Evaluating the Risk of Water Main Failure Using a Hierarchical Fuzzy Expert System

by

Hussam A. Fares

A Thesis in

the Department of Building, Civil, and Environmental Engineering

in Partial Fulfillment of the Requirements

for the Degree of Master of Applied Science at CONCORDIA UNIVERSITY

Montreal, Quebec, Canada

April, 2008

Your file Votre référence
ISBN: 978-0-494-40866-7
Our file Notre référence
ISBN: 978-0-494-40866-7

ABSTRACT

Evaluating the Risk of Water Main Failure Using a Hierarchical Fuzzy Expert System

Hussam A. Fares

Water distribution systems are the most expensive part of the water supply infrastructure system. In Canada and the United States, there are 700 water main breakae every day, and there have been more than 2 million breaks since the beginning of this century, which have cost more than 6 billion Canadian dollars in repairs costs for the two countries. Municipalities and other authorities that manage potable water infrastructure often must prioritize the rehabilitation needs of their water main. This is a serious challenge because the current potable water networks are old (i.e. deteriorated) and require certain modifications to bring them up to acceptable reliability and safety levels within a limited budget. In other words, municipalities need to develop a balanced rehabilitation plan to increase the reliability of their water networks by rehabilitating (first) only those pipelines at high risk of failure.

The objective of this research is to develop a risk model for water main failure, which evaluates the risk associated with each pipeline in the network. This model considers four main factors: environmental, physical, operational, and post-failure factors (consequences of failure) and sixteen sub-factors which represent the main factors. Data are collected to serve two purposes: to build the model and to show its implementation to case studies. The required data are collected from literature review and through a questionnaire sent to the experts in the field of water distribution network management. From the collected data, pipe age is found to have the most significant indication of water main failure risk, followed by pipe material and breakage rate. In order to develop the risk of failure model, hierarchical fuzzy expert system (HFES) technique is used to process the input data, which is the effect of risk factors, and generate the risk of failure index of each water main. In order to verify the developed model, a validated AHP deterioration model and two real water distribution network data sets are used to check the results of the developed model. The results of the verification show that the Average Validity Percent is 74.8 %, which is reasonable considering the uncertainty involved in the collected data. Based on the developed model, an application is built that uses Excel ® 2007 software to predict the risk of failure index. At last, three case studies are evaluated using the developed application to estimate the risk of failure associated with the distribution water mains.

ACKNOWLEDGEMENT

I would like to express my sincere appreciation and thanks to my supervisor Dr. Tarek Zayed for his guidance and constructive comments. His support and encouragement far exceed his duty and the scope of this research.

I would like to thank all the faculty, staff, and friends at Concordia University who supported me throughout the course of my research.

I want to express my deep appreciation to my dear friend Abdelsalam Obeidat for his advice, encouragement, and support. I am also very grateful to Mr. Hassan Al-Barqawi, who gave me excellent technical and practical advices along with insights that helped me to succeed in this research.

I would like to express my appreciation to all the engineers and experts who facilitated my research by their positive participation.

Finally, I would like to express my deep gratitude to my parents, sister, and brother for their support and encouragement, and who provided me with the patience and determination to achieve my graduate degree.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

Chapter I: INTRODUCTION

I.1. Problem Statement

The water distribution system is considered to be the most expensive part of the water supply infrastructure system (Giustolisi *et al.* 2006). In a recent survey conducted by the United States Environmental Protection Agency, it is estimated that $77 billion will be needed to repair and rehabilitate the water main over the next 20 years (Selvakumar *et al.* 2002). In Canada and the United States, there have been more than 2 million breaks since January 2000, costing more than 6 billion Canadian dollars in repair costs on an average of 700 water main breaks every day (Infrastructure Report, 2007). Moreover, providing communities with safe water through a reliable water network has become more and more a topic of concern. Water distribution networks are buried pipelines and as a result, they have received little attention from decision makers. The breakage rate and the high associated costs of failure have reached a level that now draws the attention of both the public and the decision makers. As a result, dealing with the risk of water main failure has been undergoing a great change in concept from reacting to failure events to taking preventive actions that maintain the water main in good working condition.

The risk of failure is defined as the combination of the probability and the impact severity of a particular circumstance that negatively affects the ability of infrastructure assets to meet the objectives of the municipality (InfraGuide, 2006). Risk of water main failure factors can be divided broadly into deterioration and consequence (post failure) factors. The deterioration factors are either responsible for deterioration of the potable water distribution network or they can give an indication of the level of network deterioration. Environmental, physical, and operational factors are included within the deterioration framework. Consequence or post failure factors represent the cost of water main failure and should be considered when evaluating the risk of pipeline failure. Municipalities and other authorities must build long-term and short-term management plans that prioritize the rehabilitation of the water works within their limited budgets in order to upgrade the status of their water main networks. Thus, it is crucial to apply management strategies to upgrade, repair, and maintain the potable water network. These strategies should be built on scientific approaches that consider the risk of pipeline failure in tandem with all of the failure factors.

I.2. Research Objectives

The objectives of the current research can be summarized as follows:

- Design a risk of water main failure model to evaluate the risk associated with each pipeline in the network.
- Propose a failure risk scale that provides guidance to decision makers.
- Develop an automated tool that helps water system managers make their short and long terms management plans.

I.3. Research Methodology

The research methodology consists of several stages. It starts with a comprehensive literature review of the risk of water main failure followed by data collection (model information data and case study data). Next, a hierarchical fuzzy expert system (HFES) is developed based on the collected model information data. A failure risk scale is proposed that will guide the network operators on how to best manage their networks. Then, the developed model and application is tested and verified. After that, three case studies application is analyzed which utilize the developed model to assess the risk of failure of the water distribution network. An Excel-based application is built to allow the developed model to be used by municipalities and other authorities to manage their water main.

I.3.1. Literature Review

All the topics related to the risk of water main failure are reviewed in order to have a better overview of the topic and how to achieve the research objective. In the literature review, many topics are studied such as: water main classification, risk of pipeline failure, risk evaluation process and modeling approaches, risk of water main failure, fuzzy logic, and failure risk scales.

I.3.2. Data Collection

The data collection consists of two stages that are required to develop the water main failure risk model and to run it. In stage one, the data is collected from many sources such as the literature review and experts via a questionnaire collected from twenty experts. The data collected is the weights of various factors to be incorporated in the model and the performance of each factor. In stage two, data from case studies are collected from real networks under operation and are presented to the developed model to assess the water main failure risk.

I.3.3. Hierarchical Fuzzy Expert System Model

A model is developed to evaluate the risk of water main failure. The developed model considers four main risk factors which have sixteen sub-factors that represent both deterioration of the water distribution network and the failure consequences. In the light of the literature review, a failure risk scale is proposed to help decision makers in water resources management (i.e. companies, municipalities) make informed decisions and establish their rehabilitation plans. The developed model is analyzed and verified and the results show that the model is robust and reliable.

I.4. Thesis Organization

As stated earlier, the main objective of this research is to build a water main failure risk model using fuzzy expert system. Accordingly, the thesis is organized to achieve this objective.

The literature review is compiled and organized in Chapter II, including water main classification, risk of water main failure, risk evaluation process and modeling approaches. Fuzzy logic literature is also reviewed, along with its application in the field of pipeline management, together with expert opinions.

Chapter III gives an overview of the research methodology followed in this research. Moreover, it includes an overview of the developed hierarchical fuzzy expert system model for water main failure risk and the Excel-based application development.

Chapter IV describes the data collection process. Data are needed in order to perform two tasks: to build the model and to apply a case study using the developed model.

Chapter V explains the developed failure risk model. It introduces the risk factors incorporated into the model, the definition of the fuzzy sets for each factor, the development of the hierarchical model, the fuzzy rules extraction, and the fuzzy defuzzification. Moreover, the development of a failure risk index scale is described in this chapter. The application of the case studies to the developed system is also presented. Three case studies are introduced, processed and analyzed using the developed system. In addition, the process of sensitivity analysis and verification of the developed model/system is described in this chapter.

Chapter VI explains the developed application. It is based on Excel ® 2007 software. This section explains the different parts of the application and how to use it.

Chapter VII presents the conclusions and recommendations. It includes the limitations of the developed model and application, research contribution, research enhancement and extension of the research in the future.

Chapter II: LITERATURE REVIEW

Distribution networks often account for up to 80% of the total expenditure involved in water supply systems (Kleiner and Rajani, 2000). The breakage rates of the water main increase and their hydraulic capacity decreases as the pipelines deteriorate. Engineering systems are designed, constructed, and operated under unavoidable conditions of risk and uncertainty. In order to solve the risk of water main failure problem, many topics are reviewed as shown in Figure II.1 which illustrates the literature reviewed in this chapter.

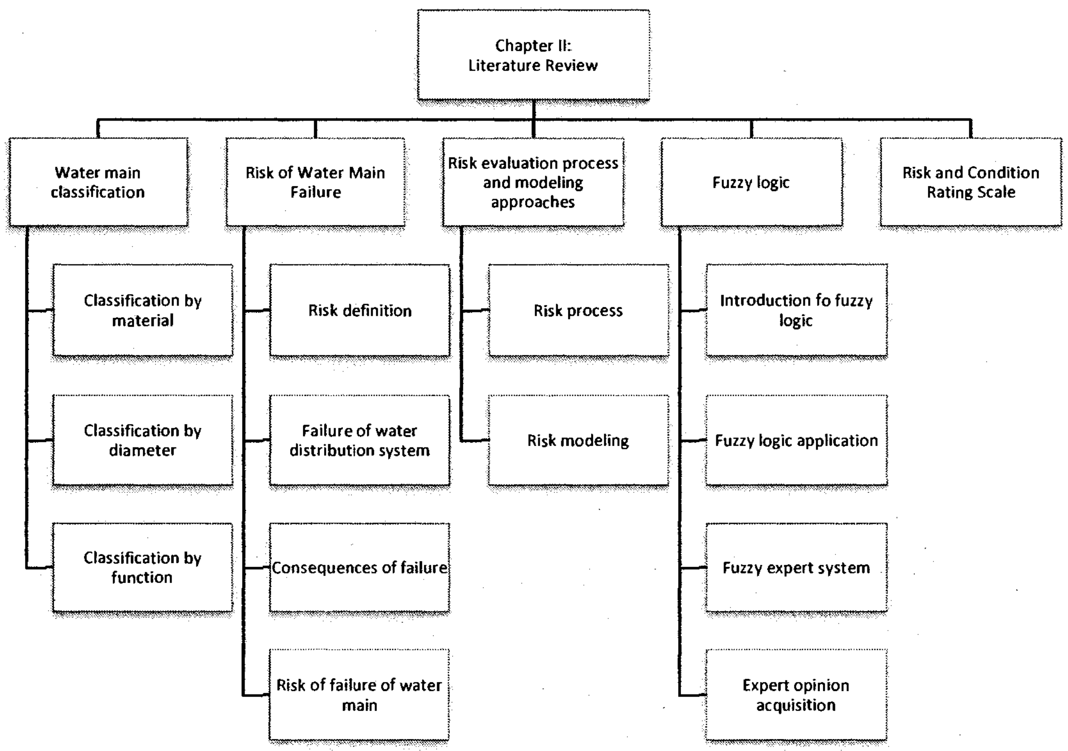


Figure II.1 – Literature review chapter layout.

II.1. Water Main Classification

Water main can be classified depending on its characteristics by following various approaches such as: by material, diameter, function. In this context, the pipelines' different classifications are shown as follows:

II.1.1. Classification by Material

Mainly, there are three main categories of pipeline materials that are used in the construction of pressurized pipelines. They are: Cement-Based pipes, Plastic pipes, and Metallic pipes. Each category of pipeline material contains a variety of materials (Najafi, 2005). Steel, cast iron (CI), ductile iron (DI), reinforced concrete (RC), pre-stressed concrete cylinder pipe (PCCP), and asbestos cement (AC) are used in the construction of large-diameter water mains, whereas more recently, polyvinyl chloride (PVC) and polyethylene (PE) pipes have been widely used, especially in the lower diameter range (Rajani *et al.* 2006). It is worth mentioning that there is another type of pipeline material, Verified Clay, which is only used in sewer pipelines due to its low tensile strength.

II.1.2. Classification by Diameter

Water main can be classified according to its diameter into three groups: small diameter (2 in. to 8 in. or 50 mm to 203 mm), medium diameter (10 in. to 30 in. or 254 mm to 762 mm), and large diameter (36 in. to 72 in. or 914 mm to 1829 mm) (Raven, 2007). Pipelines are classified depending on the structural behavior of the pipeline. For instance, large diameter pipes have more beam strength than small diameter pipes (Najafi, 2005).

II.1.3. Classification by Function

Water main can be classified according to its function into two main categories: transmission and distribution lines. The function of transmission pipelines is to transfer water from a main source to a storage system (i.e. water tanks). They are considered the most expensive part of the system because of their higher initial construction costs (i.e. material, installation, equipments). The function of distribution lines is to carry water out from the storage system to the domestic users (i.e. residential buildings or industrial factories). The minimum diameter for a distribution pipe is two inches, and the minimum diameter required for serving fire hydrants is six inches (Al Barqawi, 2006).

II.2. Risk of Water Main Failure

II.2.1. Risk Definition

Risk is defined by InfraGuide (2006) as the combination of the probability and impact severity of a particular circumstance that negatively affects the ability of infrastructure assets to meet the objectives of the municipality. Moreover, the probability is defined as the likelihood of an event occurring.

There are many other risk definitions that share the same concept. Some are given by Kirchhoff and Doberstein (2006) as:

- The potential for the realization of unwanted, adverse consequences to human life, health, property, and/or the environment.
- A measure of economic loss or human injury in terms of both the incident likelihood and the magnitude of the loss or injury.
- A measure of the probability and magnitude of adverse consequences.

The same authors also stated some definitions of risk assessment as:

- The process by which the form, dimension, and characteristics of risks are estimated, and
- The process of gathering information about adverse effects in a structured way and the forming of a judgment about them.

Risk assessment is also defined by (InfraGuide, 2006) as the analysis of the severity of the potential loss and the probability that the loss will occur, leading to the quantification of impacts.

There is not yet a single international consent about the definition of risk or risk assessment terminology; however, all of the definitions have almost the same implicit meaning. Several risk assessment methods are used in the industry. The selection of a method depends on many factors, such as: system complexity, availability of historical data, and the validity required by the analysis. The most common are failure probability methods and ranking systems (Mohitpour *et al.* 2003).

The difference between risk assessment and risk management is that risk assessment tries to answer the following questions:

- What can go wrong?
- What is the likelihood that it would go wrong?
- What are the consequences?

While risk management continues this process by additionally attempting to answer the following questions:

- What can be done and what options are available?
- What are the associated trade-offs in terms of the costs, benefits, and risks?
- What are the impacts of current management decisions on future options? (Haimes, 2004).

II.2.2. Failure of a Water Distribution System

Failure of a pipeline can be defined as the *unintentional release of pipeline contents* or *loss of integrity*. However, a pipeline can fail in other ways that do not involve a loss of contents. *Failure to perform a pipeline's intended function* is a more general definition of pipeline failure (Muhlbauer, 2004). The more precise definition of pipeline failure is *the inability to satisfy basic requirements from the distribution system, failure to satisfy customer demand or failure to maintain pressures within specific limits*. The types of water distribution failure can be categorized into: 1) performance failure and 2) mechanical failure (Ozger, 2003).

There are many causes of performance failure. A principle cause is when the actual demand on the network exceeds the network capacity (design demand). Another cause is when the hydraulic capacity of the network is reduced below the actual demand due to the network's deterioration with age. Performance failure is also called demand variation failure or demand failure. The second general type of failure is mechanical failure, which is associated with the failure of components of the distribution system such as pipes, pumps, control valves, treatment plant, and supervisory and data acquisition system. The most common type of mechanical failure is pipe failure (Ozger, 2003). The causes of pipe failure are categorized into time-dependant (dynamic), static, and operational factors. Examples of non-static factors are pipe age, soil moisture, temperature, and soil electrical resistivity. Examples of static factors are pipe material, pipe diameter, surrounding soil type, and internal pressure, whereas replacement rates, cathodic protection, and water pressure are examples of operational factors. These factors are shown in Table II.1 (Kleiner and Rajani, 2000; Kleiner and Rajani, 2002; Kleiner *et al.* 2006; Pelletier *et al.* 2003). InfraGuide (2003) summarized the structural failure modes for each of the common water main materials as shown in Table II.2. Failure occurs mainly when structural deterioration of a pipe reduces its capacity to resist stresses imposed on it by external and internal factors (Sadiq *et al.* 2004).

Table II.1 – Factors affecting pipe breakage rates (Kleiner and Rajani, 2002).

Static	Dynamic	Operational
Material	Age	Replacement rates
Diameter	Temperatures (soil, water)	Cathodic protection
Wall thickness	Soil moisture	Water pressure
Soil (backfill)	Soil electrical	
characteristics	Resistivity	
Installation	Bedding condition	
	Dynamic loadings	

Table II.2 – Structural failure modes for common water main materials (InfraGuide, 2003).

Water Main Material	Structural Failure Modes
Cast Iron (CI) Small diam (<375 mm) Large diam (>500 mm) Medium diam (375-500 mm)	• Circumferential breaks, split bell, corrosion through holes • Longitudinal breaks, bell shear, corrosion through holes • Same as small, plus longitudinal breaks and spiral cracking, blown section
Ductile Iron (DI)	• Corrosion through holes
Steel	• Corrosion through holes, large diameter pipes are susceptible to collapse
Polyvinyl Chloride (PVC)	• Longitudinal breaks due to excessive mechanical stress • Susceptible to impact failure in extreme cold condition (i.e. far north)
High Density Polyethylene (HDPE)	• Joint imperfections, mechanical degradation from improper installation methods, susceptible to vacuum collapse for lower pressure ratings
Asbestos Cement (AC)	• Circumferential breaks, pipe degradation in aggressive water • Longitudinal splits
Concrete Pressure Pipe (CPP)	• Pipes with pre-stressed wires may experience ruptures due to loss of pre-stressing upon multiple wire failure. • Pipe degradation in particularly aggressive soils, corrosion of pipe canister, concrete damage due to improper installation methods

Sources of failure can be categorized into five groups (InfraGuide, 2006):

1- Natural occurring events: like fire, storm, flood, and earthquake. The timing of these types of events is unknown and uncontrollable but their probability and severity can be statistically predicted.

2- External impacts: as a result of failure by an outside party such as power failure, spills, labor strike. This source of risk is unpredictable making it difficult to calculate the probability of failure. However, the consequences can be mitigated by management plans.

3- External aggressions: deliberate acts of terrorism that results in destruction of assets. The consequences of failure can be reduced through security and protection programs to the strategically important facilities.

4- Aging infrastructure and physical deterioration: the condition of the infrastructure and its deterioration can be predicted and determined. Many factors contribute to a pipeline's deterioration. These factors are categorized into three groups: Physical factors, Environmental factors, and Operational factors as shown in Table II.3 (InfraGuide, 2003).

5- Operation risk of failure: this category arises as a result of the way the infrastructure is designed, managed, and operated to meet the organizational objectives. It includes design standards risks, management policies, and operator behavior. This risk can be reduced through pro-active condition and performance assessment and inspection of assets at regular intervals and through preventive maintenance programs.

Table II.3 – Factors that Contribute to Water System Deterioration (InfraGuide, 2003).

Factors		Explanation
Physical	Pipe material	Pipes made from different materials fail in different ways.
	Pipe-wall thickness	Corrosion will penetrate thinner walled pipe more quickly.
	Pipe age	Effects of pipe degradation become more apparent over time.
	Pipe vintage	Pipes made at a particular time and place may be more vulnerable to failure.
	Pipe diameter	Small diameter pipes are more susceptible to beam failure.
	Type of joints	Some types of joints have experienced premature failure.
	Thrust restraint	Inadequate restraint can increase longitudinal stresses.
	Pipe lining and coating	Lined and coated pipes are less susceptible to corrosion.
	Dissimilar metals	Dissimilar metals are susceptible to galvanic corrosion.
	Pipe installation	Poor installation practices can damage pipes, making them vulnerable to failure.
	Pipe manufacture	Defects in pipe walls produced by manufacturing errors can make pipes vulnerable to failure. This problem is most common in older pit cast pipes.
Environmental	Pipe bedding	Improper bedding may result in premature pipe failure.
	Trench backfill	Some backfill materials are corrosive or frost susceptible.
	Soil type	Some soils are corrosive; some soils experience significant volume changes in response to moisture changes, resulting in changes to pipe loading. Presence of hydrocarbons and solvents in soil may result in some pipe deterioration.
	Groundwater	Some groundwater is aggressive toward certain pipe materials.
	Climate	Climate influences frost penetration and soil moisture. Permafrost must be considered in the north.
	Pipe location	Migration of road salt into soil can increase the rate of corrosion.
	Disturbances	Underground disturbances in the immediate vicinity of an existing pipe can lead to actual damage or changes in the support and loading structure on the pipe.
	Stray electrical currents	Stray currents cause electrolytic corrosion.
	Seismic activity	Seismic activity can increase stresses on pipes and cause pressure surges.
Operational	Internal water pressure, transient pressure	Changes to internal water pressure will change stresses acting on a pipe.
	Leakage	Leakage erodes pipe bedding and increases soil moisture in the pipe zone.
	Water quality	Some water is aggressive, promoting corrosion
	Flow velocity	Rate of internal corrosion is greater in unlined dead-ended mains.
	Backflow potential	Cross connections with systems that do not contain potable water can contaminate water distribution systems.
	O&M practices	Poor practices can compromise structural integrity and water quality.

II.2.3. Consequences of Failure

A judgment of the potential consequences is inherent in any risk evaluation. This is the answer to the question of: if something goes wrong, what are the consequences? Consequence implies a loss of some kind. Losses can be quantified, in terms of damaged buildings, vehicles, and other property; costs of service interruption; cost of the lost product; cost of the cleanup; and so on. The consequences of failure are categorized into two groups: direct and indirect consequences as shown in Table II.4 (Muhlbauer, 2004; Bhave, 2003).

Table II.4 – Categories of failure consequences.

Direct consequences	Indirect consequences
• Property damages • Damages to human health • Environmental damages • Loss of product • Repair costs • Cleanup and remediation costs	• Litigation and contract violations, • Customer dissatisfaction, • Political reactions, • Loss of market share, and • Government fines and penalties.

Some of these consequences are monetized in a straightforward process. However, for indirect consequences and environmental damages, it is more difficult to quantify the consequences with a monetary value (Muhlbauer, 2004).

The consequences of failure are different among pipelines and vary with time relative to a business cycle. They are also affected by pipeline flow load and by the generated revenue from a pipeline (Nikolaidis *et al.* 2005).

Skipworth *et al.* (2001) made a comparison between the Whole Life Costing (WLC) approach and the Risk Score-Based approach. Both approaches have the same definition of risk as probability multiplied by consequences; however the main difference between the two is how consequence is measured. In the WLC approach, the consequence is to be defined in absolute values (in monetary units) whereas in the Risk Score-Based approach, the consequence is defined by a scoring system.

II.2.4. Water Main Failure Risk

This section provides an overview of the researches and various efforts related to water main failure risk. In their efforts to assess the risk or the probability of pipeline failure, researchers have used a broad variety of techniques. Some of the techniques are: fuzzy logic (Marshall *et al.* 2005; Kleiner *et al.* 2006; Rajani *et al.* 2006), hierarchical holographic modeling (Ezell *et al.* 2000), first order reliability modeling using Monte Carlo Simulation (Sadiq *et al.* 2004), the analytical hierarchy process (Bandyopadhyay *et al.* 1997; Al Barqawi, 2006), statistical non-homogeneous Poisson modes (Moglia *et al.* 2006; Rogers, 2006), fault tree analysis (Yuhua and Datao, 2005), probability density function (Souza and Chagas, 2001), artificial neural networks (Christodoulou *et al.* 2003; Al Barqawi, 2006), the non-homogenous Markov process (Kleiner *et al.* 2004), multi-criteria decision making (Yan and Vairavamoorthy, 2003), and the Bayesian belief network expert system (Hahn *et al.* 2002). These efforts are thoroughly explained in the following paragraphs.

Marshall *et al.* (2005) evaluated the risk of failure of large diameter pre-stressed concrete cylinder pipelines. A simplified strength model was developed to evaluate the remaining strength of pre-stressed concrete pipe as it ages. This model is derived from a process of inspecting pipelines using direct observations and non-destructive tests. Many parameters are included in the fuzzy risk model, such as: parameters affecting the rate of deterioration, parameters affecting repair time, and the consequences of failure.

Kleiner *et al.* (2006) developed a methodology to evaluate pipeline failure risk using the fuzzy logic technique. The model consists of three parts: possibility of failure, consequence of failure and a combination of these two to obtain failure risk. In the possibility of failure part, a seven-grade fuzzy set is used to describe the asset condition rating and a nine-grade possibility of failure is used to reflect the possibility of failure. The failure condition rating is fuzzified (remapped) on the nine-grade possibility of failure. In the consequences of failure part, the severity of an asset failure consequence is described in a nine-severity grade. The consequences of failure can be in the form of direct cost, indirect cost, and social cost. The risk of failure is assessed by combining the probability of failure with the consequences of failure in nine fuzzy triangular subsets.

Rajani *et al.* (2006) used a fuzzy synthetic evaluation technique to translate observations from visual inspection and non-destructive tests into water main condition ratings. The process involves three steps, (1) *fuzzification* of raw data (measurements of the distress indicators), (2) *aggregation* of distress indicators towards the condition rating, and (3) *defuzzification* that adjusts the condition rating to a practical crisp format.

Ezell *et al.* (2000) introduced the Probabilistic Infrastructure Risk Analysis model (IRAM). This system is developed for small community water supply and treatment systems. It consists of four phases. In phase 1, the infrastructure failure threats are identified by means of system decomposing. The target of phase 2 is to provide information that describes the state of consequences for a scenario executed against the system under study. An event tree is used together with expert opinion to determine the failure probability of each path in the tree and the inherent consequence. In phase 3, the consequence and the probability of failure are combined together to identify the high risk factors, which are used to manage the infrastructure in phase 4 by setting the acceptance risk level.

Sadiq *et al.* (2004) developed a method for evaluating the time-dependent reliability of underground grey cast iron water mains, and for identifying the major factors that contribute to water main failures. The first-order reliability method is used, which employs a Monte Carlo simulation. In a Monte Carlo simulation, a set of random values is generated for each input parameter of the model (assumed to be independent), in accordance with a predefined probability density function. However, the consequence of failure, which is a part of risk calculation, is ignored and here the term "risk" refers solely to the probability of failure. A sensitivity analysis showed that two of the parameters of the corrosion model (the scaling constant for pitting depth and the corrosion rate inhibition factor) were the largest contributors to the variability in the pipe's time to failure.

Bandyopadhyay *et al.* (1997) established a cost-effective maintenance program for a petroleum pipeline through risk analysis. The Analytical Hierarchy Process (AHP) is used to carry out the risk assessment. The methodology followed in this study starts with risk factors identification which can be listed as: corrosion, external interference, acts of god, construction and materials defects, and other reasons such as human error, operational error and equipment malfunction. The second step is to formulate the risk structure model using AHP, which determines the relative severity and probability of each risk factor. Then, maintenance/inspection strategy requirements are determined in order to mitigate the risk. Cost of failure is classified into four categories according to the intensity and is estimated using the Monte Carlo simulation. A cost/benefit analysis is carried out at the end to justify an investment proposal.

Moglia *et al.* (2006) explained the uses of the Decision Support System PARMS-PLANNING which was developed to support the long term assessment of costs and the implications of different management and operational asset management strategies. Risks associated with different scenarios are assessed using a standard risk management approach where risk is calculated by combining the output of failure prediction models with the output of cost assessment models. The failure prediction models use both a statistical Non-Homogeneous Poisson model and a physical/probabilistic model that provide failure rates and failure probabilities for each year into the future. The cost model is based on user input. The specific costs are classified into: pipeline renewal, valve insertions, pipe repairs, supply interruptions, and failure consequences. In the risk calculation model, a risk-based approach is based on the calculation of risk for different actions and scenarios where risk refers to an uncertain event with unwanted consequences. This risk is calculated as the statistical expectation of future costs caused by failure.

Yuhua and Datao (2005) analyzed the failure risk of oil and gas pipelines using fault tree analysis. They divided the causes of failure of gas and oil pipelines into 44 failure causes, which are categorized using fault tree analysis (FTA). The steps to be followed using FTA are: 1) Select experts to form evaluation committee, 2) Convert linguistic terms to fuzzy numbers, 3) Convert fuzzy numbers into fuzzy possibilities, and 4) Transform fuzzy possibility scores into fuzzy failure probability (FFP). This method uses expert opinion to evaluate the possibility of each event causing a failure. Next, the possibility of failure is converted into a fuzzy possibility score and then into the fuzzy possibility of failure. It is worth to note that the methodology explained above calculates the possibility of failure for oil and gas pipelines due to each failure-causing event and does not consider the actual condition of the pipeline in service.

Souza and Chagas (2001) applied the probability theory to evaluate and quantify the risk associated with water pollution. This involves identifying risk sources, failure probability, and the consequences of failure. The probability theory is useful for a system with a consistent set of data. However, for systems without a consistent set of data, the possibility of good results will be limited, because the probability density function for all the sets of random variables is required in this probability theory in order to measure the risk of any environmental system. Fuzzy Set Theory could be a better means to determine this sort of risk when only inconsistent data sets are available.

Christodoulou *et al.* (2003) used Artificial Neural Networks (ANN) to analyze the preliminary water main failure risk in an urban area with historical break data spanning two decades. The type of ANN used in this study is the backpropagation algorithm. The outputs of backpropagation ANN are the age to failure of each pipe segment, the observation outcome (a break or a non-break), and the relevant weights of the risk factors. Their study indicates that the number of previous breaks, the material, diameter, and length of pipe segments are the most important risk factors for water main failures.

Al Barqawi (2006) designed two condition rating models for water mains using artificial neural networks (ANN) and the analytical hierarchy process (AHP). In his research, he considered only the deterioration factors (physical, operational, and environmental). Using the ANN model, he concluded that the most important factors are breakage rate and pipe age. However, when using an integrated ANN/AHP model, pipe age, pipe material, and breakage rate are the most effective factors in evaluating the current condition of water mains. He proposed a condition rating scale from 0 to 10 divided into 6 regions which describe the status of the water main.

Kleiner *et al.* (2004) used a fuzzy rule-based, non-homogeneous Markov process to model the deterioration process of buried pipes. The deterioration rate at a specific time is estimated based on the asset's age and condition state using a fuzzy rule-based algorithm. Then, the possibility of failure is estimated for any age of the pipeline based on the deterioration model. The possibility of failure is coupled with the failure consequence through a matrix approach to obtain the failure risk as a function of the pipe's age.

Rogers (2006) developed a model to assess water main failure risk. He used the Power Law form of a Non-Homogeneous Poisson Process (NHPP) and Multi-Criteria Decision Analysis (MCDA) based on the Weighted Average Method (WAM) to calculate the probability of failure. Moreover, the developed model considers the consequence of failure using "what-if" infrastructure investment scenarios. The probability of failure and the consequences are directly related by a multiplication operation in order to determine the associated risk.

Yan and Vairavamoorthy (2003) proposed a methodology to assess pipeline condition using Multi-Criteria Decision Making (MCDM) techniques which combine the available pipe condition indicators into one single indicator. Both fuzzy set theory and its arithmetic corollaries are incorporated in the Composite Programming to form Fuzzy Composite Programming (FCP). The model starts by converting the linguistic variables into fuzzy numbers. These factors are normalized to allow them to be combined and aggregated after assigning weights to the different indicators. The output of the model is a fuzzy number that reflects the condition of each pipeline, which is ranked accordingly.

Hahn (2002) developed a knowledge-based expert system to predict the criticality of sewer pipelines. The expert system considers information about the environment and the state of a sewer line through an extensive set of relationships that describe failure impact mechanisms. The author used a Bayesian belief network to develop the expert system denoted as SCRAPS "Sewer Cataloging, Retrieval and Prioritization System". Six failure mechanisms that contribute to the likelihood of failure, and two consequences mechanisms that contribute to the consequences of failure are included in the model. The developed model was evaluated through three approaches: the consequence of failure, the likelihood of failure, and both the consequence and the likelihood of failure. The output of the model is the pipe line criticality or risk of failure and is categorized into three ranges and groups (high, moderate, and low).

There are other research efforts that were undertaken in disciplines other than water/sewer main pipelines such as reactor pipelines (Vinod *et al.* 2003), petroleum pipelines (Bandyopadhyay *et al.* 1997; Yuhua and Datao, 2005), and pipes transferring hydrogen sulfide (Santosh *et al.* 2006). These efforts are explained in the following paragraphs.

Vinod *et al.* (2003) developed a study aiming at finding the realistic failure frequency of pipe segments based on the degradation mechanisms to be employed in Risk Informed In-Service Inspection in Pressurized Heavy Water Reactor (PHWR) pipes. The primary PHWR piping is made of carbon steel operating at around 300 ^{o}C. The model starts with erosion-corrosion rate calculation and then applies this rate to the First Order Reliability Method to determine the piping failure probability. After that, a Markov model is developed to estimate the realistic failure probability, incorporating the effects of In-Service Inspection which yields the failure probability to be used in Risk-Informed In-Service Inspection.

Santosh *et al.* (2006) performed a study to utilize failure probability in a risk-based inspection of pipelines transferring hydrogen sulphide. This involves categorizing these pipelines based on their orders of failure probabilities. Two steps are followed in this study: 1) estimation of the remaining strength of a pipeline, and 2) evaluating the limit state function of a pipeline that defines the failure criteria

II.3. Risk Evaluation Process and Modeling Approaches

II.3.1. Risk Process

There are five steps in a risk process: 1) Risk modeling, 2) Data collection and preparation, 3) Segmentation, 4) Assessing risks, and 5) Managing risks (Muhlbauer, 2004).

1) Risk modeling: a pipeline risk assessment model is a set of algorithms or rules that use available information and data relationships to measure levels of risk along a pipeline.
2) Data collection and preparation: the collection of all the required information about the pipeline, including inspection data, original construction information, environmental conditions, operating and maintenance history, past failures, and so on. It results in data sets that are ready to be used directly by the risk assessment model.

3) Segmentation: the process of dividing the pipeline into segments with constant risk characteristics, or into measurable pieces. This is required because the risk is rarely constant along a pipeline's length.
4) Assessing risks: the available risk model is applied to the data set in order to evaluate the risk associated with each pipeline segment.
5) Managing risks: this step comprises the decision support process and provides the tools needed to best optimize the allocation of resources.

It is worth mentioning that sometimes risk modeling and data collection is done in the reverse order of this process (Muhlbauer, 2004).

II.3.2. Risk modeling

There are two types of risk assessment approaches -- either quantitative or qualitative. In a quantitative approach, the quantification of the probability and severity of a particular hazardous event can be assessed and the risk is calculated as the product: risk = probability × severity. The quantitative risk assessment approach includes many methods such as *Bayesian inference, fault tree analysis, Monte Carlo analysis,* and *fuzzy arithmetic* as a semi-quantitative method. In a qualitative approach, the probability of an event may not be known, or not agreed upon, or even not recognized as hazardous. Qualitative risk assessment includes many methods such as *Preliminary Risk/Hazard analysis (PHA), Failure Mode and Effects analysis (FMEA), Fuzzy Theory,* etc. (Kirchhoff and Doberstein, 2006; Lee M. , 2006).

Generally, there are three types of risk models. They are matrix, probabilistic, and indexing models (Muhlbauer, 2004).

i. Matrix models

Matrix models are one of the simplest risk assessment structures. This model ranks pipeline risks according to the likelihood and the potential consequences of an event by a very simple scale or a numerical scale (low to high or 1 to 5). Expert opinion or a more complicated application might be used in this approach to rank risks associated with pipelines. A simple risk matrix example is shown in Figure II.2 (Muhlbauer, 2004).

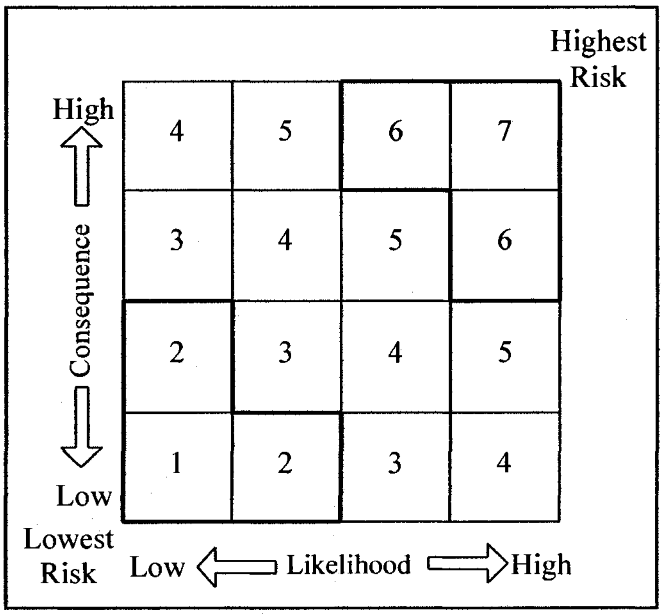


Figure II.2 – Simple risk matrix (Muhlbauer, 2004).

ii. Probabilistic models

Probabilistic risk assessment (PRA), sometimes called Quantitative Risk Assessment (QRS) or Numerical Risk Assessment (NRA), is the most complex and rigorous risk model. It is a rigorous mathematical and statistical technique that relies heavily on historical failure data and event-tree/fault-tree analysis.

This technique is very data intensive. The result of the model is the absolute risk assessments of all possible failure events (Muhlbauer, 2004).

iii. Indexing models

Indexing models and similar scoring models are the most popular risk assessment techniques. In this technique, scores are assigned to important conditions and activities on the pipeline system that contribute to the risk, and weightings are assigned to each risk variable. The relative weight reflects the importance of the item in the risk assessment and is based on statistics where available or on engineering judgment (Muhlbauer, 2004).

II.4. Fuzzy Logic

II.4.1. Introduction to Fuzzy Logic

Lotfi Zadeh developed fuzzy logic in the mid-1960s to solve the problem of representing approximate knowledge that cannot be represented by conventional, crisp methods. A fuzzy set is represented by a membership function. Any "element" value in the universe of enclosure of the fuzzy set will have a grade of membership which gives the degree to which the particular element belongs to the set (Karray and de Silva, 2004). Fuzzy theory relies on four main concepts: (1) ***fuzzy sets***: sets with non-crisp, overlapping boundaries; (2) ***linguistic variables***: variables whose values are both qualitatively and quantitatively described with fuzzy sets; (3) ***possibility distributions***: constraints on the value of a linguistic variable imposed by assigning it a fuzzy set; and (4) ***fuzzy if-then rules***: a knowledge representation scheme for describing a functional mapping or a logic formula that generalizes two-valued logic (Del Campo, 2004). More information is included in "Appendix A: Introduction to Fuzzy Expert System".

II.4.2. Fuzzy Logic Application

Fuzzy logic has been used in many areas in water resources. Bogardi and Duckstein (2002) listed some of them as follows:

- Fuzzy Regression: used where a casual relationship exits with few data points.
- Hydrologic forecasting: Kalman filtering and fuzzy logic are used for short-term and medium term forecasting.
- Hydrologic modeling: where traditional rainfall runoff models can be replaced by fuzzy-rule systems with similar performance.
- Fuzzy-set geostatistics: can be used where imprecise and indirect measurements and small data sets are combined in spatial statistical analysis.
- Incorporation of spatial variability into groundwater flow and transport modeling with fuzzy logic.
- Regional water resources management: when selecting among many alternative management schemes under small data sets and with imprecisely known or modeled objectives.

- Multi-criterion decision-making under uncertainty: can be used when there are multiple and conflicting criteria and when the criteria corresponding to alternative systems are imprecisely known.
- Fuzzy rule-based modeling: used in the classification of spatial hydrometeorological events, climate modeling of flooding, modeling of groundwater flow and transport, forecasting pollutants' transport in surface water.
- Reservoir operation planning: applies fuzzy logic to derive operation rules.
- Fuzzy risk analysis: used in evaluating uncertainty in any or all elements of risk analysis (load, capacity, and consequence).

Fuzzy-based methods have been increasingly applied to civil and environmental engineering problems in recent years, especially when the available information (measured data or expert opinion) is vague and too imprecise to justify the use of numbers. As a solution, fuzzy logic provides a language with syntax and semantics to translate qualitative knowledge into numerical reasoning. Fuzzy systems are used where crisp probabilistic models do not exist (Najjaran *et al.* 2004).

II.4.3. Fuzzy Expert Systems

i. Definition

Usually, systems that can process knowledge are called knowledge-based systems. One of the most popular and successful knowledge-based systems is the expert system (Jin, 2003). Fuzzy logic can be used as a tool to deal with imprecision and the qualitative aspects that are associated with problem solving and in the development of expert systems. Fuzzy expert systems use the knowledge of humans, which is qualitative and inexact. In many cases, decisions are to be taken even if the experts may be only partially knowledgeable about the problem domain, or data may not be fully available. The reasons behind using fuzzy logic in expert systems may be summarized as follows (Karray and de Silva, 2004):

- The knowledge base of expert systems summarizes the human experts' knowledge and experience.
- Fuzzy descriptors (e.g., large, small, fast, poor, fine) are commonly used in the communication of experts' knowledge, which is often inexact and qualitative.
- Problem description by the user may not be exact.
- Reasonable decisions must be taken even if the experts' knowledge base may not be complete.
- Educated guesses need to be made in certain situations.

ii. Fuzzy Expert System Application

An expert system consists of a knowledge base in the form of rules representing specific domains of knowledge, plus a database (Jin, 2003). In this section, the use of expert systems in the field of water resources is reviewed.

Nasiri *et al.* (2007) proposed a fuzzy multiple-attribute decision support expert system to compute the water quality index and to provide an outline for the prioritization of alternative plans.

Najjaran *et al.* (2004) developed a fuzzy logic expert system in order to establish a criterion for predicting the deterioration of a cast/ductile iron water main using soil properties. The fuzzy model determines the relationships between the output and inputs of a system using *antecedent* and *consequent* propositions in a set of IF-THEN rules. The input variables used in this model are selected from soil properties. The output is *corrosivity potential*. The developed fuzzy model is imprecise for a certain range of *corrosivity potential* either because the number of fuzzy rules in the rule base is insufficient or the input and output partitions are not appropriately tuned in some range of their universe of discourse.

iii. Hierarchical fuzzy expert system

Acquiring knowledge for fuzzy rule-based systems can be achieved from human experts or from experimental data using several methods (see A.8.2. Fuzzy Knowledge Rules Acquisition). Mainly, there are three different approaches (Jin, 2003):

- Indirect Knowledge Acquisition
- Direct Knowledge Acquisition
- Automatic Knowledge Acquisition

Reducing the total number of rules and their corresponding computation requirements is one of the most important issues in subjective fuzzy logic systems where the knowledge base rules are solicited from experts in contract to the objective fuzzy systems where the knowledge base rules are extracted from data. The "Curse of dimensionality" is an attribute of subjective fuzzy systems since the number of rules and thus the complexity increases exponentially with the number of variables involved in the model. To minimize this problem, the hierarchical fuzzy system is proposed, where the overall system is divided into a number of low-dimensional fuzzy systems. This has the advantage that the total number of rules increases linearly with the number of the input variables. The number of rules is greatly reduced by using a hierarchical fuzzy system (Lee *et al.* 2003). There are several approaches to deal with hierarchical structures. One is that the output of the last layer as a crisp value can be used as the input of the next layer in the hierarchical fuzzy system. The advantage of this approach is that it will reduce the uncertainty of the new result by reducing the number of fired rules in the new layer, but at the expense of the information of uncertainty that is lost. Another approach is to consider the fuzzy output of the last layer as the fuzzy input of the next layer. The advantage of this approach is that it preserves the information about uncertainty. However, if the fuzzy set is too wide, it will trigger too many rules in the new layer resulting in a very uncertain result. Another approach is to decompose the defuzzification of the output that is used as the input in the new layer into two or more crisp singletons (Gentile, 2004).

II.4.4. Expert Opinion Acquisition

Collecting information from experts will often require the use of qualitative descriptive terms. Verbal labeling has some advantages, including ease of explanation and familiarity. There are many emerging techniques of artificial intelligence systems that aim at solving problems involving incomplete knowledge and the use of descriptive terms through better use of human reasoning. Fuzzy logic as an artificial intelligence system makes use of natural language in risk modeling (Muhlbauer, 2004).

Expert opinions can be acquired with several methods; some of which are (Cooke and Goossens, 2004):

1. Point Values: This method is used in the earlier expert systems like the Delphi method where experts are asked to guess the values of unknown quantities as a form of single point estimates. However, it has many disadvantages which restrict the use of this method such as: scale dependency, no indication of uncertainty in the assessments, and the methodology of processing and combining the experts' judgment as physical measurements.
2. Paired Comparisons: Experts are asked to rank alternatives pair-wise according to certain criteria. Some of the disadvantages of this method are: redundancy of the judgment data and no assessment of uncertainty.
3. Discrete Event Probabilities: Experts are asked to assess the probability of occurrence of uncertain events as a point in the [0, 1] interval. Disadvantages of this method are: careless formulations can easily introduce confusion, and large finite populations are needed to adequately measure the variable.
4. Distributions of Continuous Uncertain Quantities: Experts may be asked to give a unique real value and to give a subjective probability distribution. The probability distribution can be in the form of a cumulative distribution function, density or mass function, or other information such as the mean and standard deviation.

5. Conditionalisation and Dependence: Relevant variables must be specified in the background information and the failure to specify background information can lead experts to conditionalise their uncertainties in different ways, which can introduce "noise" into the assessment process. Variables whose values are not specified in the background information can produce dependencies in the uncertainties of target variables. Experts should be asked to address the dependencies in their subjective distributions for these variables.

II.5. Risk and Condition Rating Scale

The risk of failure scale derives its importance from the need for a common language to be spoken among the different authorities and municipalities. It is also used to compare the condition and status of the infrastructure through a standardization process and to allow decision makers to make informed decisions about the needs of an infrastructure to be maintained or rehabilitated. Moreover, one of the important benefits of the failure risk scale is to track the deterioration process of an asset over time. Developing a failure risk scale usually depends on experts' opinions and experience or on the common practice followed in managing the asset.

II.5.1. Different Types of Scales

Any risk scale consists mainly of three parts: numerical scale, linguistic scale, and sometimes the associated corrective actions or maintenance plan. There are many types of scales. Certain types of scales are more popular and used more in the domain of one specific type of asset than others. This is due to standardization efforts and the studies made in that domain of assets. For example, a five-point scale is widely used in the condition assessment process of underground sewer pipes because it has been adopted by many codes of practice in that area. Some of the more commonly used types of scales are summarized in Table II.5 (Rahman, 2007). It should be mentioned that this table only shows examples and does not include all of the types of scales since it is possible to modify almost any scale to fit an organization's needs.

Table II.5 – Scales types (adapted from Rahman, 2007).

Scale Type	Example
2 Point Scale	Unsatisfactory-satisfactory
3 Point Scale	Poor-fair-good, poor-adequate-good
4 Point Scale	Poor-fair-satisfactory-good
5 Point Scale	Poor-fair-average-good-excellent Unserviceable-poor-fair-good-very good failed-poor-fair-good-new 1-5 scale
6 Point Scale	0-5 scale
10 Point Scale	1-10 rating; 0-9 rating
11 Point Scale	0-10 rating

Al Barqawi (2006) developed a condition rating scale for underground pipelines. The scale is divided into 6 categories ranging numerically form 0 to 10 and linguistically from "critical to excellent" as shown in Figure II.3. He extracted this scale from experts' opinions via a questionnaire.

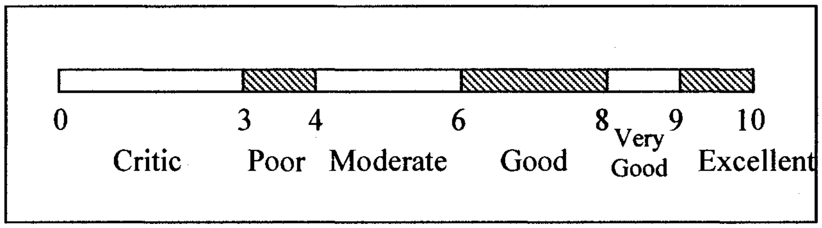


Figure II.3 – Underground pipelines condition rating scale.

Similar to Al Barqawi (2006), Rahman (2007) proposed a scale from 0 to 10 divided into six condition grades (Figure II.4). The purpose of this scale is to fit the results of a condition assessment of different elements of a drinking water treatment plant.

