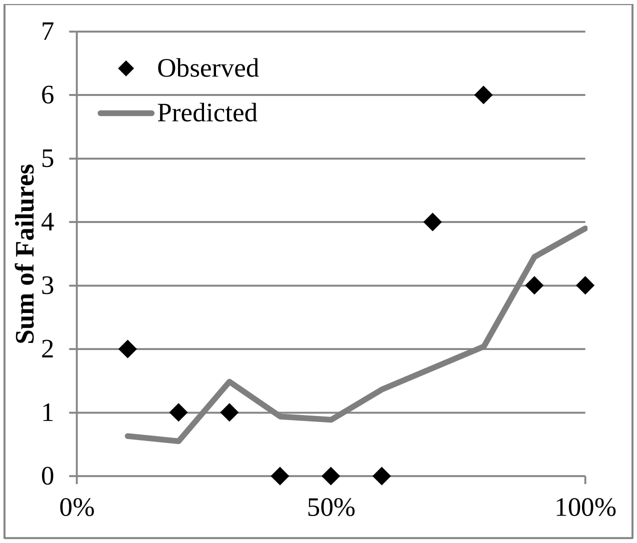
(a) CU Base Model

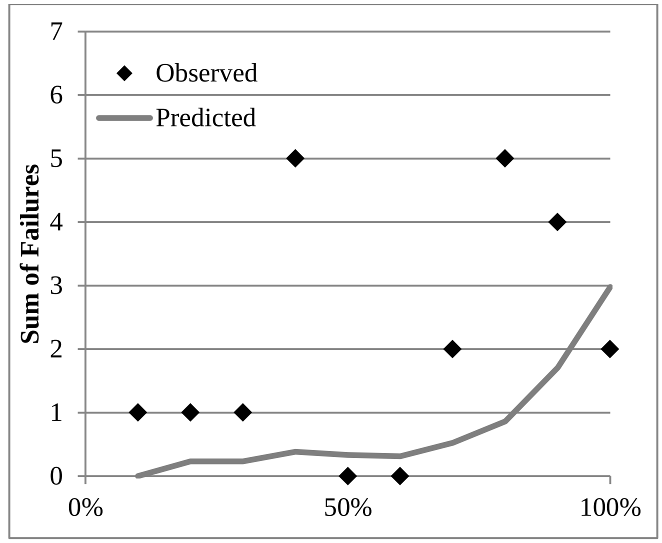


(b) DI Base Model

FIGURE 6.1: UTILITY A VALIDATION

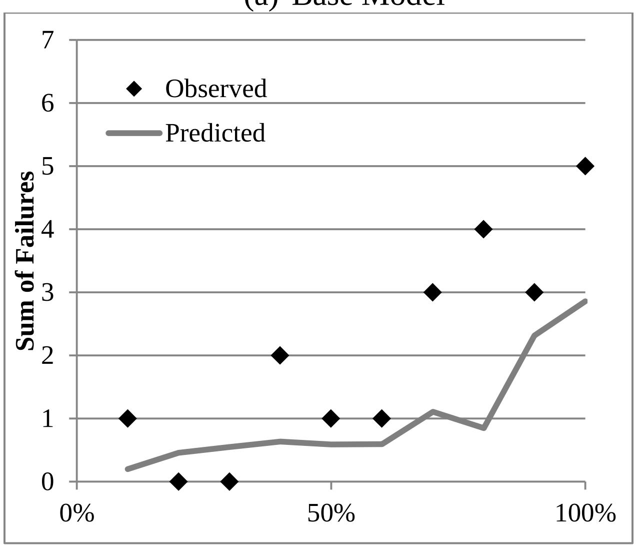
In contrast, Figures 6.2a to Figure 6.5a show the validation charts for models calibrated for Utilities B and C. The jaggedness of the predicted failure lines correlating with the disagreement between predicted and observed deciles shows that the survival models are not valid. Though the models perform reasonably well at predicting the appropriate number of failures, they do not accurately predict failures in the correct classes of pipe.

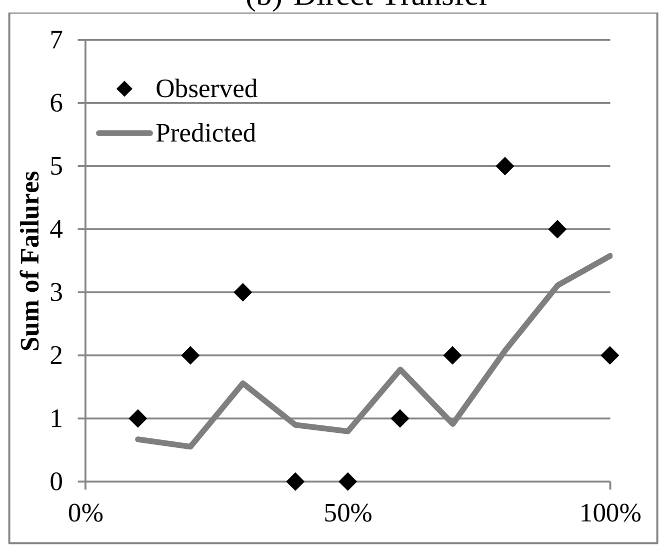




(a) Base Model

(b) Direct Transfer

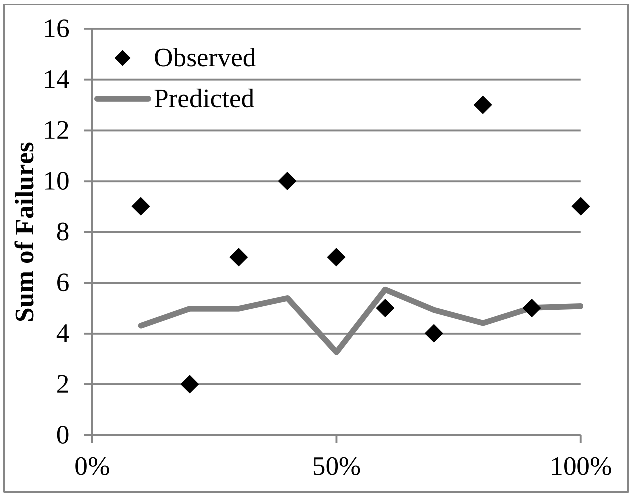




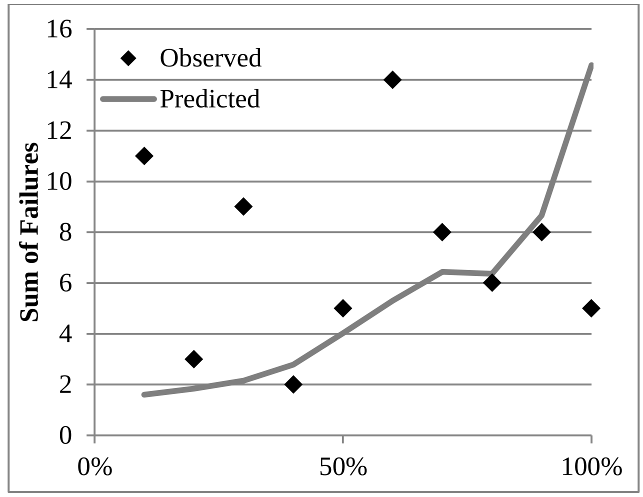
(c) Joint Context Estimation

(d) Combined Transfer Estimation

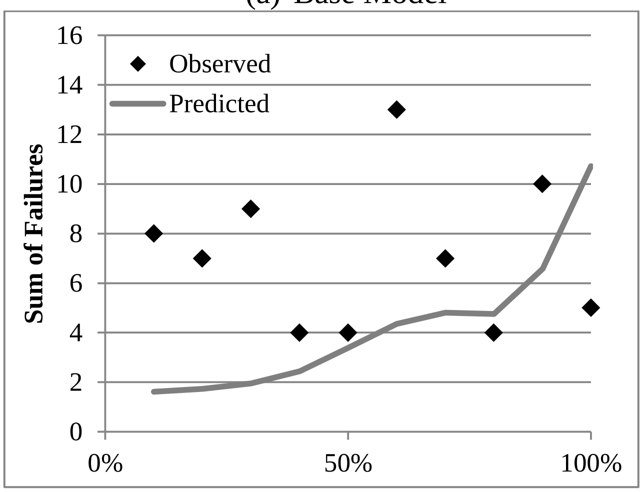
FIGURE 6.2: UTILITY B – DI PIPE VALIDATION



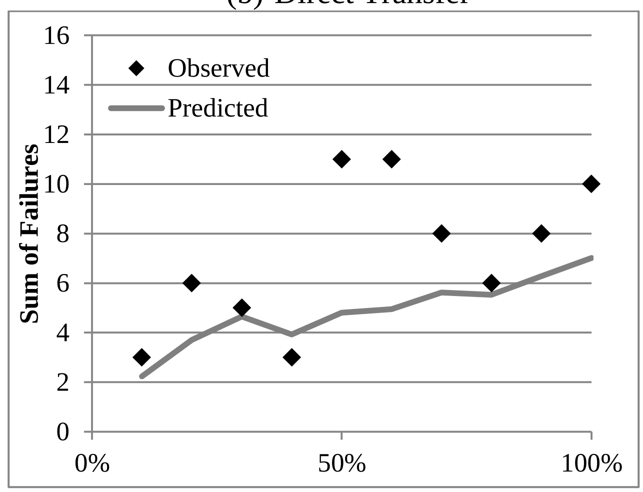
(a) Base Model



(b) Direct Transfer

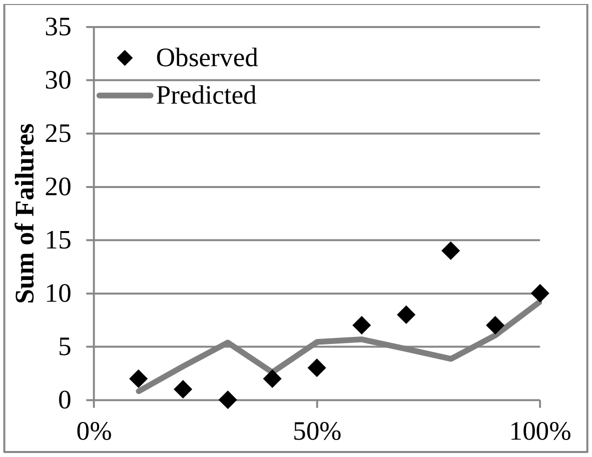


(c) Joint Context Estimation

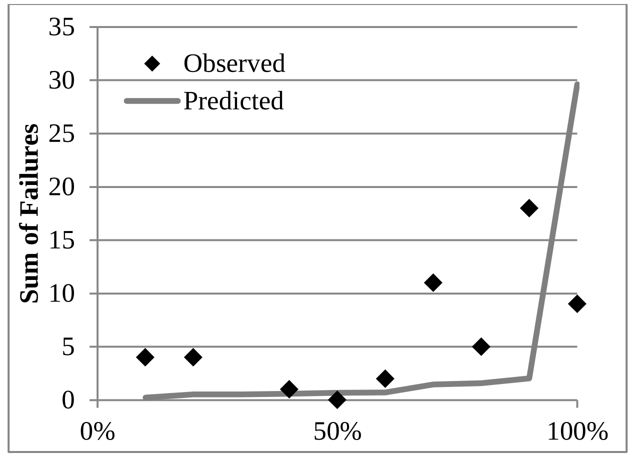


(d) Combined Transfer Estimator

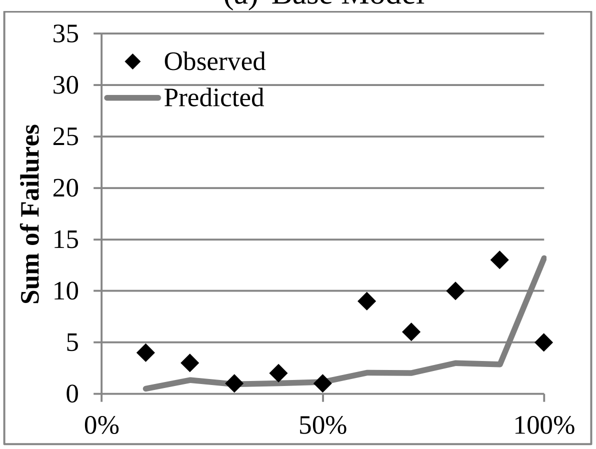
FIGURE 6.3: UTILITY C—CU PIPE VALIDATION



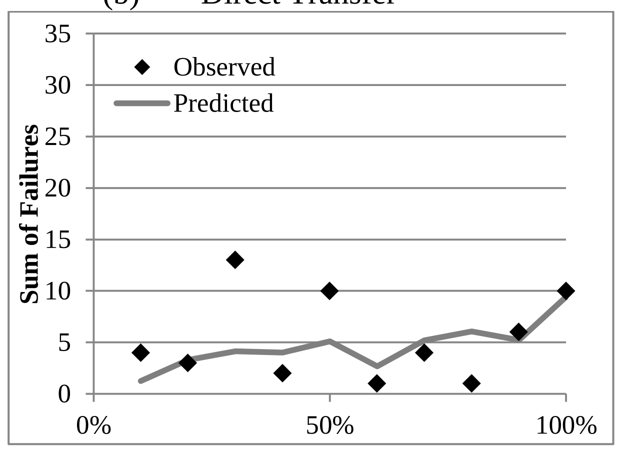
(a) Base Model



(b) Direct Transfer



(c) Joint Context Estimation



(d) Combined Transfer Estimator

FIGURE 6.4: UTILITY C—DI PIPE VALIDATION

6.4.2 Direct Transfer Models

The direct transfer models result in worse performance than the base models, evidenced by both the base comparison to the observed failures and the cohort analysis of predicted versus observed failures. Potential reasons for the lack of direct transferability of the models could be explained by differences in the data sets. The conclusion that the model parameter information from the larger utilities is not valuable to the smaller utilities cannot yet be made. The direct transfer method ignores any differences in the utilities, and the importance of the data from the smaller utilities. The model transfer methods explored below allow for contribution from the smaller utility in the survival model.