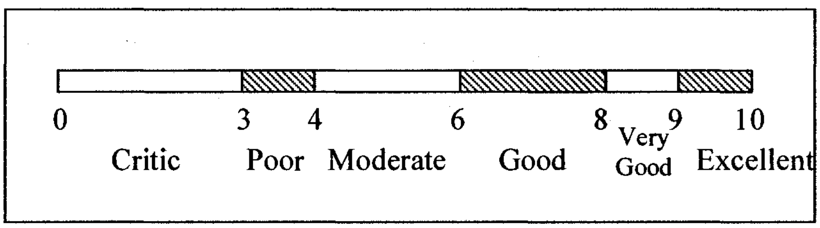


Figure II.4 – Drinking water treatment plant condition rating scale.

Chughtai (2007) used a scale for sewer pipeline condition assessment which consists only of integers and does not allow for intervals, as shown in Figure II.5. The reasoning behind this is that the results of the condition assessment model are only integers and thus there is no need for an interval scale.

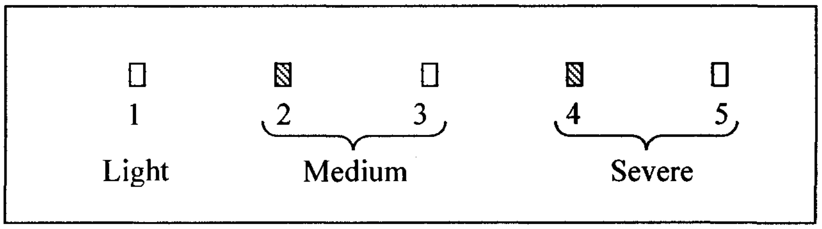


Figure II.5 – Sewer pipeline condition assessment scale.

II.6. Summary

This chapter reviewed many topics that give an overview on how to approach the stated problem. The topics included are: water main classification, risk of water main failure, the risk evaluation process and modeling approaches, and fuzzy logic.

From the literature review, it is clear that the works that have addressed the problem of water main failure risk have certain limitations, and therefore research that addresses the problem with a broad, concrete, and robust approach is still needed. Certain researchers have approached the problem in too shallow fashion, considering very few risk factors which sometimes were limited to only the deterioration factors (condition rating) and/or they did not consider the consequence of failure. Moreover, some of these researches were so complicated in their derivation and usage that different municipalities and authorities management teams are reluctant to use and depend on them. Other efforts were too specific to certain conditions (such as pipe material, diameter, function, etc…) and thus are not applicable to different water distribution networks. Some examples of these researches were done by: Christodoulou *et al.* (2003), Yan and Vairavamoorthy (2003), Kleiner *et al.* (2004), Sadiq *et al.* (2004), Kleiner *et al.* (2006), Rajani *et al.* (2006) and Al Barqawi (2006). The most relevant and solid research was done by Rogers (2006); however, there are some limitations inherent to his research such as: the model uses the weighted average method which does not address the uncertainty and the model is too sensitive to the weights of the factors. Moreover, Rogers' failure consequence model is not well-established and depends solely on the input of the model user. In addition, some of the risk factors are derived from a specific data set and seem to be more reflective of that data set instead of reflecting the state of the art.

There are other research works that address the risk of failure of pipelines other than water mains. Some of these researches cover areas such as: risk of failure of large diameter pre-stressed concrete cylinder pipelines (Marshall *et al.* 2005), small community water supply and treatment systems (Ezell *et al.* 2000), failure risk of sewers (Hahn, 2002), pressurized heavy-water reactor pipelines (Vinod *et al.* 2003), and pipelines transferring hydrogen sulphide (Santosh *et al.* 2006). There have been more efforts to address the risk associated with pipelines transferring petroleum materials (oil and gas) due to the catastrophic consequence of their failure. Some of these studies are by Bandyopadhyay *et al.* (1997), and Yuhua and Datao (2005).

Based on this exhaustive literature review, it is clear that there is a need to address the problem of water main failure risk using a technique -- such as fuzzy logic -- that considers the uncertainty usually associated with risk factors.

Chapter III: RESEARCH METHODOLOGY

The research methodology consists of eight stages as shown in Figure III.1. It starts with a comprehensive literature review of the risk of water main failure followed by data collection, which in itself consists of two parts. A hierarchical fuzzy expert system (HFES) is developed using model information data which is then underwent a verification process. The next part of the research methodology is to develop a risk of failure scale which will guide the network operators to best manage their networks. The HFES model is used to assess the case study data collected from two municipalities. The developed fuzzy expert application is based on MS ® Excel 2007 software.

III.1. Literature Review

All of the subjects related to the risk of water main failure are reviewed in order to have a better overview of the topic, how others approached the problem, and how to best solve the stated problem. In this research, the reviewed topics are water main classifications, risk definition, different types of failure, different sources of risk associated with water main failure, and consequences of failure. Different approaches in evaluation and modeling the risk of failure such as matrix models, probabilistic models, and indexing models are reviewed.

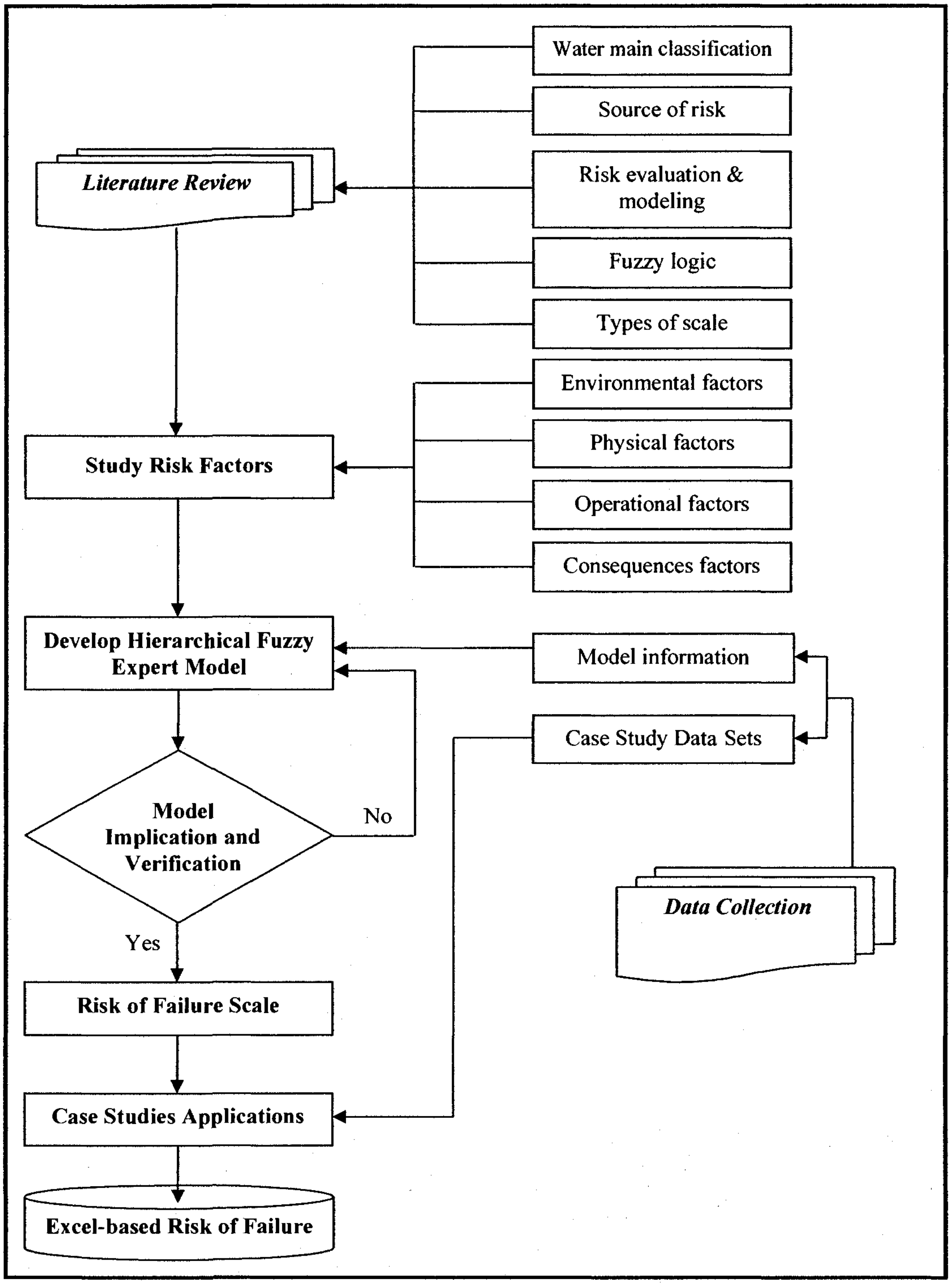


Figure III.1 – Research methodology.

Some topics related to fuzzy logic are reviewed and can be listed as: literature review of fuzzy logic and related research in the field of municipal networks, fuzzy logic process and operation, and fuzzy expert systems. The use of experts' opinions in developing fuzzy expert systems and hierarchy applications topics are reviewed as well. At last, different scales used in the field of failure risk and condition assessment are reviewed.

III.2. Data Collection

The data collected for this research are used to develop the model and to apply the HFES to case studies. The data needed to develop the model is collected from the literature and by a questionnaire which is sent to experts in the field of water main infrastructure management as shown in (Appendix B). Data for case studies are collected from two municipalities that operate water mains. Three case studies data sets are introduced. The first is the data set collected from the City of Moncton, New Brunswick. It contains data for only seven out of the sixteen factors considered in this research; mainly the quantitative factors. The second data set is constructed out of the first data set. Since the first data set has some unavailable information about many risk factors (qualitative in nature and do not require exact measurements or calculations), the second data set is completed by randomly assuming the unavailable factors. The third data set is collected from the City of London, Ontario. It contains data about nine failure risk factors.

III.3. Hierarchical Fuzzy Expert System for Water Main Failure Risk

A hierarchical fuzzy expert system (HFES) is developed to estimate the risk of water main failure. A Mamdani's fuzzy rules system is used as implication operation in the fuzzy model. The hierarchical fuzzy model consists of four main models which correspond to the four main factors: Environmental, Physical, Operational, and Post-failure models. The results of these four models as crisp values are used as crisp observations (input) to the fifth model which calculated the risk of failure of water mains. Each of these branches has its own sub-factors as shown in the hierarchy in Figure III.2.

In this research, in order to build the model knowledge base, the indirect knowledge acquisition method (by means of a questionnaire and the available literature) is used to gather the required information. However, a methodology is proposed that is different from the traditional rules-building methodology, as explained in detail in Section IV.1.3. Expert knowledge base.

The first step in processing the data in the model is the fuzzification process. This step uses the factors' membership functions to convert the real number into a fuzzy number of a value in the period [0, 1]. In order to do this, each membership function (linguistic variable) can be represented in an equation as shown in Equation III.1 (Del Campo, 2004), which converts the input number as a real number (x) into a fuzzy number.

$$\mu_A(x) = \begin{cases} 0 & x < a \\ \left(\frac{x-a}{b-a}\right) & a \le x \le b \\ 1 & b < x < c \\ \left(\frac{c-x}{c-d}\right) & c \le x \le d \\ 0 & x > d \end{cases}$$

Equation III.1

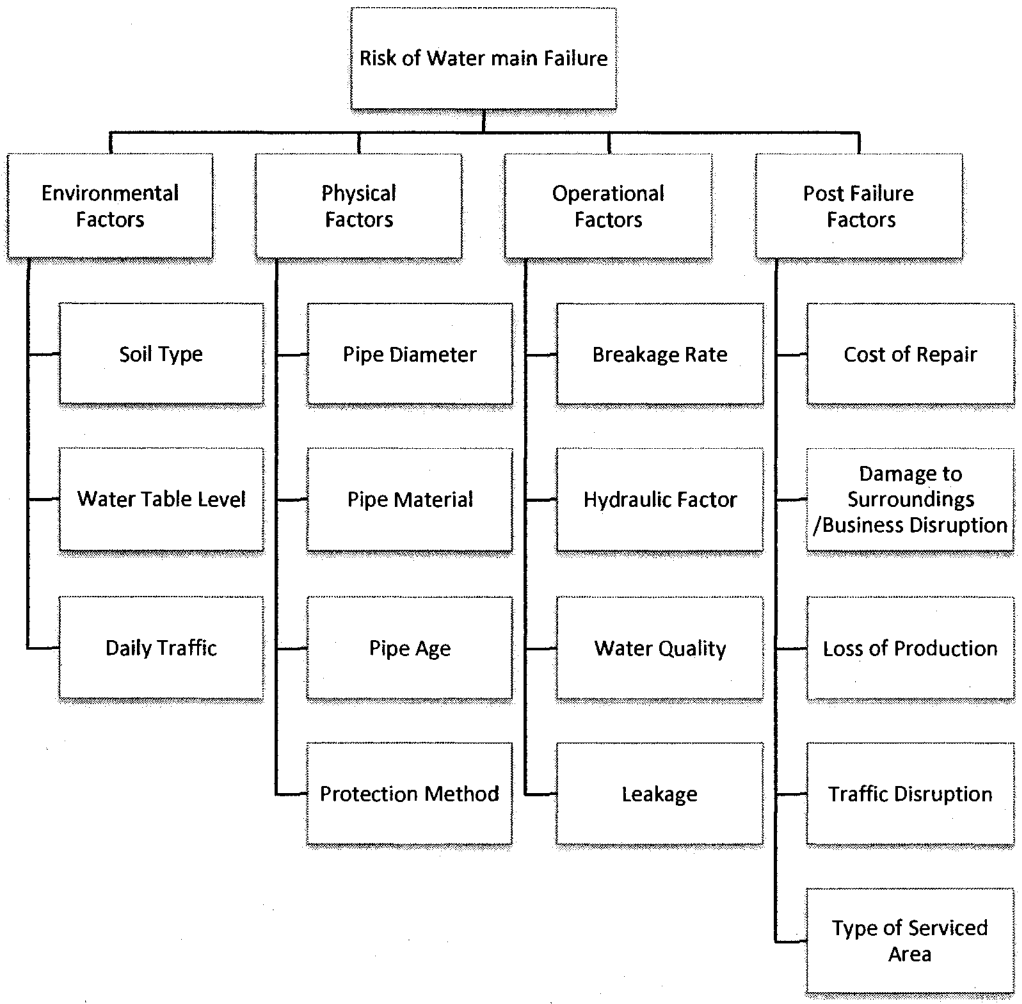


Figure III.2 – Hierarchical risk factors of water main failure.

After converting the real numbers of all the factors into fuzzy numbers, each of the knowledge base rules is evaluated. The general form of the knowledge base rules is as follows (Jin, 2003):

$$R^j: If\ x_1\ is\ A_1^j\ and\ x_2\ is\ A_2^j\ and\ x_3\ is\ A_3^j\ and \ldots x_n\ is\ A_n^j\ THEN\ y\ is\ B^j$$

Where R^j is the j-th rule, A_i^j $(\mathrm{j} = 1, 2, \ldots \mathrm{N}, \mathrm{i} = 1,2, \ldots \mathrm{n}), B^j$ are the fuzzy subsets of the inputs and outputs respectively.

This rule can be written mathematically as Equation III.2 (Jin, 2003):

$$\mu_{R^j}(x_1, x_2, x_3, \ldots, x_n, y) = \mu_{A_1^j} \wedge \mu_{A_2^j} \wedge \mu_{A_3^j} \ldots \wedge \mu_{A_n^j} \wedge \mu_B \qquad \text{Equation III.2}$$

Where $\wedge$ denotes the minimum operator.

The consequent linguistic variable B^j is to be chosen from a standard list of seven linguistic variables (Extremely low, Very low, Moderately Low, Medium, Moderately High, Very High, and Extremely High).

After evaluating each rule in the knowledge base, the membership value of each consequent function (output linguistic variable) is aggregated using a maximum operation as shown in Equation III.3 (Jin, 2003). In other words, the maximum membership value of the B^j consequent variable is used to truncate that consequent membership function for later use in the defuzzification of the fuzzy output.

$$\mu_R(x_1, x_2, x_3, \ldots, x_n, y) = \bigvee_{j=1}^{N} [\mu_{R^j}(x_1, x_2, x_3, \ldots, x_n, y)] \qquad \text{Equation III.3}$$

Where $\vee$ denotes the maximum operation, R represents each of the consequent membership functions as standardized to the list of (Extremely low, Very low, Moderately Low, Medium, Moderately High, Very High, and Extremely High).

The next step is to defuzzify the consequent membership functions into one crisp number. The defuzzification method used is the Center of Sum. It calculates the center of gravity of each function individually and then average-weights them by their areas. It has the advantage of being simple to program, requiring minimal computer resources, and it gives reasonable results. The crisp output of risk can be found using Equation III.4.

$$Crisp\ Risk\ Output =$$

$$= \sum_{n=extremely\ low}^{extremely\ high} (Truncated\ Area_n \times Centeriod_n) \Bigg/ \sum_{n=extremely\ low}^{extremely\ high} Truncated\ Area_n$$

Equation III.4

The procedure described above is generic and is applicable to every branch (model) of the hierarchy (environmental, physical, operational, post-failure factors branches (models), and risk of failure model which combines the four models) to generate crisp output. Figure III.3 shows the data processing flow in the physical risk model. The data processing flow of other factors' models are identical to the physical factor model (only the sub-factors, their associated membership functions (linguistic variables), and knowledge rules are different). These crisp numbers are used as input to the risk of failure model and processed as explained above. Figure III.4 shows the data processing in the Risk of Failure model. It starts with the risk input from the four main models (branches of the hierarchy) and then the model converts them into fuzzy numbers which then are evaluated using the fuzzy knowledge base rules. The triggered fuzzy rules are aggregated and defuzzified into a crisp number which represents the Risk of Failure of the water main.

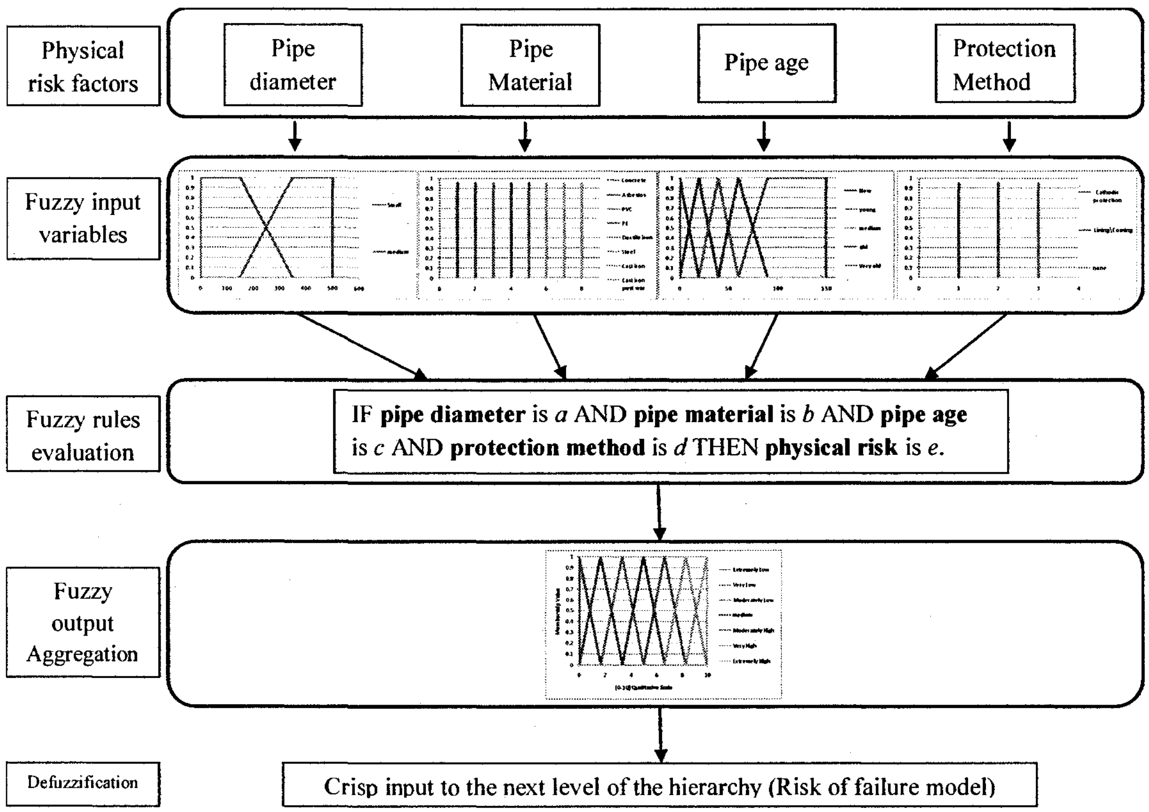


Figure III.3 – Physical risk model structure.

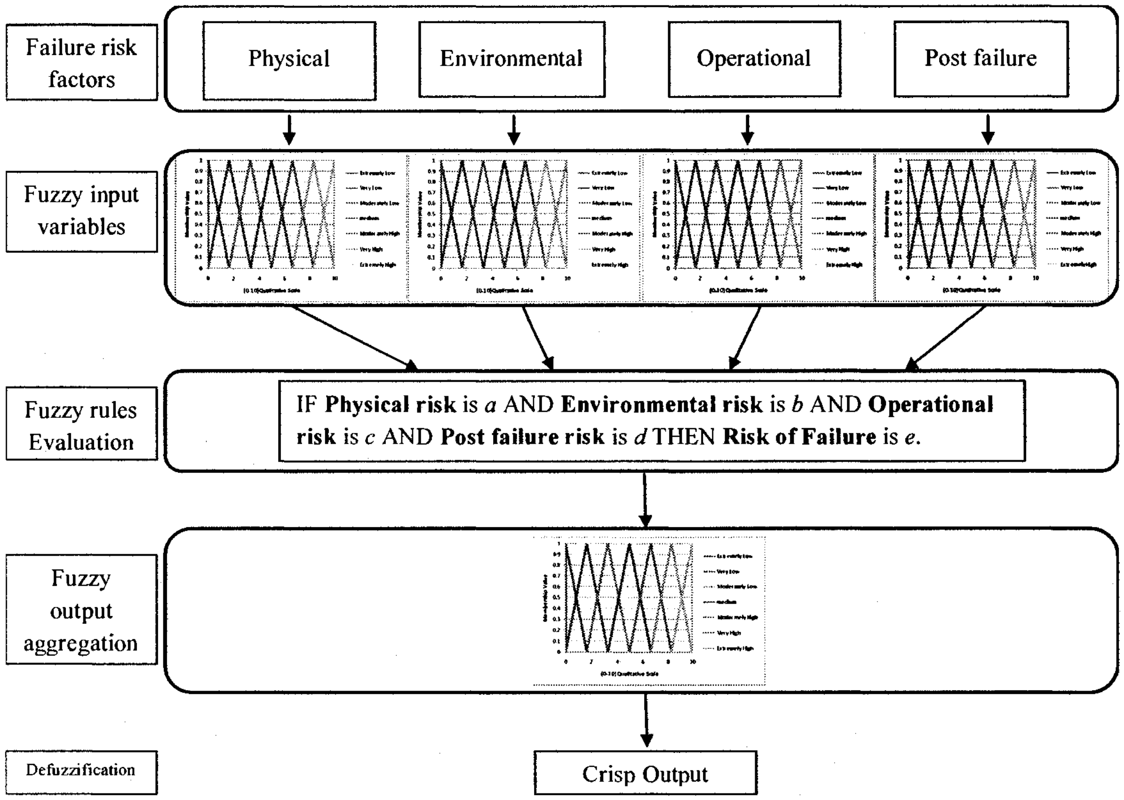


Figure III.4 – Risk of failure model structure.

III.4. Model Analysis and Verification

A sensitivity analysis is carried out in order to test the sensitivity of the model and its stability, and to insure that it is performing as expected. The sensitivity analysis is done by assuming many scenarios; testing most weighted factors, testing, the effect of weights on the model, testing the model performance from the least risky to highest risky status of the factors, etc.... Model verification is carried out by using a validated model to compare the results of the proposed model with the validated model. The model used for verification is the AHP deterioration model developed by Al Barqawi (2006) which can only evaluate the deterioration part of the proposed model.

III.5. Risk of Failure Scale

In light of the reviewed literature, a risk of failure scale is proposed to help decision makers in water resources management companies/municipalities make an informed decision. The scale ranges numerically from 0 to 10, where 10 indicates the most risky condition of the pipeline and 0 indicates the least risky condition. Linguistically, the scale is divided into five groups or regions which describe the risk of pipeline failure and the required corrective actions to be taken if needed, as explained later in Section "V.8. Proposed Risk of Failure Scale". The number of proposed groups and their ranges and associated corrective actions are likely to be changed to best suit municipality's strategies and risk tolerance.

III.6. Case Studies Application

Three case studies data sets are collected and used to show the application of the developed model. The case studies results are studied and analyzed and some recommendations are proposed depending on the results of the model. In case study one, the results show that the condition of the network is fair (66% of the network) with some parts of the network requiring mitigation action in the short-term plan. However, the results of case study two show that the condition of the network is risky (50%) to fair (47%) with some parts of the network requiring immediate mitigation action. In case study three, the condition of the network is fair (50%) to risky (45%) with some parts of the network require immediate mitigation action.

III.7. Excel-Based Application Development

An application is built to implement the developed model. The application is based on MS© Excel 2007. The application consists of many spreadsheet files. It is controlled through an Excel file called *"Navigation"*. This file contains all the step by step instructions that will guide the user to the easy use of the application. Other spreadsheet files represent the four branches of the hierarchy, the risk of failure model, etc... A full review of the application is explained in "Chapter VI: Excel-Based Application Development".

III.8. Summary

This chapter presented the research methodology followed in this thesis. This methodology includes the literature review of the risk of water main failure, data collection (which consists of model information data and case studies data) HFES model development, the risk of failure scale, case studies evaluation, and the development of an Excel-based application.

Chapter IV: DATA COLLECTION

The data collection consists of two stages which are required to develop and run the fuzzy expert system. In stage one, the information needed to build the model is collected. In stage two, real network characteristics are gathered and analyzed by the developed model. The process of data collection and its parts is shown in Figure IV.1.

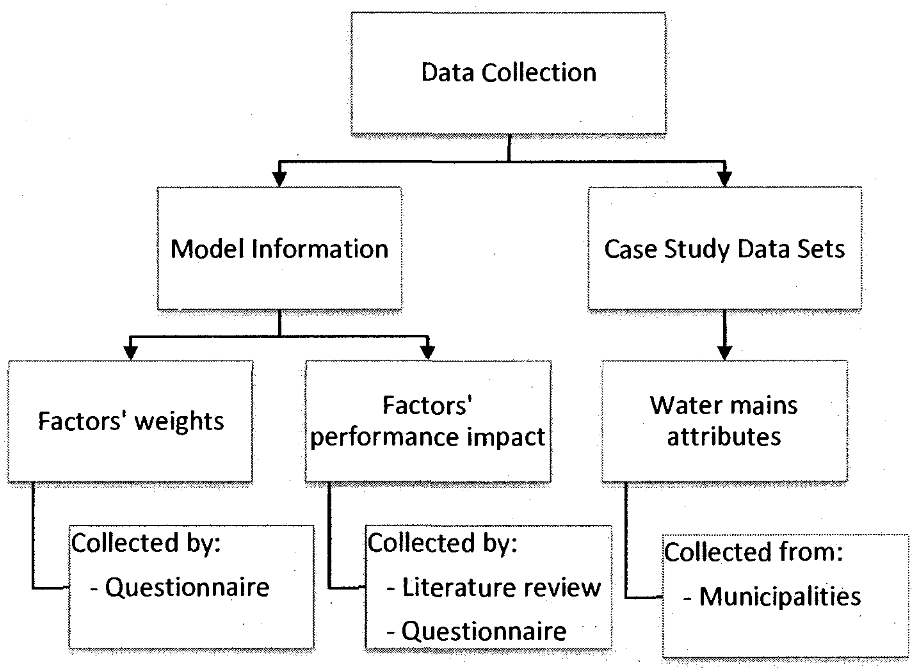


Figure IV.1 – Water main data collection process.

IV.1. Model Information

The information needed to develop the model consists of two parts: factors' weights and factors' performance impact. The majority of information is gathered from the literature. The information that can not be collected from the literature is collected via a questionnaire. A sample questionnaire is shown in Appendix B. The two parts of the model information are as follows:

IV.1.1. Weights of Factors

In this section, the relative weight of each factor at each level of the hierarchy is collected. This could be the answer to the question of "What is the strength of the factor in contributing to the failure event?" This information is collected by a questionnaire. The questionnaire was sent to fifty-eight experts (designers, operators, consultants, researchers), and feedback was received from only twenty, giving an average response of 34% as shown in Figure IV.2.

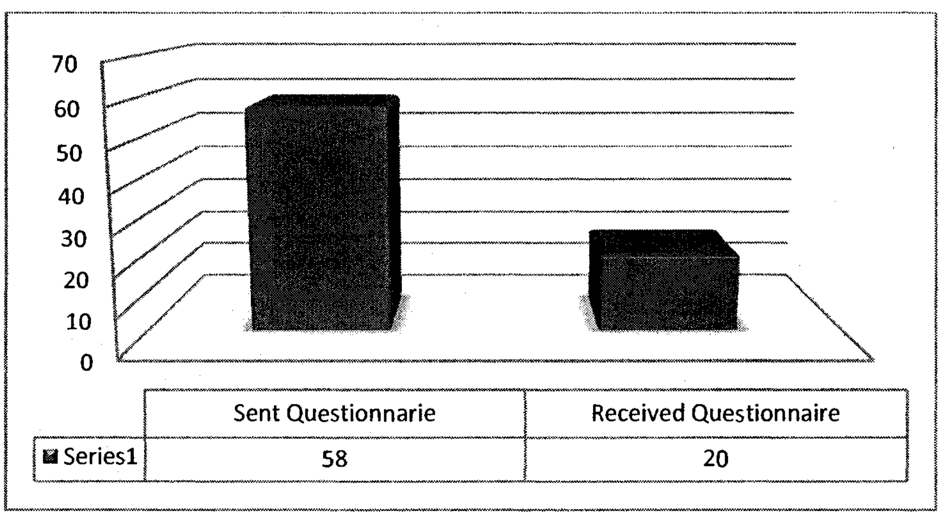


Figure IV.2 – Questionnaire statistics.

Geographically, the received responses can be summarized according to their locations as follows: Quebec 4 responses, Alberta 6, Ontario 6, British Colombia 2, New Brunswick 1, and Saskatchewan 1 response. Figure IV.3 shows the percentages of received responses from each participating province.

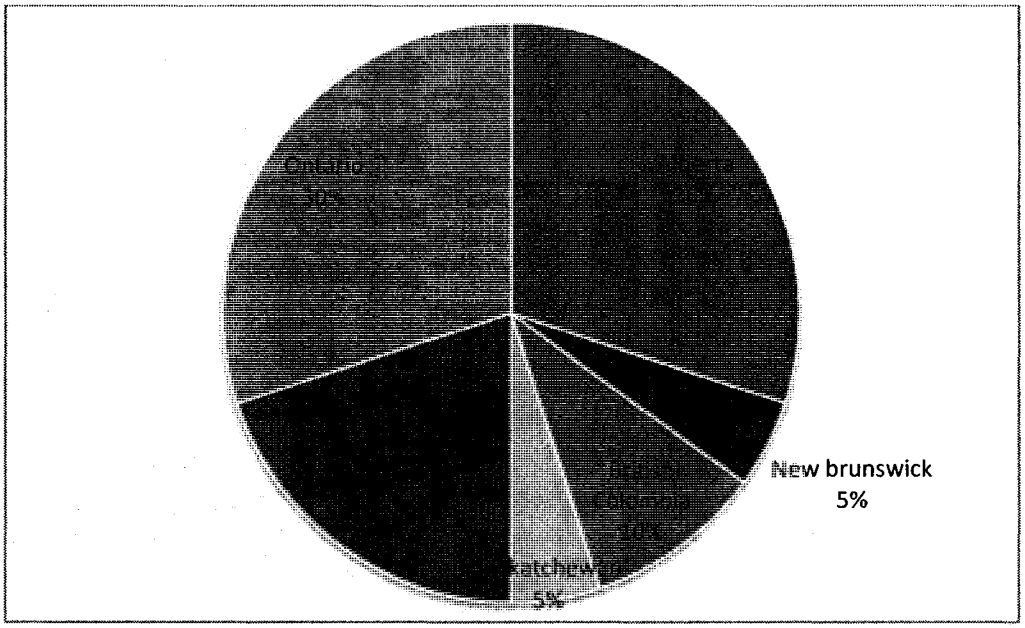


Figure IV.3 – Percentages of received questionnaires from each Canadian province.

Table IV.1 shows the collected factors' weights. The local weight of each risk sub-factor shown in that column is comparable only to its group of risk factor (hierarchy branch). The weights of the sub-factors can be compared to those in other groups by multiplying the sub-factors' local weights by the main factor weight of where they belong, as shown in the Global Weights column. It is worth mentioning that the developed model can use either the local or global weights since each main risk factor is treated and modeled separately. It is obvious that pipe age has the highest weight and thus it has the most effect on the model. Figure IV.4 shows a graph of the normalized global weights of the failure risk factors. From this graph, it can be deduced that pipe age has the highest effect among the other factors, followed by pipe material and breakage rate. By reviewing the values of standard deviation, the highest value is 20, the lowest is 4, and the average is 10, which are acceptable values.

Table IV.1 – Risk of failure factors' weights.

Main Risk Factor	Factor Weight	Risk sub-factor	Sub-factors local weights	Standard deviation	Global weights	Normalized Global weights
Environmental Factors	19	Soil Type	42	9.2	798	46
		Average Daily Traffic	13	5.0	247	14
		Water Table Level	20	6.8	380	22
Physical Factors	43	Pipe Material	30	12.1	1290	75
		Pipe Diameter	19	9.1	817	48
		Pipe Age	40	16.1	1720	100
		Protection Method	15	8.2	645	38
Operational Factors	28	Breakage Rate	35	10.8	980	57
		Hydraulic Factor	13	6.6	364	21
		Water Quality	17	4.4	504	29
		Leakage	20	5.9	560	32
Post-Failure Factors	20	Cost of Repair	20	7.1	400	23
		Damage to surroundings	21	5.7	420	24
		Loss of Production	18	6.0	360	21
		Traffic Disruption	17	5.3	340	20
		Type of Serviced Area	14	5.6	280	16

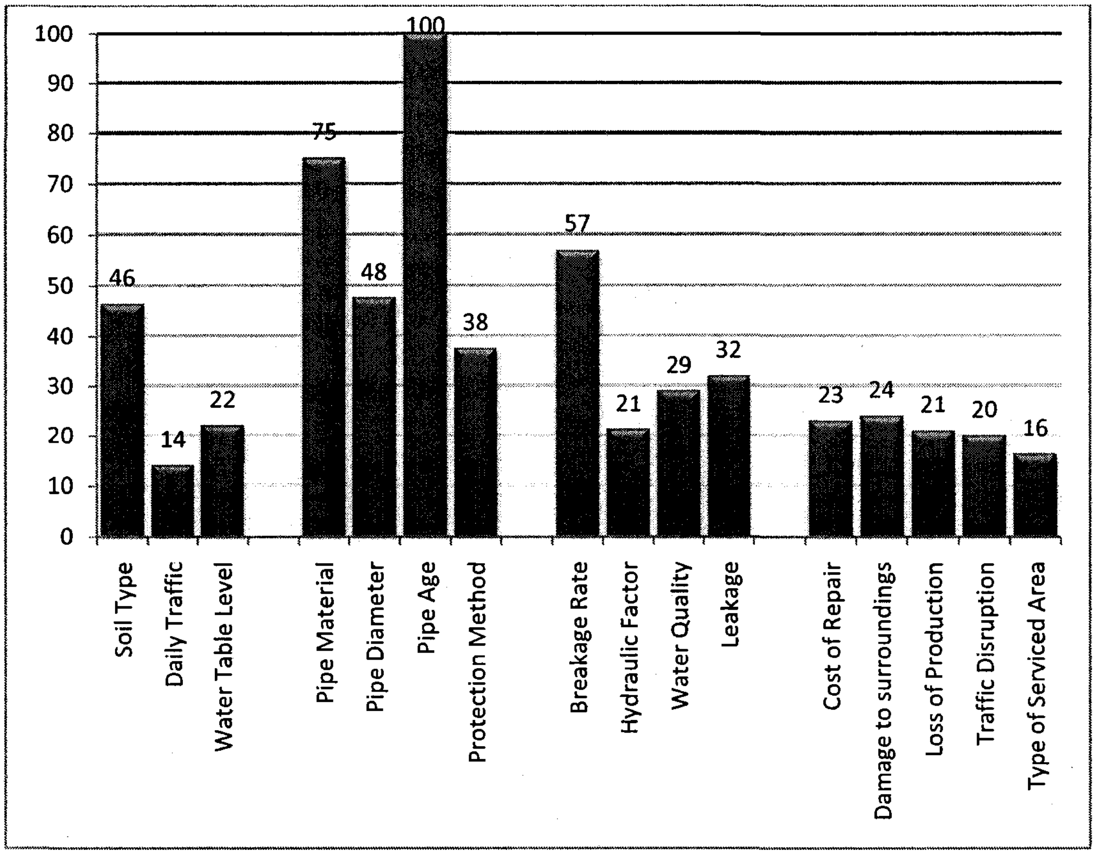


Figure IV.4 – Risk factors normalized global weights.

IV.1.2. Factors Performance Impact

In this section, performance assessment of the different factors considered in this project (as shown in Figure III.2) is collected mainly from the literature. The information about the factors where the performance is not clear or is missing is collected via a questionnaire. The performance or behavior of 13 out of 16 factors are collected from the literature review: soil type, daily traffic, pipe diameter, pipe material, pipe age, protection method, breakage rate, hydraulic factor, water quality, leakage rate, cost of repair, traffic disruption, and type of serviced area. The remaining required information is collected in the form of a questionnaire sent to experts in order to gather their opinion and experience about the behavior and characteristic of the water main network. The performance impact of only three factors are collected via the questionnaire: water table level, damage to surroundings/business disruption, and loss of production. This information is gathered in the form of (if-then) or (cause-effect) where the answer is standardized to the following list: "*Extremely High, Very High, Moderately High, Medium, Moderately Low, Very Low, Extremely Low*"

The factors' performance impact is shown below:

i. Environmental Factors

The environmental factors are type of soil, water table level, and average daily traffic. The performance of the type of soil and average daily traffic factors are collected from the literature and are shown in Table IV.2 and Table IV.4 respectively. Water table level factor performance is collected by a questionnaire and shown in Table IV.3.

Table IV.2 – Type of soil factor performance.

Factor performance		**Impact on risk**		
if soil is	Very lightly deteriorative	then	the risk of failure is	Extremely Low
if soil is	Lightly deteriorative	then	the risk of failure is	Very Low
if soil is	Moderately deteriorative	then	the risk of failure is	Medium
if soil is	Highly deteriorative	then	the risk of failure is	Very High
if soil is	Very highly deteriorative	then	the risk of failure is	Extremely High

Table IV.3 – Water table level factor performance.

Factor performance		**Impact on risk**		
if WT is	rarely present	then	the risk of failure is	Extremely Low
if WT is	seasonally present	then	the risk of failure is	Extremely High
if WT is	always present	then	the risk of failure is	Moderately High

Table IV.4 – Daily traffic factor performance.

Factor performance		Impact on risk		
if ADT is	Very light	then	the risk of failure is	Extremely Low
if ADT is	Light	then	the risk of failure is	Very Low
if ADT is	Moderate	then	the risk of failure is	Medium
if ADT is	Heavy	then	the risk of failure is	Very High
if ADT is	Very heavy	then	the risk of failure is	Extremely High

ii. Physical Factors

This category of factors includes pipe diameter, material, age, and pipe protection method. The performance of the different factors are shown in Table IV.5 through Table IV.8. These factors' performance is collected from the literature.

Table IV.5 – Pipe diameter factor performance.

Factor performance		Impact on risk		
if dia. is	small (≤ 250 mm)	then	the risk of failure is	Very High
if dia. is	medium (between 250 to 500)	then	the risk of failure is	Very Low

Table IV.6 – Pipe material factor performance.

Factor performance		Impact on risk		
if ADT is	Concrete	then	the risk of failure is	Medium
if ADT is	Asbestos	then	the risk of failure is	Moderately High
if ADT is	PVC	then	the risk of failure is	Very Low
if ADT is	PE	then	the risk of failure is	Extremely Low
if ADT is	Ductile iron	then	the risk of failure is	Very Low
if ADT is	Steel	then	the risk of failure is	Very Low
if ADT is	Cast iron	then	the risk of failure is	Very High
if ADT is	Cast iron post war	then	the risk of failure is	Extremely High

Table IV.7 – Pipe age factor performance.

Factor performance		Impact on risk		
if age is	new (0 yrs < Age ≤ 10 yrs)	then	the risk of failure is	Extremely Low
if age is	young (10 yrs < Age ≤ 30 yrs)	then	the risk of failure is	Very Low
if age is	medium (30 < Age ≤ 50)	then	the risk of failure is	Medium
if age is	old (50 yrs < Age ≤ 70 yrs)	then	the risk of failure is	Very High
if age is	very old (> 70 yrs)	then	the risk of failure is	Extremely High

Table IV.8 – Protection methods factor performance.

Factor performance		Impact on risk		
if pipe has	Cathodic protection	then	the risk of failure is	Very Low
if pipe has	Lining/Coating	then	the risk of failure is	Extremely Low
if pipe has	none	then	the risk of failure is	Medium

iii. Operational Factors

Operational factors include breakage rate, hydraulic factor (hazen-william coefficient), water quality, and leakage rate. The factors' performance is collected from the literature and are shown in Table IV.9 through Table IV.12.

Table IV.9 – Breakage rate factor performance.

Factor performance		Impact on risk		
if breakage rate is	low (< 0.5 brk/km/yr)	then	the risk of failure is	Extremely Low
if breakage rate is	average (bet 0.5 and 3)	then	the risk of failure is	Medium
if breakage rate is	high (> 3 brk/km/yr)	then	the risk of failure is	Extremely High

Table IV.10 – Hydraulic factor performance.

Factor performance		Impact on risk		
if pipe is	very rough (C-factor $\leq$ 40)	then	the risk of failure is	Extremely High
if pipe is	rough (60 $\geq$ C-factor > 40)	then	the risk of failure is	Very High
if pipe is	medium (80 $\geq$ C-factor > 60)	then	the risk of failure is	Medium
if pipe is	smooth (100 $\geq$ C-factor > 80)	then	the risk of failure is	Very Low
if pipe is	very smooth (C > 100)	then	the risk of failure is	Extremely Low

Table IV.11 – Water quality factor performance.

Factor performance		Impact on risk		
if WQ is	very good	then	the risk of failure is	Extremely Low
if WQ is	good	then	the risk of failure is	Very Low
if WQ is	acceptable	then	the risk of failure is	Medium
if WQ is	bad	then	the risk of failure is	Very High
if WQ is	very bad	then	the risk of failure is	Extremely High

Table IV.12 – Leakage rate factor performance.

Factor performance		Impact on risk		
if leakage is	very low	then	the risk of failure is	Extremely Low
if leakage is	low	then	the risk of failure is	Very Low
if leakage is	medium	then	the risk of failure is	Medium
if leakage is	high	then	the risk of failure is	Very High
if leakage is	very high	then	the risk of failure is	Extremely High

iv. Post failure Factors

This category of factors represents the consequence of failure. It includes five factors. The performance of cost of repair, traffic disruption, and type of serviced area and are shown in Table IV.13, Table IV.16, and Table IV.17 are collected from literature. Damage to surrounding and loss of productions factors' performance is collected by a questionnaire as shown in Table IV.14 and Table IV.15.

Table IV.13 – Cost of repair factor performance.

Factor performance		**Impact on risk**		
if repair cost is	very low	then	the risk of failure is	Extremely Low
if repair cost is	low	then	the risk of failure is	Very Low
if repair cost is	medium	then	the risk of failure is	Medium
if repair cost is	high	then	the risk of failure is	Very High
if repair cost is	very high	then	the risk of failure is	Extremely High

Table IV.14 – Damage to surroundings factor performance.

Factor performance		**Impact on risk**		
if failure in	Industrial area	then	the risk of failure is	Extremely High
if failure in	Commercial area	then	the risk of failure is	Moderately High
if failure in	Residential are	then	the risk of failure is	Extremely Low

Table IV.15 – Loss of production factor performance.

Factor performance		**Impact on risk**		
if pipe is	small & redundant	then	the risk of failure is	Very Low
if pipe is	medium & redundant	then	the risk of failure is	Moderately Low
if pipe is	small & not redundant	then	the risk of failure is	Medium
if pipe is	medium & not redundant	then	the risk of failure is	Very High

Table IV.16 – Traffic disruption factor performance.

Factor performance		**Impact on risk**		
if failure is	Very lightly disruptive	then	the cost of failure is	Extremely Low
if failure is	Lightly disruptive	then	the cost of failure is	Very Low
if failure is	Moderately disruptive	then	the cost of failure is	Medium
if failure is	Highly disruptive	then	the cost of failure is	Very High
if failure is	Very highly disruptive	then	the cost of failure is	Extremely High

Table IV.17 – Type of serviced area factor performance.

Factor performance		Impact on risk		
if failure	stops service to Industrial area	then	the cost of failure is	Very High
if failure	stops service to Commercial area	then	the cost of failure is	Medium
if failure	stops service to Residential are	then	the cost of failure is	Moderately Low

v. Risk of failure Factors

This contains the four main risk of failure factors as shown in Figure III.2. It includes environmental, physical, operational and post failure factors. There performance is shown in Table IV.18 to Table IV.21.

Table IV.18 – Environmental factor performance.

Factor performance		Impact on risk		
if environmental risk is	Extremely Low	then	the risk of failure is	Extremely Low
if environmental risk is	Very Low	then	the risk of failure is	Very Low
if environmental risk is	Moderately Low	then	the risk of failure is	Moderately Low
if environmental risk is	Medium	then	the risk of failure is	Medium
if environmental risk is	Moderately High	then	the risk of failure is	Moderately High
if environmental risk is	Very High	then	the risk of failure is	Very High
if environmental risk is	Extremely High	then	the risk of failure is	Extremely High

Table IV.19 – Physical factor performance.

Factor performance		Impact on risk		
if physical risk is	Extremely Low	then	the risk of failure is	Extremely Low
if physical risk is	Very Low	then	the risk of failure is	Very Low
if physical risk is	Moderately Low	then	the risk of failure is	Moderately Low
if physical risk is	Medium	then	the risk of failure is	Medium
if physical risk is	Moderately High	then	the risk of failure is	Moderately High
if physical risk is	Very High	then	the risk of failure is	Very High
if physical risk is	Extremely High	then	the risk of failure is	Extremely High

Table IV.20 – Operational factor performance.

Factor performance		Impact on risk		
if operational risk is	Extremely Low	then	the risk of failure is	Extremely Low
if operational risk is	Very Low	then	the risk of failure is	Very Low
if operational risk is	Moderately Low	then	the risk of failure is	Moderately Low
if operational risk is	Medium	then	the risk of failure is	Medium
if operational risk is	Moderately High	then	the risk of failure is	Moderately High
if operational risk is	Very High	then	the risk of failure is	Very High
if operational risk is	Extremely High	then	the risk of failure is	Extremely High

Table IV.21 – Post failure factor performance.

Factor performance		Impact on risk		
if post failure risk is	Extremely Low	then	the risk of failure is	Extremely Low
if post failure risk is	Very Low	then	the risk of failure is	Very Low
if post failure risk is	Moderately Low	then	the risk of failure is	Moderately Low
if post failure risk is	Medium	then	the risk of failure is	Medium
if post failure risk is	Moderately High	then	the risk of failure is	Moderately High
if post failure risk is	Very High	then	the risk of failure is	Very High
if post failure risk is	Extremely High	then	the risk of failure is	Extremely High

IV.1.3. Expert knowledge base

The next step is to combine the collected factors' weights and factors performance in a form that can be used in an expert system and represents the knowledge of the experts. To do so, a rules building methodology will be used which uses weighted average method to combine the factors' performance depending on the weights. The reason behind using this methodology over the traditional methodology of directly soliciting the rules from experts is that the traditional methodology requires the expert to evaluate the performance of a huge number of rules; this process is exhausting and time consuming and the human expert will most probably not carry out this task. To overcome this major drawback to using a traditional methodology in rules extraction, a new methodology is proposed. The proposed new methodology will ask the expert to evaluate the performance of each factor independent of other factors as collected in IV.1.2. Factors Performance Impact (e.g. IF pipe age is old THEN the risk of failure is High). Moreover, the expert is also asked to give a weight to each factor which reflects the contribution of each factor to the risk of water main failure as collected in IV.1.1. Weights of Factors. The general outlines of the proposed method are explained in Figure IV.5. The proposed method shares an idea with the method originally developed by Shaheen (2005). The shared idea is the use of the impact of each factor individually together with its weight (importance) to generate the equivalent impact of the rules. The main difference between the proposed method and the one developed by Shaheen (2005) is the use of the average weighted method to choose the equivalent impact of the rule instead of using a normalization process. Moreover, both the factors' performance impact scale and equivalent ranges of impact scale are derived from the performance of the model (fuzzy model to be built in the next chapter) instead of using two different scales and using a normalization process. Using the weights of factors, the combined performance impact of the different factors is calculated using the weighted average method. The followings shows the steps followed in finding the equivalent impact of different combinations of each four branches of the hierarchy and the main level of the hierarchy.

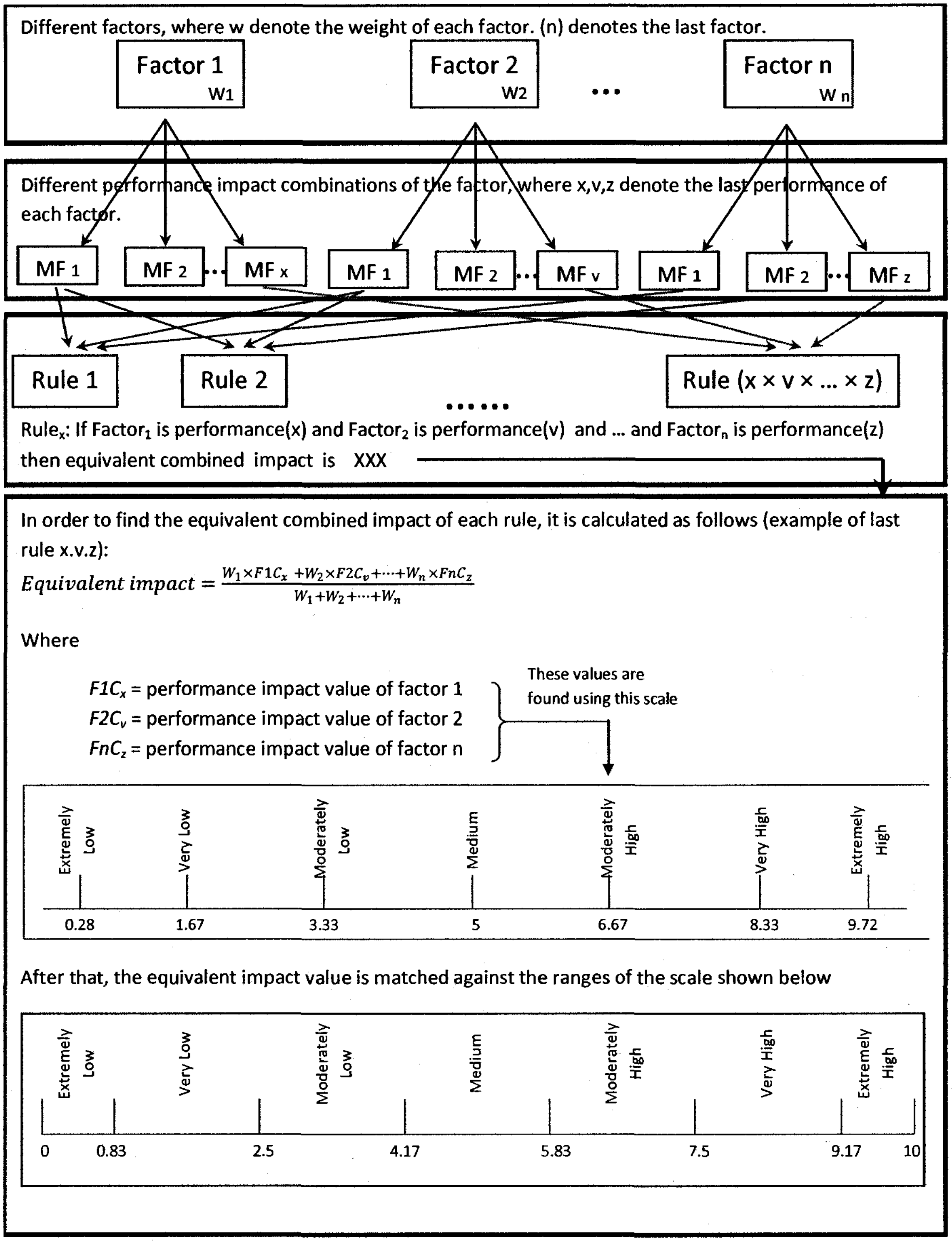


Figure IV.5 – Proposed methodology for fuzzy rules extraction.

i. Environmental Factors

1) Number of performance combination

The number of performance combination can be calculated for the environmental factors by multiplying the number of performance stages of the factors. In this research, the number of different performance combinations needed to cover all the combination possibilities can be found as:

$$\begin{aligned} &number\ of\ environmental\ rules \\ &= 5\ (soil\ type) \\ &\times 5\ (average\ daily\ traffic) \\ &\times 3\ (water\ table\ level) = 75\ rules \end{aligned}$$ Equation IV.1

> Example: IF Type of soil is *moderately deteriorative* AND Average daily traffic is *extremely high* AND Water table level is *rarely present*.

2) Weights of factors

This is discussed in section IV.1.1. Weights of Factors. The normalized global weights or the local weights can be used in this process as shown in Table IV.22.

Table IV.22 – Environmental factors weights.

Main Risk Factor	Risk sub-factor	Sub-factors local weights	Global weights	Normalized Global weights
Environmental Factors	Soil Type	42	798	46
	Average Daily Traffic	13	247	14
	Water Table Level	20	380	22

3) Factors performance impacts

The factors performance impacts at different stages (collected in section IV.1.2. Factors Performance Impact) is measured by a scale as shown in Figure IV.6.

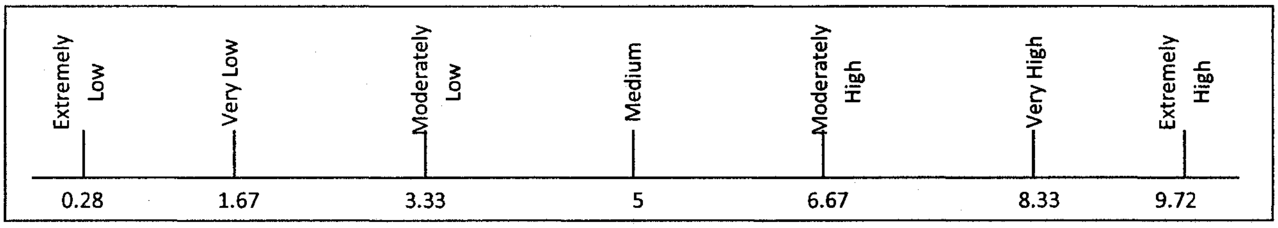


Figure IV.6 – Factor performance impact scale

To illustrate how this scale works, an example is given below:

- If soil type is moderately deteriorative then risk of failure is medium (5 impact value)
- If average daily traffic is very heavy then risk of failure is extremely high (9.72 impact value)
- If water table level is rarely present then the risk of failure is extremely low (0.28 impact value)

4) Factors performance combined impact

In order to find the combined impact of the different factors' performance, weighted average method is used as shown in Equation IV.2.

$$Equivalent\ impact = \frac{\sum(impact \times Weight)\ \text{of each factor}}{\sum factors\ Weights} \qquad \text{Equation IV.2}$$

This equivalent impact is matched against another scale that is divided into ranges of different equivalent impact as shown in Figure IV.7.

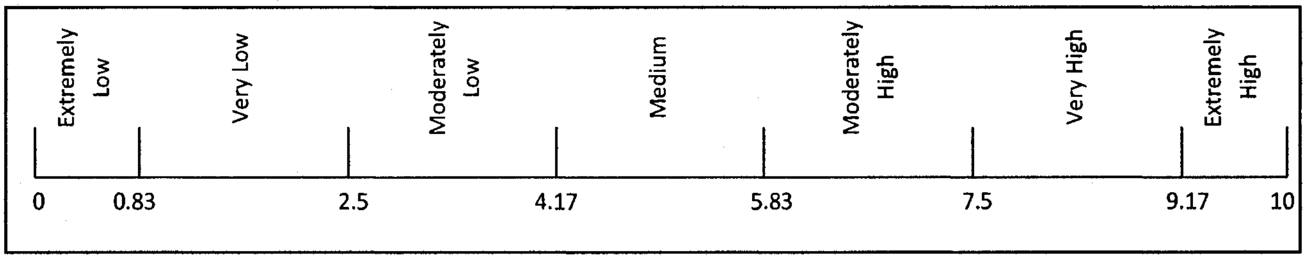


Figure IV.7 – Equivalent ranges of combined impact.

Both impact values scale and impact ranges scale are derived from the shape of membership functions of the consequent of the fuzzy system to be developed in next chapter. Continuing the example shown above, the combined impact is found and the rule is established. So,

$$Equivalent\ impact = \frac{(5 \times 42) + (9.73 \times 13) + (0.28 \times 20)}{(42 + 13 + 20)} = 4.56 \qquad \text{Equation IV.3}$$

Matching the value of *equivalent impact* against the scale in Figure IV.7 yields that the linguistic equivalent impact is medium. So the rule will be completed as:

IF Type of soil is *moderately deteriorative* AND Average daily traffic is *extremely high* AND Water table level is *rarely present* THEN risk of failure is *Medium*.

All other factors' performance impacts combinations (75 combinations) can be found using this approach. The developed form of rule can be used in an expert system such as fuzzy expert system. A sample of these rules are shown in Table IV.23.

Table IV.23 – Sample environmental factors performance combined impact.

Rule No.	Factors' performance			Combined Impact
	Soil	Traffic	Water Table	linguistic
1	Very highly deteriorative	Very heavy	always present	Very High
2	Very highly deteriorative	Very heavy	seasonally present	Extremely High
3	Very highly deteriorative	Very heavy	rarely present	Moderately High
4	Very highly deteriorative	Heavy	always present	Very High
5	Very highly deteriorative	Heavy	seasonally present	Extremely High
6	Very highly deteriorative	Heavy	rarely present	Moderately High
7	Very highly deteriorative	Moderate	always present	Very High
:	:	:	:	:
:	:	:	:	:
73	Very lightly deteriorative	Very light	always present	Very Low
74	Very lightly deteriorative	Very light	seasonally present	Moderately Low
75	Very lightly deteriorative	Very light	rarely present	Extremely Low

ii. Physical Factors

1) Number of performance combination

The number of performance combination can be calculated for the physical factors by multiplying the number of performance stages of the factors. In this research, the number of different performance combinations needed to cover all the combination possibilities can be found as:

$$number\ of\ physical\ rules$$

$$= 2\ (diameter) \times 8\ (pipe\ material) \times 5\ (pipe\ age) \times 3\ (protection\ method) = 240\ rules \quad \text{Equation IV.4}$$

Example: IF pipe diameter is *small* AND pipe material is *ductile iron* AND pipe age is *old* AND pipe protection method is *Cathodic protection.*

2) Weights of factors

This is discussed in section IV.1.1. Weights of Factors. The normalized global weights or the local weights can be used in this process as shown in Table IV.24.

Table IV.24 – Operational factors weights.

Main Risk Factor	Risk sub-factor	Sub-factors local weights	Global weights	Normalized Global weights
Physical Factors	Pipe Material	30	1290	75
	Pipe Diameter	19	817	48
	Pipe Age	40	1720	100
	Protection Method	15	645	38

3) Factors performance impacts

The factors performance impacts at different stages (collected in section IV.1.2. Factors Performance Impact) is measured by a scale as shown in Figure IV.6.

To illustrate how this scale works, an example is given below:

• If pipe diameter is small then risk of failure is very high (8.33 impact value)

- If pipe material is ductile iron then risk of failure is very low (1.67 impact value)
- If pipe age is old then the risk of failure is very high (8.33 impact value)
- If pipe protection method is cathodic protection then the risk of failure is very low (1.67 impact value)

4) Factors performance combined impact

In order to find the combined impact of the different factors' performance, weighted average method is used as shown in Equation IV.2 which is matched against another scale that is divided into ranges of different equivalent impact as shown in Figure IV.7.

Continuing the example shown above, the combined impact is found and the rule is established. So,

$$\begin{aligned} &Equivalent\ impact \\ &= \frac{(8.33 \times 19) + (1.67 \times 30) + (8.33 \times 40) + (1.67 \times 15)}{(19 + 30 + 40 + 15)} \quad \text{Equation IV.5} \\ &= 5.4 \end{aligned}$$

Matching the value of *equivalent impact* against the scale in Figure IV.7 yields that the linguistic equivalent impact is medium. So the rule will be completed as:

IF pipe diameter is *small* AND pipe material is *ductile iron* AND pipe age is *old* AND pipe protection method is *Cathodic protection* THEN risk of failure is *Medium*.

All other factors' performance impacts combinations (240 combinations) can be found using this approach. The developed form of rule can be used in an expert system such as fuzzy expert system. A sample of these rules are shown in Table IV.25.

Table IV.25 – Sample physical factors performance combined impact.

Rule No.	Factors' performance				Combined impact
	pipe type	pipe dia	pipe age	protection	linguistic
1	Concrete	Small	Very old	Cathodic protection	Moderately High
2	Concrete	Small	Very old	Lining\Coating	Moderately High
3	Concrete	Small	Very old	none	Moderately High
4	Concrete	Small	old	Cathodic protection	Moderately High
5	Concrete	Small	old	Lining\Coating	Moderately High
6	Concrete	Small	old	none	Moderately High
7	Concrete	Small	medium	Cathodic protection	Medium
8	Concrete	Small	medium	Lining\Coating	Medium
9	Concrete	Small	medium	none	Medium
:	:	:	:	:	:
:	:	:	:	:	:
238	Cast iron post war	medium	New	Cathodic protection	Moderately Low
239	Cast iron post war	medium	New	Lining\Coating	Moderately Low
240	Cast iron post war	medium	New	none	Moderately Low

iii. Operational Factors

1) Number of performance combination

The number of performance combination can be calculated for the operational factors by multiplying the number of performance stages of the factors. In this research, the number of different performance combinations needed to cover all the combination possibilities can be found as:

$$number\ of\ operational\ rules = 3\ (breakage\ rate) \times 5\ (hydraulic\ factor) \times 5\ (water\ quality) \times 5\ (leakage\ rate) = 375\ rules \quad \text{Equation IV.6}$$

Example: IF breakage rate is *high* AND hydraulic factor (roughness test) is *rough* AND water quality is *very bad* AND leakage rate is *medium*.

2) Weights of factors

This is discussed in section IV.1.1. Weights of Factors. The normalized global weights or the local weights can be used in this process as shown in Table IV.26.

Table IV.26 – Operational factors weights.

Main Risk Factor	**Risk sub-factor**	**Sub-factors local weights**	**Global weights**	**Normalized Global weights**
Operational Factors	Breakage Rate	35	980	57
	Hydraulic Factor	13	364	21
	Water Quality	17	504	29
	Leakage	20	560	32

3) Factors performance impacts

The factors performance impacts at different stages (collected in section IV.1.2. Factors Performance Impact) is measured by a scale as shown in Figure IV.6.

To illustrate how this scale works, an example is given below:

- If breakage rate is high then risk of failure is extremely high (9.17 impact value)
- If hydraulic factor (roughness test) is rough then risk of failure is very high (8.33 impact value)
- If water quality is very bad then the risk of failure is extremely high (9.17 impact value)
- If leakage rate is medium then the risk of failure is medium (5 impact value)

4) Factors performance combined impact

In order to find the combined impact of the different factors' performance, weighted average method is used as shown in Equation IV.2 which is matched against another scale that is divided into ranges of different equivalent impact as shown in Figure IV.7.

Continuing the example shown above, the combined impact is found and the rule is established. So,

$$\text{Equivalent impact} = \frac{(9.17 \times 35) + (8.33 \times 13) + (9.17 \times 17) + (5 \times 20)}{(35 + 13 + 17 + 20)} \qquad \text{Equation IV.7}$$

$$= 8.06$$

Matching the value of *equivalent impact* against the scale in Figure IV.7 yields that the linguistic equivalent impact is very high. So the rule will be completed as:

IF breakage rate is *high* AND hydraulic factor (roughness test) is *rough* AND water quality is *very bad* AND leakage rate is *medium* THEN risk of failure is *very high.*

All other factors' performance impacts combinations (375 combinations) can be found using this approach. The developed form of rule can be used in an expert system such as fuzzy expert system. A sample of these rules are shown in Table IV.27.

Table IV.27 – Sample operational factors performance combined impact.

Rule No.	Factors' performance				combined impact
	breakage	roughness	W. Quality	Leakage	linguistic
1	low	very rough	Very good	very high	Moderately Low
2	low	very rough	Very good	high	Moderately Low
3	low	very rough	Very good	medium	Moderately Low
4	low	very rough	Very good	low	Very Low
5	low	very rough	Very good	very low	Very Low
6	low	very rough	Good	very high	Medium
7	low	very rough	Good	high	Moderately Low
8	low	very rough	Good	medium	Moderately Low
9	low	very rough	Good	low	Very Low
:	:	:	:	:	:
:	:	:	:	:	:
373	high	very smooth	Very bad	medium	Moderately High
374	high	very smooth	Very bad	low	Moderately High
375	high	very smooth	Very bad	very low	Moderately High

iv. Post failure Factors

1) Number of performance combination

The number of performance combination can be calculated for the post failure factors by multiplying the number of performance stages of the factors. In this research, the number of different performance combinations needed to cover all the combination possibilities can be found as:

$$\begin{aligned} &number\ of\ post\ failure\ rules \\ &\quad = 5\ (repair\ cost) \\ &\quad \times 3\ (damage\ to\ surrounding) \\ &\quad \times 4\ (loss\ of\ production) \\ &\quad \times 5\ (traffic\ disruption) \\ &\quad \times 3\ (type\ of\ serviced\ area) \\ &\quad = 900\ rules \end{aligned}$$

Equation IV.8

Example: IF repair cost is *low* AND damage to surrounding is *in industrial area* AND loss of production is *in small redundant pipe* AND traffic disruption is *lightly disruptive* AND type of serviced area is *industrial*.

2) Weights of factors

This is discussed in section IV.1.1. Weights of Factors. The normalized global weights or the local weights can be used in this process as shown in Table IV.28.

Table IV.28 – Operational factors weights.

Main Risk Factor	Risk sub-factor	Sub-factors local weights	Global weights	Normalized Global weights
Post-Failure Factors	Cost of Repair	20	400	23
	Damage to surroundings	21	420	24
	Loss of Production	18	360	21
	Traffic Disruption	17	340	20
	Type of Serviced Area	14	280	16

3) Factors performance impacts

The factors performance impacts at different stages (collected in section IV.1.2. Factors Performance Impact) is measured by a scale as shown in Figure IV.6.

To illustrate how this scale works, an example is given below:

- If repair cost is low then risk of failure is very low (1.67 impact value)
- If damage to surrounding is in industrial area then risk of failure is extremely high (9.72 impact value)
- If loss of production is in small redundant pipe then the risk of failure is very low (1.67 impact value)
- If traffic disruption is lightly disruptive then the risk of failure is very low (1.67 impact value)
- If type of serviced area is industrial then the risk of failure is very high (8.33 impact value)

4) Factors performance combined impact

In order to find the combined impact of the different factors' performance, weighted average method is used as shown in Equation IV.2 which is matched against another scale that is divided into ranges of different equivalent impact as shown in Figure IV.7.

Continuing the example shown above, the combined impact is found and the rule is established. So,

$$Equivalent\ impact = \frac{(1.67 \times 20) + (9.72 \times 21) + (1.67 \times 18) + (1.67 \times 17) + (8.33 + 14}{(20 + 21 + 18 + 17 + 14)} \quad \text{Equation IV.9}$$

$$= 4.58$$

Matching the value of *equivalent impact* against the scale in Figure IV.7 yields that the linguistic equivalent impact is medium. So the rule will be completed as:

IF repair cost is *low* AND damage to surrounding is *in industrial area* AND loss of production is *in small redundant pipe* AND traffic disruption is *lightly disruptive* AND type of serviced area is *industrial* THEN risk of failure is *medium.*

All other factors' performance impacts combinations (900 combinations) can be found using this approach. The developed form of rule can be used in an expert system such as fuzzy expert system. A sample of these rules are shown in Table IV.29.

Table IV.29 – Sample post failure factors performance combined impact.

Rule No.	Factors' performance					Combined impact
	Repair cost	Damage to Surrounding	Loss of Production	Traffic Disruption	Type of Service area	Linguistic
1	very high	Industrial	Small-Redundant	Very heavy	Industrial	Very High
2	very high	Industrial	Small-Redundant	Very heavy	Commercial	Moderately High
3	very high	Industrial	Small-Redundant	Very heavy	Residential	Moderately High
4	very high	Industrial	Small-Redundant	Heavy	Industrial	Very High
5	very high	Industrial	Small-Redundant	Heavy	Commercial	Moderately High
6	very high	Industrial	Small-Redundant	Heavy	Residential	Moderately High
7	very high	Industrial	Small-Redundant	Moderate	Industrial	Moderately High
:	:	:	:	:	:	:
:	:	:	:	:	:	:
998	very low	Residential	medium-not Redundant	Very light	Industrial	Moderately Low
999	very low	Residential	medium-not Redundant	Very light	Commercial	Moderately Low
900	very low	Residential	medium-not Redundant	Very light	Residential	Very Low

v. Risk of failure Factors

1) Number of performance combination

The number of performance combination can be calculated for the main risk factors by multiplying the number of performance stages of the factors. In this research, the number of different performance combinations needed to cover all the combination possibilities can be found as:

$$\begin{aligned} &number\ of\ post\ failure\ rules \\ &\quad = 7\ (environmantal\ factor) \\ &\quad \times 7\ (physical\ factor) \\ &\quad \times 7\ (operational\ factor) \\ &\quad \times 7\ (post\ failure\ factor) = 2401\ rules \end{aligned}$$

Equation IV.10

Example: IF environmental factor is *moderately low* AND physical factor is *very high* AND operational factor is *moderately high* AND post failure factor is *very low*.

2) Weights of factors

This is discussed in section IV.1.1. Weights of Factors. The normalized global weights or the local weights can be used in this process as shown in Table IV.28.

Table IV.30 – Operational factors weights.

Main Risk Factor	Risk sub-factor	Sub-factors local weights	Normalized weights
Main risk factors	Environmental factor	19	44
	Physical factor	43	100
	Operational factor	28	65
	Post failure factor	20	46

3) Factors performance impacts

The factors performance impacts at different stages (collected in section IV.1.2. Factors Performance Impact) is measured by a scale as shown in Figure IV.6.

To illustrate how this scale works, an example is given below:

- If environmental factor is moderately low then risk of failure is moderately low (3.33 impact value)
- If physical factor is very high then risk of failure is very high (8.33 impact value)

- If operational factor is moderately high then the risk of failure is moderately high (6.67 impact value)
- If post failure factor is very low then the risk of failure is very low (1.67 impact value)

4) Factors performance combined impact

   In order to find the combined impact of the different factors' performance, weighted average method is used as shown in Equation IV.2 which is matched against another scale that is divided into ranges of different equivalent impact as shown in Figure IV.7.

   Continuing the example shown above, the combined impact is found and the rule is established. So,

$$\begin{aligned} &\mathit{Equivalent\ impact} \\ &= \frac{(3.33 \times 19) + (9.72 \times 43) + (6.67 \times 28) + (1.67 \times 20)}{(19 + 43 + 28 + 20)} \qquad \text{Equation IV.11} \\ &= 6.38 \end{aligned}$$

   Matching the value of *equivalent impact* against the scale in Figure IV.7 yields that the linguistic equivalent impact is moderately high. So the rule will be completed as:

IF environmental factor is *moderately low* AND physical factor is *very high* AND operational factor is *moderately high* AND post failure factor is *very low* THEN risk of failure is *moderately high*.

All other factors' performance impacts combinations (2401 combinations) can be found using this approach. The developed form of rule can be used in an expert system such as fuzzy expert system. A sample of these rules are shown in Table IV.31.

Table IV.31 – Sample risk of failure factors performance combined impact.

Rule No.	Factors' performance				Combined impact
	physical	environmental	operational	consequence	Linguistic
1	Extremely Low	Extremely Low	Extremely Low	Extremely Low	Extremely Low
2	Extremely Low	Extremely Low	Extremely Low	Very Low	Extremely Low
3	Extremely Low	Extremely Low	Extremely Low	Moderately Low	Very Low
4	Extremely Low	Extremely Low	Extremely Low	medium	Very Low
5	Extremely Low	Extremely Low	Extremely Low	Moderately High	Very Low
6	Extremely Low	Extremely Low	Extremely Low	Very High	Very Low
7	Extremely Low	Extremely Low	Extremely Low	Extremely High	Very Low
8	Extremely Low	Extremely Low	Very Low	Extremely Low	Extremely Low
9	Extremely Low	Extremely Low	Very Low	Very Low	Very Low
:	:	:	:	:	:
:	:	:	:	:	:
2399	Extremely High	Extremely High	Extremely High	Moderately High	Very High
2400	Extremely High	Extremely High	Extremely High	Very High	Extremely High
2401	Extremely High	Extremely High	Extremely High	Extremely High	Extremely High

IV.2. Case Study Data Sets

In this stage, the performance data is collected from real networks under operation. Three sets of data are collected from municipalities in New Brunswick and Ontario.

IV.2.1. Data Set One

The data of this case study is collected from the City of Moncton, New Brunswick, Canada. The City of Moncton operates a water supply and distribution system which provides water to 95% of its population. The approximate length of the water main is 448 km. It serves more than 58,000 people. Cast iron water mains account for about 39% of all the water main, followed by ductile iron with 31%. PVC water mains account for 19%. Asbestos cement (3%) accounts for a much smaller part of the system (Dillon Consulting and Harfan Technologies, 2003).

The factors included in dataset one are: pipe material, pipe diameter, installation year, protection method, number of breakage, Hazen-William factor, and loss of production (pipe diameter). The number of records in this data set is only 544. The actual data is much bigger than this number, however, these 544 records are the only records that have information about their current status (breakage rate, Hazen-William coefficient ...etc). Some statistics about the 544 records data set are shown in Table IV.32. The percentages of the pipe material used in the Moncton system is shown in Figure IV.8, which shows that the most used pipe material is Post War Cast Iron (built after World War II).

Table IV.32 – Moncton data set statistics.

Pipe Material	Length, *m*	Diameter	Length, *m*	Protection Method	Length, *m*
Asbestos	6,578	Small	107,770	Lining	2,401
Cast Iron (Pre WW II)	28,024	Medium	45,274	Cathodic protection	0
Cast Iron Post War	79,483			None	150,643
Ductile Iron	35,869				
PVC	3,091				

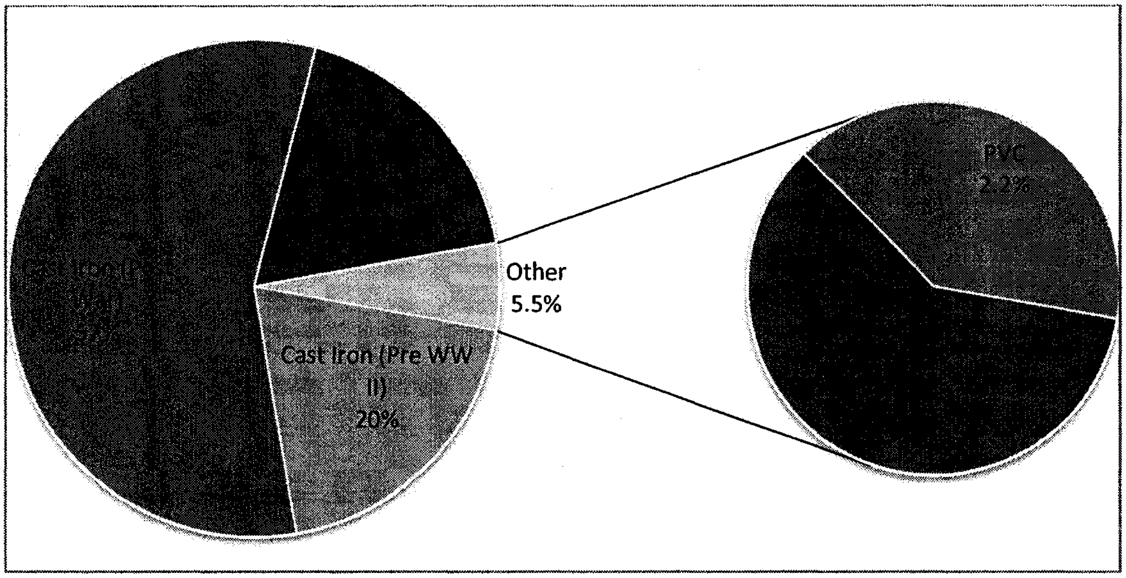


Figure IV.8 – Percentages of pipe materials used in Moncton.

IV.2.2. Data Set Two

This case study is derived from Case Study one by assuming the unavailable qualitative factors. This data set is built to show and study the results of the model when information about all the factors incorporated in the model is available to the management team. The factors assumed in this dataset are: type of soil, average daily traffic, water table level, water quality, leakage rate, cost of repair, damage to surroundings, traffic disruption, and type of serviced area. The data is randomly assumed and does not fit any distribution. However, the values are assumed in the higher risk performance of the factors. The number of records in this data set is 544.

IV.2.3. Data Set Three

The data in this case study is collected from the City of London, Ontario, Canada. The information included in this data set is: type of soil, average daily traffic, pipe material, pipe diameter, installation year, protection method, number of breaks, hydraulic factor, and loss of production. The number of records in this data set is 1702. Some statistics about this 1702-record database are shown in Table IV.33. The percentages of pipe material used in the London database are shown in Figure IV.9, which shows that the most used pipe material is Post War Cast Iron (built after WW II) and which accounts for 66% of the network.

Table IV.33 – London data set statistics.

Pipe Material	Length , *m*	Diameter	Length , *m*	Protection Method	Length , *m*
Cast Iron (Pre WW II)	56,022	Small	244,894	Lining	0
Cast Iron (Post War II)	190,083	Medium	29,881	Cathodic protection	13,418
Ductile Iron	26,419			None	261,357
PVC	2,084				
PE	40				
Concrete	171				

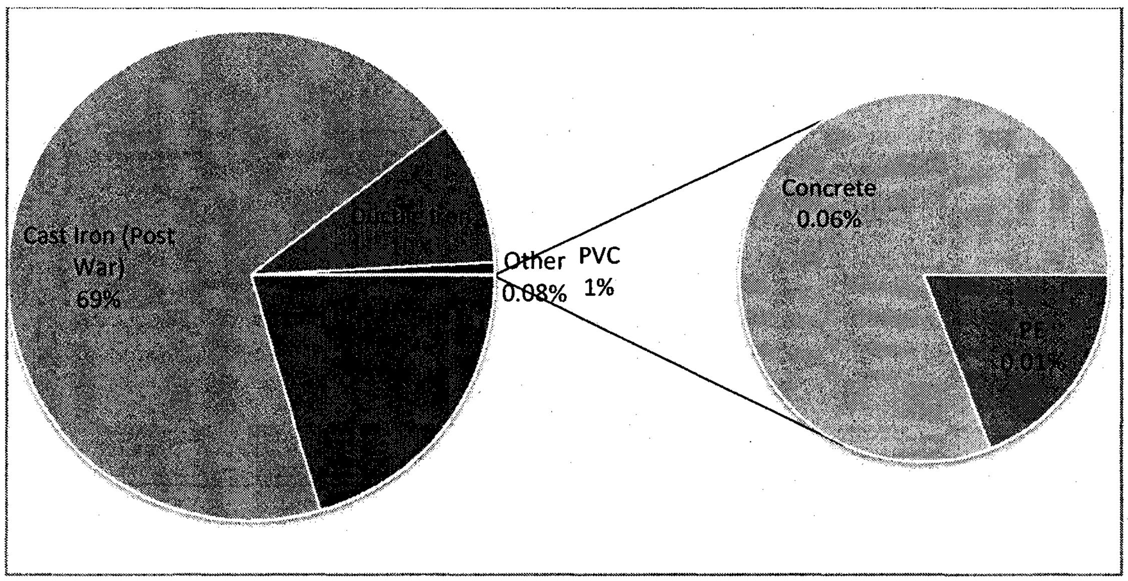


Figure IV.9 – Percentages of pipe materials used in London.

Chapter V: HIERARCHICAL FUZZY EXPERT SYSTEM FOR RISK OF WATER MAIN FAILURE

Maintaining a water main in good condition requires the adoption of a maintenance/repair plan which prioritizes and ranks the most critical (risky) pipelines. This can be done by using an expert system that makes use of the expert opinions and experiences in the specified field. In this section, the development of a fuzzy expert system for water main failure risk is explained and discussed. Figure V.1 shows the different topics covered in this chapter that explain the steps followed in building the hierarchical fuzzy system.

V.1. Risk Factors Incorporated in the Model

In this step, the failure risk factors are identified and selected. Sixteen factors are incorporated in this model, which represents the deterioration and post-failure factors. These factors are extracted from Table II.3 which lists the deterioration factors that contribute to the pipeline failure event and from section "II.2.3. Consequences of Failure", which lists some of the consequences (cost) of a failure event. The deterioration factors chosen to be incorporated in this model are selected based on the ease of gaining the required attributes of the water main by the facility managers. These attributes can be gathered from different types of documents such as: design information, visual inspection reports, maintenance reports, etc. The cost of failure (consequence) factors are difficult to quantify and thus a qualitative approach will be followed. The factors selected to be incorporated in the pipeline failure risk model are shown in Figure III.2.

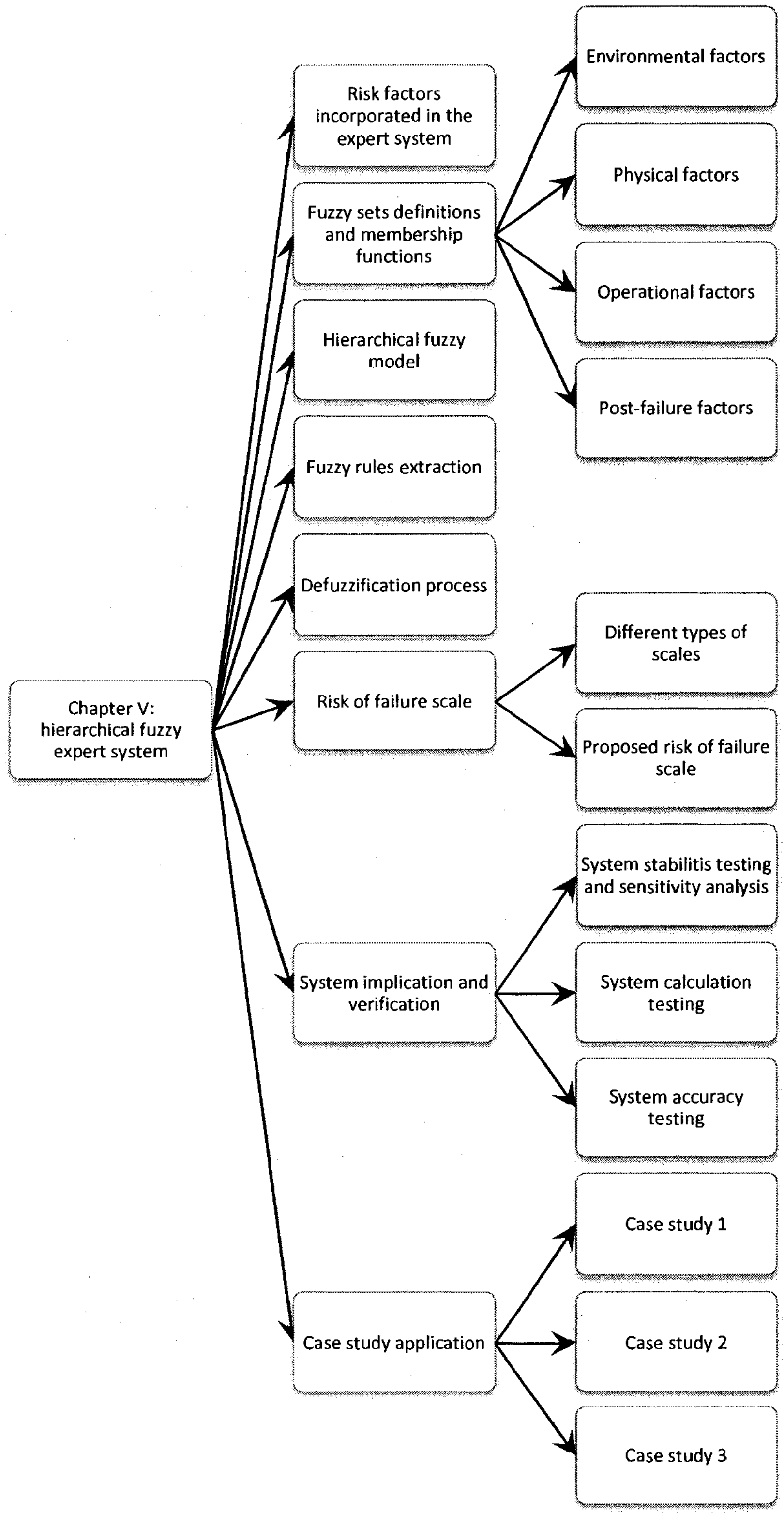


Figure V.1 – Chapter layout of hierarchical fuzzy expert system for water main failure risk.

V.2. Hierarchical Fuzzy Model

The hierarchical fuzzy model structure consists of four branches (models) which correspond to the four main factors and another model that combines the results of the four branches of the hierarchy to produce risk of failure. These are: Environmental, Physical, Operational, Post-failure branches and risk of failure model as shown in Figure V.2. This hierarchical structure will facilitate the creation of a pre-failure model which combines three factors (environmental, physical, and operational) to produce a pre-failure index or the possibility of failure index on a scale of 0 to 10 as shown in Figure V.3. The post-failure model represents the consequence(s) of failure. The fuzzy structure of each of the five models is identical and only the membership functions of each factor in each model and the knowledge base rules of each model are different. The full view of the hierarchical fuzzy model is shown in Figure V.4 which shows the processing of the observations characteristics of the water main network.

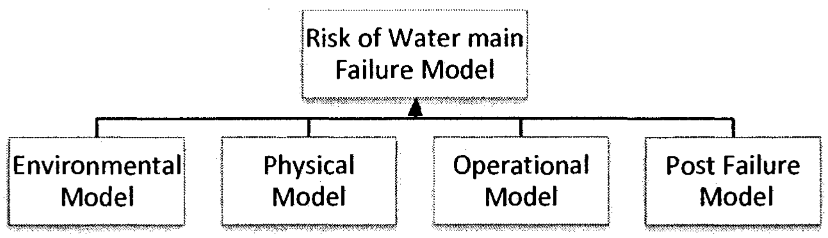


Figure V.2 – Hierarchical fuzzy failure risk Model.

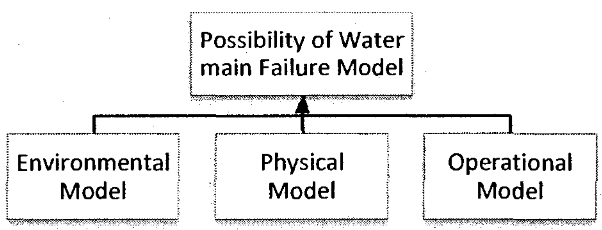


Figure V.3 – Hierarchical fuzzy possibility of failure models.

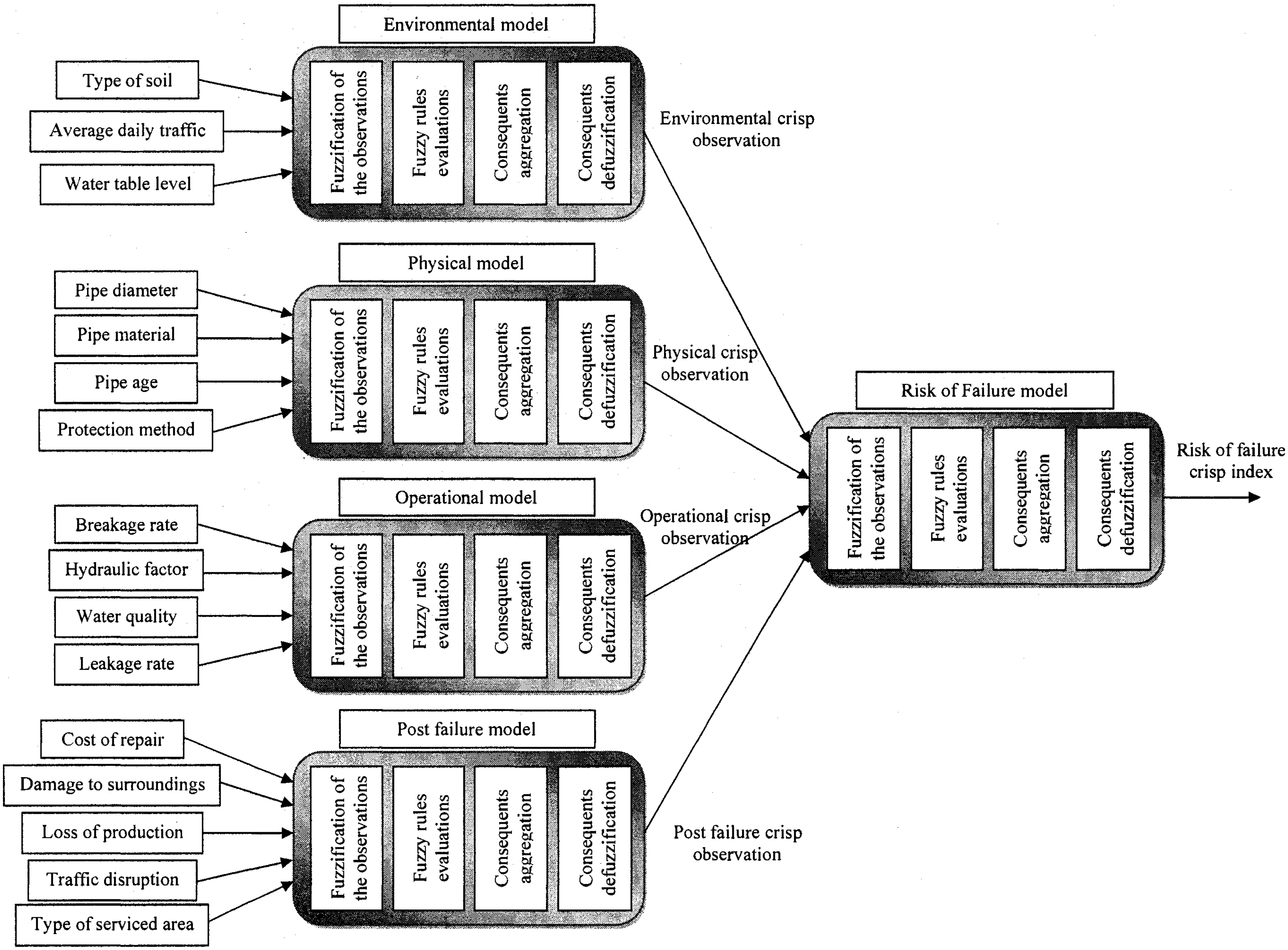


Figure V.4 – Full view of the model components

The crisp defuzzified results of the four models (environmental, physical, operational, and post-failure) are combined together through a risk of failure model which calculates the risk of failure index of a water main as shown in Figure III.4. The use of a hierarchical fuzzy expert system is a key to reducing the total number of required expert rules. In this model, if a hierarchical fuzzy system is not used, then the total number of rules required to cover all of the possible factor performance combinations is calculated by the simple multiplication of the number of factors performances of each of the sixteen factors. Section A.10. Hierarchical fuzzy expert system in Appendix A includes more information about this topic.

V.3. Fuzzy Sets Definitions and Membership Functions

The membership functions of the different factors are built based on the information gathered from the literature, such as the characteristics of each factor, and the effects of these characteristics on the risk of failure. The qualitative factors are evaluated on a 0-10 scale and assigned a standard five membership functions. In this section, the established fuzzy sets for each of the factors are shown as follows:

V.3.1. Environmental Model

It includes soil type, water table level, and average daily traffic.

Soil Type

Soil type affects the external corrosion rate of metallic pipes and thus is considered one of the mort important factors. Specific types of soil can lead to biochemical, electrochemical, and physical reactions which can degrade the pipe material and make it vulnerable to structural degradation, which then results in thinning or weakening of the pipe material, causing the material to lose its ability to resist the forces in the surrounding soil (Hahn *et al.* 2002). Soil is typically classified by grain size according to the Unified Soil Classification System as coarse grained and fine grained which in their turn are classified as Gravel, Sand, Clay and Silt with liquid limit > 50, and Clay and Silt with liquid limit < 50. However, the most important soil characteristic for water mains is the presence of chemicals that deteriorate pipeline material, and the interaction between the soil and the pipe material. Thus, soil is classified according to potential corrosiveness as highly corrosive, moderately corrosive, and low corrosive (Al Barqawi, 2006). Soil uniformity is also considered an important factor. When the pipe is in contact with dissimilar soil types, localized corrosion cells can be developed which contribute to metallic pipe material corrosion. Moreover, soil pH is considered a good indicator of external corrosion because corrosion occurs in a certain range of pH (Najafi, 2005). There are many soil characteristics that play a role in the deterioration process and thus make studying their effects complex and beyond the scope of this research. Therefore, for this research, the soil is classified into five subjective groups according to the strength of deterioration action as very highly deteriorative, highly deteriorative, moderately deteriorative, lightly deteriorative, and very lightly deteriorative. The membership functions and their characteristics are shown in Figure V.5. The data type to be used for this factor is numerical from 0 to 10 where 0 indicates the least deteriorative soil condition and 10 indicates the highest deteriorative soil condition.

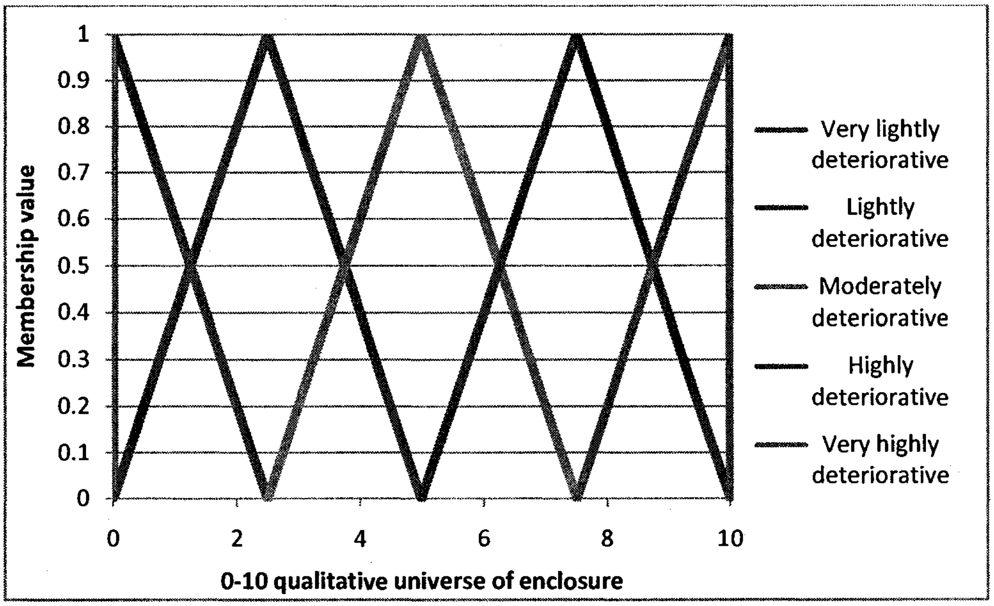


Figure V.5 – Soil type membership functions.

Water Table Level

The effect of water table on pipeline materials is due to the presence of certain salts and other corrosive materials dissolved in the water. Another adverse effect of the presence of groundwater is the tendency of the water to cause corrosion of metallic pipelines. Al Barqawi (2006) classified the groundwater level as high, moderate, or low. There is little work that covers the effect of groundwater on water mains in comparison to sewer mains. In regard to sewer mains, the presence of groundwater is classified as above or below invert and whether it is stable or varies seasonally (Hahn *et al.* 2002). Moreover, the rate of frost heaves which bears a load on pipelines is controlled by the presence of groundwater

(Najafi, 2005). The easily observed characteristic of groundwater which can be monitored is the presence of it. So it can be classified as always present, seasonally present, and rarely present. The membership functions and their characteristics are shown in Figure V.6. The membership functions of the water table level are discrete and a 0.95 confidence level (certainty) is assumed. The data type to be used for this factor is linguistic and chosen from this list: rarely present, seasonally present, always present.