6.4.3 Joint Context Estimation Models

Table 6.3 shows the joint context model estimates, which are very different from both the parameters estimated for Utility A and the parameters estimated for the smaller utilities. The joint context estimation models could not pass validation tests for the small utilities. The data is being fit to the larger utility, and the small utility influences the estimates, but serves more as noise. The CTE methodology allows for greater influence of the smaller data set on the regression parameters, by minimizing the bias between the parameters for the large data set and the smaller data set.

Table 6.3. Joint Context Parameter Estimates

Utility	**A,B,C**		**A,C**	
Material	**DI**		**CU**	
Model	NOKPF=0	NOKPF>0	NOKPF=0	NOKPF>0
Intercept	-42.0592	0.596822548	35.03107	0.716367
Scale	0.955053	1.063302848	0.99486	1.005168
Parameters				
Pipe Diameter	0.081981	--	0.141488	--
Installation Date	0.024444	--	-0.01568	--
Assumed Installation Date	-0.90639	--	-1.0493	--
Assumed Material	-0.39358	--	1.872616	--
Average Break Rate	-0.0022	--	-0.00278	--

6.4.4 Combined Transfer Estimator Models

Even with the bias minimization technique, the CTE method did not result in improved prediction performance. The variance between the model parameters for the utilities was so large, that the contribution from the large utility was minimal. Table 6.4 shows the model parameters which are only minimally different from the base model parameters.

Table 6.4 Combined Transfer Estimator Parameters

Utility	B		C			
Material	DI		CL		DI	
Model	NOKPF=0	NOKPF>0	NOKPF=0	NOKPF>0	NOKPF=0	NOKPF>0
Intercept	6.728369	6.07507	1.263821	0.397665	7.036349	0.38682
Scale	0.92356	0.92141	0.700301	1.276072	1.190305	1.11641
Parameters						
Diameter	--	--	--	--	-0.09804	--
Installation Date	--	--	--	--	--	--
Assumed Installation Date	--	--	--	--	--	--
Assumed Material	--	--	--	--	27.817	--
Average Break Rate	-0.00839	--	--	--	--	--

6.4.5 Discussion

Statistically valid models for DI and CU pipes from Utilities B and C could not be developed using the prescribed model form, due to lack of recorded failures. Consultants has previously emphasized to utilities that at least 3 to 5 years of asset management data is needed to develop valid statistical models [19], or utilities must have more than 200 pipes [111]. The utilities studied met both of these requirements, yet valid models could not be produced. Evidence of this observation is given by model results for PVC pipe in Utility B, which was not evaluated for model transferability, due to the lack of prevalence of this material in the other networks. With over 100 recorded failures, a valid model for PVC pipe was produced. More research is needed to examine the minimum number of breaks required to produce valid statistic models.

Additionally, research is needed to discern if machine learning models are more appropriate for predicting pipe failures in medium and small utilities with little variance in pipe-intrinsic properties. This is especially true with respect to DI pipe in utility B, which is very new to the network. Most of this pipe was installed within the past ten years. The Weibull-based survival model is appropriate for predicting the accelerated failure rate at the end of life of an infrastructure asset, yet does not adequately predict infant mortality rate, or the premature failures of a pipe. Given that DI pipe has a design life of over 100 years [116], one can infer that DI pipe in this network still in infant mortality or early wear out stage, and the Weibull model is probably not an appropriate model form for DI pipe in this network; whereas PVC is reaching the end of its useful life and the Weibull model accurately describes the increased failure rate.

Several observations can be drawn from the model transferability investigations. The validation metrics used for Utility A show that models are spatially transferable across the same network. Contrary to suggestions made in literature, model transferability is not always demonstrated across other similar networks. Potential reasons for the lack of transferability of the models including differing construction practices, environmental conditions, and operating conditions. Though the differences in construction practices and environmental conditions were minimized by analyzing three neighboring utilities, operational and maintenance differences still exist. With respect to operation, the number of pressure zones, maximum operating pressure, and pressure variance from low to high vary for all three utilities. Varying pressure is cause of breaks, as water hammers can degrade the structural integrity of pipes.

Also contributing to network reliability, replacement and maintenance activities differ amongst utilities. For example, differences in maintenance activities such as line flushing programs can cause an increase or decrease the longevity of pipes. Also, the aggressiveness of pipe replacement programs post failures additionally influences the overall level of service for the network, as new pipes are often expected to perform better than repaired pipes.

Differences in digital asset management could also explain the lack of model transferability. Data quality is a concern for all utilities, but the sources and levels of uncertainty in both the distribution system properties and failure records can influence transferability. Mentioned previously, three utilities referenced in this study all maintained failure records differently, yet each data set was lacking basic information regarding pipe material, installation date, and/or location. Pipe segmentation within the GIS model differed for all utilities, which is another possible explanation for the lack of transferability. Since pipe definitions in GIS vary and do not always reflect the actual pipe segments, grouping pipes into cohorts prior to analysis can have an impact on the predictive performance of modes [18]. Though pipe length is not included, segmentation can inflate the number of pipes with properties that influence regression parameter estimates.

6.5 CONCLUSION

The results of the analysis of data from three utilities first show that record keeping and GIS models vary significantly. These differences might be one explanation for the lack of failure model transferability. This paper investigated survival model transfer techniques suggested in literature as ways of developing valid survival models for medium and small utilities with limited recorded failure data. In addition, a model transfer technique not yet suggested as a solution to developing survival models was investigated. The results of the study suggest that pipeline survival models are not transferable across utilities, even utilities with similar environmental characteristics, contractors, and material sourcing. The results of this study show that further research is needed to investigate the minimum number of recorded failures needed to train a statistically significant, regression based survival model. For failure sets with too few data to train survival models, more case studies of applications of pipe prioritization frameworks including machine based learning models are needed.

CHAPTER VII – RISK-BASED OPTIMIZATION OF MR&R ACTIVITIES FOR MEDIUM AND SMALL UTILITIES

7.1 Introduction

The previous chapter examined using Weibull based models to predict pipe failures in medium and small utilities. The results of these efforts showed that even when exploring the use of model transfer, valid prediction models could not be developed using the specified WHRM form. Alternative, machine learning based pipeline MR&R prioritization frameworks need to be developed for medium and small utilities with sparse data. When developing such frameworks, the researcher needs to balance accuracy and value of information gained from machine learning models with the ease of implementation by utilities. The less computationally intensive and the more intuitive the model, the more likely it is to be adopted by utilities.

Though several machine learning models have been introduced to prioritize pipe replacement, these models are complex, require advanced knowledge to implement. Validation metrics show the accuracy of some of these models is less than desirable.

As an alternative to other computationally intensive models, researchers have implemented and improved clustering algorithms to identify network regions with higher than average breakage rates. Cluster analysis can be used to identify potential reasons for increased break rates and to identify ways to mitigate future failures.

This chapter presents a background into machine learning models and clustering algorithms. A methodology for identifying clusters using the algorithm DBSCAN and heuristic knowledge of network connectivity is described. A case study is presented that synthesizes criticality analyses and cost estimates with cluster analysis to prioritize MR&R projects for Utility B. In the final section, conclusions, recommendations, and limitations are discussed.

7.2 Background

Though statistical pipeline performance prediction models have been utilized extensively over the past several decades, research is moving towards machine learning models. Though statistical models are easier to analyze with respect to visualizing the impact of parameters on the overall performance of pipe, machine learning models are often more accurate and better model the underlying failure process because they include fewer assumptions about the model structure [74]. Machine learning models can take various forms including ANN, fuzzy sets, Bayesian updating models, and data mining based models. Each has been utilized in varying capacity for prioritizing water and sewer infrastructure for replacement.

The review of both statistical and machine learning models in Chapter 2 shows that pipe break rate or previous breaks and pipe age are the most important parameters in determining future failures. Cluster analysis models can be used to identify areas of high break rate, and attributes that may play a role in the elevated breakage rate. Since the number of clusters in a zone is not readily realized, a clustering algorithm that is ignorant of the number of clusters is more prudent. Rather, algorithms like DBSCAN define clusters based on the number of points that define a cluster, and the search radius between points.

Summarized in Chapter 2, Oliveira et al. [53] expand the use of DBSCAN to identify clustered pipe break regions. The improved DBSCAN algorithm is used to define large clusters. Subclusters within larger clusters are determined by refining the algorithm to use a smaller search radius and minimum points to define a cluster. The subcluster analysis is not extensive, and does not investigate explanatory parameters beyond the presence of bus routes in the cluster regions.

A more detailed analysis of environmental and operational conditions within clusters can potentially lead to the identification of MR&R activities to decrease break rates in cluster zones. Additionally, these zones can be evaluated for the potential consequences associated with failures, allowing for risk-based prioritization of MR&R activities.

7.3 Methodology

The methodology presented expands upon the cluster analysis frameworks described previously, as it incorporates heuristic knowledge to refine clusters and identify MR&R activities. In order to define initial clusters, the DBSCAN algorithm is utilized. DBSCAN is dependent upon an arbitrarily defined minimum number of clusters to define a failure, *minpts*¸ and a search radius, ε. The algorithm is dependent upon the concept of density reachability. Consider a core point, p, and an alternative point, q. A cluster is formed when the minimum number of point to reach q from point p is contained within the search radius of p. Alternatively, q is density reachable to p when other points within the search area are reachable given the search radius and minimum number of points.

To identify failure clusters, the MATLAB function developed by Kovesi [117] is used. Failures coordinates are input in state plane coordinates. The search radius is defined in feet. The output of the function is the failure number a number associated with a cluster. Failures considered to be noise, outside of clusters, are assigned a failure number of 0.

The failure clusters are mapped in ArcGIS. Using knowledge of network connectivity, pipes within cluster regions are identified. In some instances, failure points identified as clusters are outside of hydraulically connected regions. These failures are subsequently excluded from the cluster regions, and reclassified as noise.

Next, the clusters are analyzed to identify attributes that might explain the elevated break rate in cluster zones. Geo-processing must be performed in order to gather some of the necessary data including estimating traffic above the pipe and assessing pressure variances. After performing this analysis, comparisons can be made between clusters and the rest of the network, using histograms.

After this analysis has been made, and potential causes of failure have been identified, projects to mitigate the risk of pipe failure are proposed. Expert opinion is used to subdivide operations large clusters into smaller, more realistic projects. The cost of each project is then estimated.

After identifying potential projects, a consequence analysis is conducted to evaluate the operational, economic, and environmental impacts of pipe failure. The total consequence score is considered as the weighted total of the categorical scores shown in Figures 7.1to Figure 7.3. This score is normalized to 1,000 L.F. based on GIS recorded pipe segment length.

Finally, a cost-benefit analysis is performed by dividing the total cost of the project by the normalized total consequence score for all pipes within a project. This cost-benefit ratio is coupled with expert analysis to prioritize projects.