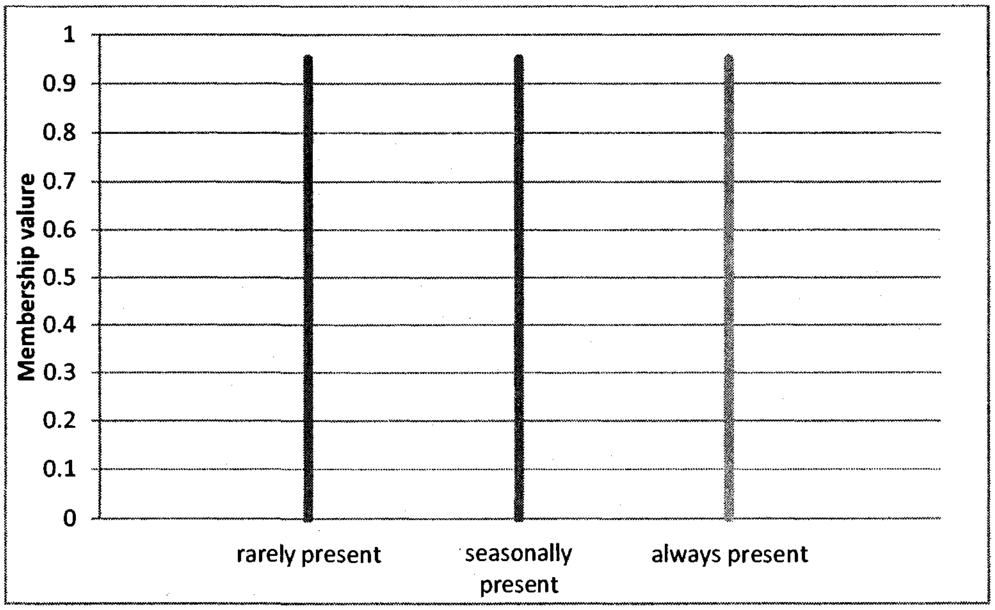


Figure V.6 – Water table lever membership functions.

Average Daily Traffic

The daily traffic on the road above a buried pipeline creates a dynamic load on the pipeline. Dynamic forces that cause structural defects are either large, one-time events or smaller cyclic events that occur at a variety of frequencies (daily or seasonally). Large, one-time events include periods of heavy surface construction, in-ground utility construction, or non-construction events, such as earthquakes or landslides. This type of dynamic load is beyond the scope of this research and only the smaller cyclical dynamic loads caused by routine truck, machinery, bus, or train traffic is considered in this research (Hahn *et al.* 2002). Moreover, the depth of pipelines plays a role in transferring the dynamic surface load into pipes structure (the greater the depth, the lesser the load transferred to the pipe structure). Al Barqawi (2006) classified this factor into high, moderate, and low according to average daily traffic. Raven (2007) classified road types into (1) paved, (2) low/moderate traffic, and (3) high traffic. In this research, daily traffic is classified into 5 subjective groups as very heavy, heavy, moderate, light and very light average daily traffic. The membership functions and their characteristics are shown in Figure V.7. The data type to be used for this factor is numerical from 0 to 10 where 0 indicates the lightest average daily traffic condition and 10 indicates the highest average daily traffic condition.

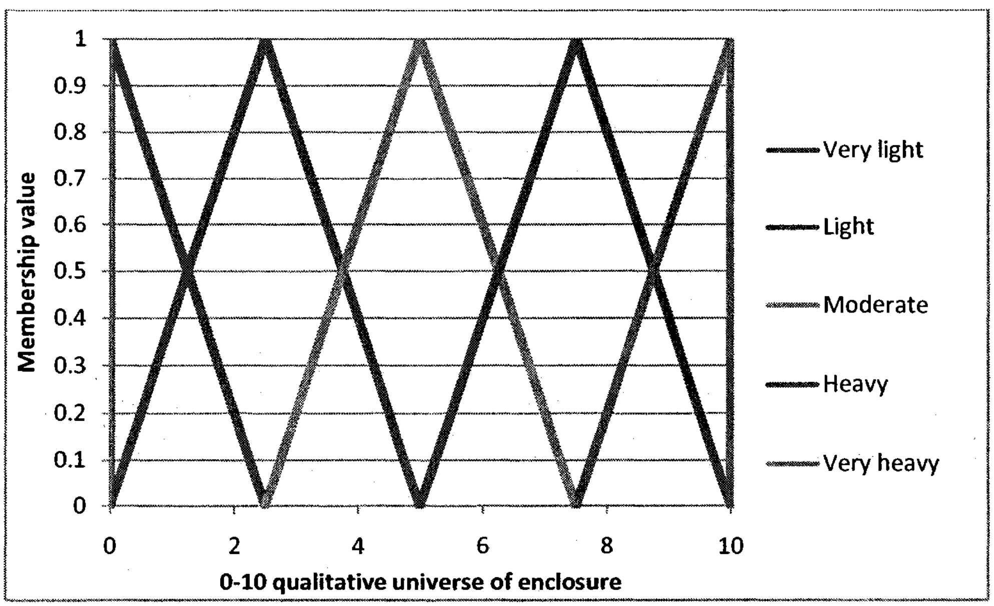


Figure V.7 – Average daily traffic membership functions.

V.3.2. Physical Model

This model includes pipe diameter, pipe material, pipe age, and protection method factors.

Pipe Diameter

According to Al Barqawi (2006), pipe size is one of the most important factors that contribute to the pipeline failure. In his investigation of risk factors in urban pipeline failure, Raven (2007) classified pipeline diameter into three groups: group 1 (4 in. to 8 in.), group 2 (10 in. to 30 in.), and group 3 (36 in. to 72 in.). Ozger (2003) developed a regression model to estimate water main breakage rate -- one of his findings is that the breakage rate of pipelines decreases as the pipe diameter increases. This is because larger diameter pipes have more beam strength than smaller diameter pipes (Najafi, 2005). In light of the above review, the pipe diameter factor is classified into 2 groups as small (less than 250 mm) and medium (250 mm to 500 mm). The large diameter pipelines (greater than 500 mm) are not considered here, since they are used in transmission water mains, which are beyond the scope of this research. The membership functions and their characteristics are shown in Figure V.8. The data type to be used for this factor is pipeline diameter, up to 500 millimeters.

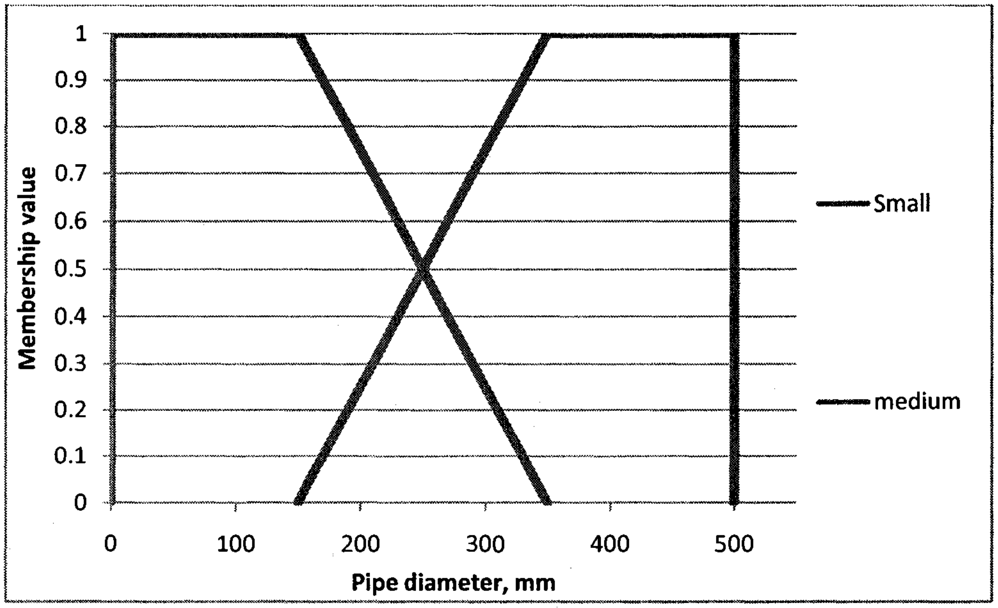


Figure V.8 – Pipe diameter membership functions.

Pipe Material

There are three main categories of pipeline materials that are used in the construction of pressurized pipelines: cement-based, plastic, and metallic. Each category of pipeline material includes a variety of materials. The pipeline materials considered in this research are summarized in Figure V.9. There are other types of pipeline material which are not considered in this research; Verified Clay pipes are only used in sewer pipelines due to their low tensile strength, and Glass-Reinforced Plastic (Fiberglass) pipe, which is traditionally used in industrial applications and large diameter (transmission) municipal water mains (Najafi, 2005).

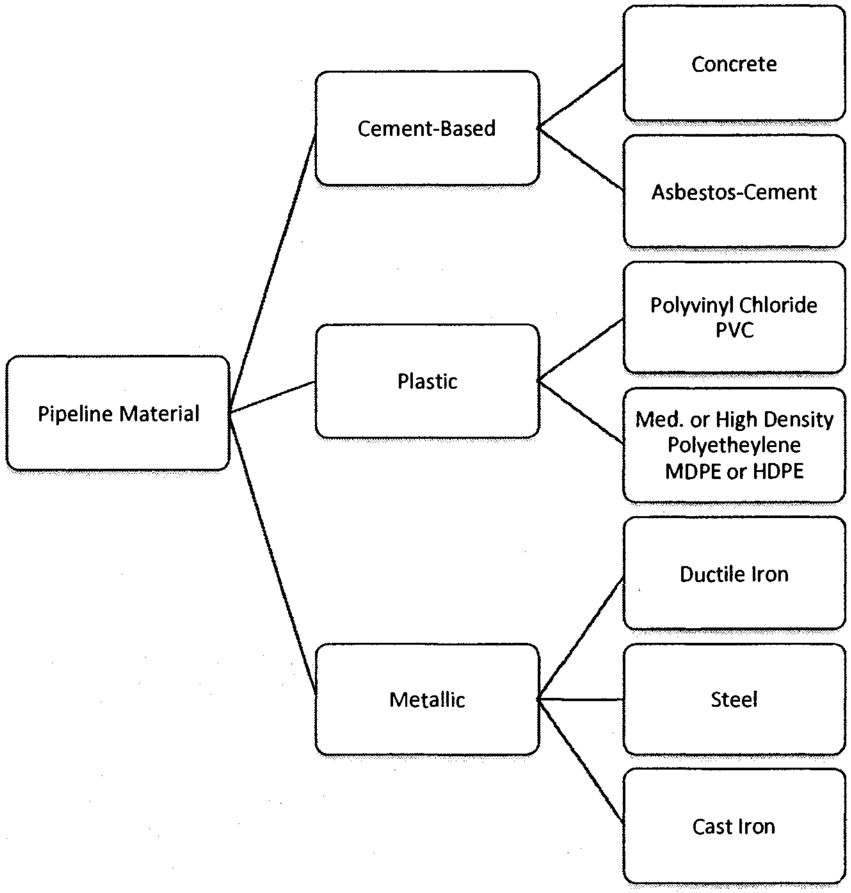


Figure V.9 – Pipeline materials.

The risk of exterior pipe deterioration depends on the pipe material, which is susceptible to acidic substances and galvanic corrosion. Acidic soils or groundwater attack unprotected cementious or metallic pipe materials, while stray currents in the ground cause galvanic corrosion with metal or metal reinforced pipes. Erosion is often a problem in concrete, asbestos cement, and metallic pipes (Hahn *et al.* 2002). Another aspect that should be considered regarding pipe material is the pipe vintage. This concept is especially related to cast iron water mains. Water mains made of cast iron were produced using two different casting methods; before and after the Second World War. Post-war cast iron pipes (made by open casting) are more vulnerable to failure in long-term performance. For this reason, cast iron pipes are categorized as pre-war and post-war cast iron (between 1950 and 1970) (Dillon Consulting and Harfan Technologies, 2003). According to Al Barqawi's (2006) findings, pipeline material is considered the second most important factor in the pipeline deterioration process. The membership functions and their characteristics are shown in Figure V.10. The membership functions of pipe materials are discrete and a 0.95 confidence level (certainty) is assumed. Examining the effects of different confidence levels (80 to 100) shows that the model is not very sensitive to this value. The data type to be used for this factor is linguistic and chosen from this list: Concrete, Asbestos, PVC, PE, Ductile iron, Steel, Cast iron, and Cast iron post war.

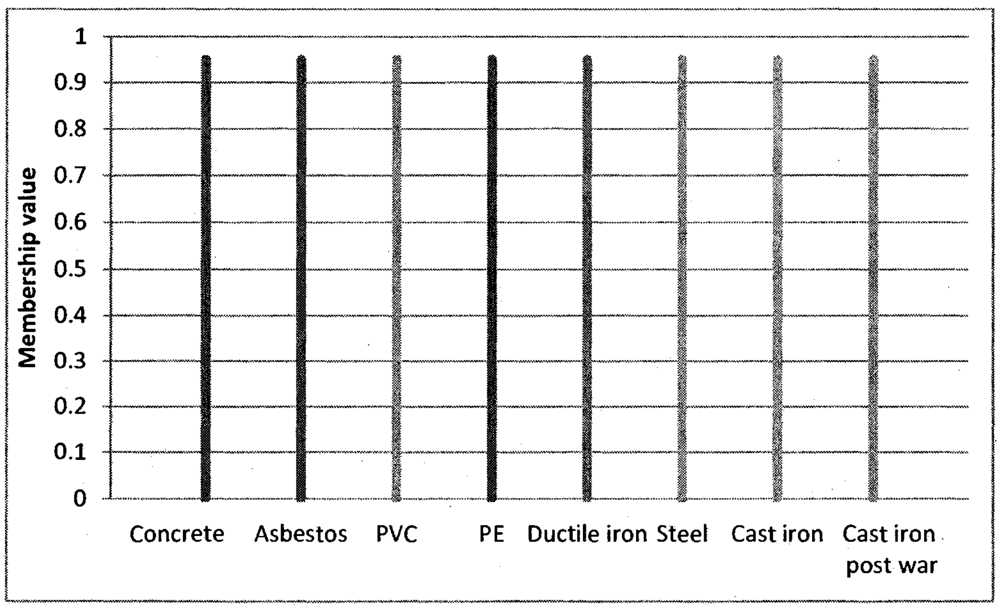


Figure V.10 – Pipe material membership functions.

Pipe Age

Pipe age is considered the most important factor in indicating the level of pipeline deterioration (Al Barqawi, 2006). Pipelines usually have a 'bathtub' rate of failure relative to the age of the pipes as shown in Figure V.11. Early failure is due to human factors in the actual laying of the pipe, such as manufacturing faults. The second part of the curve has a low failure rate. In the third part of the curve, the failure rate increases exponentially as the pipeline approaches the end of its' useful life (Najafi, 2005).

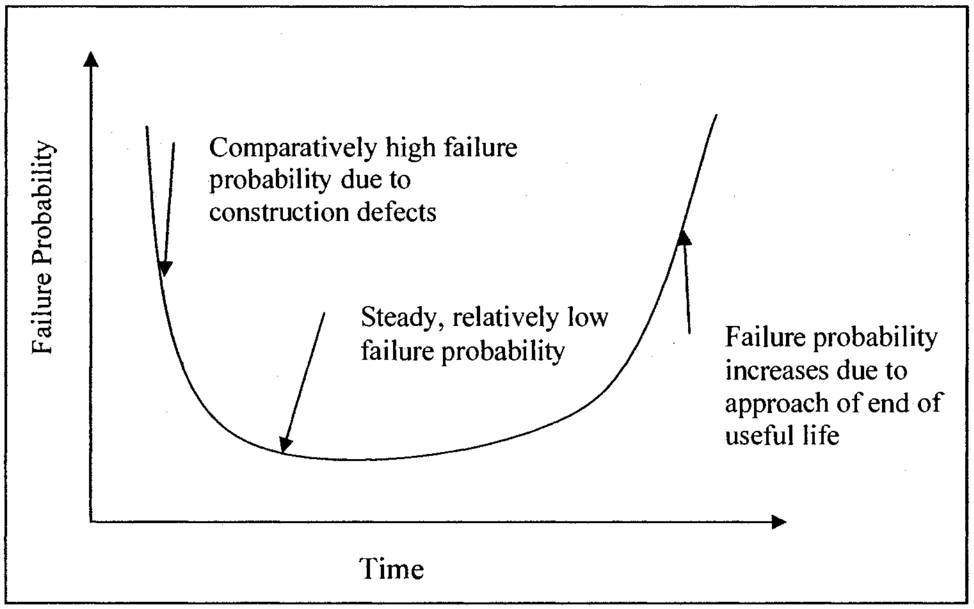


Figure V.11 – Bathtub curve of pipe performance with age (Najafi, 2005).

Kleiner *et al.* (2004) divided the age range into five membership functions: new, young, medium, old, and very old. In this research, similar assumptions are used. The age membership functions and their characteristics are shown in

Figure V.12. The data type to be used for this factor is the installation year of the pipeline – the model will automatically calculate its age.

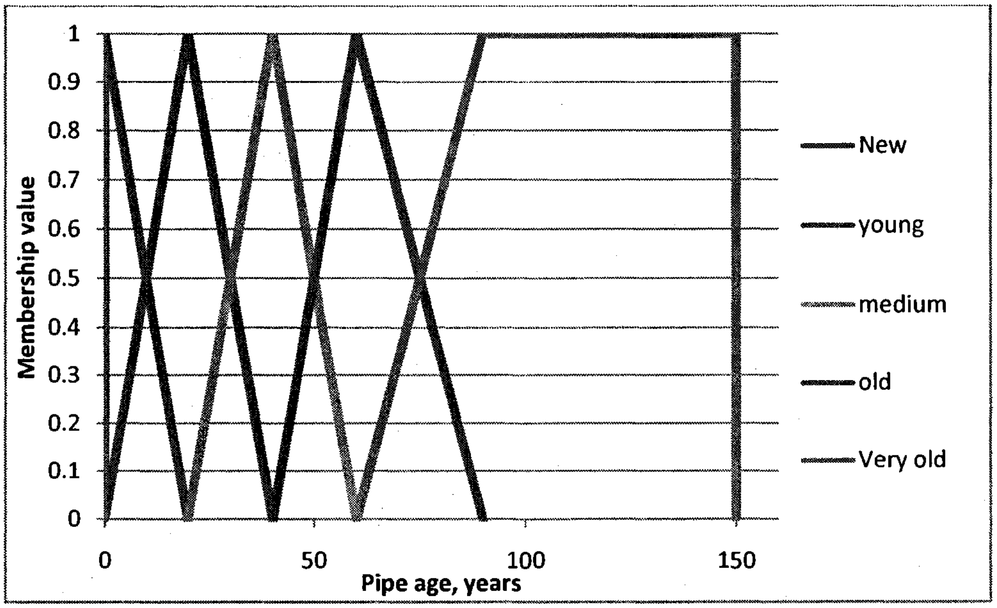


Figure V.12 – Pipe age membership functions.

Protection Method

Pipelines should be protected against potential corrosion and deterioration. Many corrosion protection methods are being applied in the field of pipeline protection especially for pipelines made of iron and steel materials. Other types of pipe materials such as reinforced concrete, plastics, and composites also undergo forms of corrosion or different environmental or stress-related deterioration. For instance, Polyvinyl Chloride (PVC) pipes have a high resistance to deterioration and can be used in very corrosive environments, but they are likely to be affected by deterioration if they are exposed to weather, chemical attack, or mechanical degradation arising from improper installation methods (Al Barqawi, 2006). However, protection methods are mainly applied to iron and steel pipes due to their high vulnerability to corrosion. Some of the protection methods are internal cement mortar lining, polyurethane lining, polyethylene encasement or wrapping, tape coating, coal tar enamel coatings, epoxy or polyurethane coatings, and cathodic protection, which is the most effective protection method for steel pipes (Najafi, 2005). In this research, the protection methods are classified as cathodic protection, lining/coating, and not applied. The membership functions and their characteristics are shown in Figure V.13. The membership functions are discrete and a 0.95 confidence level (certainty) is assumed. Examining the effects of different confidence levels (80 to 100) shows that the model is not very sensitive to this value. The data type used for this factor is linguistic, chosen from this list: Cathodic protection, Lining\Coating, and none.

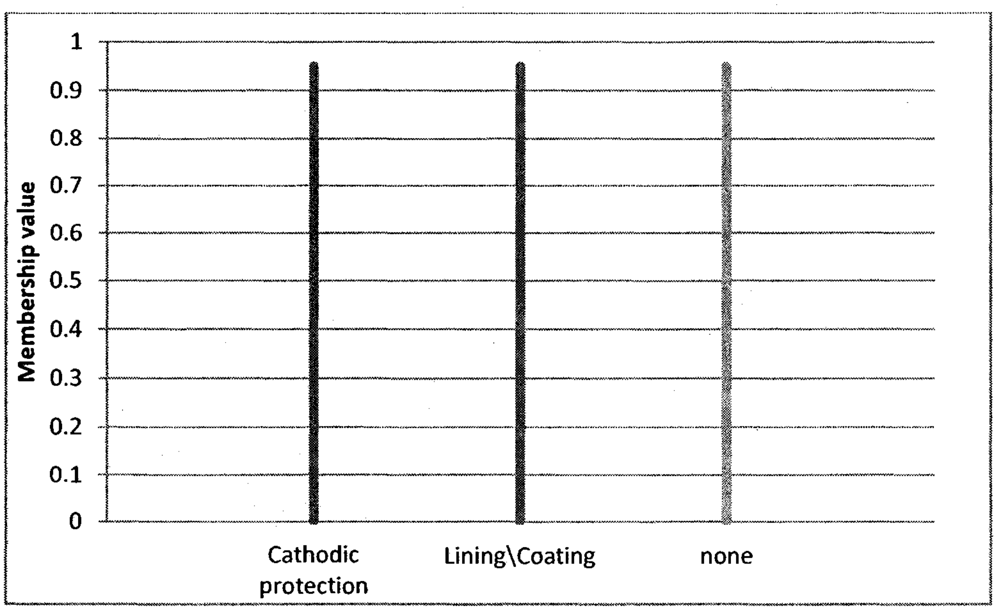


Figure V.13 – Protection methods membership functions.

V.3.3. Operational Model

This model includes breakage rate, hydraulic factor, water quality, and leakage rate factors.

Breakage Rate

The breakage rate is measured as the number of breakage per one kilometer of pipeline length per year. This factor actually gives an indication of the current overall status of the pipeline rather than contributing exclusively to its condition. The breakage rate is considered the third most important factor that indicates material deterioration and thus the probability of failure of the pipeline (Al Barqawi, 2006). However, from closely studying Al Barqawi's results and findings, the breakage rate as a risk factor can be classified into three ranges: low (0 to 0.5), average (0.5 to 3), and high (> 3). Figure V.14 shows of the breakage rate versus the condition rating scale developed by Al Barqawi (2006) which is used to divide the ranges of the breakage rate membership functions. According to the curve analysis, the breakage rate factor changes its behavior at values of 0.5 and 3 brk/km/yr. The membership functions and their characteristics are shown in Figure V.15. The data type used for this factor is the number of water main breaks per one kilometer pre year with a maximum of 10 brks/km/yr.

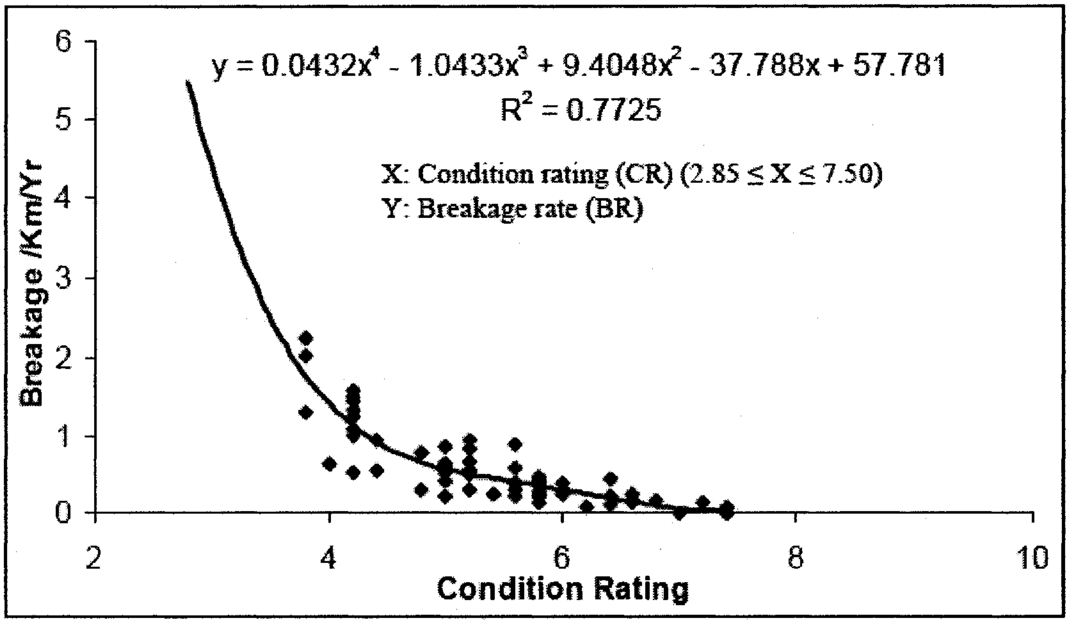


Figure V.14 – Breakage rate vs. condition rating of Cast Iron (Al Barqawi, 2006).

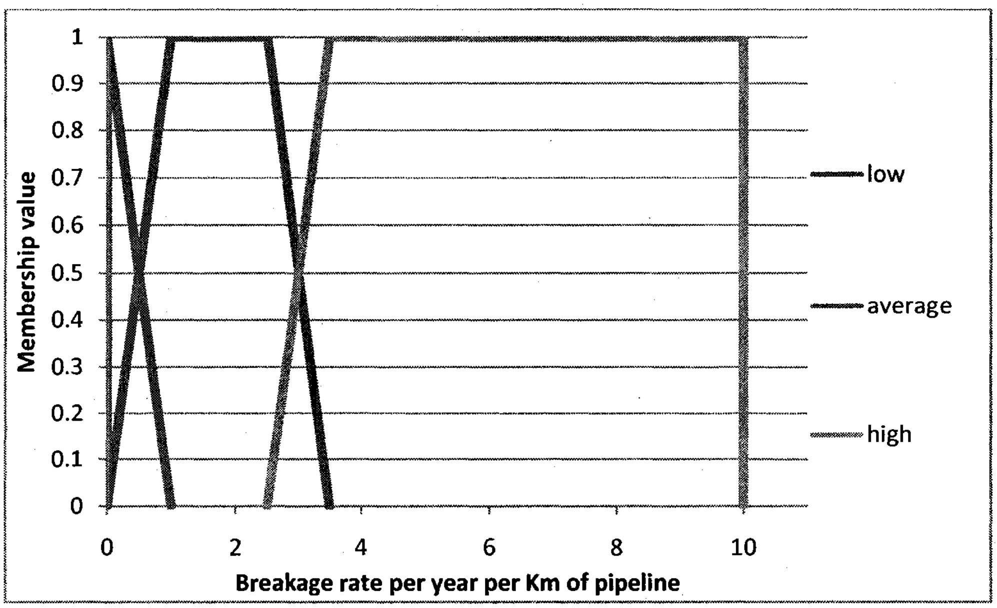


Figure V.15 – Breakage rate membership functions.

Hydraulic Factor

The hydraulic factor is used to measure the current network performance and is usually indicated by the C-factor of the pipelines. A preliminary investigation of the hydraulic capacity of a distribution system can be done by analyzing the results of low-pressure complaints, hydrant-flow, rusty color occurrence, and visual inspection of interior pipe tests, which will give the trend of the distribution system's hydraulic capacity change over time and how it varies spatially. Detailed investigation of the hydraulic factor is carried out using Hazen-William factor tests (roughness test) (InfraGuide, 2003). Hydraulic factor is one of the important factors that give a good indication of the status of the network, and it is classified according to the Hazen-William factor into five groups (Al Barqawi, 2006). The membership functions and their characteristics are shown in Figure V.16. The data type to be used for this factor is the value of the Hazen-Willam factor with a maximum of 150 (maximum value of a new installation).

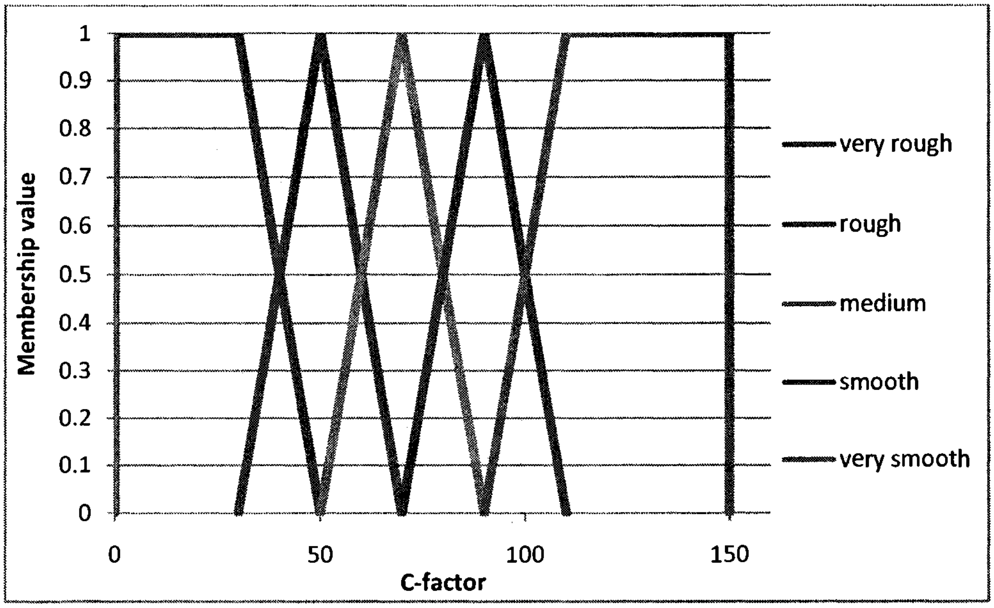


Figure V.16 – Hydraulic factor membership functions.

Water Quality

Measuring water quality in the water distribution network gives a good indication about the internal condition of a pipelines' network. The preliminary data collected in order to assess the water quality in a distribution system is based on analyzing the water quality complaint records and the routine water quality monitoring data. Water quality can be measured based on the concentrations of chlorine residuals and iron in metallic mains. When chlorine residuals are decreased in some areas of a water system, it indicates that these areas are deteriorating. An increasing concentration of iron in the water indicates the degree of internal corrosion of unlined metallic mains (InfraGuide, 2003). In this research, water quality is considered as a subjective factor and classified into five groups: very good, good, acceptable, bad, and very bad. The membership functions and their characteristics are shown in Figure V.17. The data type used for this factor is numerical from 0 to 10 where 0 indicates the best water quality and 10 indicates the worst water quality.

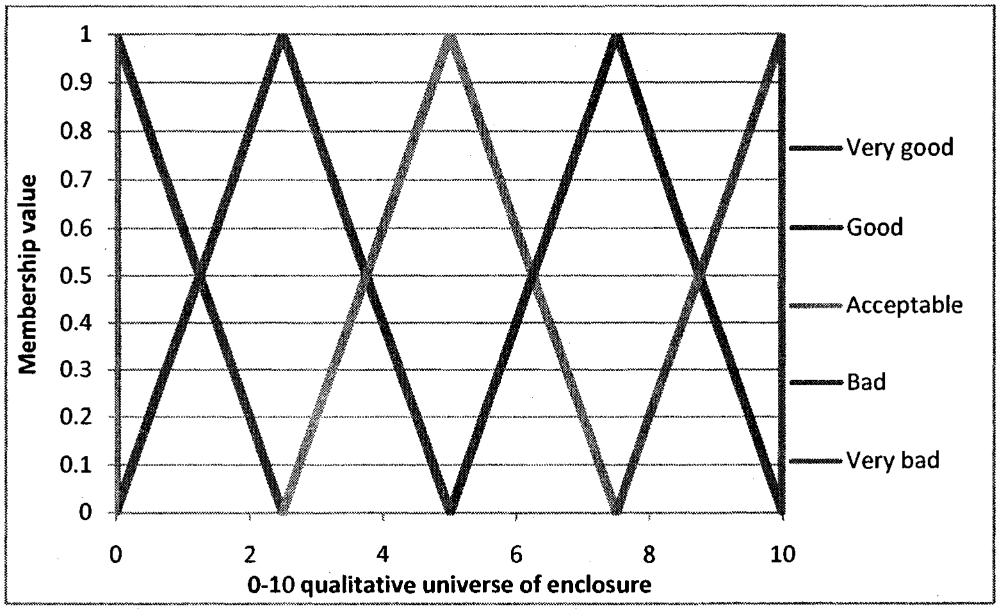


Figure V.17 – Water quality membership functions.

Leakage rate

Leakage in the pipelines indicates the presence of cracks and/or joint failure. This can give a strong indication about the status of the network. There are many tests that can assess the network leakage. Hydrostatic leakage tests and water audits are the most common methods used to detect leakage in the water system (Al Barqawi, 2006). Leakage erodes pipe bedding and increases soil moisture in the pipe zone (InfraGuide, 2003). In this research, due to lack of information about leakage rate evaluation and rating, it is considered as a subjective factor and classified into five groups: very high, high, medium, low, and very low. The membership functions and their characteristics are shown in Figure V.18. The data input to be used for this factor is numerical from 0 to 10, where 0 indicates the least leakage rate and 10 indicates the worst leakage rate.

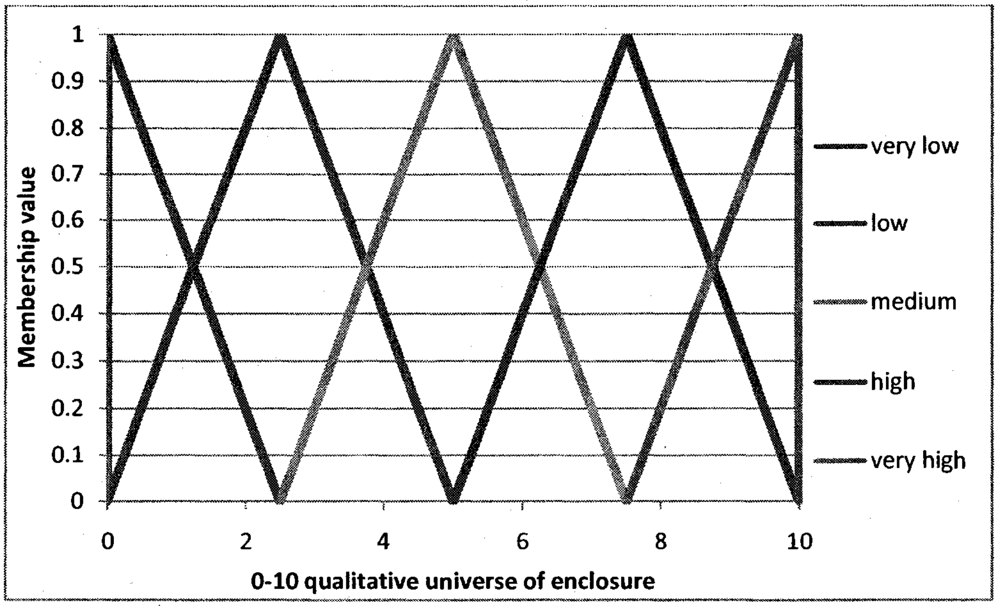


Figure V.18 – Leakage rate membership functions.

V.3.4. Post Failure Model

Estimating the consequences of pipeline failure is a complicated process. A simplification of the process is sought, therefore, and a qualitative (subjective) approach rather than a quantitative (objective) approach will be followed. Five factors are considered in this research: cost of repair, damage to surroundings/business disruption, loss of production, traffic reduction, and type of area serviced.

Cost of Repair

The cost of repair is the direct cost due to a burst pipeline. However, it is difficult to comprehend this factor since it varies depending on the magnitude of failure, the area of failure, time of failure, country of failure, etc (Bandyopadhyay *et al.* 1997). The main factors which contribute to the repair cost are the cost of the repair material and the cost of labor. The cost of repair material is dependent on the original pipeline material and its characteristics, for example: steel pipes can be repaired by welding a steel sleeve to the position of pipeline burst, but it is not applicable to plastic pipes which require another type of repair material. On the other hand, the cost of labor is dependent on the time consumed and the number of laborers involved in the repair. In its turn, the time consumed is dependent on the pipeline cover material and depth, the presence of other buried utilities such as electricity power lines, telephone lines, gas lines, etc, the location of the failure (accessibility) (Pickard, 2007). In this research, cost of repair is classified into five subjective groups (on a scale of 0 to 10) as very high, high, medium, low, and very low. The membership functions and their characteristics are shown in Figure V.19. The data type to be used for this factor is numerical from 0 to 10 where 0 indicates the lowest and 10 indicates the highest cost of repair.

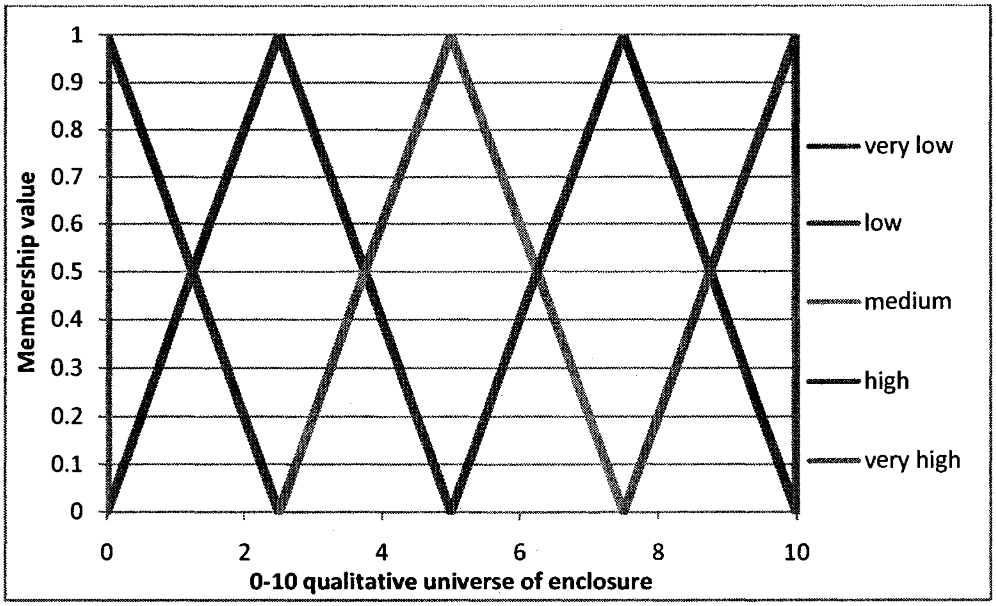


Figure V.19 – Cost of repair membership functions.

Damage to surroundings/Business Disruption

The most visible impact associated with a water main break is the occurrence of flooding affecting structures. Flooding causes quantifiable damage to structures and their contents which is dependent on the specific structure type, value, regional location and use. The cost associated with flooding is building structure damage and building content damages and even damage to property surrounding building such as gardens and sheds (Cromwell *et al.* 2002). In this research, the damage to surroundings and business disruption is classified into three groups according to the location of the pipeline failure as residential, commercial, and industrial. The membership functions and their characteristics are shown in Figure V.20. The data type to be used for this factor is linguistic and chosen from this list: Industrial, Commercial, and Residential.

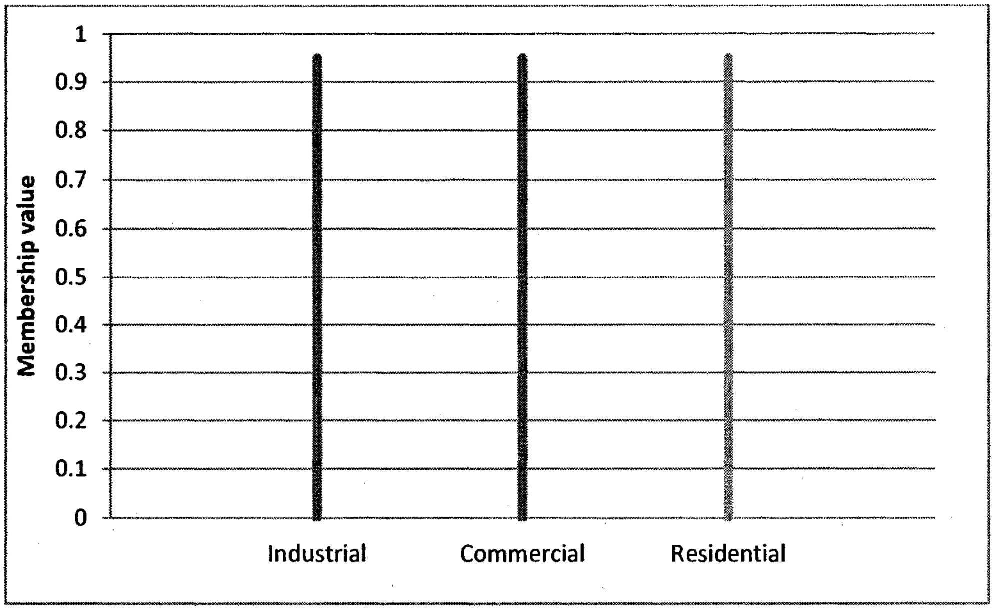


Figure V.20 – Damage to surroundings membership functions.

Loss of Production

Loss of production means the loss of profit from normal service. Loss of production is usually dependent on the size of the pipeline, the duration from time of failure to time of service resumption, and the location of the pipeline and whether it is redundant or not. Redundancy of the water network is achieved by duplicating elements in the network in order to eliminate the effects of any single point of failure. For this reason, the loss of production is classified here according to pipeline size and redundancy status as small redundant, small not-redundant, medium redundant, and medium not-redundant. The membership functions and their characteristics are shown in Figure V.21. The redundant and not-redundant pipelines share the same membership functions (overlaid each other). The data type to be used for this factor is pipeline diameter, up to 500 millimeters, and the redundancy condition to be chosen from Redundant or Not-Redundant.

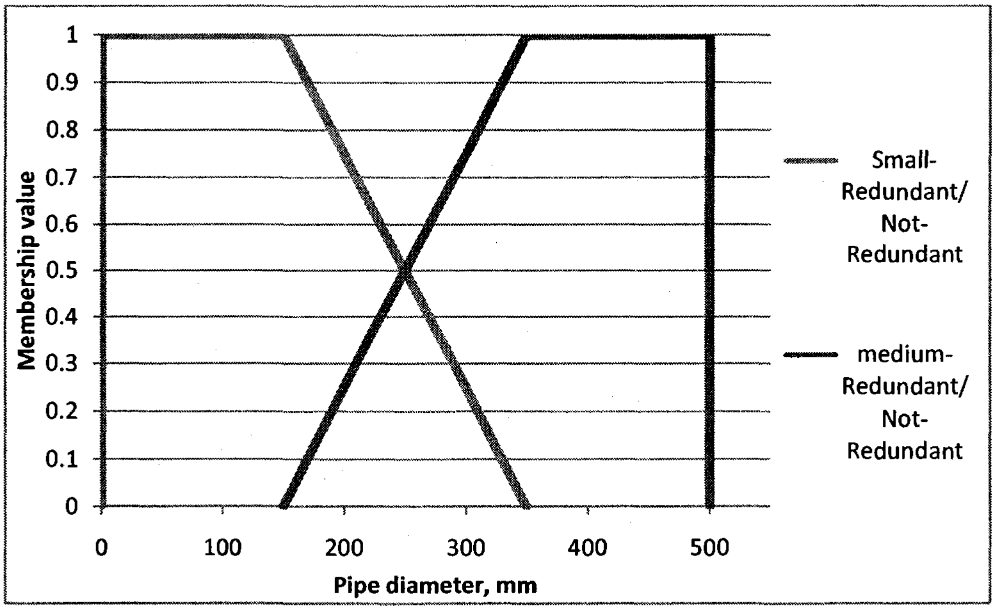


Figure V.21 – Loss of production membership functions.

Traffic Disruption

In the event of water main failure, mild to severe traffic disruption can occur. Traffic disruption causes inconveniences for the travelling public and can disrupt different businesses, in terms of customers and with freight and package delivery. The cost of traffic disruption is dependent on the increase of travel time and the value of travel time (Cromwell *et al.* 2002). However, the value of travel time is dependent on many factors by its turn. The increase of travel time is dependent on the type of road and traffic above the failed pipeline. In this research, a qualitative approach will be followed and traffic disruption as a cost will be classified as a subjective factor into 5 groups as: very disruptive, disruptive, moderate, light, and very light. The membership functions and their characteristics are shown in Figure V.22. The data type to be used for this factor is numerical from 0 to 10, where 0 indicates the least traffic disruption and 10 indicates the highest traffic disruption.

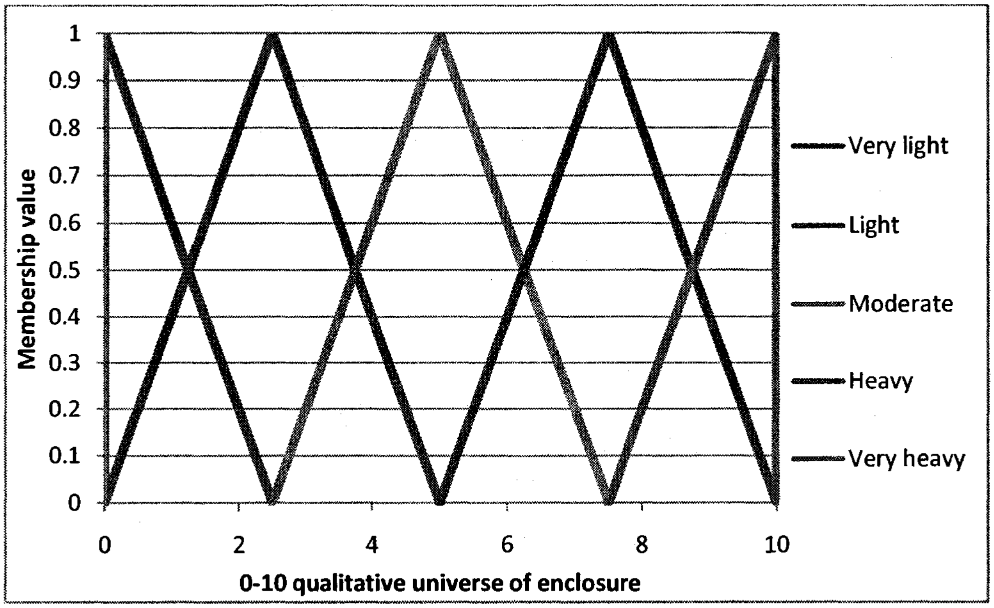


Figure V.22 – Traffic disruption membership functions.

Type of Serviced Area

In the case of pipeline failure, the water supply will stop serving a targeted destination. Hence, numerous businesses in the destination area will be negatively affected by the failure. The end users start to complain when they don't receive the service they need, which will damage the operator's reputation. Usually, water main networks are designed in a way to keep delivering water even if a failure occurs by using other paths (redundancy) (Oppenheimer, 2004). However, a drop in water pressure is also considered a failure. Depending on the area serviced, this factor can be classified into residential, commercial, and industrial (Al Barqawi, 2006). The membership functions and their characteristics are shown in Figure V.23. The data type to be used for is linguistic and chosen from this list: Industrial, Commercial, and Residential.

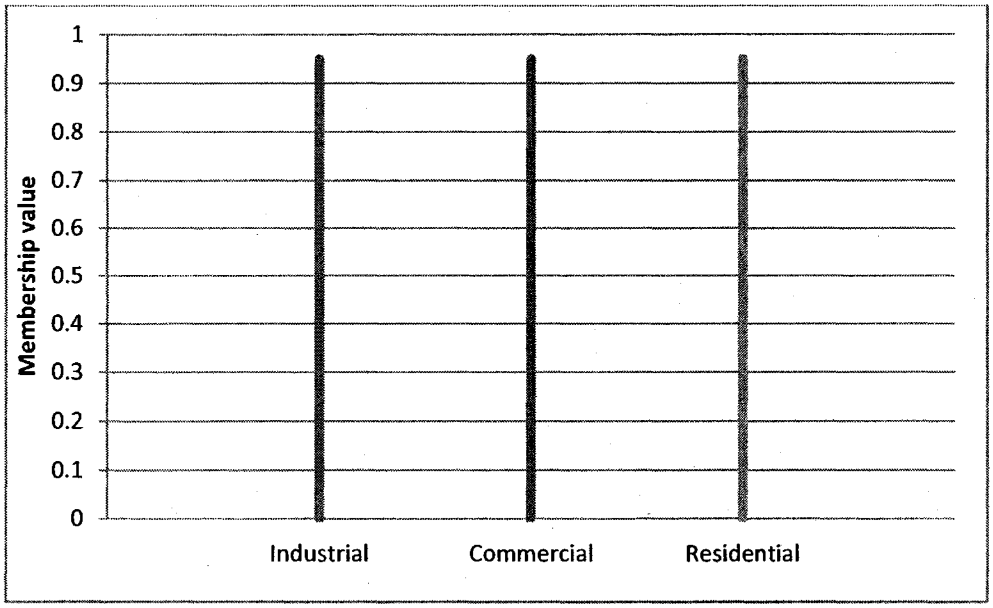


Figure V.23 – Type of serviced area membership functions.

V.3.5. Risk of failure Model

This model combines the results of the previous four models to generate the risk of failure of water main. Thus, the membership functions of the four main factors (environmental, physical, operational, and post failure) are identical to the standardized consequent membership functions of the four models (as will be explained in Section V.5. Consequents Aggregation). It consists of seven membership functions (extremely low, very low, moderately low, medium, moderately high, very high, and extremely high) on a qualitative scale from 0 to 10 as shown in Figure V.24.

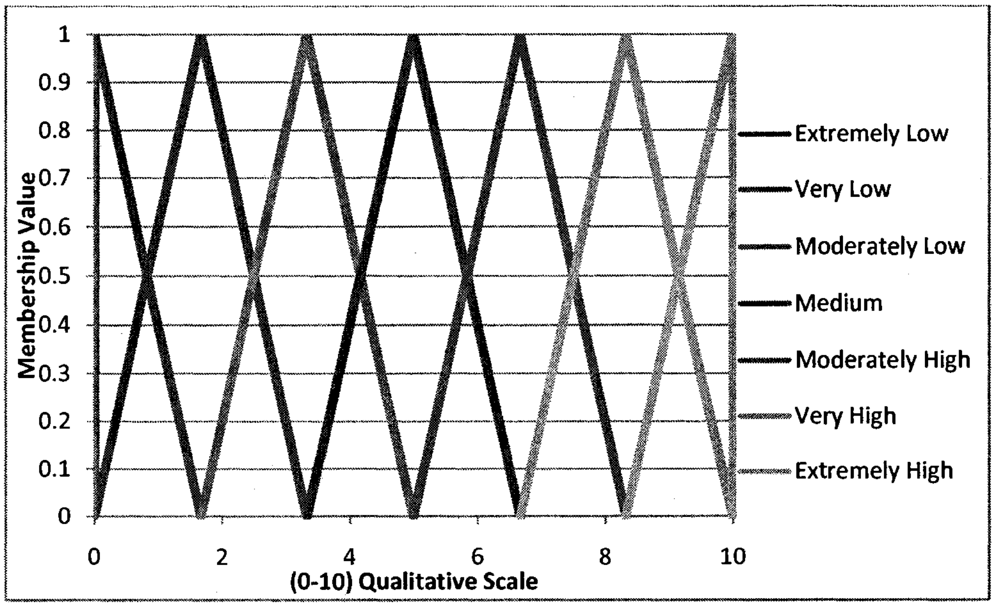


Figure V.24 – Fuzzification membership functions of the four main factors in the risk of failure model.

V.4. Fuzzy Inference

In this research, the indirect knowledge acquisition method (by means of a questionnaire and the available literature) is used to develop the knowledge base of the risk of water main failure model as shown in IV.1.3. Expert knowledge base.

The Mamdani fuzzy rules system type is used in the fuzzy model, which has an advantage over the Takagi-Sugeno-Kang (TSK) method of being easier to understand and the consequents of the system is defined in terms of fuzzy sets as explained in Section "A.6.1. Mamdani Method". The Mamdani method is based on a simple structure of Min operations as follows:

$$R^j{:}\,If\ x_1\ is\ A_1^j\ and\ x_2\ is\ A_2^j\ and\ x_3\ is\ A_3^j\ and\ \dots x_n\ is\ A_n^j\ THEN\ y\ is\ B^j$$

Where R^j is the j-th rule, A_i^j $(\mathrm{j} = 1, 2, \dots \mathrm{N}, \mathrm{i} = 1,2, \dots \mathrm{n}), B^j$ are the fuzzy subsets of the inputs and outputs respectively.

This rule can be written mathematically as Equation V.1:

$$\mu_{R^j}(x_1, x_2, x_3, \dots, x_n, y) = \mu_{A_1^j} \wedge \mu_{A_2^j} \wedge \mu_{A_3^j} \dots \wedge \mu_{A_n^j} \wedge \mu_B \qquad \text{Equation V.1}$$

Where $\wedge$ denotes the minimum operator.

In this research, the consequent linguistic variable B is standardized on a list of seven linguistic variables (Extremely low, Very low, Moderately Low, Medium, Moderately High, Very High, and Extremely High as shown in Figure V.25). This is applicable to each model of the five models (environmental, physical, operational, post failure, and risk of failure).

V.5. Consequents Aggregation

After evaluating each rule in the knowledge base, the membership value of each consequent membership function (output linguistic variable) is aggregated using a maximum operation as shown in Equation V.2. In other words, the maximum membership value of any consequent membership function (shown in Figure V.25) is used to truncate that consequent membership function for later use in the defuzzification of the fuzzy output.

$$\mu_R(x_1, x_2, x_3, \dots, x_n, y) = \bigvee_{j=1}^{N} [\mu_{Rj}(x_1, x_2, x_3, \dots, x_n, y)] \qquad \text{Equation V.2}$$

Where $\vee$ denotes the maximum operation, R represents each of the consequent membership functions as standardized to the list of (Extremely low, Very low, Moderately Low, Medium, Moderately High, Very High, and Extremely High).

This is also applicable to each model of the five models (environmental, physical, operational, post failure, and risk of failure).

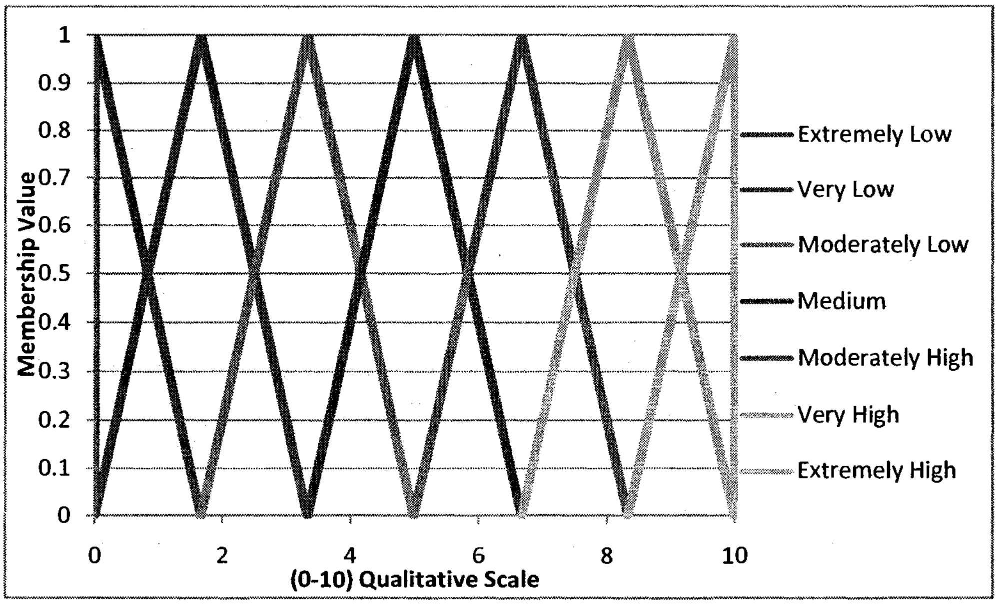


Figure V.25 – Consequent membership functions.

V.6. Defuzzification Process

There are many defuzzification methods that convert the fuzzy consequents of all of the triggered fuzzy rules to a crisp value. The method used in this research is the Center of Sum as shown in Equation V.3.

$$Crisp\ Risk\ Output =$$

$$= \sum_{n=extremely\ low}^{extremely\ high} Truncated\ Area_n \times Centeriod_n \Bigg/ \sum_{n=extremely\ low}^{extremely\ high} Truncated\ Area_n$$

Equation V.3

This equation calculates the center of gravity of each truncated consequent membership function found from the previous step (neglecting the union operation) and then average-weights them by their areas. It has the advantage of being simple to program, requiring less computer resources, and it gives reasonable results. Section A.7. Defuzzification Methods in Appendix A contains more information.

This is also applicable to each model of the five models (environmental, physical, operational, post failure, and risk of failure).

V.7. System Analysis and Verification

Two different approaches are used to test and verify the developed model and system. The first approach is system sensitivity analysis (stability testing), which tests the effects of the different factors on the behaviour of the model. The second approach is model accuracy testing which uses a validated AHP deterioration model to verify the results of the developed model.

V.7.1. Sensitivity Analysis and System Stability Testing

The sensitivity and stability of the model have to be tested in order to insure that the model is performing as expected under different model parameters. Therefore, several scenarios are assumed and applied to the model and the results are examined for any illogicality. The scenarios are as follows:

1. Lowest and highest risk of failure (ID numbers 1 and 2 in Table V.1). By analyzing the results, the maximum theoretical risk of failure that the model can generate is 9.4 and the minimum is 0.6. This is due the fact that the method used in defuzzification is Center of Sum, which calculates the center of the area under the triangular membership functions and thus limits the risk index to a maximum of 9.4 and a minimum of 0.6. However, the actual maximum risk of failure that can be generated by the model is 8.8 and the minimum is 1.7, due to the behavior of certain physical factors since there is no Extremely High Risk output membership functions used in the model for the Pipe Diameter and Protection Method factors, due to the performance conflicts of different factors (e.g. cathodic protection can not accompany PE or PVC pipes, pipe diameter performance conflicts with loss of production performance).
2. Sensitivity analysis by increasing the risk of failure values. This is done by increasing the adverse effect (riskier performance) of the factors (one at a time), starting from the factor that has the highest weight among the considered factors and ending with the factor that has the lowest weight (ID number 3 to 112 in Table V.1 and Appendix C). By analyzing this scenario, it is noticeable that the risk of failure index is changing at a quicker pace in the early stages (for the first

factors) and at a slower pace in the later stages (in the last factors) as shown in Figure V.26. This is due to the fact that the factors with the highest weight are examined first and the factors with the lowest weights are examined at the end, where they don't have enough power to make a noticeable change. Figure V.26 shows jump steps in the results. This occurs for several reasons, such as the change in the factor status alone is not enough to make a change in the final result. Other reasons are the rules evaluation and aggregation, and the use of the center of sum defuzzification method.

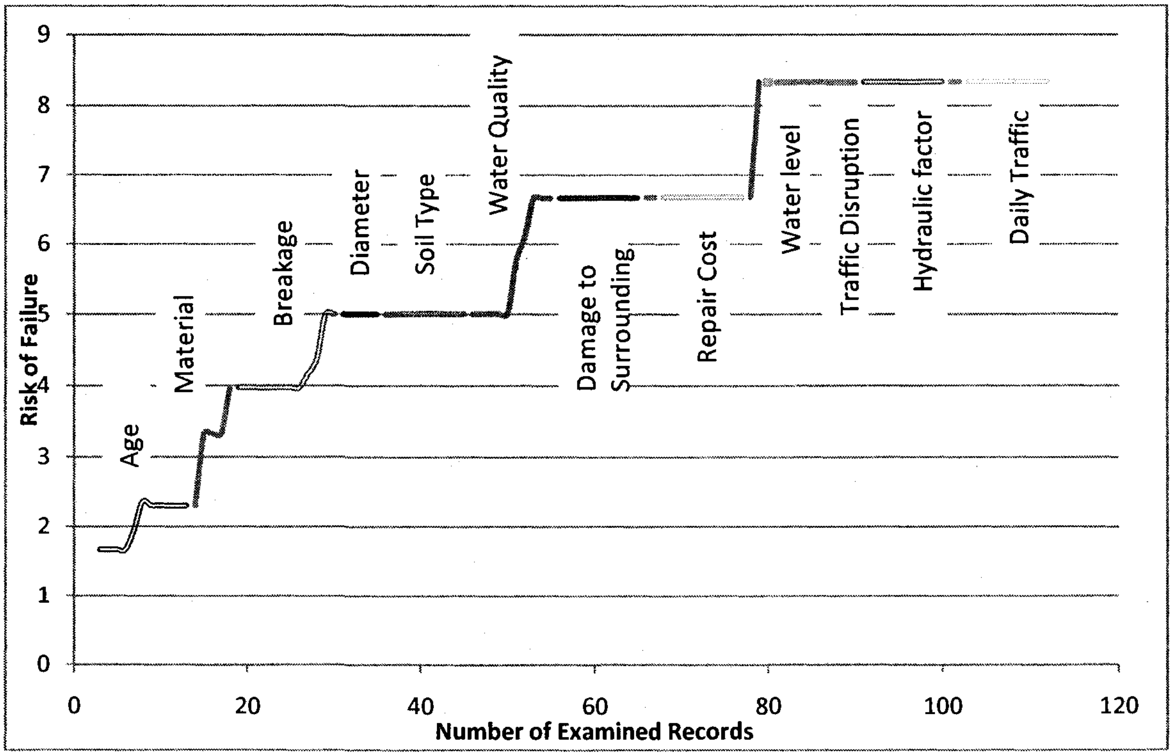


Figure V.26 – Sensitivity analysis of the model.

Table V.1 – Risk of failure model sensitivity analysis sample results.

	Notes	Environmental			Physical				Operational				Consequences						Results				
ID number	Scenarios	Type of Soil	Average Daily Traffic	Water Table Level	Type of Pipe	Pipe Diameter	Installation Year	Protection Method	Number of Break	Hydraulic Factor	Water Quality	Leakage	Cost of Repair	Damage to Surroundings	Loss of Production	Traffic Disruption	Serviced Area	Redundancy	Environmental	Physical	Operational	Consequence	Risk of Failure
1	most risky	10	10	seasonally present	Cast iron	100	1900	none	6	30	10	10	10	Industrial	100	10	Industrial	Not Redundent	9.4	8.3	9.4	8.3	8.3
2	least risky	0	0	rarely present	PE	500	2007	none	0	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	1.7	0.6	1.7	1.7
3	age 1	0	0	rarely present	PE	500	2007	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	1.7	0.6	1.7	1.7
4	age 2	0	0	rarely present	PE	500	1997	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	1.7	0.6	1.7	1.7
5	age 3	0	0	rarely present	PE	500	1987	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	1.8	0.6	1.7	1.7
6	age 4	0	0	rarely present	PE	500	1977	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	2.6	0.6	1.7	1.7
7	age 5	0	0	rarely present	PE	500	1967	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	3.5	0.6	1.7	1.9
8	age 6	0	0	rarely present	PE	500	1957	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	4.2	0.6	1.7	2.4
9	age 7	0	0	rarely present	PE	500	1947	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	5.0	0.6	1.7	2.3
10	age 8	0	0	rarely present	PE	500	1937	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	5.0	0.6	1.7	2.3
11	age 9	0	0	rarely present	PE	500	1927	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	5.0	0.6	1.7	2.3
12	age 10	0	0	rarely present	PE	500	1917	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	5.0	0.6	1.7	2.3
13	age 11	0	0	rarely present	PE	500	1907	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	5.0	0.6	1.7	2.3
14	type of material 2	0	0	rarely present	PVC	500	1907	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	5.0	0.6	1.7	2.3
15	type of material 3	0	0	rarely present	Concrete	500	1907	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	6.7	0.6	1.7	3.3
16	type of material 4	0	0	rarely present	Asbestos	500	1907	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	6.7	0.6	1.7	3.3
17	type of material 5	0	0	rarely present	Cast iron	500	1907	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	6.7	0.6	1.7	3.3
18	type of material 6	0	0	rarely present	t iron post	500	1907	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	0.6	1.7	4.0
19	breakage 2	0	0	rarely present	t iron post	500	1907	none	0.25	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	1.1	1.7	4.0
20	breakage 3	0	0	rarely present	t iron post	500	1907	none	0.50	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	1.3	1.7	4.0
21	breakage 4	0	0	rarely present	t iron post	500	1907	none	0.75	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	1.5	1.7	4.0
22	breakage 5	0	0	rarely present	t iron post	500	1907	none	1.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	1.7	1.7	4.0
23	breakage 6	0	0	rarely present	t iron post	500	1907	none	2.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	1.7	1.7	4.0
24	breakage 7	0	0	rarely present	t iron post	500	1907	none	2.25	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	1.7	1.7	4.0
25	breakage 8	0	0	rarely present	t iron post	500	1907	none	2.50	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	1.7	1.7	4.0
26	breakage 9	0	0	rarely present	t iron post	500	1907	none	2.75	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	2.2	1.7	4.0
27	breakage 10	0	0	rarely present	t iron post	500	1907	none	3.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	2.5	1.7	4.2
28	breakage 11	0	0	rarely present	t iron post	500	1907	none	3.25	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	2.8	1.7	4.4
29	breakage 12	0	0	rarely present	t iron post	500	1907	none	3.50	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	3.3	1.7	5.0
30	breakage 13	0	0	rarely present	t iron post	500	1907	none	4.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	3.3	1.7	5.0
31	pipe diameter 2	0	0	rarely present	t iron post	350	1907	none	4.00	130	0	0	0	Residential	350	0	Residential	Redundent	0.6	8.3	3.3	1.7	5.0
32	pipe diameter 3	0	0	rarely present	t iron post	300	1907	none	4.00	130	0	0	0	Residential	300	0	Residential	Redundent	0.6	8.3	3.3	1.7	5.0
33	pipe diameter 4	0	0	rarely present	t iron post	250	1907	none	4.00	130	0	0	0	Residential	250	0	Residential	Redundent	0.6	8.3	3.3	1.7	5.0
34	pipe diameter 5	0	0	rarely present	t iron post	200	1907	none	4.00	130	0	0	0	Residential	200	0	Residential	Redundent	0.6	8.3	3.3	1.7	5.0
35	pipe diameter 6	0	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	0.6	8.3	3.3	1.7	5.0
36	type of soil 2	1	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	1.3	8.3	3.3	1.7	5.0
37	type of soil 3	2	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	1.5	8.3	3.3	1.7	5.0
38	type of soil 4	3	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	2.1	8.3	3.3	1.7	5.0
39	type of soil 5	4	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	2.6	8.3	3.3	1.7	5.0
40	type of soil 6	5	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	3.3	8.3	3.3	1.7	5.0
41	type of soil 7	6	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	4.1	8.3	3.3	1.7	5.0
42	type of soil 8	7	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	4.5	8.3	3.3	1.7	5.0
43	type of soil 9	8	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	5.0	8.3	3.3	1.7	5.0
44	type of soil 10	9	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	5.0	8.3	3.3	1.7	5.0
45	type of soil 11	10	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	5.0	8.3	3.3	1.7	5.0

3. Close testing of the factor that has the highest weight in the physical model. This is the Age factor at an increment of 10 years, as shown in Table V.1. Analyzing the results, which tests the Age factor (fixing the rest of the physical factors at medium consequents) and draws a curve for the physical risk values (Figure V.27), it shows a steady stage at the start (10 to 30 years) and at the end (50 to 80 years) and a smooth increasing stage in the middle of the curve (30 to 50 years). This is because the model maps five input membership functions in this factor to seven output membership functions which will cause unevenly distributed results (two close membership functions at the beginning and two at the end of the curve and one membership function in the middle of the curve). However, by analyzing the effect of age factor on the risk of failure values, one can observe that it is difficult to make a change in the risk of failure value by only changing the performance of one factor, since the other fifteen factors try to resist the change in the risk values, as shown in Figure V.26. Testing the Age factor when the other factors are absent will give an indication of how the system performs when there is no resistance from other factors. Figure V.28 shows a graph of the physical index of the age factor alone. The stable part at the very end of the graph is due to the shape of the fuzzy membership functions of the age factor (a trapezoidal membership function).

Figure V.27 – Sensitivity analysis of Age factor on physical risk assuming other physical factors are present.

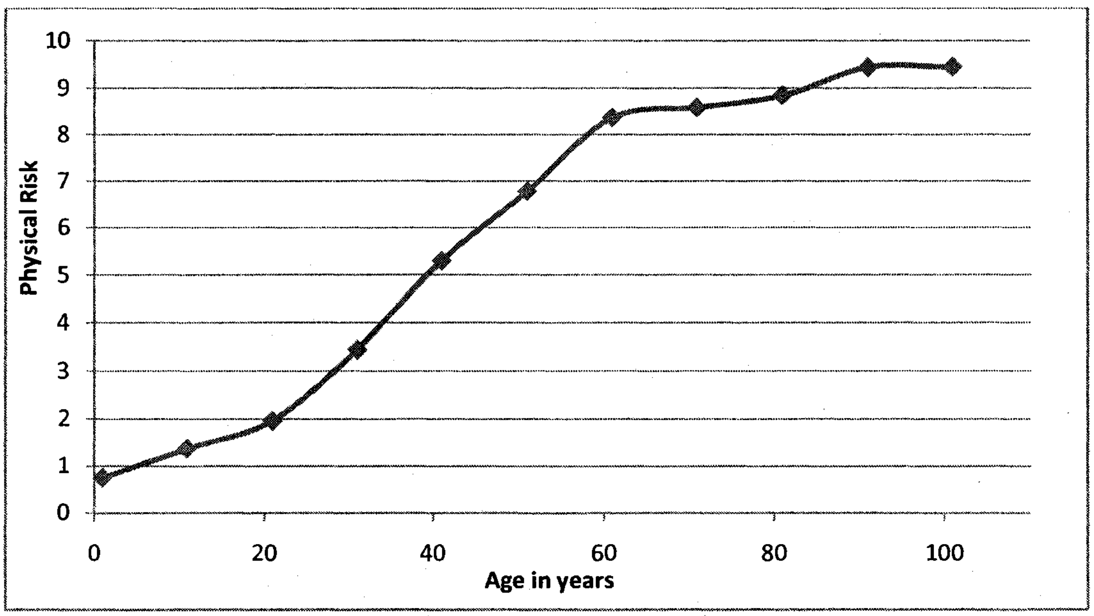


Figure V.28 – Sensitivity analysis of Age factor on physical risk assuming other physical factors are absent.

4. Close testing the physical model (chosen here because the two factors under study are parts of this model) and risk of failure model sensitivity towards change in the consequents of the two highest weighted factors (Age and Pipe material). This is done through three scenarios by fixing the values of the other physical factors at the medium consequents (all other physical factors are present), at the highest risk consequents, and at the lowest risk consequents, and at the same time changing the values of the factors under study from lowest to highest risks consequents (Table V.2). Analyzing the results of physical model sensitivity towards the age and pipe material factors and analyzing Figure V.29, Figure V.31, and Figure V.32, it is obvious that the physical model is more sensitive toward the age factor (which has the highest weight) since it causes a change in the physical risk at medium risk level from 3.5 to 6.7 with a steeper curve (compared to the other factor, pipe material). However, the pipe material factor causes a change in the physical risk at medium risk level from 4.2 to 6.0 with a smoother curve. Figure V.30 shows that due to the very close weights of the age and pipe material factors (30 and 40), the risk of failure model has the same sensitivity towards these two factors where they both cause a change in failure risk value from 5.0 to 5.8. As a result, each of the four models (environmental, physical, operational, and post failure) is more sensitive to its own factors than the risk of failure model is. This fact is due to the use of a hierarchical system where the farther the factor is in the hierarchy, the less its effectiveness (sensitivity) to the top level model is.

Table V.2 – Physical and risk of failure models sensitivity analysis of age and pipe material factors.

ID number	Scenarios	Factors under study	Type of Pipe	Pipe Diameter	Installation Year	Protection Method	Physical Risk	Risk of Failure
1	Medium Risk	age 1	Concrete	250	2007	none	3.5	5.0
2		age 2	Concrete	250	1982	none	4.2	5.0
3		age 3	Concrete	250	1957	none	5.8	5.4
4		age 4	Concrete	250	1932	none	6.0	5.8
5		age 5	Concrete	250	1907	none	6.7	5.8
6		type of material 1	PE	250	1967	none	4.2	5.0
7		type of material 2	PVC	250	1967	none	4.3	5.0
8		type of material 3	Concrete	250	1967	none	5.2	5.4
9		type of material 4	Asbestos	250	1967	none	5.8	5.8
10		type of material 5	Cast iron post war	250	1967	none	6.0	5.8
11	High Risk	age 1	Cast iron post war	50	2007	none	5.0	7.7
12		age 2	Cast iron post war	50	1982	none	5.6	7.7
13		age 3	Cast iron post war	50	1957	none	7.6	8.3
14		age 4	Cast iron post war	50	1932	none	8.3	8.3
15		age 5	Cast iron post war	50	1907	none	8.3	8.3
16		type of material 1	PE	50	1900	none	6.7	8.3
17		type of material 2	PVC	50	1900	none	6.7	8.3
18		type of material 3	Concrete	50	1900	none	6.7	8.3
19		type of material 4	Asbestos	50	1900	none	8.3	8.3
20		type of material 5	Cast iron post war	50	1900	none	8.3	8.3
21	Low Risk	age 1	PE	500	2007	Lining\Coating	0.7	1.3
22		age 2	PE	500	1982	Lining\Coating	1.7	1.7
23		age 3	PE	500	1957	Lining\Coating	2.6	1.7
24		age 4	PE	500	1932	Lining\Coating	3.3	1.7
25		age 5	PE	500	1907	Lining\Coating	3.3	1.7
26		type of material 1	PE	500	2007	Lining\Coating	0.7	1.3
27		type of material 2	PVC	500	2007	Lining\Coating	1.7	1.7
28		type of material 3	Concrete	500	2007	Lining\Coating	1.7	1.7
29		type of material 4	Asbestos	500	2007	Lining\Coating	1.8	1.7
30		type of material 5	Cast iron post war	500	2007	Lining\Coating	3.3	1.7

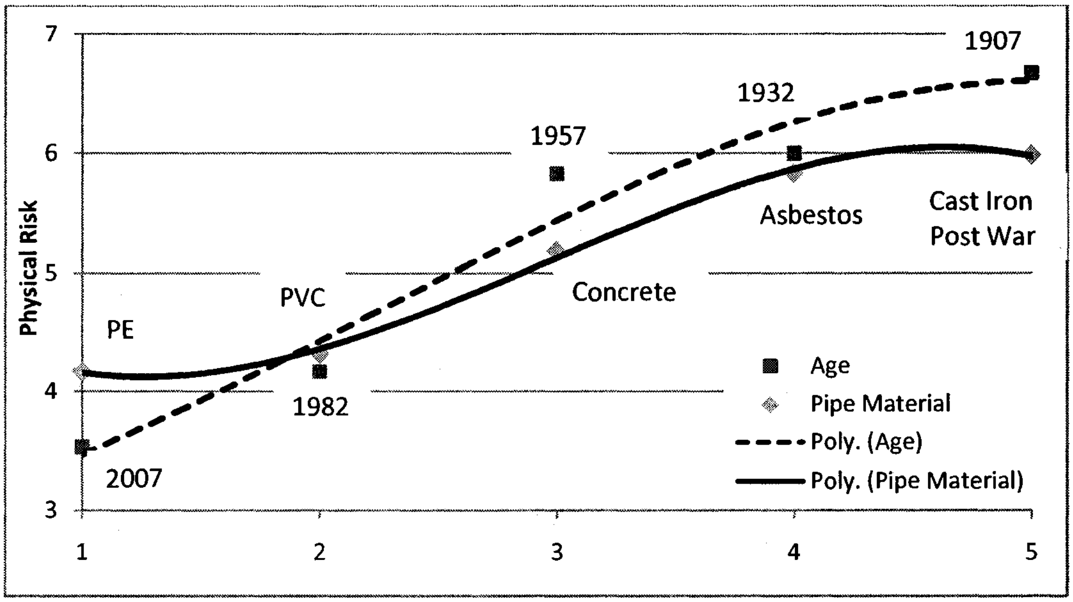


Figure V.29 – Physical model sensitivity analysis of age and pipe material factors at medium risk level.

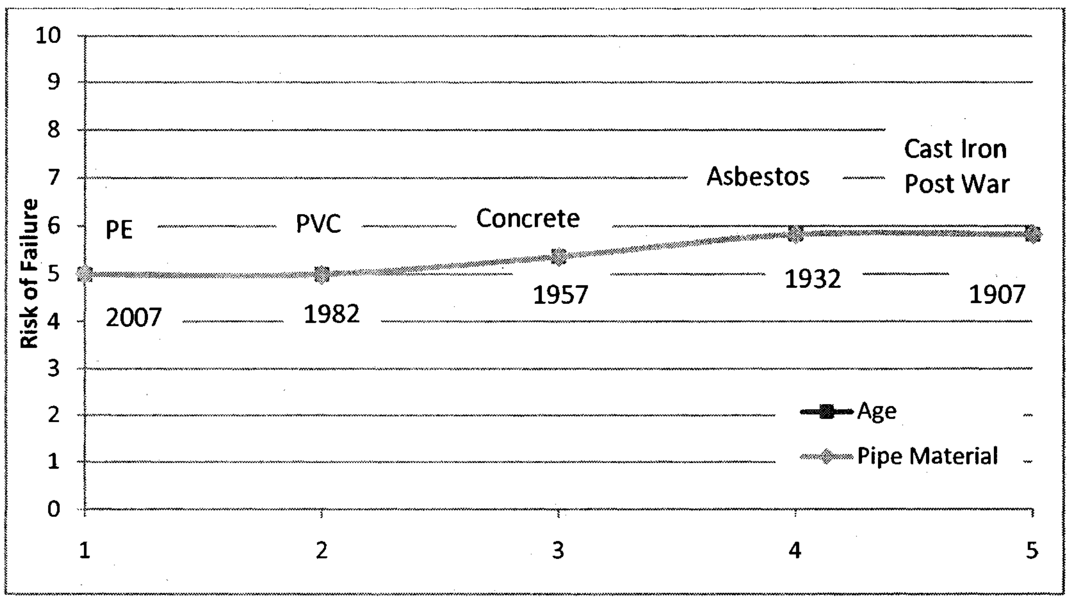


Figure V.30 – Risk of failure model sensitivity analysis of age and pipe material factors at medium risk level.

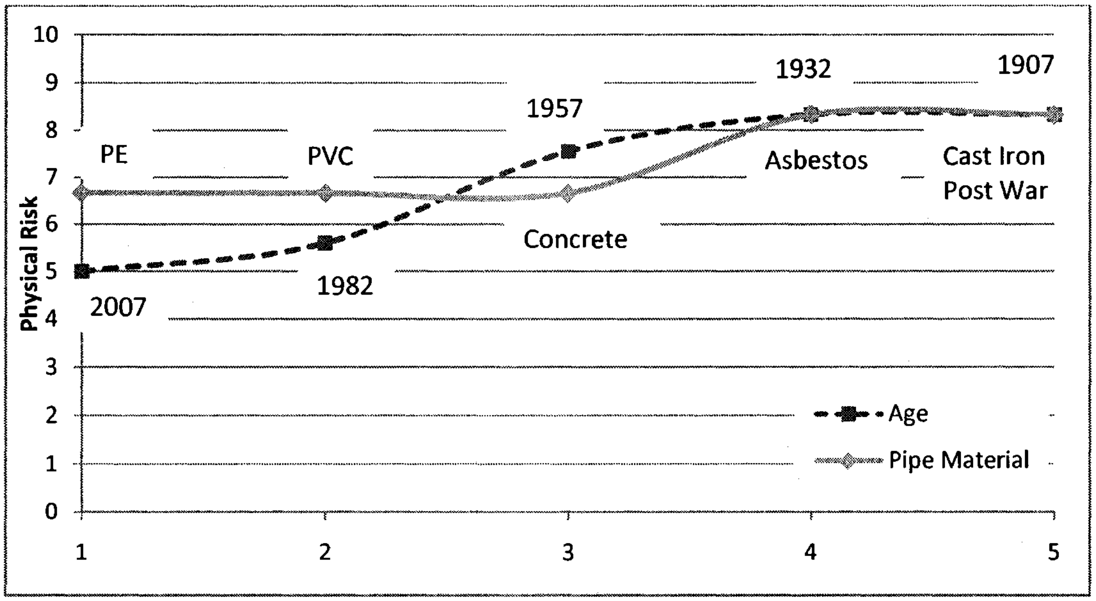


Figure V.31 – Physical model sensitivity analysis of age and pipe material factors at high risk level.

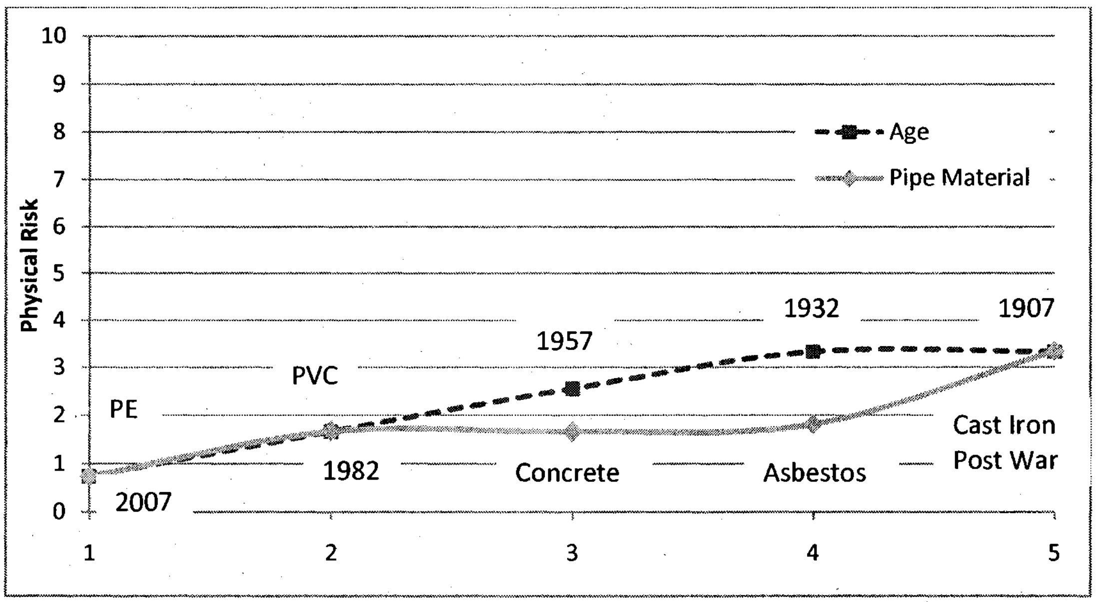


Figure V.32 – Physical model sensitivity analysis of age and pipe material factors at low risk level.

5. Close testing of the sensitivity of the physical and risk of failure models toward the weights of the two most-weighted factors (age and pipe material) within two times their standard deviation. This will examine the effect of the standard

deviation and the change of the mean values. The mean weight of the age factor is 40 with a standard deviation of 16, whereas the mean weight of pipe material is 30 with a standard deviation of 12. This test is performed by fixing the values of the other factors at their least risky effect and fixing the value of the factor under study at its most risky effect in order to get the most sensitive case. After that, the test is conducted by changing the value of the weight within the associated standard deviation (from different expert opinions as shown in Table IV.1) at steps of five units and then calculating the physical and risk of failure values as shown in Table V.3. Analyzing the results of this scenario as shown in Table V.3, Figure V.33, and Figure V.34, it can be deduced that the physical model is not very sensitive to the weight of the examined factors within the factors' standard deviations. Changing the weight of the age factor will cause a change in the physical risk from 3.3 to 5.0 and the change in pipe material weight will also make a change from 3.3 to 5.0. The change of the value of the weight from $(\mu - \sigma)$ to $(\mu + \sigma)$ will result in only 1.7 units of difference which is not a large difference. This is because of the presence of other factors that act as resistances to the change in the risk values. Figure V.33 shows that the physical model starts to be sensitive toward the change in age factor weight when the weight is reduced below 30. However, the physical model shows more sensitivity toward the pipe material weight, as shown in Figure V.34, which can be attributed to the rules evaluation and aggregation, and to the defuzzification process of the fuzzy output which results in steps in the output, as shown above in the second sensitivity scenario (Figure V.26). In addition, the risk of failure model shows low sensitivity to the change of the weight of the age and pipe material as the risk of failure changes from 1.7 to 3.3 in the case of age factor and 1.7 to 2.3 for pipe material factor, as shown in Table V.3 and Figure V.35.

Table V.3 – Sensitivity analysis of physical and risk of failure models.

ID number	Scenarios	Physical Risk	Risk of Failure
1	Age weight = 8	3.3	1.7
2	Age weight = 15	3.3	1.7
3	Age weight = 20	3.3	1.7
4	Age weight = 25	3.3	1.7
5	Age weight = 30	5.0	2.3
6	Age weight = 35	5.0	2.3
7	Age weight = 40	5.0	2.3
8	Age weight = 45	5.0	2.3
9	Age weight = 50	5.0	2.3
10	Age weight = 55	5.0	2.3
11	Age weight = 60	5.0	2.3
12	Age weight = 65	5.0	2.3
13	Age weight = 72	6.7	3.3
14	Pipe material weight = 6	1.8	1.7
15	Pipe material weight = 10	3.3	1.7
16	Pipe material weight = 15	3.3	1.7
17	Pipe material weight = 20	3.3	1.7
18	Pipe material weight = 25	3.5	1.9
19	Pipe material weight = 30	3.5	1.9
20	Pipe material weight = 35	5.0	2.3
21	Pipe material weight = 40	5.0	2.3
22	Pipe material weight = 45	5.0	2.3
23	Pipe material weight = 50	5.0	2.3
24	Pipe material weight = 54	5.0	2.3

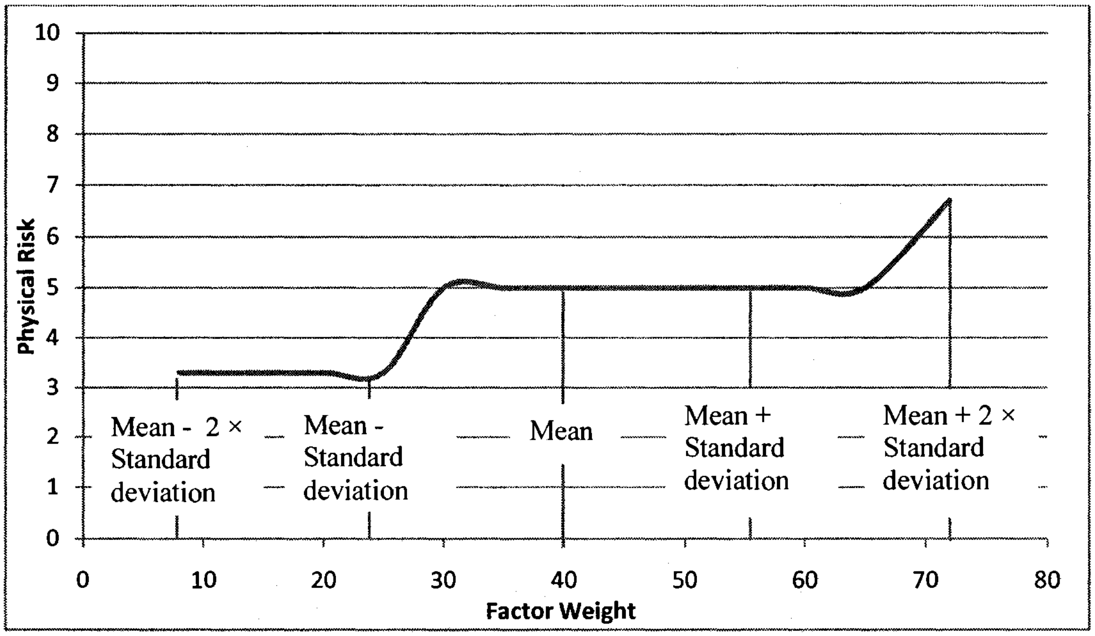


Figure V.33 – Physical model sensitivity analysis toward age factor weight.

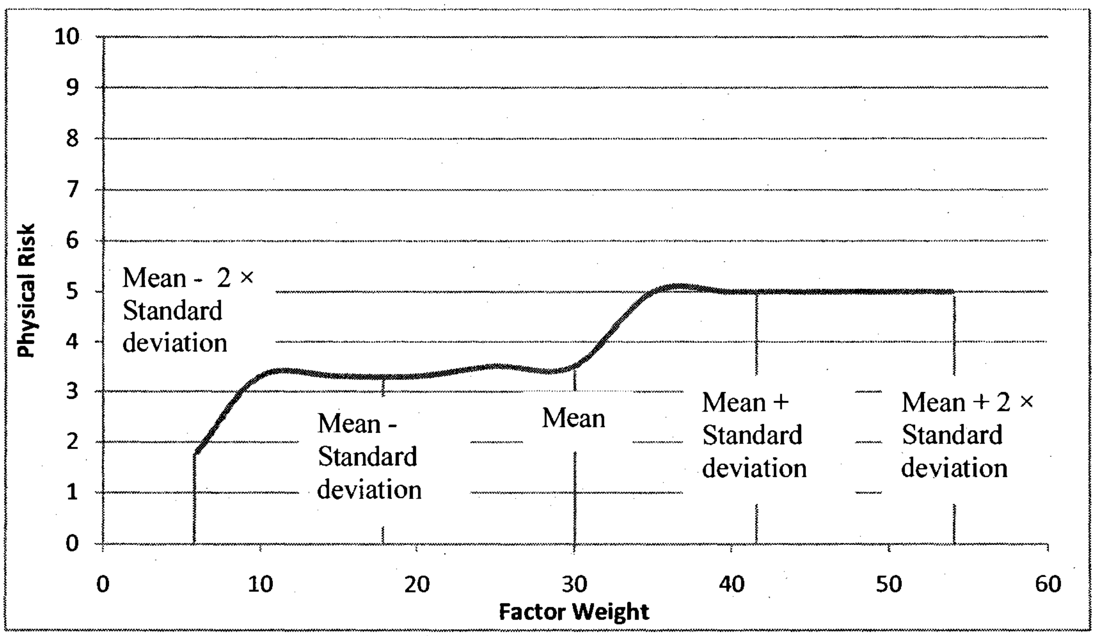


Figure V.34 – Physical model sensitivity analysis toward pipe material weight.

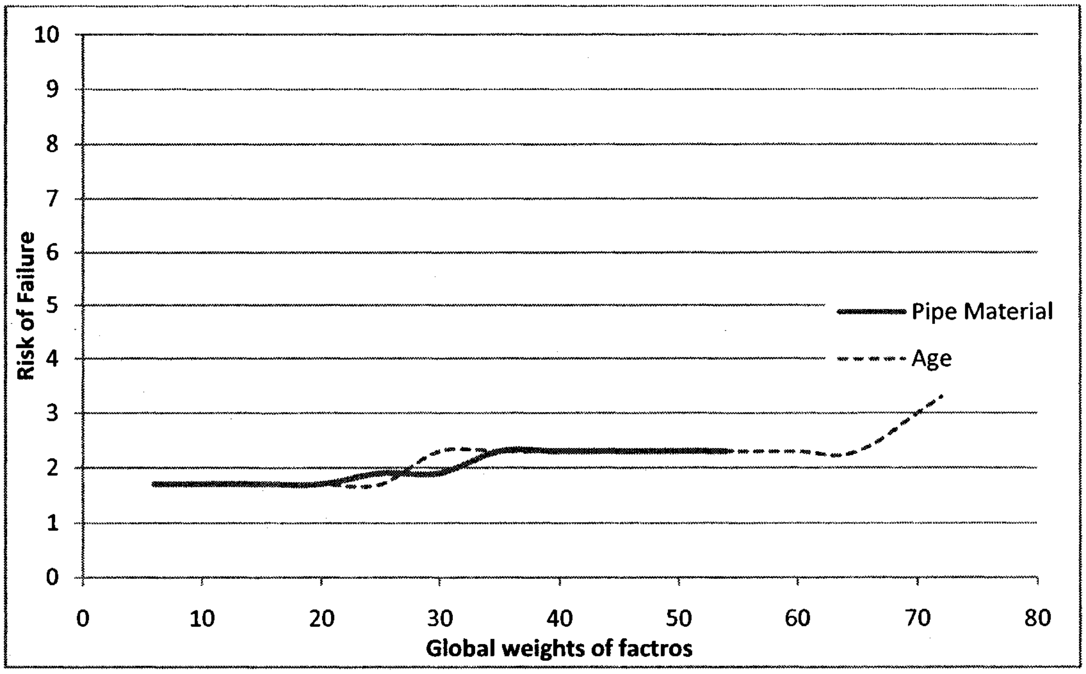


Figure V.35 – Risk of failure model sensitivity analysis toward age and pipe material weights.

6. Close testing of the sensitivity of the physical model toward the assumed value of pipe material certainty 95%. Certainty here refers to the level of confidence in the consequent part of the pipe material factor as an individual factor, as shown in Section V.3.2. Physical . This scenario is carried out by changing the certainty level from 50% to 100%. The data assumed in assessing this scenario and the differences between the assumed certainty level of 95% and varying certainty levels at 50%, 75%, and 100% are evaluated as shown in Table V.4. By analyzing these differences, it can be noticed that these values are very small and negligible. The reason behind these small differences can be attributed to many facts related to the use of the certainty level in the model. For instance, the certainty level (membership value) of the pipe material factor is not important unless it is the lowest membership value in the triggered rule compared to the

other membership valued collected from other physical factors in that rule, as minimum value is used among them (an AND operation that combines the physical factor in the knowledge rule). Moreover, the rules evaluation and aggregation and defuzzification process will reduce even more the effect of the change in certainty level of the pipe material. The maximum difference recorded is 0.0439 which is so small that it can be neglected. As a conclusion, the physical model (and other models in general) is not sensitive to the certainty level value of the pipe material, which is the second most important factor among the sixteen factors.

As a result of the analysis conducted here before, the model(s) is sound, stable, and performs as expected without any irregular, illogic results.

V.7.2. Verification of the Developed Model

In order to verify the developed model, different approaches can be used. Experts' testing and feedback is one approach. Another approach is to compare the model results to the results of another related and validated model (Shaheen, 2005). In this context, the second approach will be used, which is comparing the results of the proposed model with another model output. The most relevant model to compare to the proposed model is the AHP model developed by Al Barqawi (2006). However, some points should be kept in mind when examining the results:

Table V.4 – Physical model sensitivity toward the certainty level of the pipe material factor.