Index Weight	OPERATIONAL							
	0.33							
Category Weight	Customer Impact						Traffic Impact	
	0.7						0.3	
Variable	Critical Customer		Material		Pipe Size		Road Type	
Data Source	Tennessee POI		WM_Database		WM_Database		TIGER Streets	
Data Field	Category		MATERIAL		PIPE_SIZE		Class	
Processing	Create buffer around pipe and join POI Government and Social Services points.		None		None		Spatial intersect with buffer around streets	
Valid Entries	Value	Score	Value	Score	Value	Score	Value	Score
	Intersects With Critical Customer	100	OTH	1	0.5-2.0	1	S	10
	Doesn't Intersect	0	PVC	5	2.25-6	5	M	50
			CI, CL, COPP,CU, DIP, STEE	10	8 - 14	10	I	100
			AC, CONC	100	16 - 18	50		
					20 - 24	100		

FIGURE 7.1: OPERATIONAL CONSEQUENCE SCORING

Index Weight	**ECONOMIC**					
	0.33					
Category Weight	**REPAIR COSTS**					
	1					
Variable	**Land Use**		**Material**		**Pipe Size**	
Data Source	Tenneessee POI/ TIGER Census Blocks		WM_Database		WM_Database	
Data Field	Category/Pop10		MATERIAL		PIPE_SIZE	
Processing	Intersect pipeline buffers with property polygons. Intersect Census population densities with pipe buffer		None		None	
Notes	Land use categories must be aggregated into valid entry categories.					
Valid Entries	Value	Score	Value	Score	Value	Score
	Other	1	OTH	1	0.5-2.0	1
	Park	5	PVC	5	2.25-6	5
	Residential	10	CI, CL, COPP,CU, DIP, STEE	10	8 - 14	10
	Commerical/ Industrial	25	AC, CONC	100	16 - 18	50
	High Density	100			20 - 24	100

FIGURE 7.2: ECONOMIC CONSEQUENCE SCORING

Index Weight	**ENVIRONMENTAL**					
	0.33					
Category Weight	**ENVIRONMENTAL IMPACT**					
	1					
Variable	**Critical Habitat**		**Wetlands**		**Pipe Size**	
Weight	0.5		0.3		0.2	
Data Source	US FWS Crit. Habitat		US FWS Wetlands		WM_Database	
Data Field	CRIT_HAB Poly		CONUS_Poly		Pipe Size	
Processing	Intersect pipeline buffers with habitat polygon		Intersect pipeline buffers with wetland polygon		Only used if a pipeline intersects environmental area	
Valid Entries	VALUE	SCORE	VALUE	SCORE	VALUE	SCORE
	Intersects	100	Intersects	100	0.5-2.0	1
	Doesn't Intersect	0	Doesn't Intersect	0	2.25-6	5
					8 - 14	10
					16 - 18	50
					20 - 24	100

FIGURE 7.3: ENVIRONMENTAL CONSEQUENCE SCORING

7.4 Case Study

This section presents a case study of implementing the aforementioned prioritization framework for Utility B. Over an 8 year observation period, Utility B experienced over 140 total breaks. The coordinates of these breaks were obtained through GPS shots taken at the time of repair. These failure points were aggregated and imported into MATLAB.

The DBSCAN algorithm was run under several conditions for both the *minpts* and ε values. The final evaluation of the DBSCAN algorithm was conducted with *minpts* equal to 5 and ε equal to 1,000 ft. (305 m). This evaluation resulted in 7 clusters varying in size from 5 failures to over 20 failures. The mapped cluster regions are shown in Figure 7.4.

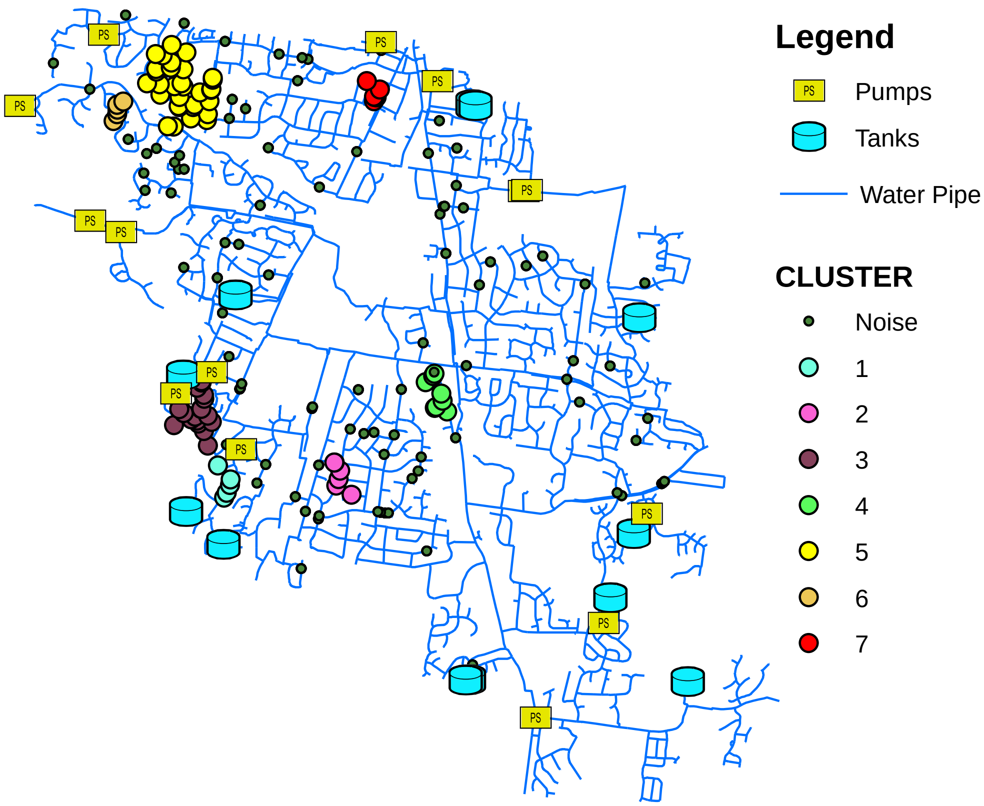


FIGURE 7.4: UTILITY B INITIAL CLUSTERS

Each cluster was visually inspected in ArcGIS for connectivity constraints. The clusters were redefined as needed. Furthermore, cluster 5 was subdivided into subclusters based on estimated project sizes. Figure 7.5 depicts an example of cluster refinement in the selection of pipes for cluster 5.

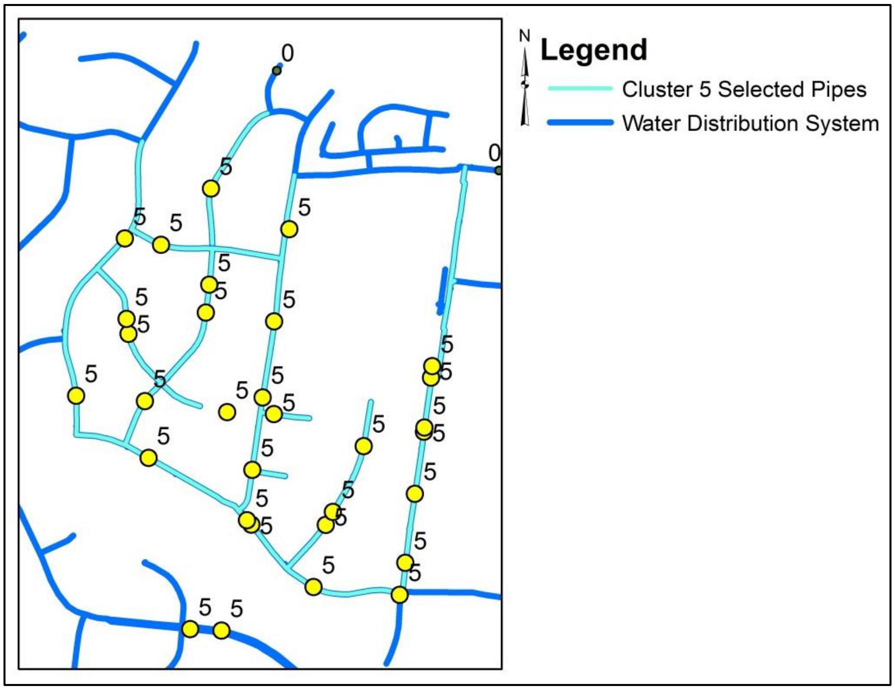
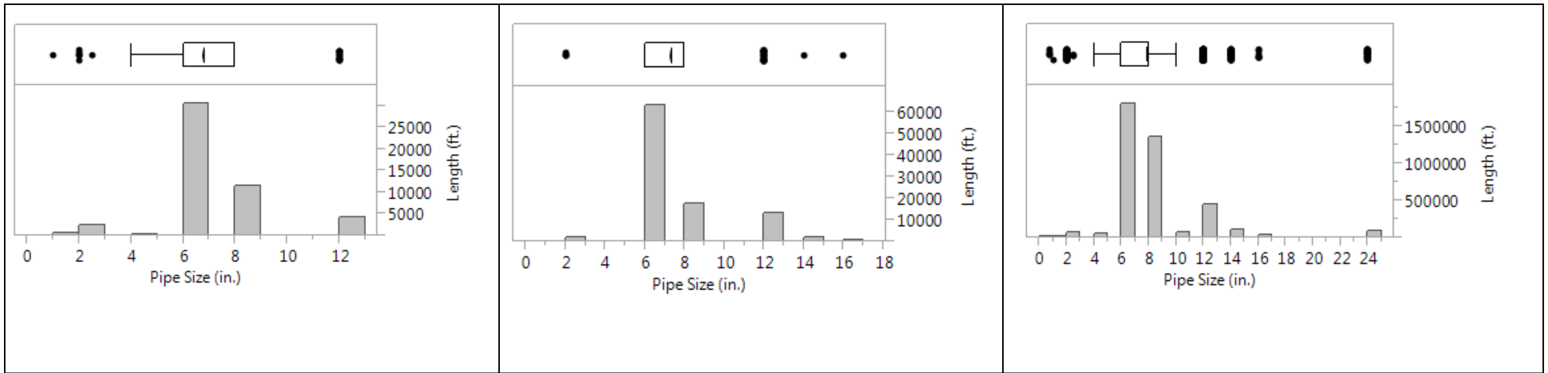


FIGURE 7.5. CLUSTER REFINEMENT AND SUB-CLUSTERING

Next, geoprocessing was performed to link the hydraulic model to the GIS model. Nodal pressures from the hydraulic model were input as points into the GIS model. A buffer was created around the pipe and spatially joined to the nodal point. This geoprocessing step allowed for the assigning of the average nodal pressure for the group of pipes between nodes. The hydraulic model nodes also contain information regarding the maximum, average, and differences in operating pressures. Knowing this hydraulic information can help the decision maker determine if pipe failures can be attributed to pressure surges. The upgrade of pump station controls could be a potential solution to mitigate pipe failures, and would be much more economical than a pipe replacement program.

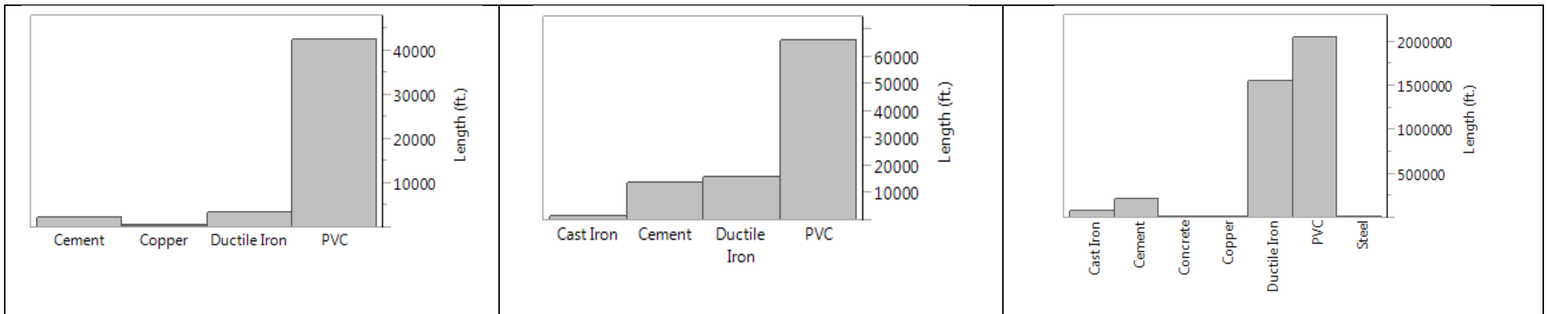
Histogram analysis was performed on the clusters, noise, and remaining network to examine potential variables influencing pipe failures. The variables investigated include pipe diameter, material, pressure difference, and velocity.