		Physical factors				Physical Module Risk at different confidence level						
ID number	Scenarios	Type of Pipe	Pipe Diameter	Installation Year	Protection Method	0.95	0.7	Difference 0.95 - 0.7	0.5	Difference 0.95 - 0.5	1	Difference 0.95 - 1.0
1	age 1	PE	500	2007	none	1.6667	1.6664	**0.0002**	1.6661	**0.0005**	1.6667	**0.0000**
2	age 2	PE	500	1997	none	1.6662	1.6662	**0.0000**	1.6661	**0.0001**	1.6662	**0.0000**
3	age 3	PE	500	1987	none	1.8152	1.8279	**-0.0127**	1.8581	**-0.0429**	1.8152	**0.0000**
4	age 4	PE	500	1977	none	2.5557	2.5557	**0.0000**	2.5302	**0.0255**	2.5557	**0.0000**
5	age 5	PE	500	1967	none	3.4821	3.4953	**-0.0131**	3.5261	**-0.0439**	3.4821	**0.0000**
6	age 6	PE	500	1957	none	4.2240	4.2240	**0.0000**	4.1985	**0.0255**	4.2240	**0.0000**
7	age 7	PE	500	1947	none	5.0000	5.0000	**0.0000**	5.0000	**0.0000**	5.0000	**0.0000**
8	age 8	PE	500	1937	none	5.0000	5.0000	**0.0000**	5.0000	**0.0000**	5.0000	**0.0000**
9	age 9	PE	500	1927	none	5.0000	5.0000	**0.0000**	5.0000	**0.0000**	5.0000	**0.0000**
10	age 10	PE	500	1917	none	5.0000	5.0000	**0.0000**	5.0000	**0.0000**	5.0000	**0.0000**
11	age 11	PE	500	1907	none	5.0000	5.0000	**0.0000**	5.0000	**0.0000**	5.0000	**0.0000**
12	type of material 1	PVC	500	1907	none	5.0000	5.0000	**0.0000**	5.0000	**0.0000**	5.0000	**0.0000**
13	type of material 2	Concrete	500	1907	none	6.6667	6.6664	**0.0002**	6.6661	**0.0005**	6.6667	**0.0000**
14	type of material 3	Asbestos	500	1907	none	6.6667	6.6664	**0.0002**	6.6661	**0.0005**	6.6667	**0.0000**
15	type of material 4	Cast iron	500	1907	none	6.6667	6.6664	**0.0002**	6.6661	**0.0005**	6.6667	**0.0000**
16	type of material 5	Cast iron post war	500	1907	none	8.3333	8.3336	**-0.0002**	8.3339	**-0.0005**	8.3333	**0.0000**
29	pipe diameter 1	Cast iron post war	350	1907	none	8.3333	8.3336	**-0.0002**	8.3339	**-0.0005**	8.3333	**0.0000**
30	pipe diameter 2	Cast iron post war	300	1907	none	8.3335	8.3336	**-0.0001**	8.3339	**-0.0004**	8.3335	**0.0000**
31	pipe diameter 3	Cast iron post war	250	1907	none	8.3339	8.3339	**0.0000**	8.3339	**0.0000**	8.3339	**0.0000**
32	pipe diameter 4	Cast iron post war	200	1907	none	8.3335	8.3336	**-0.0001**	8.3339	**-0.0004**	8.3335	**0.0000**
33	pipe diameter 5	Cast iron post war	150	1907	none	8.3333	8.3336	**-0.0002**	8.3339	**-0.0005**	8.3333	**0.0000**

- The AHP model is a deterministic model and does not consider the uncertainty in the water main performance as the proposed model does.
- Due to the defuzzification process and the characteristics of the fuzzy expert system, the results of the proposed model may show step results as explained in section V.7.1. Sensitivity Analysis and System Stability Testing.
- The AHP model only considers deterioration factors and not the consequences (post failure factors), limiting the accuracy test to only the deterioration factors of the proposed model. However, the results of this test can be generalized to the whole proposed model since same approach is used to develop the other factors that are not examined in this test.
- The AHP model is also based on experts' opinions and experience and consequently the results are expected to be close to those of the fuzzy expert system, but not exactly the same since the experts' feedback may differ on the basis of location, serving period, questionnaire interpretation, ... etc.

Two data sets are used to carry out the test: 500 Moncton data points and 1704 London data points.

- The Moncton data is selected randomly from the data set explained in section IV.2.1. Data Set One. Some data points, which have characteristics not valid in AHP model, are excluded from the test. The factors included in this test are: pipe material, pipe diameter, installation year, protection method, number of breaks, and hydraulic factor (Hazen-William coefficient). The testing data set is first evaluated using the proposed fuzzy expert model and then evaluated using the AHP model. Sample results are shown in Table V.5.

Table V.5 – Sample Moncton testing results.

	Risk Factors						Proposed model	AHP model	Difference		
ID	Pipe Material	Pipe Dia	Year	Protection	Break	Hydraulic	Risk Index	Output	value	abs %-age	AIP
1	Asbestos	152.4	1959	none	0.47	74	4.9	4.4	0.5	9%	0.10
2	Asbestos	304.8	1959	none	0.22	74	4.2	4.3	-0.1	1%	0.01
3	Asbestos	152.4	1959	none	0.73	74	5.3	4.8	0.4	8%	0.09
4	Asbestos	203.2	1959	none	0.25	74	4.3	4.4	-0.1	3%	0.03
5	Asbestos	152.4	1959	none	0.13	74	4.9	4.1	0.8	16%	0.20
6	Asbestos	304.8	1959	none	0.12	74	4.1	3.9	0.2	5%	0.06
7	Asbestos	152.4	1959	none	0.17	74	4.9	4.1	0.8	16%	0.19
8	Asbestos	152.4	1959	none	0.00	74	4.8	3.4	1.4	29%	0.42
9	Asbestos	152.4	1959	none	0.33	74	4.9	4.4	0.4	9%	0.10
10	Asbestos	254	1959	none	0.26	74	4.2	4.3	0.0	1%	0.01
11	Asbestos	254	1959	none	0.40	74	4.4	4.3	0.1	3%	0.03
12	Asbestos	152.4	1959	none	0.00	74	4.8	3.4	1.4	29%	0.42
13	Asbestos	152.4	1959	none	0.30	74	4.9	4.4	0.4	9%	0.10
14	Asbestos	152.4	1959	none	0.35	74	4.9	4.4	0.4	9%	0.10
15	Asbestos	152.4	1959	none	0.00	74	4.8	3.4	1.4	29%	0.42
16	Asbestos	152.4	1964	none	0.30	79	4.9	3.9	1.0	21%	0.26
17	Asbestos	152.4	1959	none	0.25	74	4.9	4.4	0.4	9%	0.10
18	Asbestos	152.4	1959	none	0.10	74	4.9	3.7	1.2	24%	0.31
19	Cast iron	254	1920	none	0.95	35	7.5	5.6	1.9	25%	0.33
20	Cast iron	152.4	1909	none	0.65	24	7.1	6.1	1.0	15%	0.17
21	Cast iron	254	1943	none	0.26	58	5.8	4.0	1.8	31%	0.44

In order to judge whether the model is verified or not when using results comparison as is the case here, two terms can be used to determine the validity of the model, Average Validity Percent (*AIP*) and Average Invalidity Percent (*AIP*). *AVP* represents the validation percent out of 100 and *AIP* represents the prediction error (Zayed and Halpin, 2005). These two terms are shown in Equation V.4 and Equation V.5.

$$AIP = \left(\sum_{i=1}^{n} \left|1 - \left({}^{E_i}/_{C_i}\right)\right|\right)/n$$ Equation V.4

$$AVP = 1 - AIP$$ Equation V.5

Where

AIP : Average Invalidity Percent

AVP : Average Validity Percent

E_i : Estimated value

C_i : Actual value

Using the above equations, the following statistics can be calculated, as shown in Table V.6:

Table V.6 – Model accuracy of Moncton testing results.

Statistics	
average error %-age	19.9 %
average AIP	25.7 %
AVP	74.3 %
AIP 90%	7.4%
AIP 90-80%	24.4%
AIP 80-70%	29.6%
AIP 70-0%	38.6%

From the above table, some conclusions can be drawn about the model accuracy. The average percentage difference between the outputs of the proposed model and the AHP model is 19.9 %. The Average Validity Percent is 74.3 %, which means that the proposed model is valid for predicting the output. Moreover, this test shows that about 7.4 % of the data has an AIP of more that 90%, and 24.4 % of the data fits between 80% and 90%.

– The London data contains 1704 records as explained in section V.9.3. Case Study 3. The factors included in this test are; type of soil, average daily traffic, pipe material, pipe diameter, installation year, protection method, number of breaks, and hydraulic factor (Hazen-William coefficient). The testing data set is first evaluated using the proposed fuzzy expert model and then evaluated using the AHP model. Sample results are shown in Table V.7.

Table V.7 – Sample London testing results.

	Risk Factors								Proposed model	AHP model	Difference		
ID number	Type of Soil	Average Daily Traffic	Pipe Material	Pipe Diameter	Installation Year	Protection Method	Number of Break	Hydraulic Factor	Risk Index	Output	value	absolute %age	AIP
1	6	5	Cast iron post war	150	1955	none	1.00	120	5.9	4.9	1.0	18%	0.21
2	6	2.5	Cast iron post war	150	1961	none	1.00	120	5.7	4.8	0.9	16%	0.19
3	6	2.5	Cast iron post war	150	1961	none	3.00	30	7.4	5.3	2.1	28%	0.40
4	6	2.5	Cast iron post war	150	1961	none	1.00	30	6.7	5.2	1.5	22%	0.28
5	4	2.5	Cast iron post war	150	1961	none	1.00	53	6.1	4.9	1.2	20%	0.25
6	4	2.5	Cast iron post war	150	1961	none	3.00	53	7.1	5.0	2.1	29%	0.41
7	4	2.5	Cast iron post war	150	1961	none	1.00	53	6.1	4.9	1.2	20%	0.25
8	4	2.5	Cast iron post war	150	1961	none	2.00	53	6.1	4.9	1.2	20%	0.25
9	4	10	Cast iron post war	150	1961	none	1.00	53	6.1	5.2	0.9	15%	0.18
10	2	2.5	Cast iron post war	150	1964	none	1.00	56	5.6	4.7	0.9	16%	0.19
11	4	10	Cast iron post war	150	1961	none	9.00	120	6.7	5.0	1.7	25%	0.33
12	4	2.5	Cast iron post war	150	1964	none	2.00	56	6.0	4.9	1.1	18%	0.22
13	4	10	Cast iron post war	300	1961	none	2.00	65	5.6	4.9	0.7	13%	0.15
14	4	2.5	Cast iron post war	200	1968	none	1.00	120	4.4	4.6	-0.2	5%	0.05
15	4	2.5	Cast iron post war	150	1961	none	1.00	120	5.4	4.6	0.8	15%	0.17
16	4	10	Cast iron post war	150	1961	none	2.00	120	5.7	4.9	0.8	14%	0.17
17	4	10	Cast iron post war	150	1961	none	1.00	120	5.7	4.9	0.8	14%	0.17
18	4	10	Cast iron post war	150	1961	none	1.00	120	5.7	4.9	0.8	14%	0.17
19	4	2.5	Cast iron post war	150	1961	none	1.00	53	6.1	4.9	1.2	20%	0.25

Using the same approach used in the first data set is used again to judge whether the model is valid, the following results are obtained:

Table V.8 – Model accuracy of London testing results.

Statistics	
average error %-age	19.2%
average AIP	25.2%
AVP	74.8%
AIP 90%	16.2%
AIP 90-80%	23.4%
AIP 80-70%	18.9%
AIP 70-0%	41.5%

From Table V.8, some conclusions can be drawn about the model accuracy. The average percentage difference between the outputs of the proposed model and the AHP model is 19.2 %. The Average Validity Percent is 74.8 %, which means that the proposed model is valid for predicting the output. Moreover, the test shows that about 16.2 % of the data has an AIP of more that 90%, and 23.4 % of the data fits between 80% and 90%.

Checking the results obtained from the two tests and keeping in mind the points mentioned earlier about the AHP model, one can say that the model is accurate enough to be used in the industry. Furthermore, the proposed model is rather recommended to be used than the AHP model since it considers the uncertainty of the water main parameters, it considers more risk factors especially post-failure factors (consequences factors) which are not considered in the AHP model.

V.8. Proposed Risk of Failure Scale

In light of reviewing section II.5. Risk and Condition Rating Scale, a risk of failure scale is proposed to help the decision makers in water main management companies/municipalities make an informed decision. The scale ranges numerically from 0 to 10, where 10 indicates the riskiest condition of the pipeline and 0 indicates the least risky condition. Linguistically, the scale is divided into five groups or regions that describe the risk of pipeline failure and the required corrective actions to be taken if needed. The number of proposed groups and their ranges and associated corrective actions may be changed to best suit a municipality's strategies and their risk tolerance.

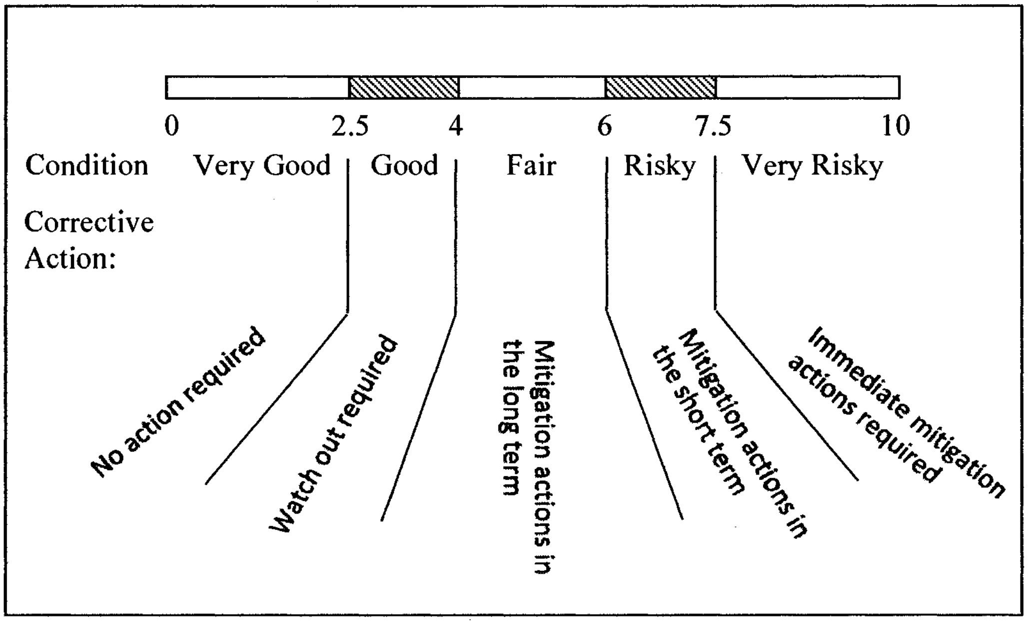


Figure V.36 – Proposed risk of failure scale.

The advantage of using a scale from 0 to10 is that it provides an easy way of making comparisons and conversions to other types of scales such as a scale from 0 to 100 or a scale from 0 to 5.

V.9. Case Study Application

In this chapter, the developed HFES model will be applied to the collected datasets to analyze the situation simulating a real management problem. Three data sets are collected from two municipalities, the City of Moncton and the City of London as explained in IV.2. Case Study Data Sets.

V.9.1. Case Study 1

Data set one is processed using HFES model and the proposed scale. Table V.9 shows a sample of the results. Table V.10 and Table V.10 summarize the results of the data set assessment using the proposed HFES model. It can be deduced that Cast Iron and Small Diameter pipes (< 250 mm) contribute most to network risk. Overall, the condition of the network is fair (66% of the network) with some parts of the network requiring mitigation action in the short-term plan as shown in Figure V.37.

Table V.9 – Sample case study 1 results.

Results				
Physical_Index	Operational_Index	Consequence_Index	Prefailure_Index	Risk_of_Failure_Index
7.5	6.0	6.7	7.5	6.3
8.3	5.3	5.1	8.3	6.6
7.5	3.8	2.5	7.5	5.7
7.3	4.5	1.7	7.4	5.7
7.3	3.2	5.1	7.4	5.7
8.3	6.7	1.7	8.3	6.7
7.8	3.8	6.1	7.7	6.0
7.7	6.7	5.1	7.6	6.8
7.6	5.1	1.7	7.6	5.9
6.9	3.6	1.7	7.0	5.3
6.8	3.1	1.7	6.9	4.7
8.3	4.7	1.7	8.3	6.2
7.1	5.4	2.2	7.2	5.6
7.3	4.0	1.7	7.4	5.7
7.6	1.7	1.7	7.6	4.3
7.5	3.4	1.7	7.5	5.2
8.0	3.3	1.7	7.9	5.0
7.8	5.5	5.1	7.7	6.1

Table V.10 – Case study 1 results summary.

Linguistic Group	Proposed Action	No. of water main	Length, *m*
Very Good	No action required	15	4,503
Good	Watch out	93	34,462
Fair	Mitigation action in long-term plan	373	101,248
Risky	Mitigation action in short-term plan	63	12,831
Very Risky	Immediate mitigation action required	0	0
	Total :	544	153,044

Table V.11 – Case study 1 pipes statistics of Fair, Risky and Very Risky status.

Pipe Characteristics		Fair		Risky		Very Risky	
		Count	**Length *m***	**Count**	**Length *m***	**Count**	**Length *m***
Dia.	Small	314	73,053	65	12,751	0	0
	Medium	57	28,196	1	80	0	0
Material	Cast Iron	56	17,057	49	10,337	0	0
	Cast Iron Post War	282	73,157	17	2,494	0	0
	Asbestos	18	6,578	0	0	0	0
	Ductile Iron	15	4,457	0	0	0	0

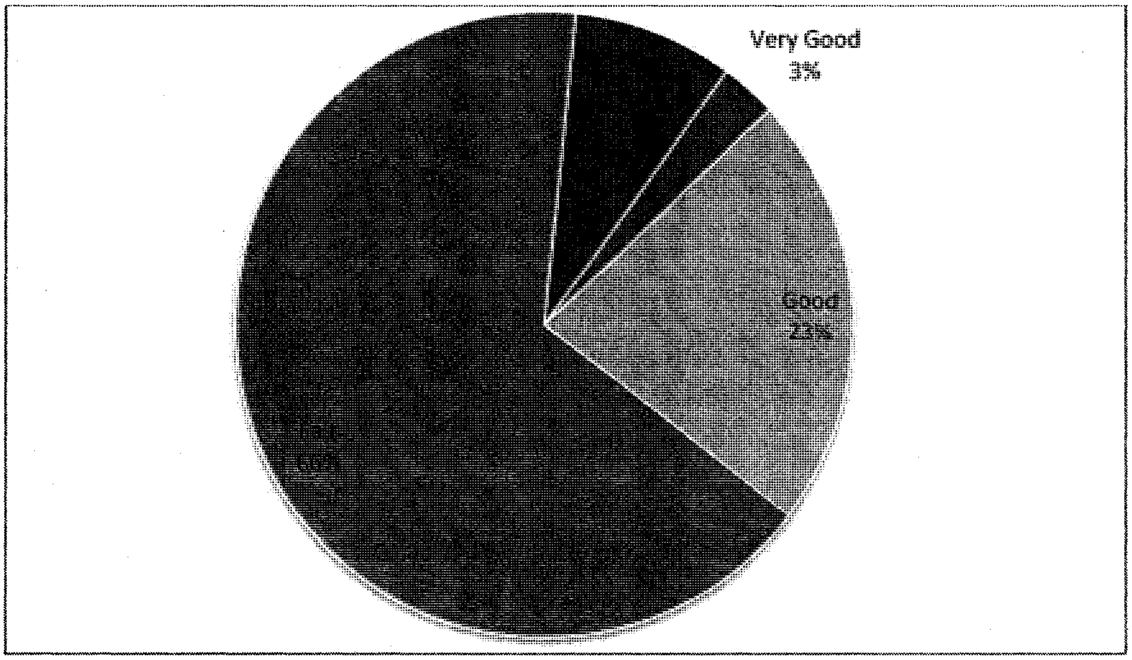


Figure V.37 – Water main risk distribution of case study 1.

The results can be further analyzed using the GIS system which provides the opportunity to locate the different pipes and ease the setup of a management plan. The pipes that are assessed using the proposed model are shown in Figure V.39. The pipes are colored and grouped according to their risk of failure score. The groups are the same proposed in the risk of failure scale: Very Good, Good, Fair, Risky, and Very Risky. After reviewing the pipelines' locations, the management team may decide to renew or rehabilitate the risky pipelines. However, due to the fact that the risky pipelines are located in an almost enclosed area, the management team may decide to include the pipelines at fair risk (which will need mitigation actions in the long-term plan) in the rehabilitation plan to save on the costs of mobilization and equipment transportation. The management team may include only the pipes at fair risk that are top ranked or may not include any fair risk pipes according to the allocated budget. Figure V.40 shows a proposed area to be included in a rehabilitation plan which includes both risky and fair pipes. The short-term rehabilitation plan can be set for every year or any other period of time depending on a management team's preference. It should be noted that not all the risky pipes are included in the plan since some are remote from the proposed area and they will require a considerable amount of money to rehabilitate them to account for the cost of mobilization and transportation, and thus the management team may be willing to carry the risk of failure by doing nothing to these pipes. Table V.12 shows a sample of the selected pipes for the short-term rehabilitation plan. Table V.13 summarizes the characteristics of the selected pipes. Figure V.38 illustrates a framework on how the decision can be taken regarding water main management using the proposed model.

Table V.12 – Sample selected pipes for short-term rehabilitation plan.

Number	Pipe Type	Year Installation	Diam	Risk of Failure	Risk Group
1	Cast Iron	1920	254	6.8	Risky
2	Cast Iron	1895	102	6.7	Risky
3	Cast Iron	1907	102	6.7	Risky
4	Cast Iron	1896	102	6.7	Risky
5	Cast Iron	1895	102	6.7	Risky

Table V.13 – Summaries of the selected pipes for short-term rehabilitation plan.

Pipe Characteristics		Fair		Risky		Very Risky	
		Count	**Length *m***	**Count**	**Length *m***	**Count**	**Length *m***
Dia.	Small	37	6,996	46	9,836	0	0
	Medium	17	6,252	0	0	0	0
Material	Cast Iron	28	8,567	31	7,639	0	0
	Cast Iron Post War	23	4,075	13	1,816	0	0
	PVC	2	314	2	381	0	0
	Ductile Iron	1	292	0	0	0	0

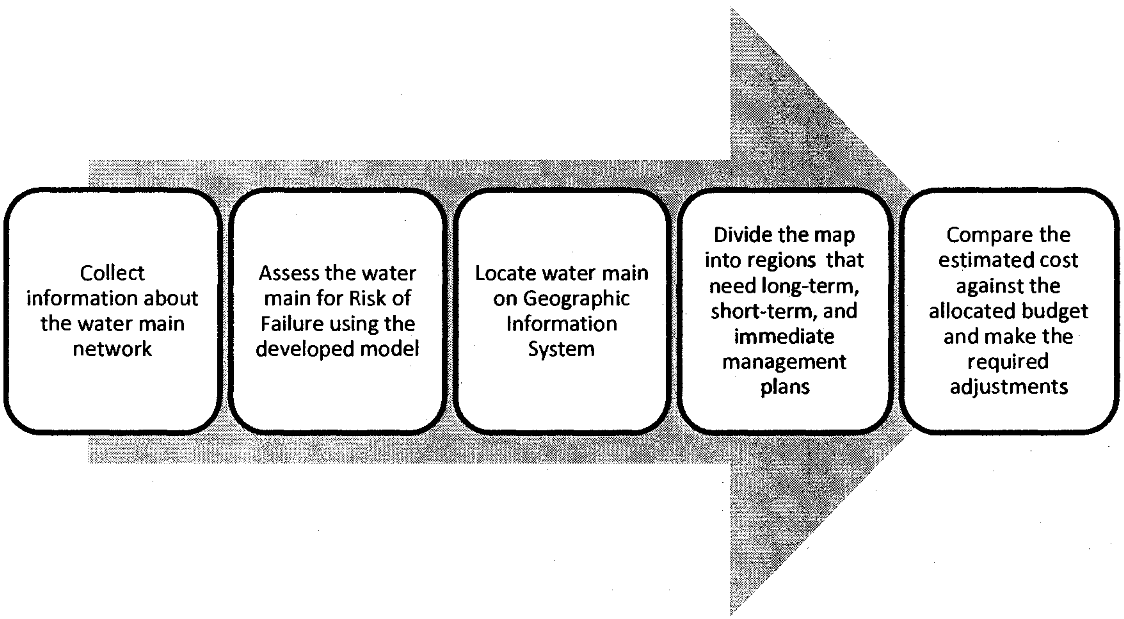


Figure V.38 – Decision making flow chart.

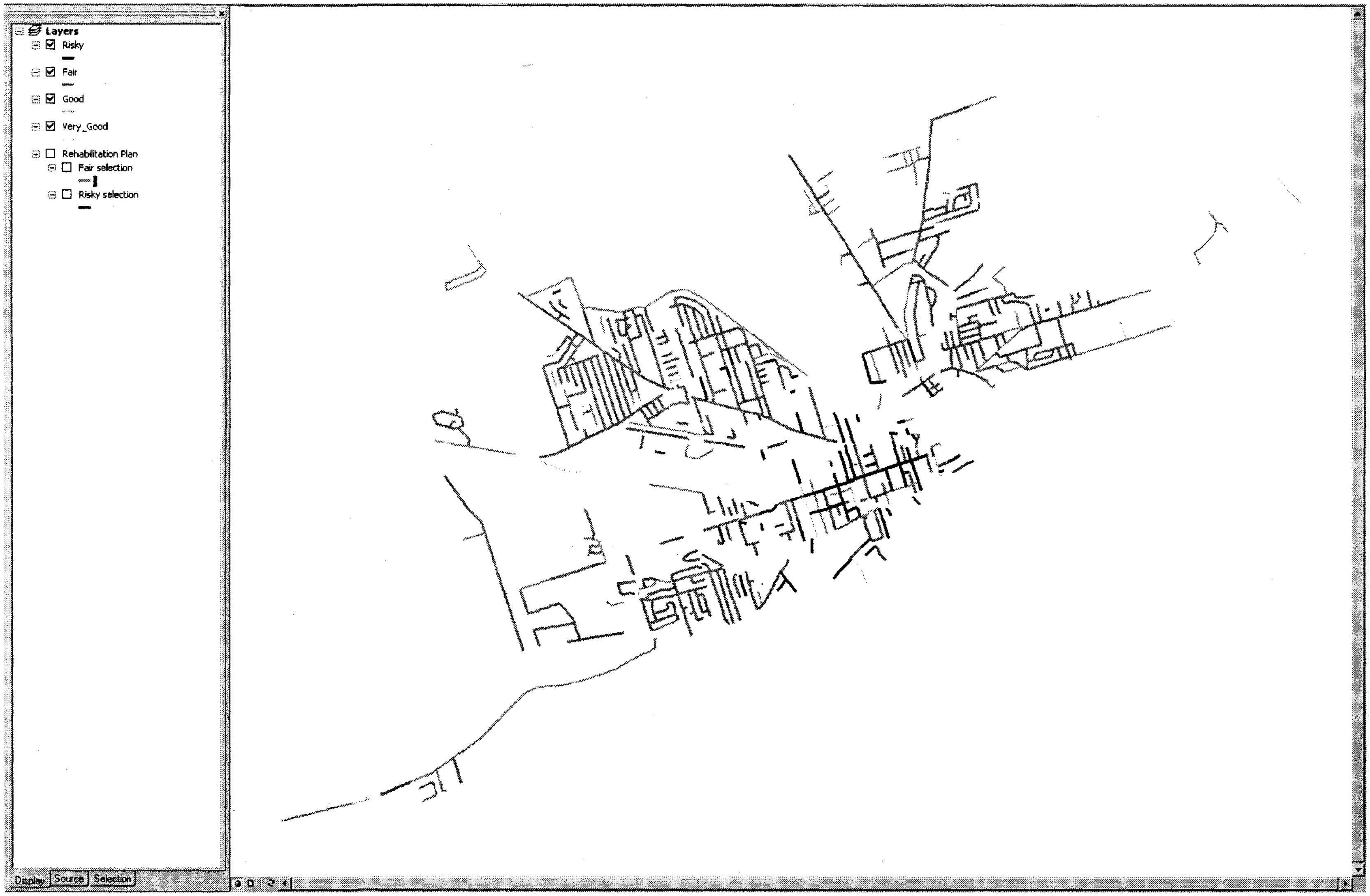


Figure V.39 – Water main assessed against the risk of failure.

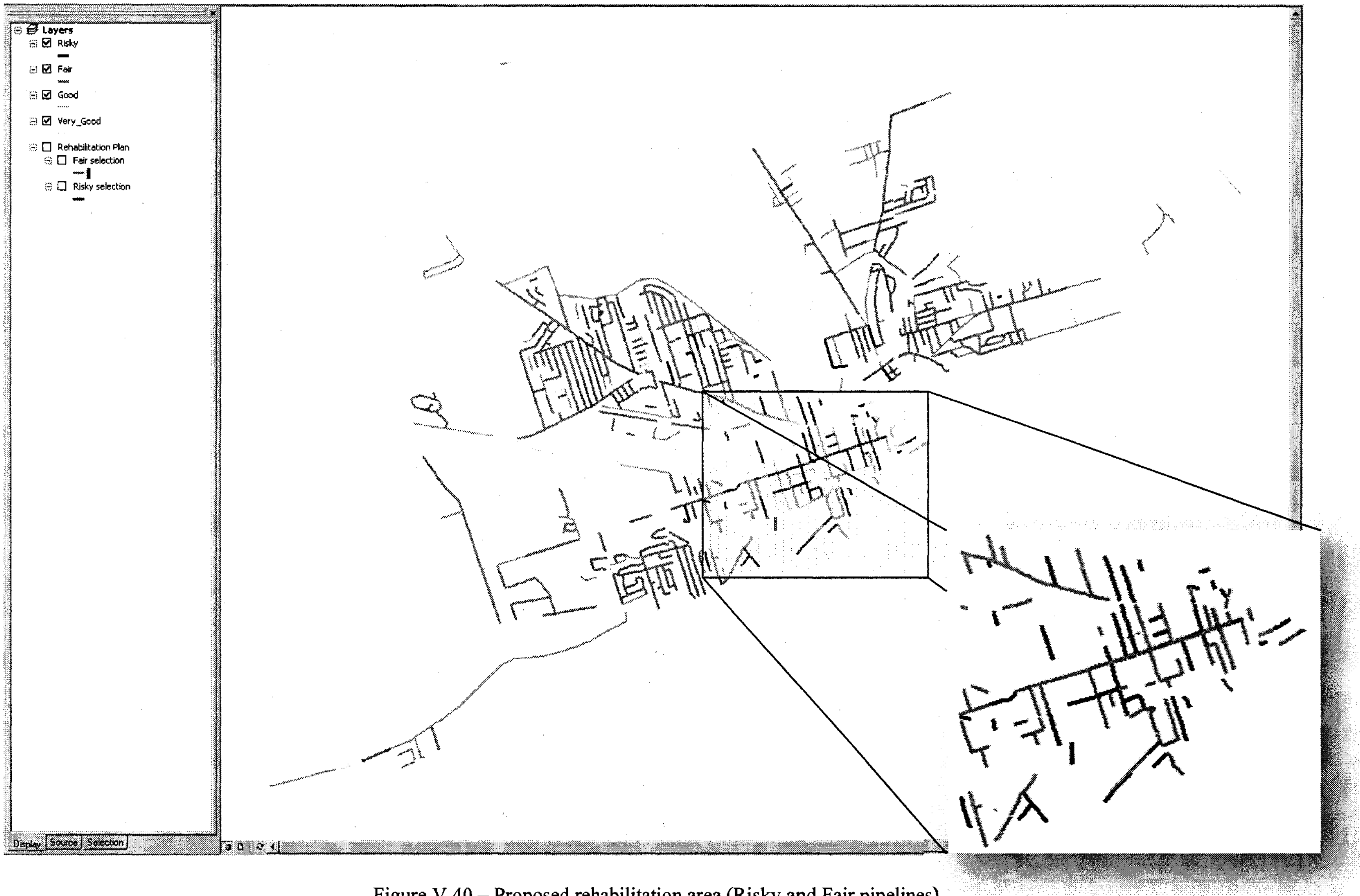


Figure V.40 – Proposed rehabilitation area (Risky and Fair pipelines).

V.9.2. Case Study 2

Using the same weights collected from experts and the proposed scale, the data set 2 is processed using the developed model. Table V.14 shows a sample of the results. Table V.15 and Table V.16 summarize the results of processing the data set, which show that the Cast Iron Post War material and Small Diameter pipes contribute most to the risky situation of the network. Overall, the condition of the network is risky (50%) to fair (47%) with some parts of the network requiring immediate mitigation action as shown in Figure V.41.

Table V.14 – Sample case study 2 results.

Results					
Environmental_Index	Physical_Index	Operational_Index	Consequence_Index	Prefailure_Index	Risk_of_Failure_Index
7.4	6.6	4.8	8.3	6.4	6.6
7.3	6.6	2.8	6.7	4.9	5.7
6.2	6.6	3.1	8.5	4.9	6.3
6.7	7.5	5.6	5.0	7.4	6.7
8.5	8.3	4.8	4.1	6.6	6.6
7.9	7.5	4.8	5.9	6.3	6.3
5.0	7.3	5.0	7.4	6.7	6.7
7.4	7.3	3.5	6.7	5.7	5.8
8.3	8.3	5.9	7.9	7.6	7.6
6.2	7.8	4.8	3.3	6.4	6.0
8.7	7.7	7.0	6.7	7.6	7.6
6.9	7.6	5.7	4.5	7.4	6.0
8.5	6.9	3.8	6.7	5.5	6.7
8.5	6.8	5.1	5.0	6.8	6.8
6.2	8.3	6.5	8.3	8.0	7.7
7.3	7.1	6.2	7.4	7.2	7.3
7.8	7.3	4.7	6.7	6.2	6.2
6.7	7.6	3.3	7.1	5.9	5.9
5.0	7.5	4.7	6.7	6.2	6.2
7.4	8.0	3.3	6.7	6.2	6.2
8.7	7.8	5.7	8.3	7.4	7.4
7.4	7.4	6.7	7.2	7.4	7.4
7.4	7.4	5.1	4.3	6.9	5.9
7.2	7.0	3.4	7.4	5.4	5.8

Table V.15 – Case study 2 results summary.

Linguistic Group	Proposed Action	No. of water main	Length, *m*
Very Good	No action required	0	0
Good	Watch out	3	512
Fair	Mitigation action in long-term plan	210	72,293
Risky	Mitigation action in short-term plan	310	75,535
Very Risky	Immediate mitigation action required	21	4,704
	Total :	544	153,044

Table V.16 – Case study 2 pipes statistics of Fair, Risky and Very Risky status.

Pipe Characteristics		Fair		Risky		Very Risky	
		Count	**Length *m***	**Count**	**Length *m***	**Count**	**Length *m***
Dia.	Small	164	41,697	277	61,389	21	4,704
	Medium	46	30,596	33	14,186	0	0
Material	Cast Iron	16	5,114	80	20,841	10	2,069
	Cast Iron Post War	87	26,179	211	50,669	11	2,635
	PVC	9	2,511	1	87	0	0
	Asbestos	9	4,167	9	2,411	0	0
	Ductile Iron	89	34,322	9	1,527	0	0

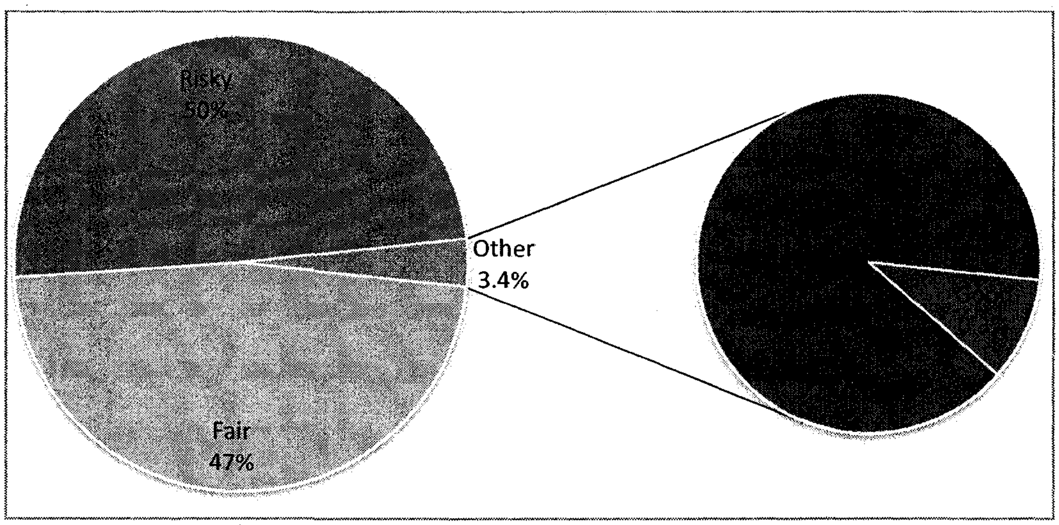


Figure V.41 – Water main risk distribution of case study 2.

V.9.3. Case Study 3

Using the same weights collected from experts and the proposed scale, the data set is processed using the developed model. Table V.17 shows a sample of the results. Table V.18 and Table V.19 summarize the results of the data set assessment using the proposed fuzzy model. The overall condition of the network is fair (50%) to risky (45%) with some parts of the network requiring immediate mitigation action as shown in Figure V.42.

Table V.17 – Sample case study 3 results.

Results					
Environmental_Index	Physical_Index	Operational_Index	Consequence_Index	Prefailure_Index	Risk_of_Failure_Index
6.4	7.7	3.3	5.0	5.9	5.9
5.7	7.3	3.3	1.7	5.7	5.0
5.7	7.3	7.6	1.7	7.4	5.9
5.7	7.3	6.7	5.0	6.7	6.7
3.6	7.3	6.3	1.7	6.1	5.7
3.6	7.3	7.3	1.7	7.1	5.7
3.6	7.3	6.3	1.7	6.1	5.7
3.6	7.3	6.3	5.0	6.1	5.7
5.2	7.3	6.3	1.7	6.1	5.7
1.5	7.1	6.1	5.0	5.6	5.6
5.2	7.3	6.7	1.7	6.7	5.7
3.6	7.1	6.1	1.7	6.0	5.6
5.2	6.3	5.5	7.3	5.6	6.1
3.6	6.1	3.3	6.1	4.4	5.0
3.6	7.3	3.3	5.0	5.4	5.4
5.2	7.3	3.3	5.0	5.7	5.7
5.2	7.3	3.3	5.0	5.7	5.7
5.2	7.3	3.3	5.0	5.7	5.7
3.6	7.3	6.3	1.7	6.1	5.7

Table V.18 – Case study 3 results summary.

Linguistic Group	Proposed Action	No. of water main	Length, *m*
Very Good	No action required	2	245
Good	Watch out	58	11,118
Fair	Mitigation action in long-term plan	823	137,453
Risky	Mitigation action in short-term plan	798	123,136
Very Risky	Immediate mitigation action required	21	2,823
	Total :	1702	274,773

Table V.19 – Case study 3 pipes statistics of Fair, Risky and Very Risky status.

Pipe Characteristics		Fair		Risky		Very Risky	
		Count	**Length *m***	**Count**	**Length *m***	**Count**	**Length *m***
Dia.	Small	729	120,265	730	111,285	21	2,823
	Medium	94	17,189	68	11,850	0	0
Material	Cast Iron	61	10,888	322	43,492	14	1,643
	Cast Iron Post War	636	109,265	475	79,300	7	1,180
	PVC	4	298	0	0	0	0
	Concrete	1	171	0	0	0	0
	Ductile Iron	121	16,832	1	344	0	0

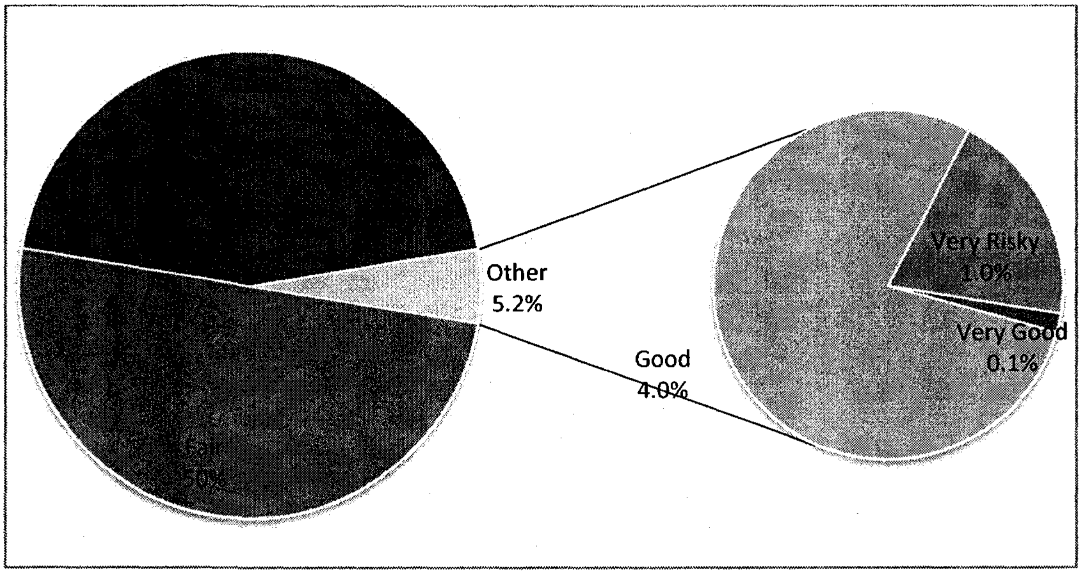


Figure V.42 – Water main risk distribution of case study 3.

To further analyze the outputs, the dataset's records are clustered according to pipe material (Cast Iron, Cast Iron post war, Ductile Iron, and PVC) and according to their scores (10 groups). Table V.20 and Figure V.43 show the percentages scores of each material (local percentages for each material are not comparable to other materials). It can be deduced that the majority score of the Cast Iron material falls in the range between

6 and 7. However, for Cast Iron Post War pipes, the scores form a triangular shape between 3 and 8 with a peak between 5 to 6 that has 48.66 % of the cast iron post war material. Ductile Iron material pipes also form a triangular shape from 2 to 7, with a peak between 4 and 5 with 63.86 % of the Ductile Iron material pipes. Figure V.44 shows the pipe material global scores percentages. It can be concluded that the Cast Iron Post War material contributes in large part to the fair-risky status of the London water main network, as 55% of that network is Cast Iron Post War with scores between 5 and 7.

Table V.20 – Pipes material percentage score.

Local Percentage	0 to 1	1 to 2	2 to 3	3 to 4	4 to 5	5 to 6	6 to 7	7 to 8	8 to 9	9 to 10
Cast Iron					0.25%	15.11%	77.58%	7.05%		
Cast iron post war				0.18%	8.13%	48.66%	35.09%	7.95%		
Ductile iron			0.60%	25.90%	63.86%	9.04%	0.60%			
PVC			47.06%	29.41%	23.53%					

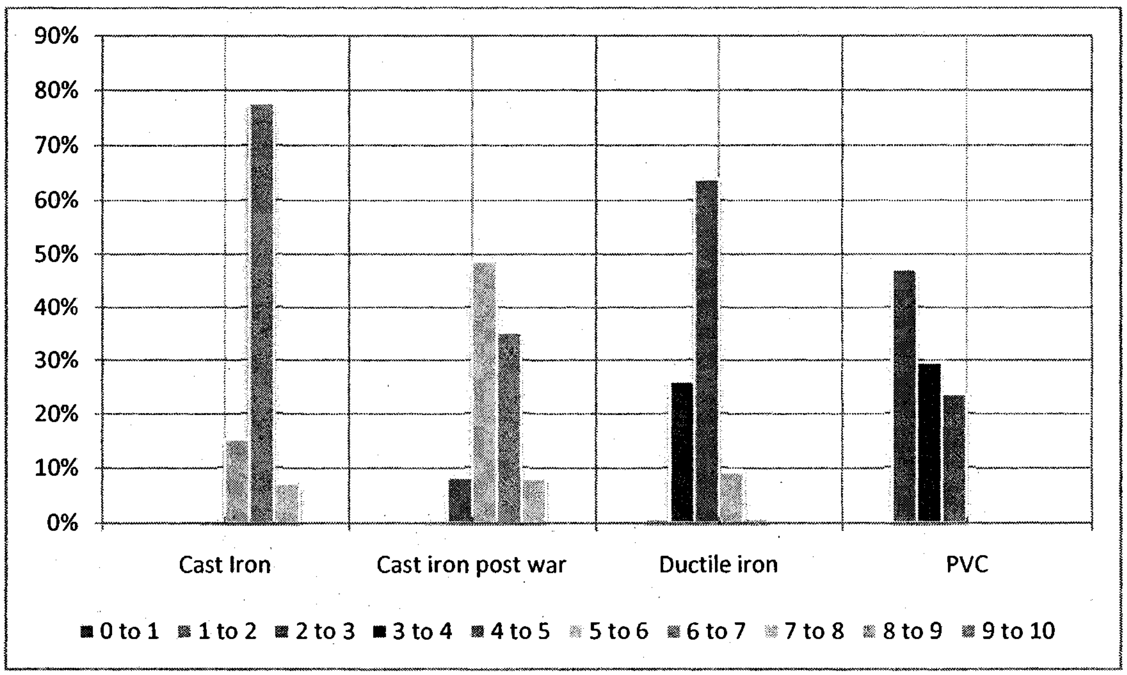


Figure V.43 – Pipes material local score percentage.

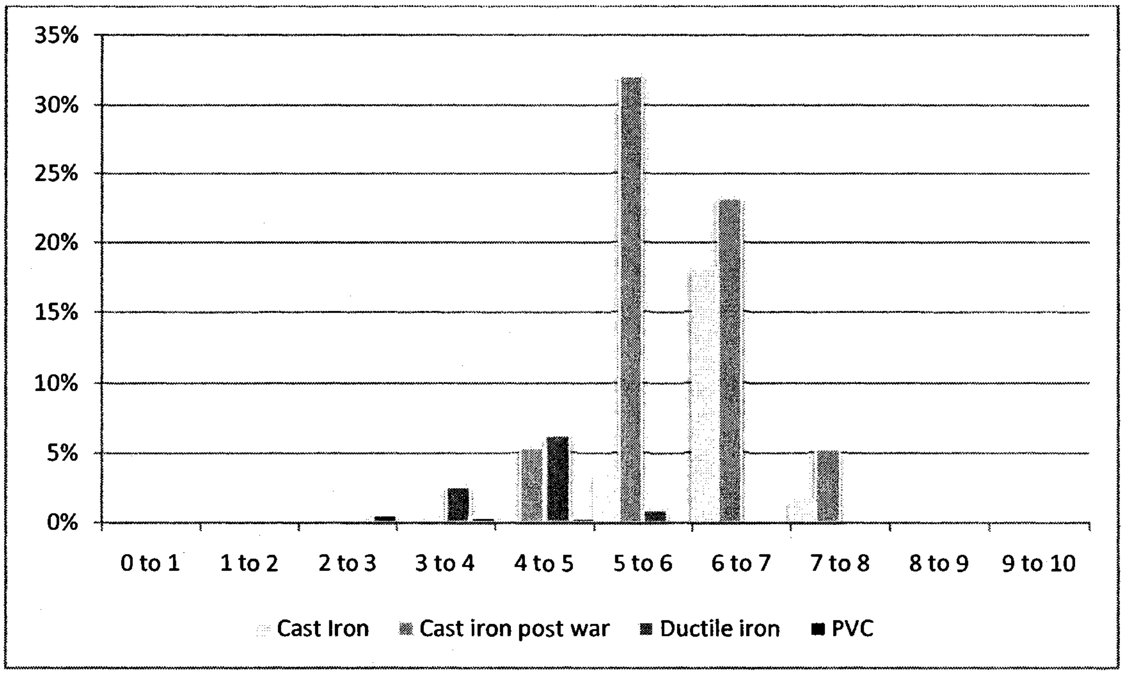


Figure V.44 – Pipes material global score percentage.

V.10. Summary

This chapter presented the work done in building the FHES model. It explains into details the different stages in the model building, starting from the considered factors and their linguistic membership functions, through the fuzzy rules extraction and evaluation process, the rules aggregation and the deffuzification process. The sensitivity of the FHES model is tested and analyzed and the results showed that the model is robust and sound. The FHES model is also verified using a validated AHP model. It also showed the proposed failure risk scale. Three case studied are evaluated and studied using the developed model in order to show the probable use of the model in the field of water main management.

Chapter VI: EXCEL-BASED APPLICATION DEVELOPMENT

VI.1. Introduction

An application was developed to implement the designed model. This application is based on MS© Excel 2007 and thus requires this program in order to run. Excel 2007 was chosen over earlier Excel versions because of the extended size of the worksheet which is required to handle all the fuzzy expert system calculations, where the earlier versions cannot handle. In addition, Excel 2007 has more options, functions, and visual aids and it results in a compressed smaller file size and a lighter load on computers.

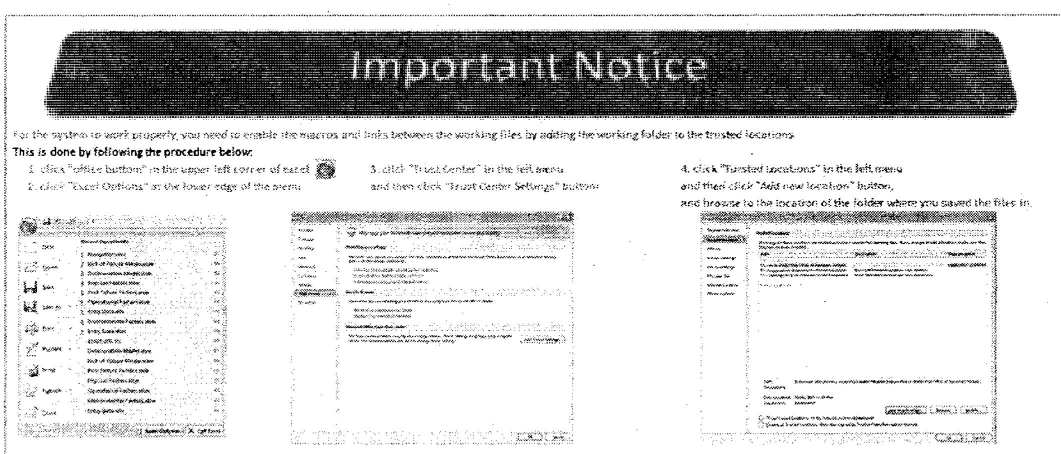
VI.2. Working Folder and Files

The working folder which contains the files is named as "Risk of Water Main Failure". This folder can be saved any where on the computer hard disk and the links between different excel files will still be functioning. The folder contains nine Excel files and these files are explained as follows:

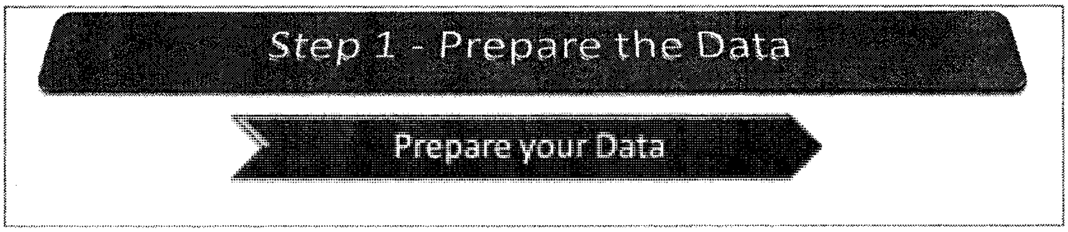
- *Navigation*: this file contains all the step-by-step instructions which will guide the user to an easy use of the application. Moreover, it allows the user to easily fine tune the factors weights and the expert knowledge base. The file contains two worksheets "Control Panel" and "Fine Tune". "Control Panel" worksheet guides the user to setup the application and enter the required information to run the model. It starts with "Important Notice" on how to deal with security issues and how to enable the macros which are needed to run the application as shown in Figure VI.1 (a).

Then, the worksheet contains three steps which will setup and run the application. The first step is "Step 1 - Prepare the Data" (Figure VI.1 (b)) which has a link to another excel file that stores the information about the different performance characteristics on the water main network. This file is named "Network Performance Data" and will be explained hereafter.

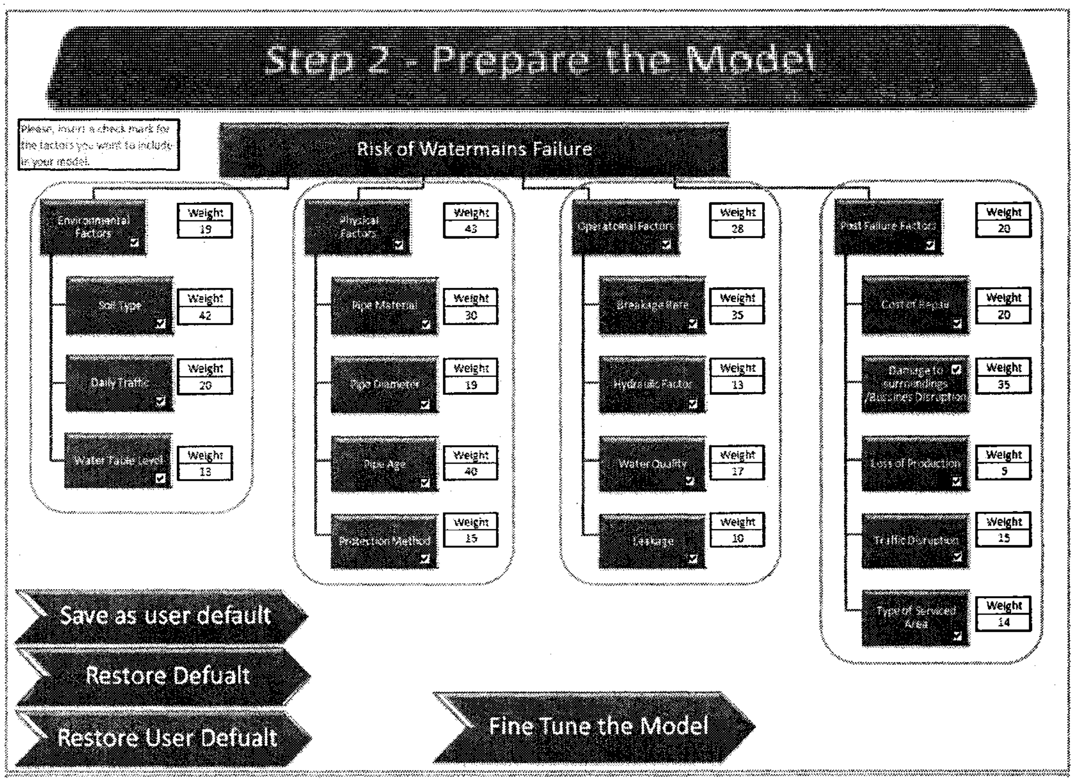
The next step is "Step 2 - Prepare the Model" which controls the factors weights and even allows the user to save the weights according to his own preferences giving the user the full flexibility as shown in Figure VI.1 (c). Moreover, the user can choose the factors to be incorporated in the risk assessment.



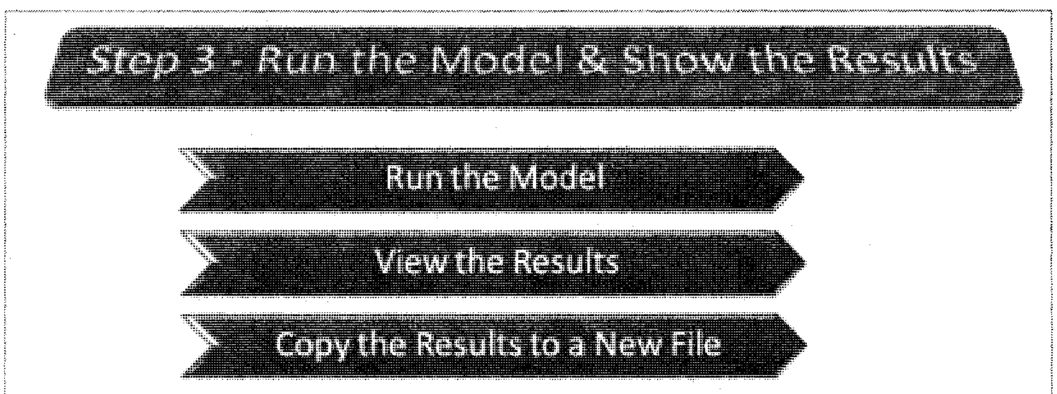
(a)



(b)



(c)



(d)

Figure VI.1 – Control panel worksheet in Navigation workbook.

In addition to that, this panel contains a button that links to the other worksheet "Fine Tune", which enables the user to change the expert knowledge base to best suit the user own experience and preferences. Figure VI.2 shows how to tune the Environmental Factors knowledge base.

Environmental Factors				
1. What is the effect of soil type on the risk of pipeline failure (qualitative factor)?				
if soil is	Very lightly deteriorative	then	the risk of failure is	Extremely Low
if soil is	Lightly deteriorative	then	the risk of failure is	Very Low
if soil is	Moderately deteriorative	then	the risk of failure is	Medium
if soil is	Highly deteriorative	then	the risk of failure is	Very High
if soil is	Very highly deteriorative	then	the risk of failure is	Extremely High
2. What is the effect of Average Daily Traffic on the risk of pipeline failure (qualitative factor)?				
if ADT is	Very light	then	the risk of failure is	Extremely Low
if ADT is	Light	then	the risk of failure is	Very Low
if ADT is	Moderate	then	the risk of failure is	Medium
if ADT is	Heavy	then	the risk of failure is	Very High
if ADT is	Very heavy	then	the risk of failure is	Extremely High
3. What is the effect of Water Table Level on the risk of pipeline failure (qualitative factor)?				
if WTL is	rarely present	then	the risk of failure is	Extremely Low
if WTL is	seasonally present	then	the risk of failure is	Extremely High
if WTL is	always present	then	the risk of failure is	Moderately High

Figure VI.2 – Fine tuning of the environmental factors attributes.

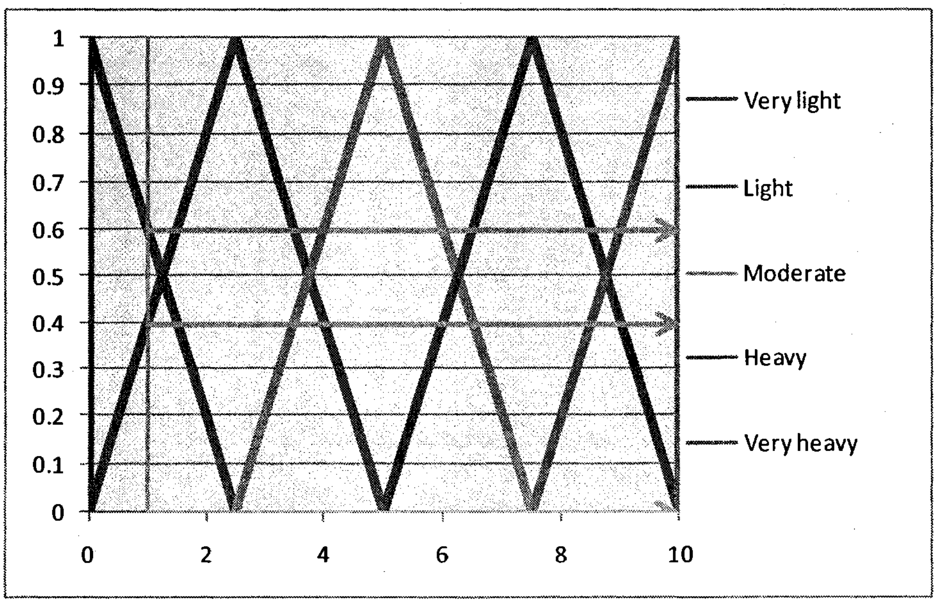
The third and last step is "Step 3 - Run the Model & Show the Results", which contains "Run the Model" button that opens the six other excel files one by one, and runs the macros to perform the required calculations. These files, which are explained later in this section, are: "Environmental Factors", "Physical Factors", "Operational Factors", "Post Failure Factors", "Pre-failure Model", and "Risk of Failure Model". The "View the Results" button opens the "Results" file and shows the calculated risk of failure for each record stored in "Network Performance Data". The "Copy the Results to a New File" button copies the risk of failure results stored temporary in the "Results" Excel file to a new file that will store the results permanently as shown in Figure VI.1 (d).

- *Network Performance Data:* This Excel file stores the attributes and performance characteristics of the water distribution network. It contains seventeen columns which correspond to the sixteen risk factors (Loss of production factor requires two columns: pipe diameter and redundancy). The total number of records which can be stored in this excel file is limited to 2000 records due to huge computer resource needed to process the large amount of data. Each of these columns has its own validation rule which will guide and restrict the user to the type and range of information to be input into the model. There are four more columns which can be used to store some notes about each record.

- *Environmental Factors, Physical Factors, Operational Factors, and Post Failure Factors Modules:* These Excel files are considered the highest level in the hierarchy shown in Figure V.2. Their files structures are similar, and thus for illustration purposes, only the "Environmental Factors" Excel file structure will be explained here. This Excel file contains two worksheets: "Environmental Model" and "Enviro Process". The "Environmental Model" worksheet contains all the calculations required to generate the dynamic rules. The weights of the factors and the expert knowledge base are directly linked to the "Navigation" Excel file and thus any changes made to "Navigation" will be directly reflected in this worksheet. This worksheet consists of two parts; part one contains the different environmental factors, their associated antecedents, membership functions and their consequents. The second part contains all the possible rules that are constructed and shown in Section IV.1.3. Expert knowledge base. The other worksheet, "Enivro Process", contains all the calculations required to process the data available about the network attributes and performance characteristics as stored in *"Network Performance Data"* file and to generate the estimated index for each record in the data. This worksheet contains many sections. It starts with fuzzification of the real data (stored in the "Network Performance Data" Excel file) in which the real network data is used to generate membership values for the membership functions of the different factors as shown in Figure VI.3 (a). The upper part of the table shows the membership functions of the factors as stored in the first worksheet "Environmental Model" which will be used to fuzzify the real data records. The results of fuzzification of the real data are shown underneath the membership functions which are identified by records serial numbers. The assessment of two records is shown in Figure VI.3.

antecedent	Very lightly deteriorative	Lightly deteriorative	Moderately deteriorative	Highly deteriorative	Very highly deteriorative	Very light	Light	Moderate	Heavy	Very heavy	rarely present	seasonally present	always present
MF	trapmf	trapmf	trapmf	trapmf	trapmf	trapmf	trapmf	trapmf	trapmf	trapmf	discrete	discrete	discrete
param1			2.5	5	7.5			2.5	5	7.5	0.95	0.95	0.95
param2		2.5	5	7.5	10		2.5	5	7.5	10			
param3		2.5	5	7.5	10		2.5	5	7.5	10			
param4	2.5	5	7.5	10	10	2.5	5	7.5	10	10			
	1	2	3	4	5	1	2	3	4	5	1	2	3
	Fuzzification of Real Data												
Serial.no	Type_of_Soil					Average_Daily_Traffic					Water_Table_Level		
1	0.2	0.8								1	0.95		
2					1	0.6	0.4						0.95

(a)



(b)

input1	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
input2	5	5	5	4	4	4	3	3	3	2	2	2	1	1	1
input3	3	2	1	3	2	1	3	2	1	3	2	1	3	2	1
input4															
input5															
output1	6	5	5	6	5	4	5	5	4	5	4	4	5	4	5
											0.4	[illegible]		0.6	[illegible]

(c)

Rule MF	Extremely Low trapmf	Very Low trapmf	Moderately Low trapmf	Medium trapmf	Moderately High trapmf	Very High trapmf	Extremely High trapmf
input1			1.67	3.33	5	6.67	8.33
input2		1.67	3.33	5	6.67	8.33	10
input3		1.67	3.33	5	6.67	8.33	10
input4	1.67	3.33	5	6.67	8.33	10	10
output1	1	2	3	4	5	6	7

Membership value of output MF	
0.2	0.66667
0.33333	0.6

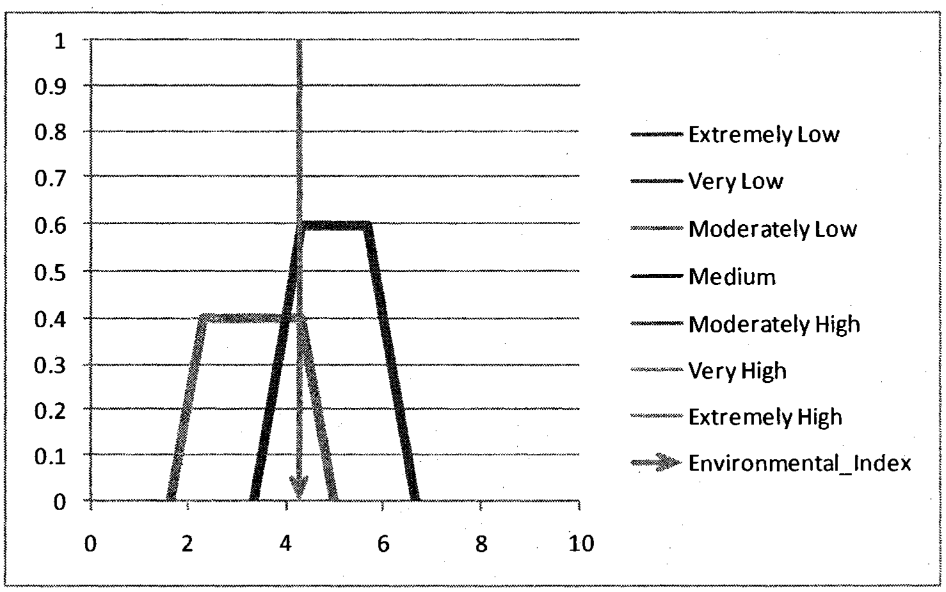
(d)

Serial.no	Extremely Low	Very Low	Moderately Low				Medium			
1			1.67	2.002	4.666	5	3.33	4.44333	5.55667	6.67
2			1.67	2.22333	4.44333	5	3.33	4.332	5.668	6.67

(e)

		Extremely Low		Very Low		Moderately Low		Medium		Moderately High		Very High		Extremely High		
9	Serial.no	Area	centriod	Area	centriod	Area	centriod	Area	centriod	Area	centriod	Area	centriod	Area	centriod	Environmental_Index
10	1					0.5994	3.33452	1.48444	5							4.5
11	2					0.925	3.33422	1.4028	5							4.3

(f)



(g)

Figure VI.3 – Data processing in the excel-based application.

Figure VI.3 (b) shows the membership functions of the Average Daily Traffic factor of the second record of the real data. Its results in a 0.6 membership value for Very light membership function and a 0.4 membership value for Light membership function. This visual aid is also included in the Excel file.

The next step is to evaluate each rule in the knowledge base. When a rule is triggered, the minimum (And operation) membership value of the considered factors from the fuzzification step is calculated and stored for later use in the aggregation process. Figure VI.3 (c) shows the different rules and the results of the triggered rules for the second record of the real data.

After assessing all of the triggered rules, the next step is to aggregate the membership values that resulted from last step according to the membership function output, and then to choose the maximum membership value (aggregation method is Maximum) for each linguistic membership function to be used in the next step, as shown in Figure VI.3 (d). It is worth mentioning that this is the same approach used in Matlab © software.

The next step is to use these aggregated membership values to truncate the output membership functions and construct new membership functions that use the Center of Sum defuzzification method to get an Environmental Index for each record in the data set. Figure VI.3 (e) shows the new membership function parameters resulting from the truncation of the original output membership functions.

The calculation of the defuzzification method and the resulting Environmental Index are shown in Figure VI.3 (f).

A visual aid (chart) is included in the Excel file to best present the calculation results, as shown in Figure VI.3 (g), where the output membership functions are truncated and new membership functions are generated accordingly. This figure also shows the Environmental Index of the second record calculated using the Center of Sum defuzzification method.

It should be observed that these four Excel files are read-only and are not to be altered. Also, it is advisable to check these files only to overview the calculations or to see the visual aids.

- *Risk of Failure Module:* The structure of this Excel file is similar to that of the "Environmental Factors" Excel file. This file is responsible for combining the results of the four main failure risk factors modules to get a crisp value of failure risk for a water main. The hierarchy is shown in Figure V.2. The partial risk indices (environmental index, physical index, operational index, and post-failure index) are processed as data inputs instead of the real data used in the second level of the hierarchy.

- *Prefailure Module:* This Excel file is a duplicate of the "Risk of Failure" Excel file. However, it uses only three instead of four inputs: Environmental index, Physical index, and Operational index. A post-failure index is precluded in this Excel file. The reason for creating such a file is to give an idea about the possibility of a failure event in the pipelines as a prefailure index as opposed to a post-failure index, as generated before. The hierarchy of the "Pre-failure Model" is shown in Figure V.3.

- *Results:* This is the last Excel file in the working folder. It displays all the indices generated in the six Excel files (Figure VI.4).

	Results					
ID number	Environmental_Index	Physical_Index	Operational_Index	Consequence_Index	Prefailure_Index	Risk_of_Failure_Index
1	4.5	4.4	4.5	7.2	4.4	5.6
2	4.3	4.1	5.9	5.0	5.0	5.0

Figure VI.4 -- The results of the application data processing.

VI.3. Testing of the developed Application's Programming

The objective of this section is to test the internal calculations, procedures and programming inside the Excel-based application using Matlab ® software. The results obtained from the developed HFES are compared against the results obtained using Matlab ©. The tested calculations are the fuzzification process, the rules triggering process, the fuzzy operations, and the defuzzification process. Two scenarios are followed to accomplish this, as explained below:

1. Using the operational model, the effects of each factor are increased one at a time from the best performance to the worst performance as shown in Table VI.1.
2. Using a risk of failure model, maximum and minimum scenarios together with eight randomly chosen values are examined as shown in Table VI.2.

By analyzing Table VI.1 and Table VI.2, one can note that the difference between the results obtained from the developed application and the Matlab ® are very small and minor, and is due to the different defuzzification methods (Matlab uses the Centriod method). Also, it can be concluded that even though the Center of Sum defuzzification method requires far less calculations, it generates fairly precise results.

Table VI.1 – Operational model Matlab testing.

		Operational				Results		
	Operational Model Matlab Testing	**Number of Break**	**Hydraulic Factor**	**Water Quality**	**Leakage**	**Matlab**	**Developed System**	**difference**
1	Breakage rate 1	0	130	0	0	0.52	0.60	0.08
2	Breakage rate 2	2	130	0	0	1.67	1.70	0.03
3	Breakage rate 3	6	130	0	0	5.00	5.00	0
4	Water Quality 1	6	130	2.5	0	5.00	5.00	0
5	Water Quality 2	6	130	5	0	5.00	5.00	0
6	Water Quality 3	6	130	7.5	0	6.67	6.70	0.03
7	Water Quality 4	6	130	10	0	6.67	6.70	0.03
8	Hydraulic Factor 1	6	90	10	0	6.67	6.70	0.03
9	Hydraulic Factor 2	6	70	10	0	8.33	8.30	-0.03
10	Hydraulic Factor 3	6	50	10	0	8.33	8.30	-0.03
11	Hydraulic Factor 4	6	20	10	0	8.33	8.30	-0.03
12	Leakage 1	6	20	10	2.5	8.33	8.30	-0.03
13	Leakage 2	6	20	10	5	8.33	8.30	-0.03
14	Leakage 3	6	20	10	7.5	9.48	9.40	-0.08
15	Leakage 4	6	20	10	10	9.48	9.40	-0.08

Table VI.2 – Risk of failure model Matlab testing.

	Risk of Failure Model Matlab Testing	Risk of Failure				Results		
		Environmental	Physical	Operational	Consequence	Matlab	Developed System	difference
1	Maximum	9.44	8.33	9.44	9.44	8.73	8.81	0.08
2	Minimum	0.56	1.67	0.56	1.67	1.67	1.66	-0.01
3	random 1	0.56	3.33	0.56	1.67	1.67	1.67	0.00
4	random 2	0.56	5.00	0.56	1.67	2.28	2.31	0.03
5	random 3	1.67	8.33	5.00	1.67	5.00	5.00	0.00
6	random 4	3.33	8.33	5.00	1.67	5.00	5.01	0.01
7	random 5	5.00	8.33	5.00	1.67	5.00	5.01	0.01
8	random 6	8.33	8.33	6.67	5.00	6.67	6.67	0.00
9	random 7	8.33	8.33	8.33	5.00	8.33	8.33	0.00
10	random 8	9.44	8.33	9.44	8.33	8.33	8.34	0.01

VI.4. Summary

An Excel-based water main failure risk assessment is developed based on the proposed hierarchical fuzzy expert system of water main risk of failure model. The model calculations and data flow is checked using Matlab software and the results show that the developed application passes the test are ready to be used. This application can be used by municipal and consultant engineers to estimate the failure risk associated with water mains in order to better manage their distribution network and spend the allocated budget more efficiently.

Chapter VII: CONCLUSIONS AND RECOMMENDATIONS

VII.1. Summary

This work has presented a methodology that addresses the challenge faced by municipalities and other authorities of prioritizing the rehabilitation of water main systems. It offers a model to evaluate the risk of water main failure that considers many risk factors, which can be divided broadly into deterioration factors that lead to the failure event and consequence factors that result from a failure event (failure impact). Sixteen failure risk factors are incorporated in the model (11 deterioration factors and 5 consequence factors). A hierarchical fuzzy expert system (HFES), which takes into account the uncertainty in the water main attributes, is used to build this model. The use of hierarchy allows the number of knowledge base rules required to construct the model to be reduced. The model is verified using a validated AHP deterioration model and two different data sets (from the cities of Moncton, NB and London, ON). A water main failure risk scale is proposed, which ranges numerically from 0 to 10 where 0 indicates the lowest risk situation and 10 the highest risk situation. Linguistically, the scale is divided into five zones: "Very good, Good, Fair, Risky, and Very Risky". Each of these zones proposes appropriate actions to mitigate the risk, as appropriate. Three case studies, from different potable water networks, are assessed using the developed model. The results of risk assessment of these case studies are analyzed and rehabilitation plans are proposed accordingly.

Based on the developed model and the proposed failure risk scale, an Excel®-based application is developed to assess and evaluate the risk of failure associated with the water main and advises the management team of some proposed mitigating actions. Municipal water main managers, consultants, and contractors can use the developed application to assess the risk of water main failure and to plan their rehabilitation works accordingly. The application provides a high level of flexibility to adapt to management preferences and the outlook of each authority.

VII.2. Conclusions

This research offers a HFES model to assess the risk of failure of water mains. During the course of the research, many points can be noticed and concluded such as:

- HFES model is recommended to assess the risk of failure associated with water main since it can deal with the vague and uncertain characteristics (factors) of the water main.
- From the collected questionnaire, it can be deduced that pipe age has the highest effect on risk of water main failure (100 units of global weight), followed by pipe material (75 units) and breakage rate (57 units).
- The more the data collected about the water main is, the more the HFES results accuracy is. However, it is so advisory to keep records and collect data about the

most important factors (most weighted) (i.e. Pipe age, material, breakage rate). This will result in more reliable management plans.

- The model is more sensitive toward the most weighted factors (i.e. pipe age, pipe material, and breakage rate).
- Due to the rules evaluation and aggregation, and to the defuzzification process, the model tends to produce results at some certain numbers and will not be distributed in a smooth curve, which is undesirable.
- It is difficult to make a change in the Risk of Failure value by only changing the performance of one factor, since the other fifteen factors try to resist the change in the risk values.
- Each of the four main branches of the hierarchy (environmental, physical, operational, and post failure) is sensitive to their own factors more than the risk of failure model is. This fact is due to the use of a hierarchical system where the farther the factor is in the hierarchy, the less its effectiveness (sensitivity) is.
- It can be deduced that the model is not very sensitive to the weight of the factors when changed within the factors' standard deviations.
- The Average Validity Percent of the model is 74.8 %, which means that the proposed model is valid for predicting the output. Moreover, the test shows that about 16.2 % of the data has an Average Invalidity Percent of more that 90%, and 23.4 % of the data fits between 80% and 90%.

VII.3. Research Contributions

The developed fuzzy expert model solved the problem of assessing the risk of failure associated with water mains. It contributes to the state of the art of sustainably managing water main infrastructure by achieving the following:

- A water main failure risk model.
- An automated tool (Excel-based application) that helps water main network managers build their short-term management plans and estimate their requirements for long-term plans.
- A failure risk scale that will provide guidance to decision makers to make the best-informed decisions.

VII.4. Limitations

VII.4.1. Model Limitations

The developed model uses hierarchical fuzzy expert system technique to assess the water main risk of failure. There are some limitations inherent in the model such as:

- The number of collected questionnaires is twenty. The model accuracy can be improved by increasing the number of experts involved in building the knowledge base rules of the fuzzy expert system.
- The model considers only eleven factors that indicate or contribute to the failure event and five factors that represent failure consequences.

- The input data membership functions in the model are limited to triangular and trapezoidal.
- A 95% confidence level is assumed as membership constant value for all of the linguistic input factors (i.e. pipe material, protection method, business disruption,...).
- The number of output membership functions (risk of failure membership functions) is limited to seven. Increasing this number will increase the accuracy of the model.
- The proposed risk of failure scale only gives some recommendation on how to manage the water mains at different risk stages. However, it is not built on a sound bases and can be improved in the future work.

VII.4.2. Application Limitations

An Excel-based application was built based on the developed model which uses hierarchical fuzzy expert system. However, there are some limitations to using this application:

- Even though the application leaves a flexible space for the user to choose among the sixteen considered factors, some other factors may come into the play, which are not considered in this model.
- Expert systems are built on expert opinions and thus the knowledge database represents the consensus among the experts -- however, some users may prefer to consider other opinions (e.g. to adapt for hot or cold climate or specific local

conditions). The flexibility to modify the knowledge database is therefore provided in the application which on the other hand causes the sensitivity analysis conducted in this research to be invalid.

- Due to technical restraints (computer resources), the input membership functions are limited to a maximum of five and the output membership functions are limited to seven.
- The developed application is based on MS Excel® version 2007, and thus it requires this software to operate and function.

VII.5. Recommendations and Future Works

Some of the recommendation and future works that can enhance the model and the research in general are listed below:

VII.5.1. Research Enhancement

The developed model can be enhanced by:

- More factors can be considered in the model; environmental, physical, operational factors and especially the consequence of failure factors.
- Some of the qualitative factors can be quantified. However, this step requires more effort to study in more details the different aspects of each factor.
- The consequences of failure factors may require more research since they have not conceived enough attention or understanding in the industry field and in the common practice.

- A detailed risk of failure scale and the associated corrective actions can be built that best utilize the results of the model to establish water main rehabilitation plans.

VII.5.2. Research Extension

Future efforts on this research could:

- Consider a third level of the hierarchy. This will lead to a better understanding and evaluation of the risk factors of the second level of the hierarchy (the sixteen risk factors considered in this research). As a result, the accuracy of the model output will be improved.
- Incorporate Geographic Information System in the research as the rehabilitation plans can also consider grouping water mains that are in the same area in order to more efficiently use allocated budgets.
- Adapt the HFES model to assess the risk of failure of different underground infrastructure such as transmission water main and sewer main.

Chapter VIII: REFERENCES

Al Barqawi, H. (2006). "Condition Rating Models for Underground Infrastructure: Sustainable Water-Mains." MS Thesis, Concordia University, Montreal, QC.

Bandyopadhyay, T., Dey, P., and Gupta, S. (1997). "A cost-effective maintenance program through risk analysis." *AACE International Transactions,* AACE, Dallas, TX, 84-89.

Bhave, P. (2003). *Optimal Design of Water Distribution Networks.* Alpha Science Int'l Ltd., Oxford.

Bogardi, I., and Duckstein, L. (2002). "The Fuzzy Logic Paradigm of Risk Analysis." *Risk-Based Decisionmaking in Water Resources*, ASCE, Santa Barbara, CA, 12-22.

Christodoulou, S., Aslani, P., and Vanrenterghem, A. (2003). "Risk Analysis Framework for Evaluating Structural Degradation of Water Mains in Urban Settings Using Neurofuzzy Systems and Statistical Modeling Techniques." *World Water and Environmental Resources Congress,* ASCE, Philadelphia, PA, 1-9.

Chughtai, F. (2007). "An Integrated Approach to Sewer Pipeline Condition Assessment." MS Thesis, Concordia University, Montreal, QC, Canada.

Cooke, R., and Goossens, L. (2004). "Expert judgement elicitation for risk assessments of critical infrastructures." *Journal of Risk Research* , 7(6), 643–656.

Cromwell, J., Reynolds, H., Pearson, N., and Grant, M. (2002). *Costs of Infastructure Failure,* Awwa Reseach Foundation and American Water Works Association, Washington DC.

Del Campo, E. (2004). "Design of Fuzzy Systems Using Knowledge Extracted Via Neural Networks." MS Thesis, University of Texas, El Paso, TX.

Dillon Consulting, L., and harfan Technologies, I. (2003). *Watermain Transmission and Distribution System Needs Assessment Study,* City of Moncton, Moncton.

Ezell, B., Farr, J., and Wiese, I. (2000). "Infrastructure Risk Analysis Model." *Journal of Infrastructure Systems* , 6(3), 114-117.

Gentile, M. (2004). "Development of a Hierarchical Fuzzy Model for the Evaluation of Inherent Safety." PhD Dissertation, Texas A&M University, Texas.

Giustolisi, O., Laucelli, D., and Savic, D. (2006). "Development of rehabilitation plans for water mains replacement considering risk and cost-benefit assessment." *Civil Engineering and Environmental Systems* , 23(3), 175–190.

Hahn, M., Palmer, R., and Merr, M. S. (2002). "Expert System for Prioritizing the Inspection of Sewers:Knowledge Base Formulation and Evaluation." *Journal of Water Resources Planning and Management* , 128(2), 121-129.

Haimes, Y. (2004). *Risk Modeling, Assessment, and Management,* 2nd ed., Wiley-IEEE, Hobocen, New Jersey.

Huanga, C., and Moraga, C. (2005). "Extracting fuzzy if–then rules by using the information matrix technique." *Journal of Computer and System Sciences* , 70, 26-52.

InfraGuide, N. G. (2003, April). "Deterioration and Inspection of Water Distribution Systems." *A Best Practice by the National Guide to Sustainable Municipal*

Infrastructure, <http://www.infraguide.ca/bestPractices/default_e.asp> (Jan. 14, 2007).

InfraGuide, N. G. (2006, July). "Managing Risk." *Decision Making and Investment Planning,* <http://www.infraguide.ca/bestPractices/PublishedBP_e.asp#dmip> (Jan. 17, 2007).

Infrastructure Report (2007). "The Water Main Break Clock" *Canada Free Press,* <http://www.canadafreepress.com/infrastructure.htm> (Jan. 23, 2008).

Jin, Y. (2003). *Advanced Fuzzy Systems Design and Applications.* A Springer-Verlag Company, New York.

Karray, F., and de Silva, C. (2004). *Soft Computing and Intelligent Systems Design.* Pearson Education Limited Essex, England.

Kirchhoff, D., and Doberstein, B. (2006). "Pipeline risk assessment and risk acceptance criteria in the State of São Paulo, Brazil." *Impact Assessment and Project Appraisal*, 24(3), 221–234.

Kleiner, Y., and Rajani, B. (2000). "Considering Time-dependent Factors in the Statistical Prediction of Water Main Breaks." *AWWA Infrastructure Conference,* Baltimore, MD, 1-12.

Kleiner, Y., and Rajani, B. (2002). "Forecasting Variations and Trends in Water-Main Breaks." *Journal of Infrastructure Systems*, 8(4), 122-131.

Kleiner, Y., Sadiq, R., and Rajani, B. (2004). "Modeling Failure Risk in Buried Pipes Using Fuzzy Markov Deterioration Process." *Pipeline Engineering and Construction: What's on the Horizon?,* ASCE, San Diego, CA, 1-12.

Kleiner, Y., Rajani, B., and Sadiq, R. (2006). "Failure risk management of buried infrastructure using fuzzy-based techniques." *Journal of Water Supply Research and Technology : Aqua* , 55(2), 81-94.

Lee, M. (2006). "Risk Assessment of Drinking Water supply in Cananda." MS Thesis, University of Guelph, Ontario.

Lee, M.-L., Chung, H.-Y., and Yu, F.-M. (2003). "Modeling ofhierarchical fuzzy systems." *Fuzzy Sets and Systems* , 138(2), 343–361.

Marshall, D., Zarghamee, M., Mergalas, B., and Kleiner, Y. (2005). "Tarrant Regional Water District's Risk Management Plan for Prestressed Concrete Cylinder Pipe." *The Pipeline Division Specialty Conference,* ASCE, Houston, TX, 854-861.

Moglia, M., Burn, S., and Meddings, S. (2006). "Decision support system for water pipeline renewal prioritisation." *ITcon , 11* (Special Issue Decision Support Systems for Infrastructure Management), 237-256.

Mohitpour, M., Golshan, H., and Murray, A. (2003). *Pipeline Design and Construction: A practical Approach,* 2nd ed., ASME Press, New York.

Muhlbauer, W. K. (2004). *Pipeline Risk Management Manual,* 3rd ed., Gulf Professional Publishing, Burlington.

Najafi, M. (2005). *Trenchless Technology: Pipeline and Utility Design, Construction, and Renewa,.* The McGraw-Hill Companies, New York.

Najjaran, H., Sadiq, R., and Rajani, B. (2004). "Modeling pipe deterioration using soil properties - An application of fuzzy logic expert system." *Pipeline Engineering and Construction: What's on the Horizon?,* ASCE, San Diego, CA, 1-10.

Nasiri, F., Maqsood, I., Huang, G., and Fuller, N. (2007). "Water Quality Index: A Fuzzy River-Pollution Decision Support Expert System." *Journal of Water Resourses Planning and Management*, 133(2), 95-105.

Nikolaidis, E., Ghiocel, D., and Singha, S. (2005). *Engineering design reliability handbook,* CRC Press, Florida.

Oppenheimer, P. (2004). *Top-Down Network Design,* 2nd ed., Cisco Press, New York.

Ozger, S. S. (2003). "A semi-pressure-driven approach to reliability assessment of water distribution networks." PhD Dissertation, Arizona State University, Arizona.

Pelletier, G., Mailhot, A., and Villeneuve, J.-P. (2003). "Modeling Water Pipe Breaks—Three Case Studies." *Journal of Water Resources Planning and Management 10*, 129(2), 115–123.

Pickard, M. (2007). "The Economics of Pipeline Repair" *Clock Spring Company, L.P.*, <www.clockspring.com/quarterly%2520072006/The%2520Economics%2520of%2520Pipeline%2520Repair.pdf+cost+repair+of+pipelineandhl=enandct=clnkandcd=3> (May 15, 2007).

Rahman, S. M. (2007). "Condition Assessment Tool For Elements of Drinking Water Treatment Plant." MS Thesis, Concordia University, Montreal, QC.

Rajani, B., Kleiner, Y., and Sadiq, R. (2006). "Translation of pipe inspection results into condition ratings using the fuzzy synthetic evaluation technique." *Journal of Water Supply Research and Technology: Aqua*, 55(1), 11-24.

Raven, A. (2007). "Risk Factors of Structural Degradation of an Urban Water Distribution System." *Journal of Infrastructure Systems*, 13(1), 55-64.

Rogers, P. (2006). "Failure Assessment Model to Prioritize Pipe Replacement in Water Utility Asset Management." Phd Dissertation, Colorado State University, Colorado.

Sadiq, R., Rajani, B., and Kleiner, Y. (2004). "Probabilistic risk analysis of corrosion associated failures in cast iron water mains." *Reliability Engineering and System Safety*, 86(1), 1-10.

Santosh, Vinod, G., Shrivastava, O., Saraf, R., Ghosh, A., and Kushwaha, H. (2006). "Reliability analysis of pipelines carrying H2S for risk based inspection of heavy water plant." *Reliability Engineering and System Safety*, 91, 163–170.

Selvakumar, A., Clark, R., and Sivaganesan, M. (2002). *Costs for Water Supply Distribution System Rehabilitation.* Environmental Protection Agency, New Jersy.

Shaheen, A. A. (2005). "A Framework for Integrating Fuzzy Set Theory and Discrete Event Simulation in Construction Engineering." PhD Dissertation, University of Alberta, Edmonton.

Shi, Y., and Sen, P. (2000). "A New Defuzzification Method for Fuzzy Control of Power Convertors." *IEEE*, 2(9), 1202-1209.

Shi, Y., Mizumoto, M., and Shi, P. (2002). "Fuzzy IF-THEN Rule Generation Based on NeuralL Network and Clustering Algorithm Techniques." *TENCON '02. Proceedings,* IEEE, Beijing, China, 651-654.

Skipworth, P., Cashman, A., Engelhardt, M., Saul, A., and Savic, D. (2001). "Quantification of Mains Failure Behavior in a Whole Life Costing Approach to Distribution System Management." *World Water and Environmental Resources Congress 2001,* ASCE, Orlando, FL, 1-11.

Souza, R., and Chagas, P. (2001). "Risk Analysis in Water Quality Problems." *World Water and Environmental Resources Congress,* ASCE, Orlando, FL, 1-6.

Tanaka, K. (1997). *An Introduction to Fuzzy Logic for Practical Applications,* Springer, Berlin.

Tettamanzi, A., and Tomassini, M. (2001). *Soft Computing: Integrating Evolutionary, Neural, and Fuzzy Systems,* Springer, Berlin.

Vinod, G., Bidhar, S., Kushwaha, H., Verma, A., and Srividya, A. (2003). "A comprehensive framework for evaluation of piping reliability due to erosion–corrosion for risk-informed inservice inspection." *Reliability Engineering and System Safety* , 82(2), 187-193 .

Yan, J., and Vairavamoorthy, K. (2003). "Fuzzy Approach for Pipeline Condition Assessment." *New Pipeline Technologies, Security, and Safety,* ASCE, Baltimore, MD, 467-476.

Yuhua, D., and Datao, Y. (2005). "Estimation of failure probability of oil and gas transmission pipelines by fuzzy fault tree analysis." *Journal of Loss Prevention in the Process Industries* , 18, 83–88.

Zayed, T. and Halpin, D. (2005). "Pile Construction Productivity Assessment." *Journal of Construction Engineering and Management,* 131(6), 705-714.

Zayed, T. (2005). "Selecting Pile Construction Method Using Fuzzy Approach." *Construction Research Congress 2005,* ASCE, San Diego, CA, 1-10.

APPENDIX A: INTRODUCTION TO FUZZY EXPERT SYSTEM

A.1. Fuzzy Logic

L.A. Zadel developed fuzzy logic in the mid-1960s to solve the problem of representing approximate knowledge that cannot be represented by conventional, crisp methods. A fuzzy set is represented by a membership function. Any "element" value in the universe of enclosure of the fuzzy set will have a membership grade which gives the degree to which the particular element belongs to the set (Karray and de Silva, 2004). Fuzzy theory relies on four main concepts: (1) ***fuzzy sets***: sets with non-crisp, overlapping boundaries; (2) ***linguistic variables***: variables whose values are both qualitatively and quantitatively described with fuzzy sets; (3) ***possibility distributions***: constraints on the value of a linguistic variable imposed by assigning it a fuzzy set; and (4) ***fuzzy if-then rules***: a knowledge representation scheme for describing a functional mapping or a logic formula that generalizes two-valued logic (Del Campo, 2004).

A.2. Fuzzy Sets

A.2.1. Introduction to Fuzzy Sets

The limitation of classical set theory is that a characteristic function that describes a classical (crisp) set can only assume 0 or 1. This can be represented as

$$\emptyset_A(x) = \begin{cases} 1, & if\ x \in A \\ 0, & if\ x \notin A \end{cases} \quad \text{Equation A.1}$$

In classical set theory, a definition of a concept (set) does not admit degrees. However, by allowing the characteristic function to take a value between 0 and 1, this limitation and difficulty will be removed. This can be represented by Equation A.2 where A is a set of universe U, described by the characteristic function

$$\mu_A(x): U \rightarrow [0,1] \quad \text{Equation A.2}$$

for any $x \in U, \mu_A(x) \in [0,1]$ is a function that specifies the degree to which element x belongs to set A. Set A is called a fuzzy set and the characteristic function $\mu_A(x)$ is called a function. (Jin, 2003). A fuzzy set is a set without clear or sharp (crisp) boundaries or with no binary membership characteristics. In a fuzzy set, partial membership is possible -- unlike an ordinary set where each object either belongs or does not belong to the set. A simple example that explains this concept is the variable "temperature", which easily takes a fuzzy value (e.g., cold, cool, tepid, warm, hot) (Karray and de Silva, 2004). Fuzzy sets are suitable for describing sets whose boundaries are not sharply defined. It provides an effective way of dealing with uncertainties other than the probability theory (Jin, 2003).

A.2.2. Types of Fuzzy Sets

There are two types of fuzzy sets; discrete and continuous. If the universe of enclosure is discrete with elements x_i, then the fuzzy set is notated as shown in Equation A.3 in which each element is paired with its grade of membership.

$$A = \frac{\mu_A(x_1)}{x_1} + \frac{\mu_A(x_2)}{x_2} + \cdots + \frac{\mu_A(x_i)}{x_i}$$ Equation A.3

$$A = \sum_{x,\in U} \frac{\mu_A(x_i)}{x_i}$$

If the universe of enclosure is continuous, then the notation is given as an integration symbol:

$$A = \int_{x,\in U} \frac{\mu_A(x_i)}{x_i}$$ Equation A.4

It is important to note that these two notations are symbolic shorthand forms of notation and are not real summations or integrations (Karray and de Silva, 2004).

A.3. Fuzzy Operations

Several methods are available to define the intersection and the union of fuzzy sets. In this context, only the classical methods proposed by Zadeh (the inventor of fuzzy logic) are introduced because of their simplicity and the analogy with crisp sets. They are complement, union, and intersection (Karray and de Silva, 2004).

A.3.1. Complement Operation

The complement operation corresponds to negation. The complement is given in the following equation and shown graphically in Figure A.1 (Karray and de Silva, 2004):

$$\mu_A{}'(x_i) = 1 - \mu_A(x_i) \text{ for all } x \in U$$ Equation A.5

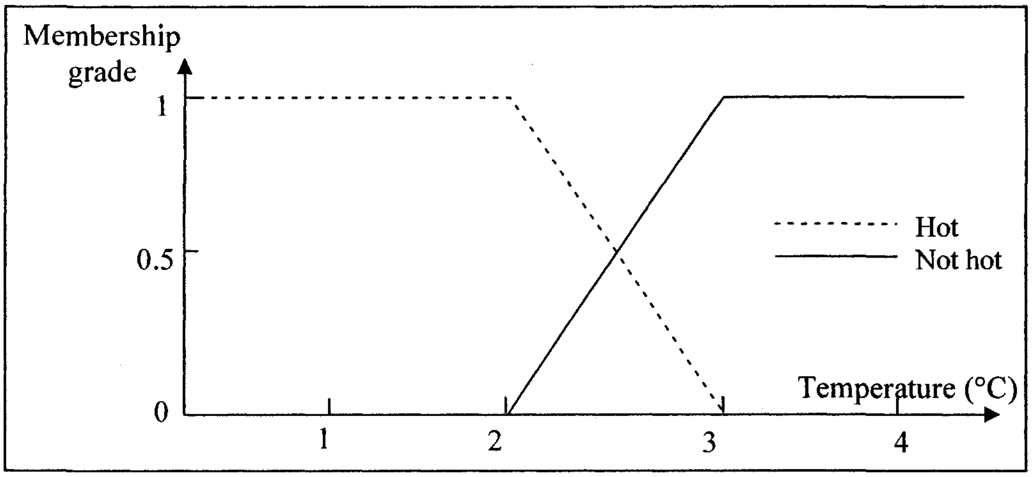


Figure A.1 – Representation of a Complement fuzzy logic operation (Karray and de Silva, 2004).

A.3.2. Union Operation

The union corresponds to a logical OR operation (called Disjunction), and is denoted by $A \cup B$, where A and B are fuzzy sets or fuzzy propositions. The union operation is shown in the equation below, and Figure A.2 shows an example (Karray and de Silva, 2004).

$$\mu_{A \cup B}(x) = max[\mu_A(x), \mu_B(x)] \quad \forall\, x \in U \qquad \text{Equation A.6}$$

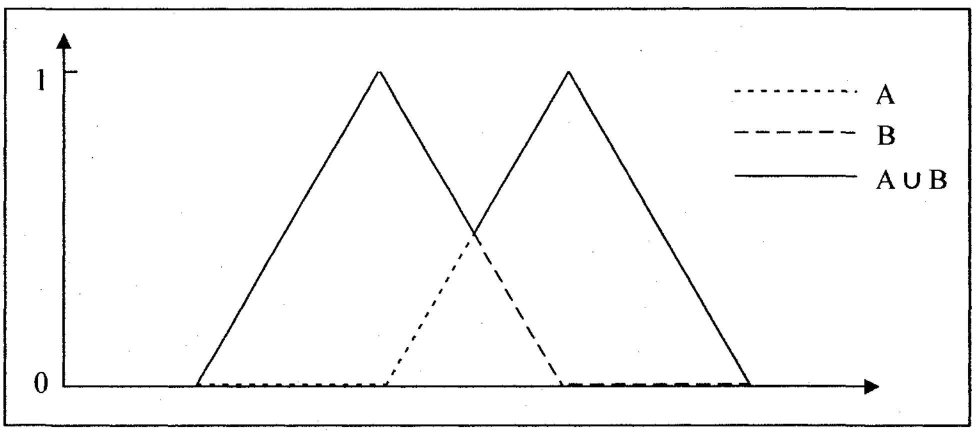


Figure A.2 – Representation of a Union fuzzy logic operation (Karray and de Silva, 2004).