The results of the histogram analysis shown in Figure 7.6 indicate that 6 inch pipe is more likely to fail than pipes of other sizes. Additionally, PVC pipe is more likely to fail than other pipe materials. This is expected as PVC is older and comprises more of the network. Typically when replacement occurs, PVC is replaced and upsized with ductile iron pipe



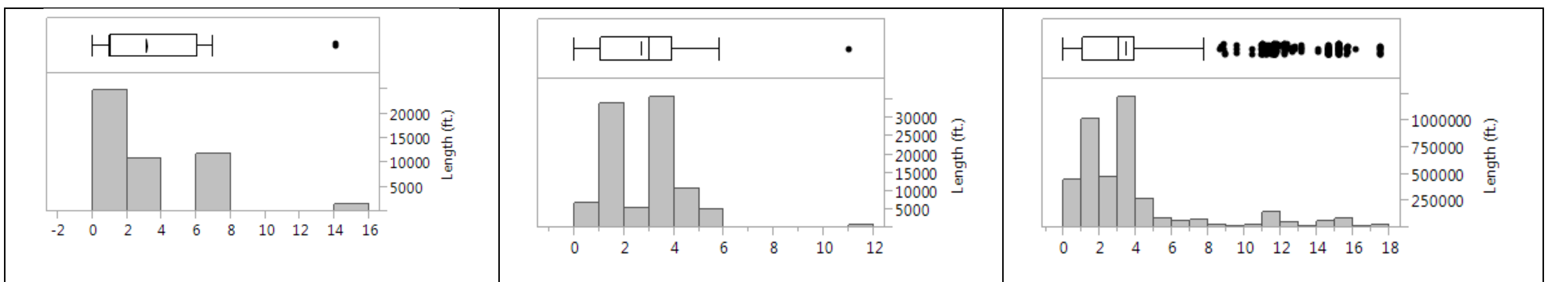
Cluster | Noise | Network

(a) Pipe Size, in.



Cluster | Noise | Network

(b) Material



Cluster | Noise | Network

(c) Pressure Difference, psi

FIGURE 7.6 HISTOGRAM ANALYSIS

Incorporating insights gained from the histogram analysis, expert opinion was utilized to identify and estimate costs for potential MR&R projects for the assets within a cluster or subcluster. For example, cluster 3 is located near a pump station. Further analysis of the cluster showed that the pressure differences in this cluster were higher than in other clusters and parts of the network.

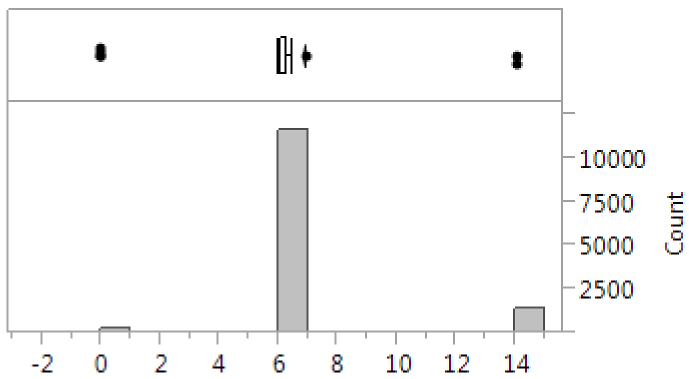


FIGURE 7.7: CLUSTER 3 PRESSURE DIFFERENCE HISTOGRAM

These pressure differences are indicative of the water hammer effect cause by the rapid changes in water velocity [118]. Changes to the pump station controls might mitigate the risk of pipe failures and prove to be a more economical solution than pipe replacement. Such changes would decrease the pressure surges experienced internally, or minimize the water hammer effect inside the pipe.

Potential pump station improvements include the addition of variable frequency drives (VFDs), soft starters, or surge tanks. Soft starters control the voltage required to start or stop a pump motor. This controlled start/stop results in a gradual change in water velocity, reducing the water hammer effects [118]. The more expensive power control alternative, VFDs can be used to control power supplied to the pump and again limit the rapid velocity changes. Though two to three times more expensive with respect to the initial capital expenditure, VFDs can result in significant energy savings, offsetting the initial purchase and installation costs [119]. Both of these operational improvements to mitigate water hammer are significantly cheaper than pipe replacement programs and should be considered as alternatives to replacement and condition assessment programs.

For pricing purposes, the average cost of a replacement was assumed to be $50 per linear foot for 8 in. pipe and $80 pera linear foot (0.3 m) for 12 in. (305 mm) pipe. Condition assessment was estimated to be half the cost of replacement. Finally, it was assumed that operational improvements at a pump would cost approximately $100,000.

Lastly, in order to evaluate the impacts of pipe failures and remediation efforts, a failure consequence score was assigned to the pipes. Figure 7.8 shows the non-normalized consequence score for the pipe network aggregated to low, medium and high consequence. The raw scores were normalized per 1,000 ft. (305 m), and were used in the benefit/cost ratio evaluation.

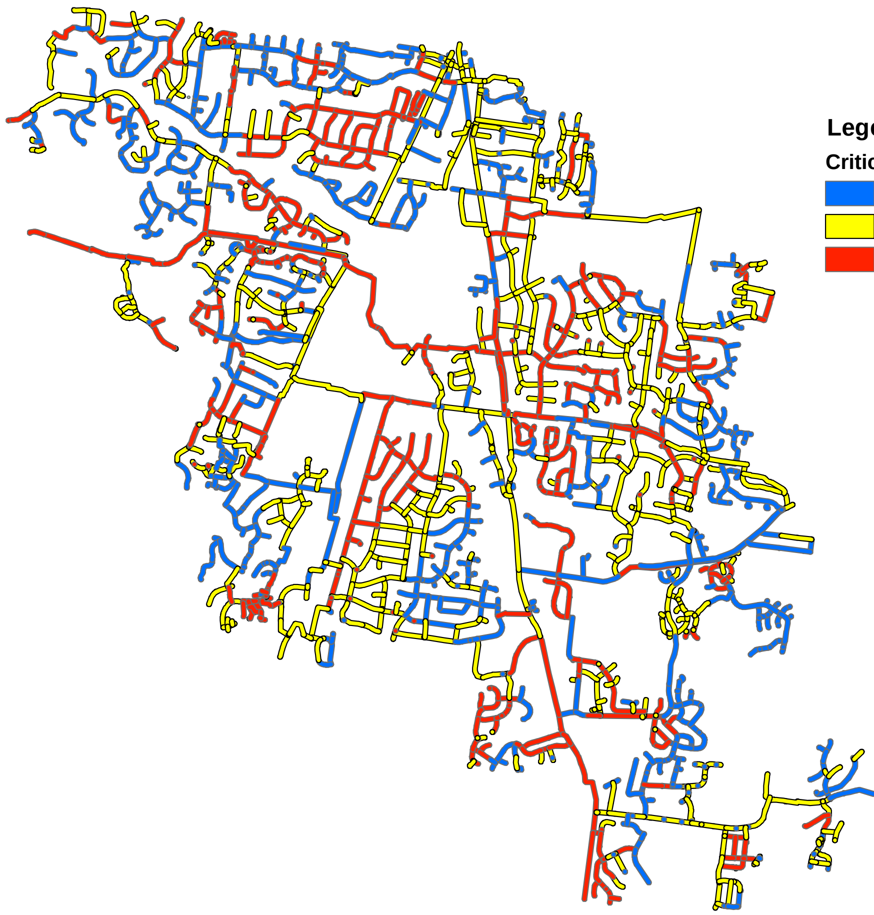


FIGURE 7.8: CONSEQUENCE SCORING

7.5 Prioritization Results

Figure 7.9 shows recommended MR&R activities and a timeline for the sequencing of these recommendations based on cost-benefit analysis.

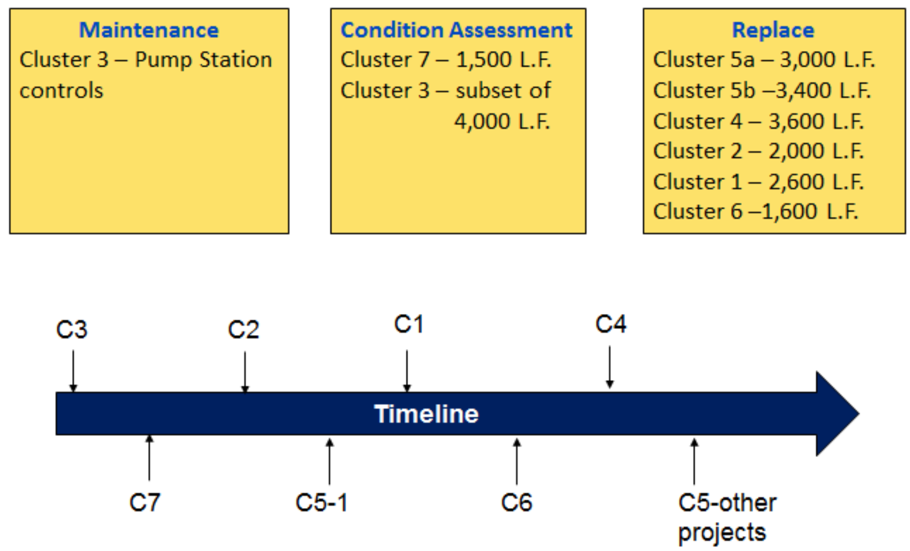


FIGURE 7.9: MR&R PRIORITIZATION SEQUENCING

7.6 Conclusion

Though many parametric models have been introduced to use for pipeline replacement prioritization, they might not be appropriate for medium and small utilities that are more homogenous and have few recorded failures with which to train models. Machine learning based models are alternatives to physical based models, yet many are cumbersome to implement and difficult and lack transparency in model development.

A review of machine learning models revealed that a primary contributing variable to degrading pipe condition is previous breaks or high break rate. As an alternative to other complex machine learning methods, cluster analysis can be used to identify areas of high break rate and potential causes for the observed failures. This chapter presented a framework for synthesizing cluster analysis results with hydraulic data, consequence analyses, and expert opinion to identify and prioritize MR&R projects.

The results of the study give credence to assumptions made in previous chapters that localized break rate in statistical models is potentially a more significant parameter than other proposed explanatory variables. Though statistical models may include many different parameters to describe operational and environmental characteristics that contribute to pipe failures, the impact of these parameters is not necessarily consistent throughout the entire network. For example, high pressure variances were not consistent in all utilities. Additionally, high traffic loadings were observed in come failure clusters, and not in others.

The framework is beneficial to utility operators for several reasons. First, the analysis showed areas where maintenance and operation improvement projects which are much cheaper than pipe replacement could be conducted to mitigate breaks. The framework also results in exhibits and planning tables that are easy to interpret. These data-driven exhibits and tables are of paramount importance when requesting funding from utility boards and customers.

This framework presented could be improved by incorporating hydraulic simulations into the consequence analysis. Determining which pipes are critical to maintaining adequate pressure through hydraulic modeling simulations could shift the consequence analysis. These pipes would result in some of the highest failure consequences as they impact the greatest number of consumers. The criticality scoring methodology would need to be shifted to account for these pipes.

CHAPTER VIII – CONCLUSIONS

8.1 Summary of Accomplishments

This dissertation examines frameworks for prioritizing MR&R activities. Examples are given for utilities of varying size. Potential reasons for the lack of adoption of pipe replacement prioritization models are discussed, and

Chapter 4 highlights the improvements made to the more commonly used and referenced Weibull-based failure prediction model. Specifically, incorporating categorical variables that account for pipe property data assumptions moderately improved the prediction performance of the Weibull based model. Incorporating a covariate describing the spatial distribution of breaks had a greater impact on prediction performance results. Of note is that this model does not include pipe length as a covariate, which could lead to prediction bias.

Chapter 5 demonstrates how to incorporate the validated failure prediction model into a framework for prioritizing inspection and replacement of water pipes. The framework includes one of the most comprehensive consequence analyses presented in literature and incorporates the use of penalty multipliers based on consequence levels. The framework also provides constraints that spatially limit proposed replacement work in zones. This framework will help decision makers decide both when to replace assets and which assets to replace with respect to minimizing the consequences of failures and identifying viable pipeline replacement capital projects.

The optimization framework is demonstrated on a subset of the large utility, Utility A. The scalability of the model is demonstrated by evaluating the model performance and run time using all data from Utility A. The results show that the Excel-based model can be a valuable tool for utilities, providing them with options to easily edit constraints and objectives.

Chapter 6 investigates using the WHRM demonstrated in Chapter 4 to predict failures for medium and small utilities. As valid models could not be produced using the data from Utilities B and C alone, model transfer techniques demonstrated and/or proposed in literature for pipeline condition assessment were tested. Neither direct transfer, joint context estimation nor combined transfer estimation of WHRM model parameters resulted in valid models for Utilities B and C. These findings are significant and drive the need to both determine the minimum number of samples required to develop valid statistical models and develop alternative, potentially machine learning based models for use by small and medium utilities.

Building upon the concept of developing alternative models for medium and small utilities, a cluster analysis framework for MR&R prioritization is introduced. Using the popular DBSCAN algorithm, failure clusters are identified and refined using heuristic knowledge. This framework is demonstrated using data from Utility B. Analysis of failure clusters facilitated the inference of root cause of failure and assessment of cost-effective failure mitigation MR&R activities.

8.2 Future Work

8.2.1. Value of Information Analysis

Resource allocation is the underlying problem associated with developing MR&R strategies for pipelines. Given limited resources, utility makers must decide the value of modeling, data collection, and condition assessment efforts as it relates to gaining more valuable information about the system to make MR&R decisions. Statistical models can serve as a first level in a hierarchy of a suite of decision support tools and methodologies. Given the results of statistical analysis, the decision maker can decide to develop more expensive physical models or machine learning based models that require laboratory testing, perform condition assessment (or multiple condition assessments), or replace an asset.

None of the models reviewed consider the cost effectiveness of employing condition assessment technologies prior to making replacement and renewal decisions. Yet utility case studies show that condition assessment is a cost effective asset management strategy, preventing unwarranted rehabilitation and replacement [12]. Additionally, utilities see a greater benefit in long-term monitoring of assets and utilizing multiple condition assessment technologies to gain a deeper insight into the state of the pipe.

Other valuable research has been performed to optimize the scheduling of condition assessment activities. Kleiner [78] introduced a model using Markov transition state probabilities to determine when to replace or inspect a pipeline. Inspection is scheduled when replacement is not warranted. The model does not address which condition assessment technology should be utilized.

Lau and Dwight [65] introduced a fuzzy-based decision support system for water pipe maintenance. The model considers three condition assessment technologies, and incorporates expert opinion to assess the probability of failure given the condition assessment observations. This model has not been tested on an entire water pipeline network, and does not evaluate the cost-benefit trade-off between performing condition assessment and replacement.

The underlying problem in prioritizing condition assessment and rehabilitation activities for infrastructure is resource allocation. With limited funds available, utilities need to decide when condition assessment is cost beneficial with respect to reducing the uncertainty associated with rehabilitation and replacement decisions. Value of Information (VOI) analyses can be a powerful tool to the decision maker when making resource allocation decisions. Originally proposed by Howard [120], VOI analysis is proven decision making framework, demonstrated in many applications including health sciences [121], economics [122], and supply chain management [123].

VOI is used to evaluate the expected value of obtaining additional information prior to making a decision compared to the outcome of making a less-informed decision. To perform a VOI analysis, the decision maker must compile a set of actions and information collection strategies, develop probabilistic models for the reliability of information collection strategies, and calculate the values for the risk outcomes [124]. These expected value (EV) problems can be solved using decision trees [125]. Commercial software programs, spreadsheet, or independently written software can be used to solve decision tree problems.

Recent applications of VOI include the study by Khader et al. [125] which uses VOI analyses to assess the value of implementing of a groundwater quality monitoring and communications system. Liu et al. [126] evaluate the influence parameter uncertainties on groundwater assessment and remediation using VOI. Messer et al. [127] developed a VOI based decision support methodology for selecting appropriate higher and lower fidelity models given unknowns about prediction accuracy. These studies address the problems of condition assessment and rehabilitation scheduling such as justifying real-time monitoring efforts, making decisions based on uncertainties, and assessing the cost-benefit between a more expensive and more reliable model or assessment technology compared to a less expensive and less reliable model or assessment technology.

With direct application to pipeline condition assessment, Osman et al [34] demonstrate the use of VOI analysis to optimize the scheduling of condition assessment activities. This framework takes into account the current condition of the asset, the accuracy of the condition assessment technology, variations in assessments when using multiple technologies, and the cost of failure. The outcome of the assessment is an optimized condition assessment policy that describes which tools to apply and the frequency in which to employ them. The application of this framework for a utility showed that estimated optimal inspection technologies and assessment frequency varied across the network with respect to the condition and criticality of assets. This work can be expanded to evaluate business cases which consider the value of condition assessment information prior to making rehabilitation and replacement decisions.

The building blocks for developing a DSS that considers both pipelines condition assessment and rehabilitation in tandem have already been established. Proven models for estimating the consequences of failure and case studies of condition assessment monitoring can be utilized to establish a basis for the cost-benefit analysis of condition assessment technologies. More research is needed; however, to link condition assessment results to potential failure modes and to reduce the uncertainty of failure probabilities assigned using condition assessment results. In the following we sections, we outline the necessary actions to develop a DSS that evaluates the cost effectiveness of condition assessment technologies on high consequence water mains.

8.2.2. Interface Development

With respect to furthering the use of the decision support system explored in this work, the workflow required to produce the desired results needs to be outlined and streamlined. A summary of this workflow, which can be followed by other large utilities wishing to implement the framework described in this document is outlined below:

1. Gather work orders/failure reports and compute start time and duration of asset management database
2. Computer time to failure, with the start time of the asset management database being zero
3. Populate failure record database with pipeline properties from GIS model
4. Make assumptions for unknown data, and account for assumptions using binary covariate
5. Calculate break rate parameter described in Chapter 4 and assign values to all assets.

6. Stratify the database with respect to material
7. For each material, create three tables. One table consisting of the first recorded failure records, the second being failure records of subsequent failures, and the third being pipes that did not fail
8. Randomly select 20% of the data from all tables to be withheld from calibration
9. Calibrate WPHM model parameters using built in calibration functions included in software such as JMP or R.
10. Write software or program Excel to perform Monte Carlo simulation outlined in Chapter 4
11. Rank assets by predicted failures and calculate decile divisions
12. Compute the predicted and observed failures for each quantile in Excel
13. Plot and validate the results
14. Computer consequence of failure using spatial tools outlined in Chapter 5
15. Identify candidate pressure zone for MR&R activities
16. Use spatial analysis tools to clip pipes within the pressure zone
17. Use the ArcGIS fishnet tool to bin the pressure zone
18. Use the spatial intersect tool to join the bins to the pipes
19. Export the pipe records and join to existing asset management database
20. Import asset management database for pipes in that pressure zone into Excel
21. Import model calibration parameter results into Excel as lookup tables
22. Create lookup table of pipe replacement and repair costs
23. Calculate cumulative failure probability for each year and costs for each year using values from the lookup tables

24. Add fields in the Excel sheet to calculate costs ratio for each year
25. Use the Excel sort feature to sort the assets by zone and descending costs ratio
26. Add fields to calculate cumulative costs ratio and replacement costs
27. Plot the costs ratio versus replacement costs and add a logarithmic trend line
28. Create a decision variables sheet that includes the bin number of the costs ratio function with respect to a capital costs decision cell
29. Start the Evolver add-in and define the optimization model to maximize the sum of the costs ratio while restraining the number of bins in which work is being done and the capital costs
30. Run the model and save results
31. For each bin and capital investment, identify the pipes that are included in the capital investment with respect to the ranking procedure
32. Create an sheet with these asset ID's and import into ArcGIS
33. Join the sheet to the pipeline shapefile in ArcGIS and display the joined features.
34. Refine projects based on connectivity

Much of the work was performed using tools that already exist within ArcGIS or add-ons to MS Excel, which was intentional to consider atomization of activities. Visual Basic and Python code can be written to automate processes in order to create a commercial tool that can be used by both consultants and large utility managers to help facilitate the increased adoption rate of risk-based asset management tools. Given the results and analysis of this work, this tool would be best used by large utilities with lined and unlined cast iron pipe that is reaching the end of its useful life. Use of this software would not be appropriate for small utilities, or areas of large utilities with large quantities of pipe installed in the past twenty years. In some cases, when the number of failures is large enough, likelihood of failure modeling efforts for large utilities can be constrained to pressure zones that do contain older pipe.

APPENDIX A. BREAK RATE KRIGING IMPLEMENTATION

The break rate distribution referenced in Chapter 4 was estimated using the Kriging spatial analysis model in ESRI ArcGIS 10. Based on the works of Krige [128] the stationary Kriging models available in ArcGIS 10 are documented by [129], [130]. The following summarizes the methodology described in the aforementioned resources. Subsequent sections show how to run a Kriging model, Kriging results from Chapter 4, and notes about improvements made in the ESRI Kriging model which are available for the newest models of ESRI ArcGIS.

A.1 Methodology

To predict a measurement or value at a location, s_o, the following equation is used:

$$\hat{Z}(s_o) = \sum_{i=1}^{N} \lambda_i Z(s_i) \tag{A.1}$$

Where λ_i = an unknown weight for the measured value at the *i*th location.

N = the number of measured values

$\boldsymbol{Z}(\boldsymbol{s_i})$ = the measured value at the *i*th location.

It is assumed that the break rate is a spatially autocorrelated process with independent random errors described by a mean and error function shown in equation A.2.

$$Z_t(s) = \mu(s) + \varepsilon_t(s) \qquad \text{[A.2]}$$

Where $\mu(s)$ is a unknown, deterministic mean value and $\varepsilon_t(s)$ is a function that accounts for random measurement and model fitting errors. The decomposition and prediction of the error function can be found in pages 262 – 263 of the Geospatial Analyst User's Manual [129].

Spatial kriging in ArcGIS is performed by creating variograms and covariance functions to estimate the autocorrelation of the measured values. Variograms are first determined by computing the difference squared between each pair of measurement/observation locations. Instead of plotting all of these location pairs, they are grouped into lag bins, $\boldsymbol{h}$, and the average semivariance is plotted.

Though most often lag bins are determined using radial functions, the geospatial analyst in ArcGIS assigns lags to a grid. Because lag vectors near the edges of bins can cause issues in determining the semivariogram, kernel functions are used to weight the semivariogram.

Once the bin semivariograms have been estimated, they are plotted with respect to distance. An empirical function is used to fit a function to the semivariogram. In examining semivariogram plots, there is a tendency for the semivariogram values to level off as distances increase. Shown in Figure A.1, this semivariance value at which this occurs is called the sill. This distance at which this occurs is called the range.

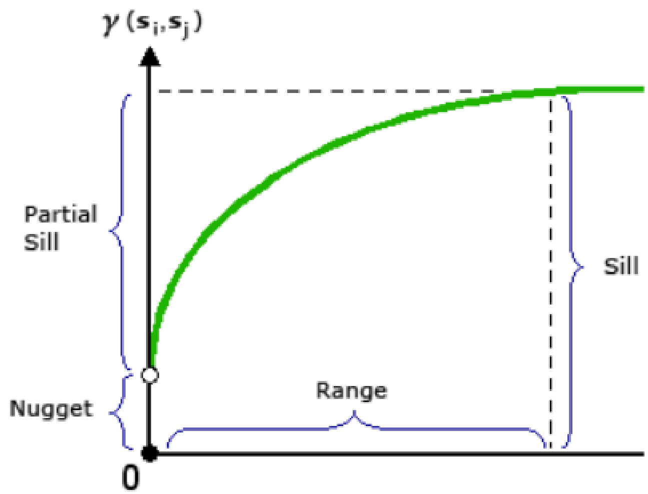


Figure A.1: Depiction of Semivariogram Range and Sill [130]

For the stationary process considered, sill relates the semivariogram to the covariance matrix as

$$C(\boldsymbol{h};\theta) = \gamma(\infty;\theta) - \gamma(\boldsymbol{h};\theta) \tag{A.3}$$

Where $\gamma(\infty;\theta)$ is the sill of the semivariogram, and $\gamma(\boldsymbol{h};\theta)$ is the semivariogram value at $\boldsymbol{h}$.

For this study, the spherical semivariogram form was chosen. The spherical function is the most widely used spatial Kriging model. The model form is shown in Figure A.2 and described by the Equation A.4.

$$\gamma(\boldsymbol{h},\boldsymbol{\theta}) = \left\{ \begin{matrix} \frac{2\theta_s}{\pi}\left[\frac{\|h\|}{\theta_r}\sqrt{1-\left(\frac{\|h\|}{\theta_r}\right)^2} + \arcsin\frac{\|h\|}{\theta_r}\right] \text{ for } 0 \leq \|h\| \leq \theta_r \\ \theta_s \text{ for } \theta_r \leq \|h\| \end{matrix} \right\} \tag{A.4}$$

Where $\theta_s \geq 0$ is the partial sill parameter and $\theta_r \geq 0$ is the range parameter.

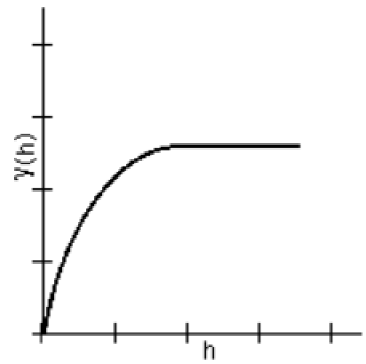


Figure A.2: Spherical Semivariogram Function [130]

The fitting algorithms for the weights are described in detail on pages 259-260 of [129]

A.2 Running a Kriging Model in ArcGIS 10.0

Kriging in ArcGIS 10.0 is performed using the Kriging tool in the Spatial Analyst Toolbox. After using the Thiessen tool to create polygons around points, the linear feet in each polygon and break rate is computed. The break rate is assigned to each failure point and used in the Kriging model.

Using the Kriging tool, the user must input the observation points and select the semivariogram model as shown in Figure A.3 The user must also specify where to store the raster project and the output cell size that defines the refinement of the raster projection surface.