A.3.3. Intersection Operation

The intersection operation corresponds to a logical AND operation (called Conjunction) and is designated as $A \cap B$, where A and B are fuzzy sets or fuzzy propositions. The intersection operation is given in the equation below, and a graphical representation is shown in Figure A.3 (Karray and de Silva, 2004).

$$\mu_{A \cap B}(x) = min[\mu_A(x), \mu_B(x)] \quad \forall\, x \in U \qquad \text{Equation A.7}$$

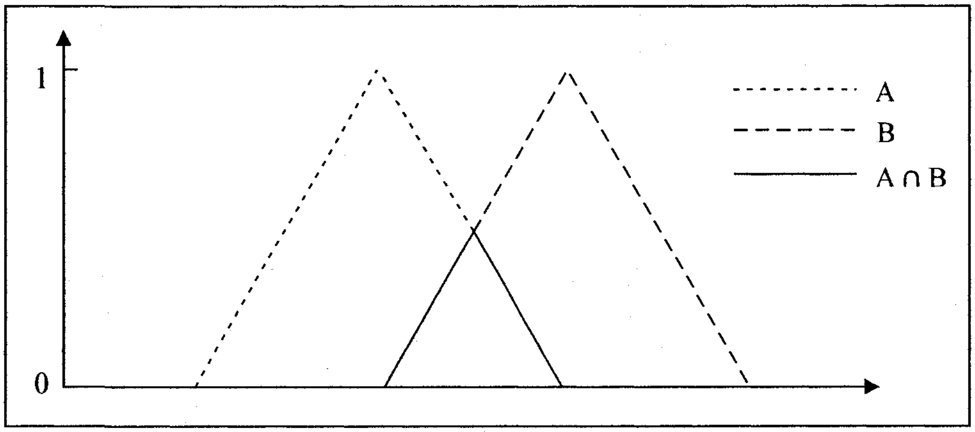


Figure A.3 – Representation of an Intersection fuzzy logic operation (Karray and de Silva, 2004).

A.4. Fuzzy Membership Functions

A membership function is what maps the input space to the output space. It is needed to smooth the transition between two regions of memberships; the region completely inside the set and the region completely outside the set. There are many forms of membership functions, such as: triangle, trapezoid, bell curve, Gaussian, and sigmoid functions (Del Campo, 2004). These are not the only available membership functions -- there are many others but these are the principle functions.

A.4.1. Triangular Membership Function

It is the simplest form of membership function. It requires only three parameters to be defined. Mathematically, it can be represented in the following equation (Del Campo, 2004):

$$\mu_A(x) = \begin{cases} 0 & x < a \\ \left(\frac{x-a}{b-a}\right) & a \le x \le b \\ \left(\frac{c-x}{c-b}\right) & b \le x \le c \\ 0 & x > c \end{cases} \qquad \text{Equation A.8}$$

The following figure represents a triangle membership function with $a = 3$, $b = 5$, and $c = 7$.

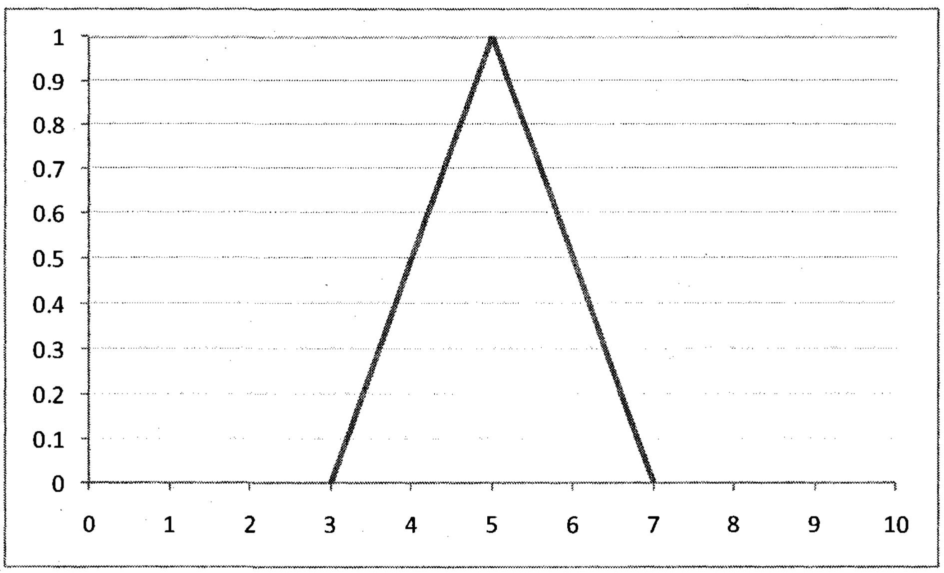


Figure A.4 – Triangular membership function.

A.4.2. Trapezoidal Membership Function

A trapezoidal membership function is characterized by four parameters (*a*, *b*, *c*, *d*) and is represented by the following equation (Del Campo, 2004):

$$\mu_A(x) = \begin{cases} 0 & x < a \\ \left(\frac{x-a}{b-a}\right) & a \le x \le b \\ 1 & b < x < c \\ \left(\frac{c-x}{c-d}\right) & c \le x \le d \\ 0 & x > d \end{cases}$$

Equation A.9

The following figure represents a trapezoidal membership function with $a = 2$, $b = 4$, $c = 6$, and $d = 8$.

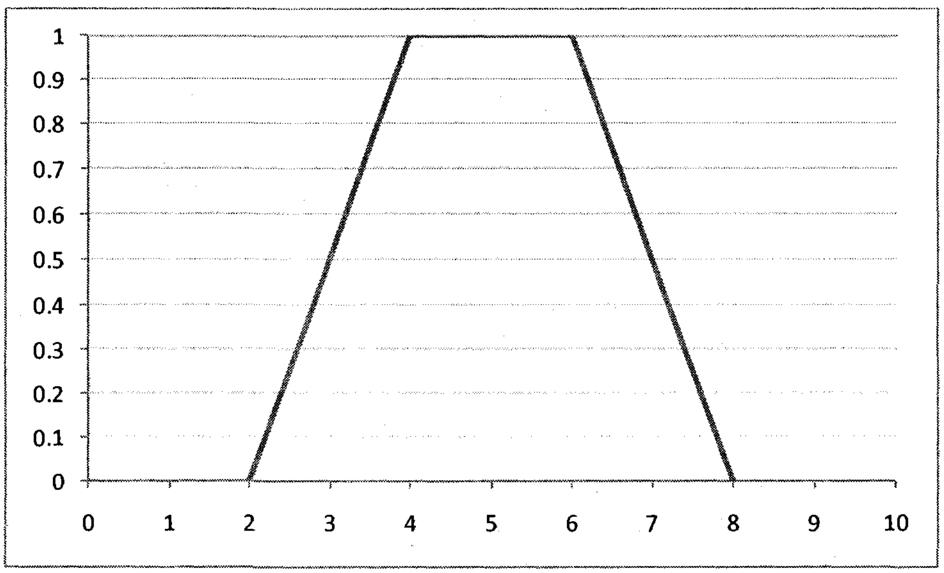


Figure A.5 – Trapezoidal membership function.

A.4.3. Gaussian Membership Function

A Gaussian membership function is defined only by two variables (c, σ), and is represented by the following equation (Del Campo, 2004):

$$\mu_A(x) = e^{-\frac{1}{2}\left(\frac{x-c}{\sigma}\right)^2} \qquad \text{Equation A.10}$$

Graphical representation of the Gaussian membership function is shown in the following figure where $c = 5$ and $\sigma = 1.25$.

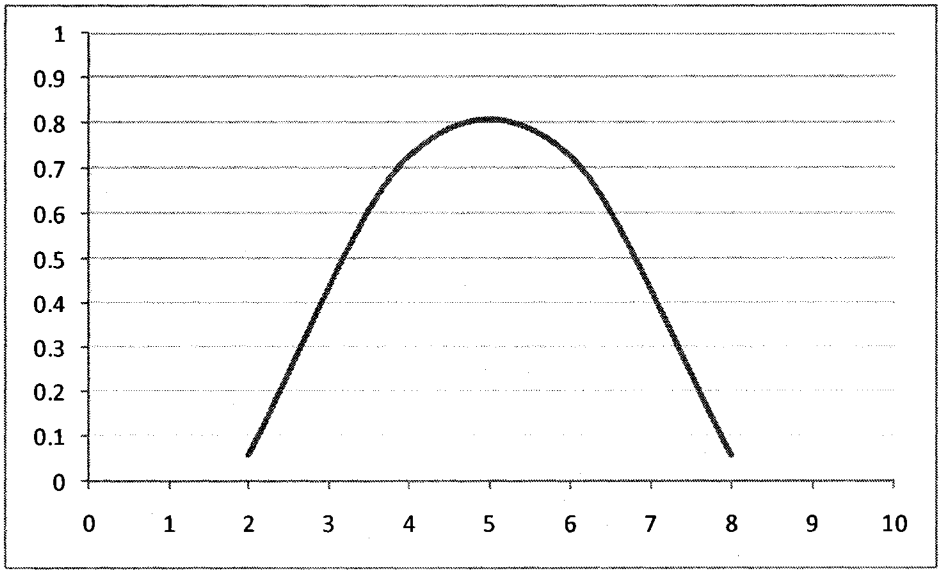


Figure A.6 – Gaussian membership function.

A.4.4. Generalized Bell Shape Membership Function

Bell shaped functions have many different forms. The one most commonly used is the Generalized Bell-shaped membership function. It has three control parameters: a controls the width of the function, b controls the slope, and c controls the center of the function (Del Campo, 2004). It is represented by the following equation and shown graphically as an example in Figure A.7 where $a = 1$, $b = 2$, and $c = 5$.

$$\mu_A(x) = \frac{1}{1 + \left|\frac{x-c}{a}\right|^{2b}} \qquad \text{Equation A.11}$$

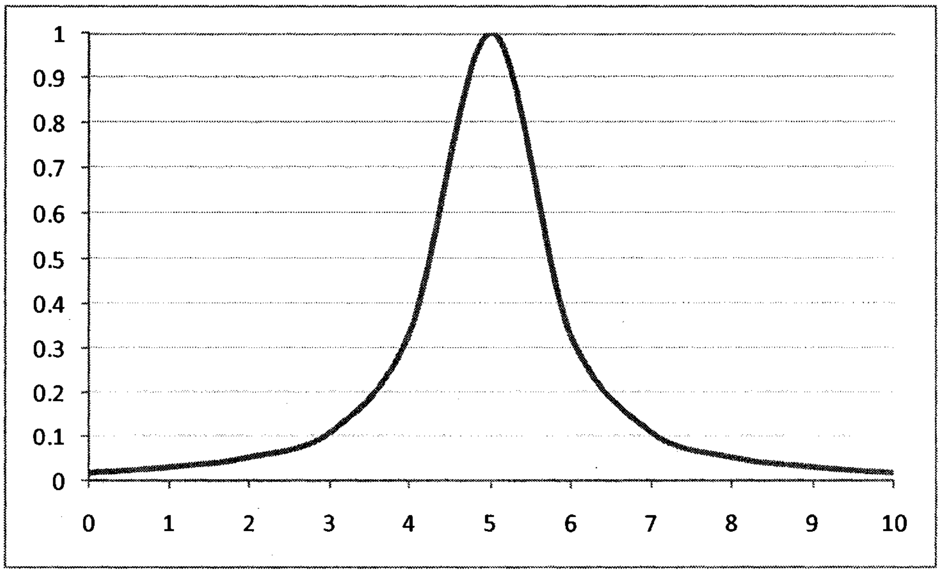


Figure A.7 – Generalized bell shape membership function.

A.4.5. Sigmoid Membership Function

Only two parameters are needed to define a sigmoid membership function. Parameter a determines the slope of the membership function and parameter c controls the shift of the sigmoid function. It can be represented by the following equation (Del Campo, 2004).

$$\mu_A(x) = \frac{1}{1 + e^{-a(x-c)}} \qquad \text{Equation A.12}$$

Graphically, it can be drawn as an example in the following figure where $a = 2$ and $c = 4$.

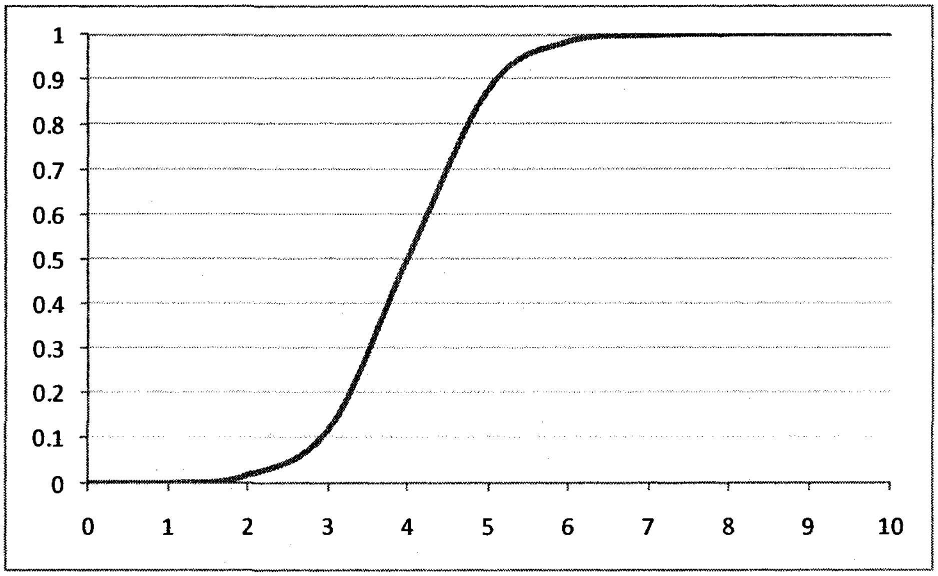


Figure A.8 – Sigmoid membership function.

A.5. Fuzzy Rule System

A rule system consists of a number of rules with a condition part and an action part:

If condition is x, then action is y.

The condition part is also known as the rule premise, or simply the IF part. The action part is also called the consequent part or the THEN part. A fuzzy rule system uses linguistic variables in the if-then relationship. Linguistic variables were defined by Zadeh as follows: "A linguistic variable is a variable whose values are sentences in a natural or artificial language". For examples, when the values of x are small, middle, large, young, not very young, old, then x is a linguistic variable. Generally, each of the fuzzy sets corresponds to one linguistic variable and this collection of fuzzy sets is called the "fuzzy partition" (the number of fuzzy sets). However, assigning a fuzzy membership set or function to a linguistic variable is a challenging matter and generally there are three methods (Jin, 2003):

- Subjective evaluation and heuristics. The membership function of fuzzy sets can be determined based on the experience or intuition of human beings since fuzzy sets are intended to model the cognitive process of human beings.
- Converted frequencies or probabilities. Membership functions can sometimes be constructed on the basis of frequency histograms or other probability curves. There are a variety of conversion methods, each with its own mathematical and methodological strengths and weaknesses.
- Learning and adaptation. Parameters of fuzzy membership functions can be learned or adapted using different optimization methods based on a set of the training data. The gradient method and genetic algorithms or reinforcement learning are a few examples. This method is the most sophisticated and objective method for the determination of membership functions (Jin, 2003).

A.6. Fuzzy Reasoning Systems

Fuzzy reasoning is expressed in the IF-THEN rules format discussed above. There are only a few types of fuzzy IF-THEN rules (reasoning). Fuzzy reasoning is classified roughly into two methods: direct and indirect. The direct method is the most popular, whereas the indirect method conducts the reasoning using truth-value space which has a relatively complex reasoning mechanism (Tanaka, 1997). This classification is shown in Figure A.9.

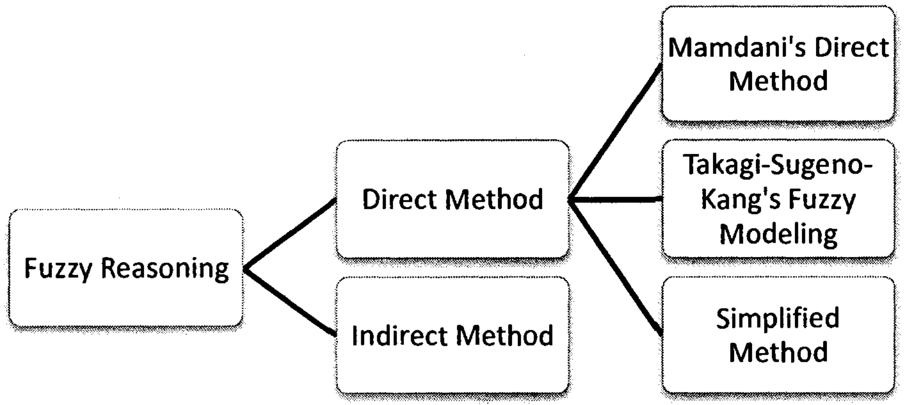


Figure A.9 – Classification of fuzzy reasoning.

A.6.1. Mamdani Method

The Mamdani method is based on a simple structure of Max and Min operations.

$$\textbf{If } x = A_1 \textbf{ and } y = B_1 \ldots \textbf{ THEN } z = C_1$$

Equation A.0.13

where x and y are input variables, A_1, and B_1 are fuzzy input linguistic values, and z is an output variable with C_1 the fuzzy output linguistic values. The Mamdani method has an advantage over the Takagi-Sugeno-Kang (TSK) method because it is easier to understand and the output of the system is defined in terms of fuzzy sets (Gentile, 2004).

A.6.2. Takagi-Sureno-Kang (TSK) Method

The TSK model is another version of the Mamdani method. Its rules' form is given as

$$\textbf{If } x = A_1 \textbf{ and } y = B_1 \ldots \textbf{ THEN } z = f(x,y,\ldots)$$

Equation A.14

where f can be any function of the input variables taking values in the output variables range. The result of the TSK method is a crisp number computed as the average of the outputs of the single rules weighted by the degrees of truth of their antecedents (Tettamanzi and Tomassini, 2001).

A.7. Defuzzification Methods

This is the last component of a fuzzy logic system. A defuzzification process is needed to convert the fuzzy output of fuzzy rules to a crisp value. There are many defuzzification strategies that can be followed to produce a crisp output. Some of them are shown below (Shi and Sen, 2000):

- Center of Area: this is the center of gravity of the output membership function.
- Center of Sum: this method ignores the union operation of membership functions. It calculates the center of gravity of each function individually and then average weights them by their areas. Thus, it is a faster defuzzification process than Center of Area.
- Height Method: the center of gravity of each membership function for each rule is first calculated and then average-weighted by their heights.
- Middle of Maxima: is the mean value of all the membership means whose membership values reach the maximum.
- First of Maxima: this uses the union of membership functions and takes the smallest value of the range with the maximum membership degree.

A.8. Use of fuzzy logic in expert systems

A.8.1. Introduction

Usually, systems that can process knowledge are called knowledge-based systems. One of the most popular and successful knowledge-based systems is the expert system. Knowledge can be represented by several forms, such as the logical knowledge representation, the procedural knowledge representation, the network knowledge representation and the structured knowledge representation. In the logical knowledge representation, knowledge is represented by an expression in formal logic. In the procedural knowledge, knowledge is described by a set of instructions or rules which can be interpreted as a procedure that achieves a goal for a given argument. Both the network and the structured knowledge representation schemes represent knowledge using graphs (Jin, 2003). Fuzzy logic can be used as a tool to deal with imprecision and qualitative aspects that are associated with problem solving and in the development of expert systems. Fuzzy expert systems use the knowledge of humans, which is qualitative and inexact. In many cases, decisions are to be taken even if the experts may be only partially knowledgeable about the problem domain, or when data may not be fully available. The reasons behind using fuzzy logic in expert systems may be summarized as follows (Karray and de Silva, 2004):

- The knowledge base of an expert system summarizes the human experts' knowledge and experience.

- Fuzzy descriptors (e.g., large, small, fast, poor, fine) are commonly used in the communication of experts' knowledge, which is often inexact and qualitative.
- The user's problem description may not be exact.
- Reasonable decisions are to be taken even if the experts' knowledge base may not be complete.
- Educated guesses need to be made in some situations.

A.8.2. Fuzzy Knowledge Rules Acquisition

Fuzzy if-then rule systems are most widely used in fuzzy knowledge representation and processing. A fuzzy knowledge system consists of a set of rules such as:

$$\text{If } x = A_1 \text{ and } y = B_1 \ldots \text{ THEN } z = C_1$$

$$\text{If } x = A_2 \text{ and } y = B_2 \ldots \text{ THEN } z = C_2$$

$$\text{If } x = A_n \text{ and } y = B_n \ldots \text{ THEN } z = C_n$$

where x, y are input variables, $A_{(1 \text{ to } n)}$, $B_{(1 \text{ to } n)}$ are fuzzy input linguistic values, z is an output variable, and $C_{(1 \text{ to } n)}$ are the fuzzy output linguistic values.

Acquiring knowledge for fuzzy rule base systems can be achieved from human experts or from experimental data using several methods. Mainly, there are three different approaches (Jin, 2003):

- Indirect Knowledge Acquisition. The designer of the knowledge-based system is not an expert and usually gathers the necessary knowledge from an expert or an experienced operator by various means, such as interviews or questionnaires.
- Direct Knowledge Acquisition. Here, the designer is an expert. The designer has rich knowledge in the related field is also able to formulate his/her knowledge in a proper fashion so that it correctly reflects the system.
- Automatic Knowledge Acquisition. Most automatic knowledge acquisition methods are developed in the field of machine learning and artificial intelligence. Specific techniques are used, such as neural networks.

A.8.3. Building Fuzzy Expert Systems

General rules can be followed in order to build a fuzzy expert system. These can be summarized as follows (Jin, 2003; Zayed, 2005):

1. Determination of the input and the output.
2. Determination the linguistic terms and their corresponding fuzzy membership functions for the input and output variables. It is necessary to determine the universe of discourse, the number of fuzzy sets and the associated fuzzy membership functions in the fuzzy partitions.
3. Extraction of fuzzy rules from expert knowledge and common sense following direct, indirect, and automatic knowledge acquisition.
4. The output can be aggregated into a single output number using one of the aggregation methods.

A.9. Fuzzy Rules Generation Techniques

As the number of inputs and outputs increase, the complexity of the fuzzy system increases and the knowledge-based rules extraction process becomes more complicated. Thus, more effort by researchers is needed to be exerted to solve this disadvantage of fuzzy logic system development. In one effort, a methodology was developed to generate fuzzy rules depending on the aggregation of the effect of the factor's "relative importance" within the rule block and its "impact on the output" and then normalizing these aggregations into the consequent part of the fuzzy rule (Shaheen, 2005). Another methodology was developed to extract the fuzzy rules from data, including noise, using unsupervised learning with normal information diffusion, called the Information Matrix Technique. This method requires the availability of data and does not depend on experts' opinions in any way (Huanga and Moraga, 2005). Another method extracts the fuzzy rules using neural network and clustering algorithm techniques. However, this method also requires the availability of data in order to extract the fuzzy rules from it (Shi *et al.* 2002; Del Campo, 2004).

A.10. Hierarchical fuzzy expert system

Reducing the total number of rules and their corresponding computation requirements is considered one of the important issues in subjective fuzzy logic systems where the knowledge base rules are solicited from experts in contrast to the objective fuzzy system where the rules are extracted from data. The "Curse of dimensionality" is an attribute of subjective fuzzy systems since the number of rules and thus the complexity increases exponentially with the number of variables involved in the model. To solve this problem (Curse of dimensionality), the hierarchical fuzzy system is proposed where the system is divided into a number of low-dimensional fuzzy systems. This has the advantage that the total number of rules increases linearly with the number of input variables. The number of rules is greatly reduced by using a hierarchical fuzzy system (Lee *et al.* 2003). Mainly, there are many approaches to deal with the output of one layer to be the input of the next layer. One is that the output of the last layer as a crisp value can be used as the input of the next layer in the hierarchical fuzzy system. The advantage of this approach is that it will reduce the uncertainty of the new result by reducing the number of the fired rules in the new layer, but at the expense of the information of uncertainty, which is lost. Another approach is to consider the fuzzy output of the last layer as the fuzzy input of the next layer, which would preserve the information about uncertainty. However, if the fuzzy set is too wide, it will fire too many rules in the new layer resulting in a very uncertain result. Another approach is to decompose the defuzzification of the output that is used as input in the new layer into two or more crisp singletons. A different approach is to use only a part of the fuzzy output with a membership degree greater than (0.4) and convert the result into a triangle membership function whose base values resulted from the membership degree cut as shown in Figure A.10 (Gentile, 2004).

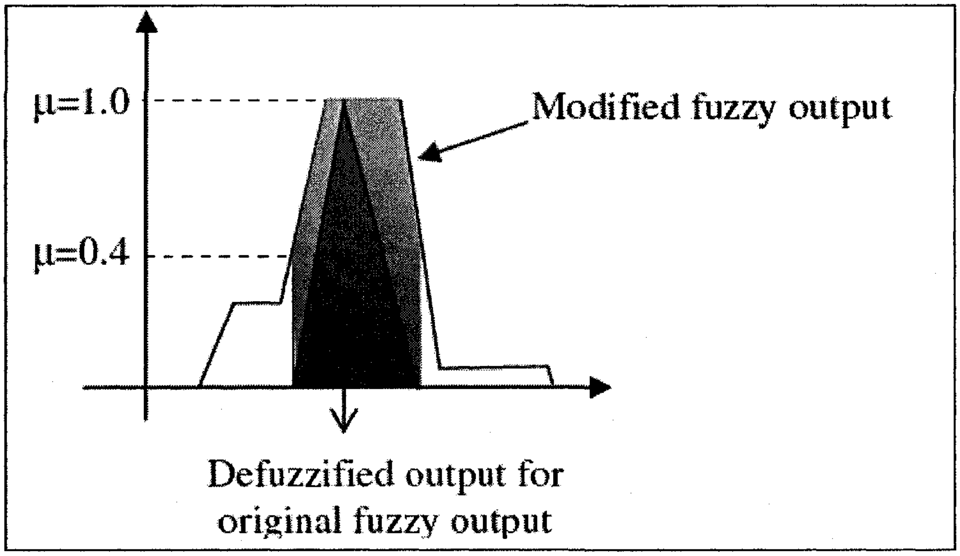


Figure A.10 – Fuzzy output used as input for the new layer.

A.11. Fuzzy Logic Advantages

Fuzzy logic has many advantages which make it a suitable technique for modeling and control problems, some of these advantages are (Lee, 2006):

1) It does not require precise inputs and the output control is a smooth control function in spite of the broad variations of the inputs.

2) Fuzzy logic is not limited to a certain number of inputs and outputs.

3) It can deal with information that would be difficult or impossible to model mathematically.

4) It is easily adjusted by simply changing the rules.

5) It saves time when compared to conventional mathematical methods.

A.12. Fuzzy Logic Disadvantages

Although Fuzzy Logic has many advantages over conventional mathematical methods, some limitations are inherent to fuzzy logic such as (Lee, 2006):

1) It lacks self-organizing and self-tuning mechanisms.

2) The knowledge base rules definition and the fuzzy system quickly becomes complex when too many inputs and outputs are used in developing a fuzzy model.

APPENDIX B: SAMPLE QUESTIONNAIRE

1. Cover Page

Risk of water main Failure Fuzzy Expert System

Risk of failure is defined as the combination of the probability and impact severity of a particular circumstance (failure) that negatively impacts the ability of infrastructure assets to meet the objectives of the municipality. Several factors play roles in water mains pipelines risk of failure. These factors are classified in this research as environmental, physical, operational, and post-failure factors. The identification of the weights and effects of these factors on water-mains risk of failure is crucial to identify the most risky water-main pipelines and to take the suitable measures to mitigate their risks. The expert opinions gathered by this questionnaire will be used in building a fuzzy expert system to predict the risk of failure index of the network pipelines.

As the expert system is mainly dependent on experts' judgment and experience, we prepared this questionnaire trying to translate and integrate your valuable judgment into our expert system. This questionnaire consists of three parts. In the first part, general information about the expert is collected. In the second part, the expert is required to give weights to the factors considered in our expert system. In the third part, the expert is asked to evaluate the performance variables (attributes) of the factors.

Your cooperation with us to advance the knowledge of water-mains infrastructure is highly appreciated.

Supervisor,

Tarek Zayed, Ph.D., P.E.

Associate Professor

Department of Building, Civil & Environmental Engineering
EV 6.401, 1515 Ste. Catherine St., Montreal, Canada H3G 1M8
Tel.: (514) 848-2424 ext. 8779
Fax: (514) 848-7965
Email: zayed@bcee.concordia.ca

Information Return:

Please, return this questionnaire to

Hussam Fares

Research Assistant,
Department of Building, Civil & Environmental Engineering,
Concordia University

Tel.:
E-mail:

---

2. Questionnaire

The following table is confidential and not to be exposed to public.

1 – Name:	
2 – Institution:	
3 – Title:	
4 – No. of years of experience:	

The following table collects the weight of the risk of water main failure factors. This can be the answer to the question "What is the strength of the factor in contributing to the failure event?" Please use a scale from 0 to 100.

Risk Factor	Factor Weights	Risk sub-factor	Sub-factors weights
Environmental Factors		Soil Type	
		Daily Traffic	
		Water Table Level	
Physical Factors		Pipe Material	
		Pipe Diameter	
		Pipe Age	
		Protection Method	
Operational Factors		Breakage Rate	
		Hydraulic Factor	
		Water Quality	
		Leakage	
Post-Failure Factors		Cost of Repair	
		Damage to surroundings	
		Loss of Production	
		Traffic Disruption	
		Type of Serviced Area	

The behaviors of only three factors of the factors listed above are not known to me and I am strongly in need of possession of this information. For each risk sub-factors attributes, please choose a number using this scale:

1	2	3	4	5	6	7
Extremely low	Very low	Moderately low	Medium	Moderately high	Very high	Extremely high

Risk Factor	**Risk Sub-factors**	**Sub-factor Attributes Criteria**	**Consequence**
Enviro. Factor	Water Table Level	rarely present	
		seasonally present	
		always present	
Post Failure Factor	Damage to surroundings	Industrial area	
		Commercial area	
		Residential area	
	Loss of Production	<= 250 mm (redundant)	
		between 250 to 500 mm (redundant)	
		<= 250 mm (not redundant)	
		between 250 to 500 mm (not redundant)	

APPENDIX C: STABILITY ANALYSIS RESULTS

	Notes	Environmental			Physical				Operational				Consequences						Results				
ID number	Scenarios	Type of Soil	Average Daily Traffic	Water Table Level	Type of Pipe	Pipe Diameter	Installation Year	Protection Method	Number of Break	Hydraulic Factor	Water Quality	Leakage	Cost of Repair	Damage to Surroundings	Loss of Production	Traffic Disruption	Serviced Area	Redundancy	Environmental	Physical	Operational	Consequence	Risk of Failure
1	most risky	10	10	seasonally present	Cast iron	100	1900	none	6	30	10	10	10	Industrial	100	10	Industrial	Not Redundent	9.4	8.3	9.4	8.3	8.3
2	least risky	0	0	rarely present	PE	500	2007	none	0	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	1.7	0.6	1.7	1.7
3	age 1	0	0	rarely present	PE	500	2007	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	1.7	0.6	1.7	1.7
4	age 2	0	0	rarely present	PE	500	1997	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	1.7	0.6	1.7	1.7
5	age 3	0	0	rarely present	PE	500	1987	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	1.8	0.6	1.7	1.7
6	age 4	0	0	rarely present	PE	500	1977	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	2.6	0.6	1.7	1.7
7	age 5	0	0	rarely present	PE	500	1967	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	3.5	0.6	1.7	1.9
8	age 6	0	0	rarely present	PE	500	1957	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	4.2	0.6	1.7	2.4
9	age 7	0	0	rarely present	PE	500	1947	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	5.0	0.6	1.7	2.3
10	age 8	0	0	rarely present	PE	500	1937	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	5.0	0.6	1.7	2.3
11	age 9	0	0	rarely present	PE	500	1927	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	5.0	0.6	1.7	2.3
12	age 10	0	0	rarely present	PE	500	1917	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	5.0	0.6	1.7	2.3
13	age 11	0	0	rarely present	PE	500	1907	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	5.0	0.6	1.7	2.3
14	type of material 2	0	0	rarely present	PVC	500	1907	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	5.0	0.6	1.7	2.3
15	type of material 3	0	0	rarely present	Concrete	500	1907	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	6.7	0.6	1.7	3.3
16	type of material 4	0	0	rarely present	Asbestos	500	1907	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	6.7	0.6	1.7	3.3
17	type of material 5	0	0	rarely present	Cast Iron	500	1907	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	6.7	0.6	1.7	3.3
18	type of material 6	0	0	rarely present	t iron post	500	1907	none	0.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	0.6	1.7	4.0
19	breakage 2	0	0	rarely present	t iron post	500	1907	none	0.25	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	1.1	1.7	4.0
20	breakage 3	0	0	rarely present	t iron post	500	1907	none	0.50	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	1.3	1.7	4.0
21	breakage 4	0	0	rarely present	t iron post	500	1907	none	0.75	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	1.5	1.7	4.0
22	breakage 5	0	0	rarely present	t iron post	500	1907	none	1.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	1.7	1.7	4.0
23	breakage 6	0	0	rarely present	t iron post	500	1907	none	2.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	1.7	1.7	4.0
24	breakage 7	0	0	rarely present	t iron post	500	1907	none	2.25	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	1.7	1.7	4.0
25	breakage 8	0	0	rarely present	t iron post	500	1907	none	2.50	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	1.7	1.7	4.0
26	breakage 9	0	0	rarely present	t iron post	500	1907	none	2.75	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	2.2	1.7	4.0
27	breakage 10	0	0	rarely present	t iron post	500	1907	none	3.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	2.5	1.7	4.2
28	breakage 11	0	0	rarely present	t iron post	500	1907	none	3.25	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	2.8	1.7	4.4
29	breakage 12	0	0	rarely present	t iron post	500	1907	none	3.50	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	3.3	1.7	5.0
30	breakage 13	0	0	rarely present	t iron post	500	1907	none	4.00	130	0	0	0	Residential	500	0	Residential	Redundent	0.6	8.3	3.3	1.7	5.0
31	pipe diameter 2	0	0	rarely present	t iron post	350	1907	none	4.00	130	0	0	0	Residential	350	0	Residential	Redundent	0.6	8.3	3.3	1.7	5.0
32	pipe diameter 3	0	0	rarely present	t iron post	300	1907	none	4.00	130	0	0	0	Residential	300	0	Residential	Redundent	0.6	8.3	3.3	1.7	5.0
33	pipe diameter 4	0	0	rarely present	t iron post	250	1907	none	4.00	130	0	0	0	Residential	250	0	Residential	Redundent	0.6	8.3	3.3	1.7	5.0
34	pipe diameter 5	0	0	rarely present	t iron post	200	1907	none	4.00	130	0	0	0	Residential	200	0	Residential	Redundent	0.6	8.3	3.3	1.7	5.0
35	pipe diameter 6	0	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	0.6	8.3	3.3	1.7	5.0
36	type of soil 2	1	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	1.3	8.3	3.3	1.7	5.0
37	type of soil 3	2	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	1.5	8.3	3.3	1.7	5.0
38	type of soil 4	3	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	2.1	8.3	3.3	1.7	5.0
39	type of soil 5	4	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	2.6	8.3	3.3	1.7	5.0
40	type of soil 6	5	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	3.3	8.3	3.3	1.7	5.0
41	type of soil 7	6	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	4.1	8.3	3.3	1.7	5.0
42	type of soil 8	7	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	4.5	8.3	3.3	1.7	5.0
43	type of soil 9	8	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	5.0	8.3	3.3	1.7	5.0
44	type of soil 10	9	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	5.0	8.3	3.3	1.7	5.0
45	type of soil 11	10	0	rarely present	t iron post	150	1907	none	4.00	130	0	0	0	Residential	150	0	Residential	Redundent	5.0	8.3	3.3	1.7	5.0
46	leakage 2	10	0	rarely present	t iron post	150	1907	none	4.00	130	0	1	0	Residential	150	0	Residential	Redundent	5.0	8.3	4.1	1.7	5.0
47	leakage 3	10	0	rarely present	t iron post	150	1907	none	4.00	130	0	2	0	Residential	150	0	Residential	Redundent	5.0	8.3	4.5	1.7	5.0
48	leakage 4	10	0	rarely present	t iron post	150	1907	none	4.00	130	0	3	0	Residential	150	0	Residential	Redundent	5.0	8.3	5.0	1.7	5.0
49	leakage 5	10	0	rarely present	t iron post	150	1907	none	4.00	130	0	4	0	Residential	150	0	Residential	Redundent	5.0	8.3	5.0	1.7	5.0
50	leakage 6	10	0	rarely present	t iron post	150	1907	none	4.00	130	0	5	0	Residential	150	0	Residential	Redundent	5.0	8.3	5.0	1.7	5.0
51	leakage 7	10	0	rarely present	t iron post	150	1907	none	4.00	130	0	6	0	Residential	150	0	Residential	Redundent	5.0	8.3	5.7	1.7	5.8
52	leakage 8	10	0	rarely present	t iron post	150	1907	none	4.00	130	0	7	0	Residential	150	0	Residential	Redundent	5.0	8.3	6.2	1.7	6.1
53	leakage 9	10	0	rarely present	t iron post	150	1907	none	4.00	130	0	8	0	Residential	150	0	Residential	Redundent	5.0	8.3	6.7	1.7	6.7
54	leakage 10	10	0	rarely present	t iron post	150	1907	none	4.00	130	0	9	0	Residential	150	0	Residential	Redundent	5.0	8.3	6.7	1.7	6.7
55	leakage 11	10	0	rarely present	t iron post	150	1907	none	4.00	130	0	10	0	Residential	150	0	Residential	Redundent	5.0	8.3	6.7	1.7	6.7

56	water quality 2	10	0	rarely present	t iron post	150	1907	none	4.00	130	1	10	0	Residential	150	0	Residential	Redundent	5.0	8.3	6.7	1.7	6.7
57	water quality 3	10	0	rarely present	t iron post	150	1907	none	4.00	130	2	10	0	Residential	150	0	Residential	Redundent	5.0	8.3	6.7	1.7	6.7
58	water quality 4	10	0	rarely present	t iron post	150	1907	none	4.00	130	3	10	0	Residential	150	0	Residential	Redundent	5.0	8.3	6.7	1.7	6.7
59	water quality 5	10	0	rarely present	t iron post	150	1907	none	4.00	130	4	10	0	Residential	150	0	Residential	Redundent	5.0	8.3	6.7	1.7	6.7
60	water quality 6	10	0	rarely present	t iron post	150	1907	none	4.00	130	5	10	0	Residential	150	0	Residential	Redundent	5.0	8.3	6.7	1.7	6.7
61	water quality 7	10	0	rarely present	t iron post	150	1907	none	4.00	130	6	10	0	Residential	150	0	Residential	Redundent	5.0	8.3	7.4	1.7	6.7
62	water quality 8	10	0	rarely present	t iron post	150	1907	none	4.00	130	7	10	0	Residential	150	0	Residential	Redundent	5.0	8.3	7.9	1.7	6.7
63	water quality 9	10	0	rarely present	t iron post	150	1907	none	4.00	130	8	10	0	Residential	150	0	Residential	Redundent	5.0	8.3	8.3	1.7	6.7
64	water quality 10	10	0	rarely present	t iron post	150	1907	none	4.00	130	9	10	0	Residential	150	0	Residential	Redundent	5.0	8.3	8.3	1.7	6.7
65	water quality 11	10	0	rarely present	t iron post	150	1907	none	4.00	130	10	10	0	Residential	150	0	Residential	Redundent	5.0	8.3	8.3	1.7	6.7
66	damage to surrounding 2	10	0	rarely present	t iron post	150	1907	none	4.00	130	10	10	0	Commercial	150	0	Residential	Redundent	5.0	8.3	8.3	3.3	6.7
67	damage to surrounding 3	10	0	rarely present	t iron post	150	1907	none	4.00	130	10	10	0	Industrial	150	0	Residential	Redundent	5.0	8.3	8.3	3.3	6.7
68	cost of repair 2	10	0	rarely present	t iron post	150	1907	none	4.00	130	10	10	1	Industrial	150	0	Residential	Redundent	5.0	8.3	8.3	3.3	6.7
69	cost of repair 3	10	0	rarely present	t iron post	150	1907	none	4.00	130	10	10	2	Industrial	150	0	Residential	Redundent	5.0	8.3	8.3	3.3	6.7
70	cost of repair 4	10	0	rarely present	t iron post	150	1907	none	4.00	130	10	10	3	Industrial	150	0	Residential	Redundent	5.0	8.3	8.3	3.8	6.7
71	cost of repair 5	10	0	rarely present	t iron post	150	1907	none	4.00	130	10	10	4	Industrial	150	0	Residential	Redundent	5.0	8.3	8.3	4.3	6.7
72	cost of repair 6	10	0	rarely present	t iron post	150	1907	none	4.00	130	10	10	5	Industrial	150	0	Residential	Redundent	5.0	8.3	8.3	5.0	6.7
73	cost of repair 7	10	0	rarely present	t iron post	150	1907	none	4.00	130	10	10	6	Industrial	150	0	Residential	Redundent	5.0	8.3	8.3	5.0	6.7
74	cost of repair 8	10	0	rarely present	t iron post	150	1907	none	4.00	130	10	10	7	Industrial	150	0	Residential	Redundent	5.0	8.3	8.3	5.0	6.7
75	cost of repair 9	10	0	rarely present	t iron post	150	1907	none	4.00	130	10	10	8	Industrial	150	0	Residential	Redundent	5.0	8.3	8.3	5.0	6.7
76	cost of repair 10	10	0	rarely present	t iron post	150	1907	none	4.00	130	10	10	9	Industrial	150	0	Residential	Redundent	5.0	8.3	8.3	5.0	6.7
77	cost of repair 11	10	0	rarely present	t iron post	150	1907	none	4.00	130	10	10	10	Industrial	150	0	Residential	Redundent	5.0	8.3	8.3	5.0	6.7
78	Water table level 2	10	0	always present	t iron post	150	1907	none	4.00	130	10	10	10	Industrial	150	0	Residential	Redundent	6.7	8.3	8.3	5.0	6.7
79	Water table level 3	10	0	seasonally present	t iron post	150	1907	none	4.00	130	10	10	10	Industrial	150	0	Residential	Redundent	8.3	8.3	8.3	5.0	8.3
80	Loss of Production 2	10	0	seasonally present	t iron post	150	1907	none	4.00	130	10	10	10	Industrial	150	0	Residential	Not Redundent	8.3	8.3	8.3	6.7	8.3
81	Traffic disruption 2	10	0	seasonally present	t iron post	150	1907	none	4.00	130	10	10	10	Industrial	150	1	Residential	Not Redundent	8.3	8.3	8.3	6.7	8.3
82	Traffic disruption 3	10	0	seasonally present	t iron post	150	1907	none	4.00	130	10	10	10	Industrial	150	2	Residential	Not Redundent	8.3	8.3	8.3	6.7	8.3
83	Traffic disruption 4	10	0	seasonally present	t iron post	150	1907	none	4.00	130	10	10	10	Industrial	150	3	Residential	Not Redundent	8.3	8.3	8.3	6.7	8.3
84	Traffic disruption 5	10	0	seasonally present	t iron post	150	1907	none	4.00	130	10	10	10	Industrial	150	4	Residential	Not Redundent	8.3	8.3	8.3	6.7	8.3
85	Traffic disruption 6	10	0	seasonally present	t iron post	150	1907	none	4.00	130	10	10	10	Industrial	150	5	Residential	Not Redundent	8.3	8.3	8.3	6.7	8.3
86	Traffic disruption 7	10	0	seasonally present	t iron post	150	1907	none	4.00	130	10	10	10	Industrial	150	6	Residential	Not Redundent	8.3	8.3	8.3	7.4	8.3
87	Traffic disruption 8	10	0	seasonally present	t iron post	150	1907	none	4.00	130	10	10	10	Industrial	150	7	Residential	Not Redundent	8.3	8.3	8.3	7.9	8.3
88	Traffic disruption 9	10	0	seasonally present	t iron post	150	1907	none	4.00	130	10	10	10	Industrial	150	8	Residential	Not Redundent	8.3	8.3	8.3	8.3	8.3
89	Traffic disruption 10	10	0	seasonally present	t iron post	150	1907	none	4.00	130	10	10	10	Industrial	150	9	Residential	Not Redundent	8.3	8.3	8.3	8.3	8.3
90	Traffic disruption 11	10	0	seasonally present	t iron post	150	1907	none	4.00	130	10	10	10	Industrial	150	10	Residential	Not Redundent	8.3	8.3	8.3	8.3	8.3
91	Hydralic factor 2	10	0	seasonally present	t iron post	150	1907	none	4.00	120	10	10	10	Industrial	150	10	Residential	Not Redundent	8.3	8.3	8.3	8.3	8.3
92	Hydralic factor 3	10	0	seasonally present	t iron post	150	1907	none	4.00	110	10	10	10	Industrial	150	10	Residential	Not Redundent	8.3	8.3	8.3	8.3	8.3
93	Hydralic factor 4	10	0	seasonally present	t iron post	150	1907	none	4.00	100	10	10	10	Industrial	150	10	Residential	Not Redundent	8.3	8.3	8.3	8.3	8.3
94	Hydralic factor 5	10	0	seasonally present	t iron post	150	1907	none	4.00	90	10	10	10	Industrial	150	10	Residential	Not Redundent	8.3	8.3	8.3	8.3	8.3
95	Hydralic factor 6	10	0	seasonally present	t iron post	150	1907	none	4.00	80	10	10	10	Industrial	150	10	Residential	Not Redundent	8.3	8.3	8.3	8.3	8.3
96	Hydralic factor 7	10	0	seasonally present	t iron post	150	1907	none	4.00	70	10	10	10	Industrial	150	10	Residential	Not Redundent	8.3	8.3	8.3	8.3	8.3
97	Hydralic factor 8	10	0	seasonally present	t iron post	150	1907	none	4.00	60	10	10	10	Industrial	150	10	Residential	Not Redundent	8.3	8.3	8.7	8.3	8.3
98	Hydralic factor 9	10	0	seasonally present	t iron post	150	1907	none	4.00	50	10	10	10	Industrial	150	10	Residential	Not Redundent	8.3	8.3	9.4	8.3	8.3
99	Hydralic factor 10	10	0	seasonally present	t iron post	150	1907	none	4.00	40	10	10	10	Industrial	150	10	Residential	Not Redundent	8.3	8.3	9.4	8.3	8.3
100	Hydralic factor 11	10	0	seasonally present	t iron post	150	1907	none	4.00	30	10	10	10	Industrial	150	10	Residential	Not Redundent	8.3	8.3	9.4	8.3	8.3
101	Type of serived area 2	10	0	seasonally present	t iron post	150	1907	none	4.00	30	10	10	10	Industrial	150	10	Commercial	Not Redundent	8.3	8.3	9.4	8.3	8.3
102	Type of serived area 3	10	0	seasonally present	t iron post	150	1907	none	4.00	30	10	10	10	Industrial	150	10	Industrial	Not Redundent	8.3	8.3	9.4	8.3	8.3
103	daily traffic 2	10	1	seasonally present	t iron post	150	1907	none	4.00	30	10	10	10	Industrial	150	10	Industrial	Not Redundent	8.3	8.3	9.4	8.3	8.3
104	daily traffic 3	10	2	seasonally present	t iron post	150	1907	none	4.00	30	10	10	10	Industrial	150	10	Industrial	Not Redundent	8.3	8.3	9.4	8.3	8.3
105	daily traffic 4	10	3	seasonally present	t iron post	150	1907	none	4.00	30	10	10	10	Industrial	150	10	Industrial	Not Redundent	8.3	8.3	9.4	8.3	8.3
106	daily traffic 5	10	4	seasonally present	t iron post	150	1907	none	4.00	30	10	10	10	Industrial	150	10	Industrial	Not Redundent	8.3	8.3	9.4	8.3	8.3
107	daily traffic 6	10	5	seasonally present	t iron post	150	1907	none	4.00	30	10	10	10	Industrial	150	10	Industrial	Not Redundent	8.3	8.3	9.4	8.3	8.3
108	daily traffic 7	10	6	seasonally present	t iron post	150	1907	none	4.00	30	10	10	10	Industrial	150	10	Industrial	Not Redundent	8.6	8.3	9.4	8.3	8.3
109	daily traffic 8	10	7	seasonally present	t iron post	150	1907	none	4.00	30	10	10	10	Industrial	150	10	Industrial	Not Redundent	9.0	8.3	9.4	8.3	8.3
110	daily traffic 9	10	8	seasonally present	t iron post	150	1907	none	4.00	30	10	10	10	Industrial	150	10	Industrial	Not Redundent	9.4	8.3	9.4	8.3	8.3
111	daily traffic 10	10	9	seasonally present	t iron post	150	1907	none	4.00	30	10	10	10	Industrial	150	10	Industrial	Not Redundent	9.4	8.3	9.4	8.3	8.3
112	daily traffic 11	10	10	seasonally present	t iron post	150	1907	none	4.00	30	10	10	10	Industrial	150	10	Industrial	Not Redundent	9.4	8.3	9.4	8.3	8.3