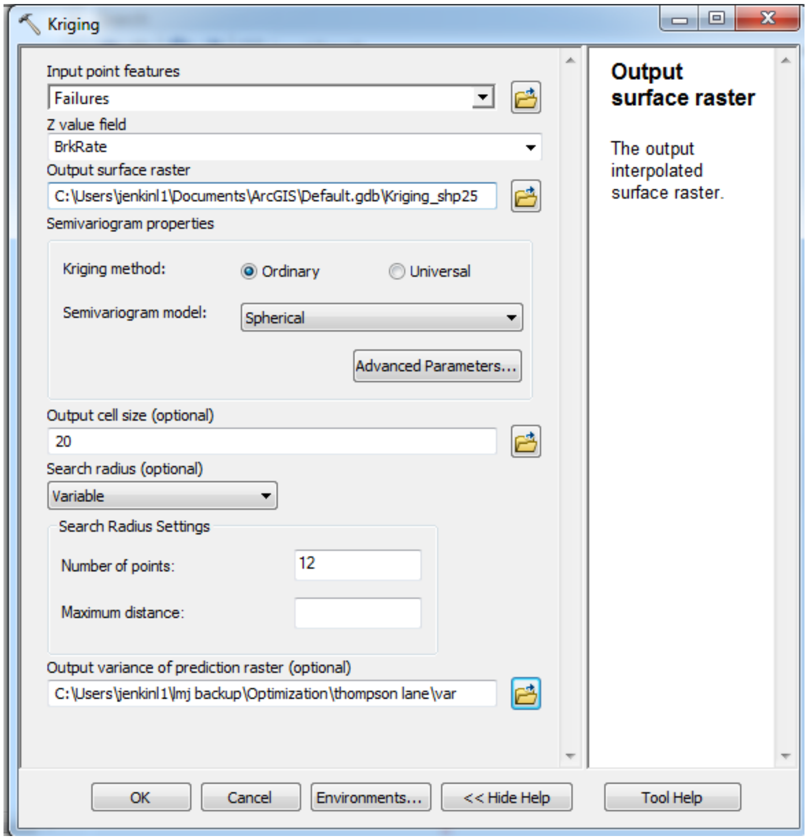


FIGURE A.3. Kriging Model Dialog Box

In order to assign the values from the raster image to the pipe segements, the raster image must be converted to a geostatistical surface. In ArcGIS 10.0, the values within this surface cannot be less than one, so the field calculator is used to multiply the kriging values by 1000. After the geostatistical surface is made, the average break rate across each pipe segment is added to the pipe asset table using the spatial join geopgrocessing tool. The results of the kriging analysis are shown in Figure A.4.

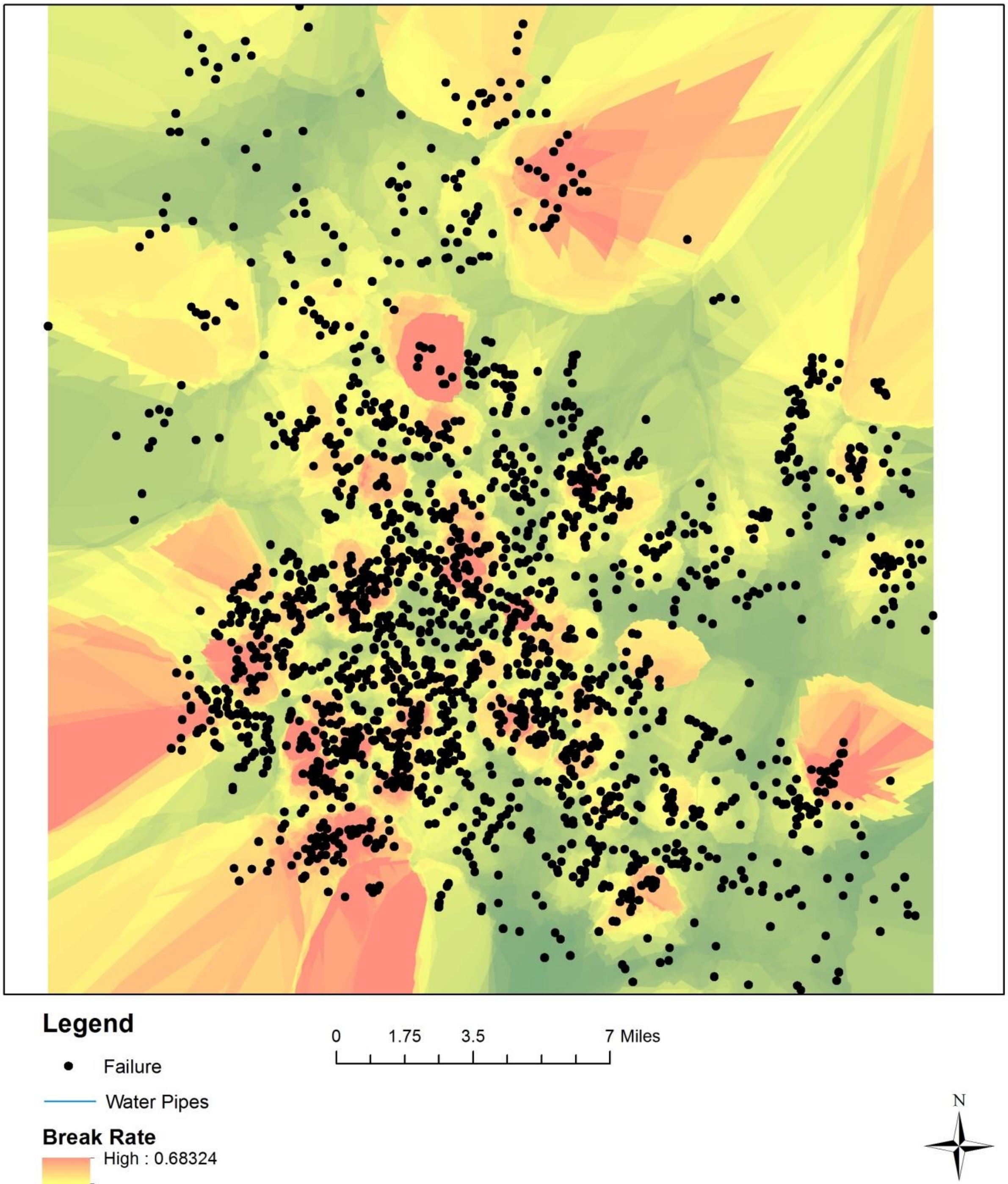


FIGURE A.4: BREAK RATE DISTRIBUTION RESULTS

A.3 Chapter 4 Kriging Results

A.4 Improvements in Kriging

In ArcGIS versions 10.1 and beyond, an improved kriging model is available. This Empirical Bayesian kriging (EBK) model [131] helps account for the error introduced in the semivariogram estimates. In the kriging model described above, it is assumed that the correlation structure defined by the estimated semivariogram, chosen prior to running the kriging model, is the true semivariogram of the observed data generated from a Gaussian distribution.

The EBK accounts for errors in these assumptions by estimating a spectrum of semivariograms that describe the semivariogram that best describes the data. Following the same procedure described in Section A.1, a semivariogram is estimated from the data. Using the estimated semivariogram, a prediction is simulated at the observation location. A new semivariogram is then estimated for the prediction data. Using Baye's rules, the new semivariogram is weighted based on the likelihood of predicting the observed data using the estimated semivariogram. The process of estimating data from the semivariogram and weighting the new semivariogram is repeated and predictions at other locations in the network are made.

The EBK methodology also includes a routine to transform data that is non-Gaussian and potentially differs in distribution across the study area using a transformation function shown in Figure A.5. The process to estimate the semivariograms is the same, with a final transformation back to the original data form.

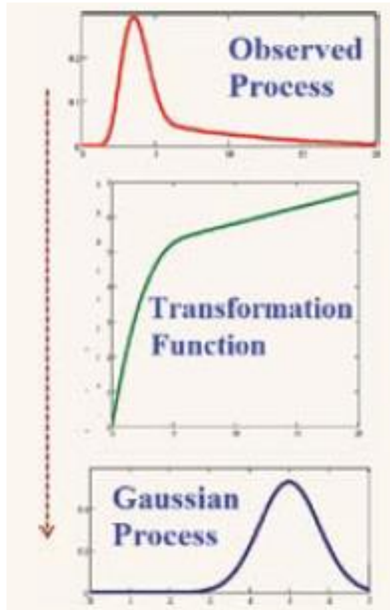


Figure A.5: Transformation of Data to Gaussian Process

This methodology was not available when the initial study was performed. It should be noted that the results prevented can be improved by utilizing this method. Future work should include an updated break rate distribution model.

APPENDIX B. LOOKUP TABLES

B.1: Repair Costs		
Pipe Material	Size	Cost ($)
CL	0.75	1000
CL	1	1000
CL	1.25	1000
CL	1.5	1000
CL	2	1000
CL	2.25	1000
CL	2.5	1000
CL	3	1000
CL	4	4000
CL	6	4000
CL	8	4000
CL	10	6000
CL	12	6000
CL	16	6000
CL	18	6000
CL	20	12000
CL	24	12000
CL	30	20000
CL	36	20000
CU	0.5	1000
CU	0.75	1000
CU	1	1000
CU	1.25	1000
CU	1.5	1000
CU	2	1000
CU	2.25	1000
CU	2.5	1000
CU	3	1000
CU	4	4000
CU	6	6000
CU	8	6000
CU	10	8000
CU	12	8000
CU	14	8000
CU	16	12000

B.2: Replacement Costs		
Pipe Material	Size	Cost ($/L.F.)
CL	0.75	30
CL	1	30
CL	1.25	30
CL	1.5	30
CL	2	30
CL	2.25	30
CL	2.5	30
CL	3	50
CL	4	50
CL	6	50
CL	8	80
CL	10	100
CL	12	120
CL	16	140
CL	18	180
CL	20	200
CL	24	200
CL	30	260
CL	36	260
CU	0.5	30
CU	0.75	30
CU	1	30
CU	1.25	30
CU	1.5	30
CU	2	30
CU	2.25	30
CU	2.5	30
CU	3	50
CU	4	50
CU	6	50
CU	8	80
CU	10	100
CU	12	120
CU	14	140
CU	16	140

B.1: Repair Costs (cont.)

Pipe Material	Size	Cost ($)
CU	18	12000
CU	20	16000
CU	24	16000
CU	30	20000
CU	36	20000
CU	48	25000
DIP	0.75	6000
DIP	1	6000
DIP	1.25	6000
DIP	1.5	6000
DIP	2	6000
DIP	2.25	6000
DIP	2.5	6000
DIP	3	6000
DIP	4	6000
DIP	6	6000
DIP	8	6000
DIP	10	8000
DIP	12	8000
DIP	16	8000
DIP	18	12000
DIP	20	12000
DIP	24	16000
DIP	30	20000
DIP	36	20000
DIP	42	25000
DIP	48	25000
DIP	60	75000

B.2: Replacement Costs (cont.)

Pipe Material	Size	Cost ($/L.F.)
CU	18	180
CU	20	200
CU	24	200
CU	30	260
CU	36	260
CU	48	300
DIP	0.75	30
DIP	1	30
DIP	1.25	30
DIP	1.5	30
DIP	2	30
DIP	2.25	30
DIP	2.5	30
DIP	3	30
DIP	4	30
DIP	6	50
DIP	8	80
DIP	10	100
DIP	12	120
DIP	16	140
DIP	18	180
DIP	20	200
DIP	24	200
DIP	30	260
DIP	36	260
DIP	42	300
DIP	48	300
DIP	60	400

The costs shown in the tables were based on repair and replacement costs presented in the literature and costs provided by engineers and contractors in the area. These costs can be further refined with utility input.

BIBLIOGRAPHY

[1] J. Cromwell, H. Reynolds, and N. Pearson, "Cost of Infrastructure Failure," 2002.

[2] P. Gaewsk and F. Blaha, "Analysis of Total Cost of Large Diameter Pipe Failure," 2007.

[3] K. R. Piratla, S. T. Ariaratnam, and M. Asce, "Criticality Analysis of Water Distribution Pipelines," *J. Pipeline Syst. Eng. Pract.*, vol. 2, no. 3, pp. 91–101, 2011.

[4] Environmental Protection Agency (EPA), "Drinking Water Infrastructure Needs Survey and Assessment Fifth Report to Congress," Washington, DC, 2013.

[5] American Water Works Association, "Buried No LoNger : Confronting America's Infrastructure Challenges," Denver, CO, 2012.

[6] Water Research Foundation, "Breaks and Leaks FAQ's." [Online]. Available: http://www.waterrf.org/knowledge/asset-management/breaks-leaks/Pages/faqs.aspx.

[7] APWA, "APWA Board Adopts Policy Supporting GASB 34," *APWA Reporter*, Feb-2001.

[8] Environmental Protection Agency (EPA), "Asset Management : A Best Practices Guide," Washington, DC, 2008.

[9] J. Matthews, W. Condit, and R. McKim, "Decision Support for Renewal of Wastewater Collection and Water Distribution Systems," 2011.

[10] A. St. Clair and S. Sinha, "State-of-the-Technology Review on Water Pipe Condition, Deterioration and Failure Rate Prediction Models," *Urban Water J.*, vol. 9, no. 2, pp. 85–112, 2012.

[11] A. Wood and B. J. Lence, "Using Water Main Break Data to Improve Asset Management for Small and Medium Utilities : District of Maple Ridge , B . C .," *J. Infrastruct. Syst.*, vol. 15, no. 2, pp. 111–119, 2009.

[12] N. Thuruthy, S. Sinha, L. Catalano, P. Harren, J. Leighton, R. Nelson, and S. Rajah, "Condition Assessment for Drinking Water Pipelines Synthesis Report," Alexandria, VA, 2013.

[13] L. Berardi, O. Giustolisi, Z. Kapelan, and D. A. Savic, "Development of pipe deterioration models for water distribution systems using EPR," *J. Hydroinformatics*, vol. 10, no. 2, pp. 113–126, Apr. 2008.

[14] A. Debón, A. Carrión, E. Cabrera, and H. Solano, "Comparing risk of failure models in water supply networks using ROC curves," *Reliab. Eng. Syst. Saf.*, vol. 95, no. 1, pp. 43–48, Jan. 2010.

[15] P. D. Rogers and N. S. Grigg, "Failure Assessment Modeling to Prioritize Water Pipe Renewal : Two Case Studies," *J. Infrastruct. Syst.*, no. September, pp. 162–171, 2009.

[16] Y. Wang, T. Zayed, and O. Moselhi, "Prediction Models for Annual Break Rates of Water Mains," *ASCE J. Perform. Constr. Facil.*, vol. 23, no. 1, pp. 47–54, 2009.

[17] UMA Engineering Ltd., "The City of Hamilton Watermain Management Framework," Mississauga, Ontario, 2008.

[18] M. Poulton, I. Kropp, and A. Vanrenterghem-Raven, "Evaluating risk of multi-segment pipes for prioritizing pipe rehabilitation," in *World Environmental and Water Resource Congress*, 2009, pp. 20–30.

[19] infraPLAN, "Data," 2014. [Online]. Available: http://www.infraplan-llc.com/data.html. [Accessed: 18-Aug-2014].

[20] Y. Kleiner, "Scheduling Inspection and Renewal of Large Infrastructure Assets," *ASCE J. Infrastruct. Syst.*, vol. 7, pp. 136–143, 2001.

[21] H. P. Hong, E. N. Allouche, and M. Trivedi, "Optimal Scheduling of Replacement and Rehabilitation," no. September, pp. 184–191, 2006.

[22] M. Tabesh, J. Soltani, R. Farmani, and D. Savic, "Assessing pipe failure rate and mechanical reliability of water distribution networks using data-driven modeling," *J. Hydroinformatics*, vol. 11, no. 1, p. 1, Jan. 2009.

[23] G. C. Dandy and M. Engelhardt, "Optimal Scheduling of Water Pipe Replacement Using Genetic Algorithms," no. August, pp. 214–223, 2001.

[24] G. J. Kirmeyer, W. Richards, and C. D. Smith, "An Assessment of Water Distribution Systems and Associated Research Needs," 1994.

[25] J. M. Makar, R. Desnoyers, and S. E. Mcdonald, "Failure Modes and Mechanisms in Gray Cast Iron Pipes," pp. 1–10, 2001.

[26] A. St. Clair, "Development of a Novel Performance Index and a Performance Prediction Model for Metallic Drinking Water Pipelines," Virginia Polytechnic Institute and State University, 2013.

[27] National Transportation Safety Board, "An Analysis of Accident Data from Plastic Pipe Natural Gas Distribution System," Washington, DC, 1980.

[28] M. M. Carroll, "Polyvinyichloride (PVC) Pipe Reliability and Failure Modes," *Reliab. Eng.*, vol. 13, no. April 1984, pp. 11–21, 1985.

[29] R. Dueck, "PVC Failures." The City of Calgary Water Resources, 2010.

[30] G. Z. Watters, "The Behavior of PVC Pipe Under the Action of Water Hammer Pressure Waves," 1971.

[31] A. K. Soares, D. I. C. Covas, and L. F. R. Reis, "Analysis of PVC Pipe-Wall Viscoelasticity during Water Hammer," *J. Hyrdaulic Eng.*, vol. 134, no. September, pp. 1389–1394, 2008.

[32] "Standard Operating Procedure for Waterline Repair," 2007.

[33] N. S. Grigg, "Condition Assessment of Water Distribution Pipes," *J. Infrastruct. Syst.*, vol. 12, no. September, pp. 147–153, 2006.

[34] H. Osman, A. Atef, and O. Moselhi, "Optimizing Inspection Policies for Buried Municipal Pipe Infrastructure," *J. Perform. Constr. Facil.*, no. June, pp. 345–352, 2012.

[35] Y. Kleiner, "Scheduling Inspection and Renewal of Large Infrastructure Assets," *J. Infrastruct. Syst.*, no. December, pp. 136–143, 2001.

[36] B. Rajani and Y. Kleiner, "Comprehensive Review of Structural Deterioration of Water Mains: Physically Based Models," *Urban Water*, vol. 3, no. 3, pp. 151–164, Sep. 2001.

[37] Y. Le Gat and P. Eisenbeis, "Using Maintenance Records to Forecast Failures in Water Networks," *Urban Water*, vol. 2, no. 2000, pp. 173–181, 2000.

[38] W. Park and G. V Loganathan, "Methodology for Economically Optimal Replacement of Pipes in Water Distribution Systems: 2. Applications," *J. Korean Soc. Civ. Eng.*, vol. 6, no. 4, pp. 545–550, 2002.

[39] W. Park and G. V Loganathan, "Methodology for Economically Optimal Replacement of Pipes in Water Distribution Systems : 1 . Theory," *J. Korean Soc. Civ. Eng.*, vol. 6, no. 4, pp. 539–543, 2002.

[40] G. V Loganathan, S. Park, and H. D. Sherali, "Threshold Break Rate for Pipeline Replacement in Water Distribution Systems," no. August, pp. 271–279, 2002.

[41] G. Pelletier, A. Mailhot, and J.-P. Villeneuve, "Modeling Water Pipe Breaks — Three Case Studies," *J. Water Resour. Plan. Manag.*, no. April, pp. 115–123, 2003.

[42] A. Vanrenterghem-Raven, "Risk factors of Structural Degradation of an Urban Water Distribution System," vol. 13, no. 1, pp. 55–64, 2007.

[43] S. Alvisi and M. Franchini, “Comparative analysis of two probabilistic pipe breakage models applied to a real water distribution system,” *J. Civ. Eng. Environ. Syst.*, vol. 27, no. 1, pp. 1–22, 2010.

[44] A. Carrión, H. Solano, M. L. Gamiz, and A. Debón, “Evaluation of the Reliability of a Water Supply Network from Right-Censored and Left-Truncated Break Data,” *Water Resour. Manag.*, vol. 24, no. 12, pp. 2917–2935, Jan. 2010.

[45] R. Sadiq, Y. Kleiner, and B. Rajani, “Water quality failures in distribution networks – risk analysis using fuzzy logic and evidential reasoning Water quality failures in distribution networks - risk analysis using fuzzy logic and evidential reasoning,” no. 5, pp. 1381–1394, 2007.

[46] Y. Kleiner and B. Rajani, “Comparison of four models to rank failure likelihood of individual pipes,” *J. Hydroinformatics*, vol. 14, no. 3, pp. 659–681, Jul. 2012.

[47] S. Park, H. Jun, N. Agbenowosi, B. Ja. Kim, and K. Lim, “The Proportional Hazards Modeling of Water Main Failure Data Incorporating the Time-dependent Effects of Covariates,” *Water Resour. Manag.*, vol. 25, no. 1, pp. 1–19, 2011.

[48] A. Malm, O. Ljunggren, O. Bergstedt, T. J. R. Pettersson, and G. M. Morrison, “Replacement predictions for drinking water networks through historical data.,” *Water Res.*, vol. 46, no. 7, pp. 2149–58, May 2012.

[49] A. Martins, J. P. Leitão, and C. Amado, “A Comparative Study of Three Stochastic Models for Prediction of Pipe Failures in Water Supply Systems,” *J. Infrastruct. Syst.*, vol. 19, no. 4, pp. 442–450, Mar. 2013.

[50] Y. Le Gat, “‘Une extension du processus deYule pour la modelisation stochastique desenements revecurrents. Application aux d´ efaillances de canalisations d’eau sous pression.,’” Paristech, France, 2009.

[51] B. Toumbou, J.-P. Villeneuve, G. Beardsell, and S. Duchesne, “General Model for Water-Distribution Pipe Breaks : Development , Methodology , and Application to a Small City in Quebec , Canada,” *J. Pipeline Syst. Eng. Pract.*, pp. 1–9, 2013.

[52] D. P. de Oliveira, D. B. Neill, J. H. Garret, and L. Soibelman, “Detection of Patterns in Water Distribution Pipe Breakage Using Spatial Scan Statistics for Point Events in a Physical Network,” *J. Comput. Civ. Eng.*, vol. 25, no. 1, pp. 21–30.

[53] D. P. de Oliveira, J. H. Garrett, and L. Soibelman, “A density-based spatial clustering approach for defining local indicators of drinking water distribution pipe breakage,” *Adv. Eng. Informatics*, vol. 25, no. 2, pp. 380–389, Apr. 2011.

[54] M. Ester, H. P. Kriegal, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Second International Conference on Knowledge Discovery and Data Mining*, 1996.

[55] M. Ankerst, M. M. Breunig, H. P. Kriegel, and J. Sander, "OPTICS: ordering points to identify the clustering structure," in *ACM SIGMOD International Conference on Management of Data*, 1999, pp. 49–60.

[56] I. Bogárdi and R. Fülüp, "A space-time probabilistic model for pipe network reconstruction planning," *Urban Water J.*, vol. 9, no. 5, pp. 333–346, 2012.

[57] S. Christodoulou, A. Gagatsis, A. Agathokleous, S. Xanthos, and S. Kranioti, "Urban Water Distribution Network Asset Management Using Spatio- Temporal Analysis of Pipe-Failure Data," in *International Conference on Computing in Civil and Building Engineering*, 2012.

[58] W.-Z. Shi, A.-S. Zhang, and O.-K. Ho, "Spatial analysis of water mains failure clusters and factors: a Hong Kong case study," *Ann. GIS*, vol. 19, no. 2, pp. 89–97, Jun. 2013.

[59] D. Achim, F. Ghotb, and K. J. Mcmanus, "Prediction of Water Pipe Asset Life Using Neural Networks," *J. Infrastruct. Syst.*, vol. 13, no. 1, pp. 26–30, 2007.

[60] Z. E. Geem and et al., "Trenchless water pipe condition assessment using artificial neural network," 2007.

[61] H. Cheng-I, M.-D. Lin, and S.-L. Lo, "Use of a GIS-based hybrid artificial neural network to prioritize the order of pipe replacement in a water distribution network.," *Environ. Monit. Assess.*, vol. 166, no. 1, pp. 177–189, 2010.

[62] H. Al-Barqawi and T. Zayed, "Infrastructure management: integrated AHP/ANN model to evaluate municipal water mains' performance," *J. Infrastruct. Syst.*, vol. 14, no. 4, pp. 305–318, 2008.

[63] M. Najafi and G. Kulandaivel, "Pipeline condition prediction using neural network models," in *Piprlines 2005: Optimizing Pipeline Design, Operations, and Maintenance in Today's Economy*, 2005, pp. 767–781.

[64] H. Osman, K. Bainbridge, M. Gibbons, C. Macey, and R. Homeniuk, "An Integrated Management Approach for Critical Water Mains," in *ASCE Pipeline Division International Conference*, 2008, pp. 1–10.

[65] H. C. W. Lau and R. a. Dwight, "A fuzzy-based decision support model for engineering asset condition monitoring – A case study of examination of water pipelines," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 13342–13350, Sep. 2011.

[66] H. Fares and T. Zayed, “Hierarchical Fuzzy Expert System for Risk of Failure,” no. February, pp. 53–62, 2010.

[67] R. Sadiq, B. Rajani, and Y. Kleiner, “Fuzzy-based method to evaluate soil corrosivity for prediction of water main deterioration,” *J. Infrastruct. Syst.*, vol. 10, no. 4, pp. 149–156, 2004.

[68] H. Fares and T. Zayed, “Hierarchical Fuzzy Expert System for Risk of Failure,” *J. Pipeline Syst. Eng. Pract.*, no. February, pp. 53–62, 2010.

[69] A. M. S. Clair and S. K. Sinha, “Development of a Fuzzy Inference Performance Index for Ferrous Drinking Water Pipelines,” vol. 5, pp. 1–11, 2014.

[70] Q. Xu, Q. Chen, W. Li, and J. Ma, “Pipe break prediction based on evolutionary data-driven methods with brief recorded data,” *Reliab. Eng. Syst. Saf.*, vol. 96, no. 8, pp. 942–948, Aug. 2011.

[71] T. G. Watson, C. D. Christian, A. J. Mason, M. H. Smith, and R. Meyer, “Bayesian-based pipe failure model,” *J. Hydroinformatics*, vol. 6, no. 4, pp. 259–264, 2004.

[72] C. Wang, Z. Niu, H. Jia, and H. Zhang, “An assessment model of water pipe condition using Bayesian inference,” *J. Zhejiang Univ. Sci. A*, vol. 11, no. 7, pp. 495–504, Jul. 2010.

[73] R. A. Francis, S. D. Guikema, and L. Henneman, “Bayesian Belief Networks for Predicting Drinking Water Distribution System Pipe Breaks,” in *PSAM/ESREL 12*, 2012, no. June.

[74] Z. Li, B. Zhang, Y. Wang, F. Chen, R. Taib, V. Whiffin, and Y. Wang, “Water pipe condition assessment: a hierarchical beta process approach for sparse incident data,” *Mach. Learn.*, Jun. 2013.

[75] L. Scholten, A. Scheidegger, P. Reichert, and M. Maurer, “Combining expert knowledge and local data for improved service life modeling of water supply networks,” *Environ. Model. Softw.*, vol. 42, pp. 1–16, Apr. 2013.

[76] L. Scholten, A. Scheidegger, P. Reichert, M. Mauer, and J. Lienert, “Strategic rehabilitation planning of piped water networks using multi-criteria decision analysis.,” *Water Res.*, vol. 49, no. 2014, pp. 124–43, Feb. 2014.

[77] S. Alvisi and M. Franchini, “Multiobjective Optimization of Rehabilitation and Leakage Detection Scheduling in Water Distribution Systems,” no. December, pp. 426–439, 2009.

[78] Y. Kleiner, “Scheduling Inspection and Renewal of Large Infrastructure Assets,” *J. Infrastruct. Syst.*, vol. 7, no. 4, pp. 136–143, Dec. 2001.

[79] Q. Xu, Q. Chen, J. Ma, and K. Blanckaert, “Optimal pipe replacement strategy based on break rate prediction through genetic programming for water distribution network,” *J. Hydro-environment Res.*, vol. 7, no. 2, pp. 134–140, Jun. 2013.

[80] O. Giustolisi and L. Berardi, “Prioritizing Pipe Replacement : From Multiobjective Genetic Algorithms to Operational Decision Support,” *J. Water Resour. Plan. Manag.*, vol. 6, no. 135, pp. 484–492, 2009.

[81] O. Giustolisi, A. Doglioni, D. A. Savic, and D. Laucelli, “A proposal for an effective multi-objective non-dominated genetic algorithm: the OPTimised Mult-Objective Genetic Algorithm - OPTIMOGA,” 2004.

[82] K. Deb, A. Member, A. Pratap, S. Agarwal, and T. Meyarivan, “A Fast and Elitist Multiobjective Genetic Algorithm :,” *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, 2002.

[83] S. Forrest, “Genetic Algorithms: Principles of Natural Selection Applied to Computation,” *Science (80-.).*, vol. 261, no. 5123, pp. 872–878, 1993.

[84] P. Ghoshal, “STUDY OF AN INITIALIZATION METHOD AND STOPPING CRITERIA FOR NONLINEAR OPTIMIZATION,” Oklahoma State University, 2008.

[85] N. Smith, “Binning: An Alternative to Point Maps,” *Mapbox Blog*, 2012. [Online]. Available: https://www.mapbox.com/blog/binning-alternative-point-maps/.

[86] Y. Kleiner and B. Rajani, “Comprehensive review of structural deterioration of water mains: statistical models,” *Urban Water*, vol. 3, no. 3, pp. 131–150, Sep. 2001.

[87] S. Yamijala, S. D. Guikema, and K. Brumbelow, “Statistical models for the analysis of water distribution system pipe break data,” *Reliab. Eng. Syst. Saf.*, vol. 94, no. 2, pp. 282–293, Feb. 2009.

[88] J. P. Dellhomme, “Kriging in the Hydrosciences,” *Adv. Water Resour.*, vol. 1, no. 5, pp. 251–266, 1978.

[89] J.-J. Lee, C.-S. Jang, S.-W. Wang, and C.-W. Liu, “Evaluation of potential health risk of arsenic-affected groundwater using indicator kriging and dose response model.,” *Sci. Total Environ.*, vol. 384, no. 1–3, pp. 151–62, Oct. 2007.

[90] S.-S. Jeon and T. D. O’Rourke, “Northridge Earthquake Effects on Pipelines and Residential Buildings,” *Bull. Seismol. Soc. Am.*, vol. 95, no. 1, pp. 294–318, Feb. 2005.

[91] E. Vazquez, G. Fleury, and E. Walter, “Kriging for indirect measurement, with application to flow measurement,” *IEEE Trans. Instrum. Meas.*, vol. 55, no. 1, pp. 343–349, 2006.

[92] S. Andreou, “Maintenance Decisions for Deteriorating Water Pipelines,” *J. Pipelines*, vol. 7, pp. 21–31, 1987.

[93] D. R. Cox, “Regression Models and Life Tables (with Discussion),” *J. R. Stat. Soc. Ser.*, vol. B, no. 34, pp. 187–220, 1972.

[94] J. D. Kalbfleisch and R. L. Prentcie, *The Statistical Analysis of Failure Time Data*. New York: Wiley, 1980.

[95] J. P. Klein and M. Moeschberger, *Survival Analysis*. New York: Springer, 1997.

[96] S. Banerjee, M. M. Wall, and B. P. Carlin, “Frailty modeling for spatially correlated survival data, with application to infant mortality in Minnesota,” *Biostatistics*, vol. 4, no. 1, pp. 123–42, Jan. 2003.

[97] L. Huang, M. Kulldorff, and D. Gregorio, “A spatial scan statistic for survival data,” *Biometrics*, vol. 63, no. 1, pp. 109–18, Mar. 2007.

[98] Y. Li and L. Ryan, “Modeling spatial survival data using semiparametric frailty models,” *Biometrics*, vol. 58, no. 2, pp. 287–97, Jun. 2002.

[99] C. Vance and J. Geoghegan, “Temporal and spatial modelling of tropical deforestation: a survival analysis linking satellite and household survey data,” *Agric. Econ.*, vol. 27, no. 3, pp. 317–332, Nov. 2002.

[100] J. Shi, J. H. Garrett, and L. Soibelman, “A Study of Associations between Hydraulic Characteristics and Pipe Breakage in Drinking Water Distribution System,” in *Proceedings of the 2012 eg-ice Workshop*, 2007, pp. 1–10.

[101] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, 2013.

[102] J. Røstum, “Statistical Modeling of Pipe Failures,” Norwegian University of Science and Technology, 2000.

[103] Y. Kleiner and B. Rajani, “I-WARP: Individual Water mAin Renewal Planner,” *Drink. Water Eng. Sci.*, vol. 3, no. 1, pp. 71–77, May 2010.

[104] A. Haldar and S. Mahadevan, *Probability, Reliability, and Statistical Methods in Engineering Design*. New York: John Wiley & Sons, Inc., 2000.

[105] A. Colombo and B. W. Karney, “Pipe breaks and the role of leaks from an economic perspective,” *Water Supply*, vol. 3, no. 1, pp. 163–169, 2003.

[106] E. Cabrera, M. A. Pardo, E. J. Cabrera, and R. Cobacho, “Optimal Scheduling of Pipe Replacement , Including Opportunity , Social and Environmental Costs,” in *Pipelines 2007*, 2007, no. 2001, pp. 1–11.

[107] M. Moglia, S. Burn, and S. Meddings, “Descision support system for water pipeline renewal prioritization,” *ITcon*, vol. 11, no. Decision Support Systems for Infrastructure Management, pp. 237–256, 2006.

[108] G. C. Dandy and M. O. Engelhardt, “Multi-Objective Trade-Offs between Cost and Reliability,” no. April, pp. 79–88, 2006.

[109] Palisade Corporation, “Evolver: The Genetic Algorithm Solver for Microsfot Excel,” Ithaca, NY, 2013.

[110] “Presentation on City of Westminster Water System Pressures,” 2013.

[111] E. Renaud, J. C. De Massiac, B. Brémond, and C. Laplaud, “SIROCO , A Decision Support System for Rehabilitation Adapted for Small and Medium Size Water Distribution Companies,” in *LESAM 2007 - 2nd Leading Edge Conference on Strategic Asset Management*, 2007, pp. 1–15.

[112] D. A. Savic, O. Giustolisi, and D. Laucelli, “Asset deterioration analysis using multi-utility data and multi- objective data mining,” *J. Hydroinformatics*, vol. 11, no. 3, pp. 211–224, 2009.

[113] M. Ben-akiva and D. Bolduc, “Short Paper The Combined Estimator Approach to Model Transferability and Updating,” *Empir. Econ.*, vol. 1, pp. 167–175, 1995.

[114] S. Hasan and S. Ukkusuri, “Transferability of Hurricane Evacuation Choice Model : Joint Model Estimation Combining Multiple Data Sources,” no. May, pp. 548–556, 2012.

[115] S. Siuhi, J. L. Mwakalonge, and J. Perkins, “Spatial Transferability : Analysis of the Regional Automobile-Specific Household-Level Carbon Dioxide (CO 2) Emissions Models,” *J. Transp. Res. Forum*, vol. 52, no. 2, pp. 7–26, 2013.

[116] Environmental Protection Agency (EPA), “Deteriorating Buried Infrastructure Management Challenges and Strategies,” Washington, DC, 2002.

[117] P. Kovesi, “DBSCAN.” The University of Western Australia, 2013.

[118] D. Barr and V. Martin, “Using Soft Starters to Control Water Hammer Without the Headache,” in *Ohio AWWA State Conference*, 2008.

[119] WEG Electronic Corporation, “Choosing a Variable Frequency Drive or Soft Starter based on Your Application Need,” 2014.

[120] R. A. Howard, “Information value theory,” *IEEE Trans. Systesm Sci. Cybern.*, vol. 2, pp. 22–26, 1966.

[121] K. Claxton, P. J. Neumann, S. Araki, and M. C. Weinstein, “Bayesian Value of Information Analyses: An Application to a Policy Model of Alzheimer’s Disease,” *Int. J. Technol. Assess. Health Care*, vol. 17, no. 1, pp. 38–55, Jan. 2001.

[122] A. Repo, “The value of information: Approaches in economics, accounting, and management science,” *J. Am. Soc. fo Inf. Sci.*, vol. 40, no. 2, pp. 68–85, 1989.

[123] H. L. Lee, K. C. So, and C. S. Tang, “The Value of Information Sharing in a Two-Level Supply chain,” *Manage. Sci.*, vol. 46, no. 5, pp. 626–643, 2000.

[124] F. Yokota and K. M. Thompson, “Value of information analysis in environmental health risk management decisions: past, present, and future.,” *Risk Anal.*, vol. 24, no. 3, pp. 635–50, Jun. 2004.

[125] a. I. Khader, D. E. Rosenberg, and M. McKee, “A decision tree model to estimate the value of information provided by a groundwater quality monitoring network,” *Hydrol. Earth Syst. Sci.*, vol. 17, no. 5, pp. 1797–1807, May 2013.

[126] X. Liu, J. Lee, P. K. Kitanidis, J. Parker, and U. Kim, “Value of Information as a Context-Specific Measure of Uncertainty in Groundwater Remediation,” *Water Resour. Manag.*, vol. 26, no. 6, pp. 1513–1535, Jan. 2012.

[127] M. Messer, J. H. Panchal, V. Krishnamurthy, B. Klein, P. D. Yoder, J. K. Allen, and F. Mistree, “Model Selection Under Limited Information Using a Value-of-Information-Based Indicator,” *J. Mech. Des.*, vol. 132, no. 12, p. 121008, 2010.

[128] D. G. Krige, “A statistical approach to some mine valuations and allied problems at the Witwatersrand,” University of Witwatersrand, 1951.

[129] K. Johnston, K. K. Ver-Hoef, and N. Lucas, “Using ArcGIS Geostatistical Analyst,” Redlands, CA, 2003.

[130] ESRI, “How Kriging Works.” [Online]. Available: http://resources.arcgis.com/en/help/main/10.1/index.html#//009z00000076000000.

[131] B. K. Krivoruchko and S. D. Team, “Empirical Bayesian Kriging Implemented in ArcGIS Geostatistical Analyst,” Redlands, CA, 2012.