INTEGRATED CONDITION ASSESSMENT MODELS FOR SUSTAINABLE SEWER PIPELINES

by

Fazal ur Rehman M Chughtai

A Thesis

in

The Department

of

Building, Civil, and Environmental Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Applied Science (Building Engineering) at
CONCORDIA UNIVERSITY
Montreal, Quebec, Canada

March, 2007

Your file Votre référence
ISBN: 978-0-494-28884-9
Our file Notre référence
ISBN: 978-0-494-28884-9

ABSTARCT

INTEGRATED CONDITION ASSESSMENT MODELS FOR SUSTAINABLE SEWER PIPELINES

Fazal ur Rehman M Chughtai

The Federation of Canadian Municipalities (FCM) reported that approximately 55% of sewer infrastructure in Canada did not meet current standards. Therefore, burden on Canadian municipalities to maintain and prioritize sewers is increasing. One of the major challenges is to develop a framework to standardize the condition assessment procedures for sewer pipelines. Lack of detailed knowledge on the condition of sewer networks escalates vulnerability to catastrophic failures. This research presents a proactive methodology of assessing the existing condition of sewers by considering various physical, environmental, and operational influence factors. Based on historic data collected from two municipalities in Canada, structural and operational condition assessment models for sewers are developed using multiple regression technique. These models are utilized to generate deterioration curves for Concrete, Asbestos Cement, and Polyvinyl Chloride (PVC) sewers.

A combined condition index (CCI) for sewers is developed, which integrates the combined effect of structural and operational conditions. The CCI is divided into 5 condition categories, ranging from "Acceptable" to "Critical". It is developed based on integrating the two major sewer condition assessment protocols adapted in Canada: WRc (Water Research Centre, UK) and CERIU (Centre for Expertise and Research on Infrastructures in Urban Areas, Canada). Unsupervised, self-organizing, neural network approach is used in order to develop the CCI and the integrated protocol.

The developed regression models show 82% to 86% accuracy when they are applied to the validation data set. The CCI and integrated protocol are verified by municipal practitioners and experts of the CERIU sub-committee for developing a unified sewer condition assessment system. Based on the developed models, a web-based sewer condition assessment tool, coded in Java (version 5.0), is developed to predict structural and hydraulic conditions as well as the CCI.

The developed models will assist municipal engineers in identifying critical sewers, prioritizing sewer inspections, and developing a unified sewer condition assessment system.

ACKNOWLEDGEMENTS

First of all, I wish to thank the Almighty Creator for giving me health, energy, and persistence for carrying out this research.

I would like to express my sincere gratitude to my mentor and supervisor, Dr. Tarek Zayed, for his continuous support, outstanding guidance, and encouragement throughout this research. The experience that he shared with me in a very friendly manner gave me an added confidence and professional knowledge.

I wish to acknowledge the cooperation of Mr. Sylvain Marcoux, chef de division infrastructures and Mr. Claude Cuillerier, Technicien municipal, municipality of Pierrefonds, Quebec, and Mr. Phil Curtis, Environmental Analyst, regional municipality of Niagara, Ontario, for providing the data used in this research

I also recognize the support and direction of Mr. Charlerick Bergeron, Technologies Director, CERIU (Centre for Expertise and Research on Infrastructures in Urban Areas), Montreal, during this research.

A very special thanks to my parents for instilling me the curiosity and inquisitiveness, and for their everlasting love that brought me this far. My never ending thanks to my brothers and their families for their unconditional love.

Further, I would like to express my deepest appreciation to my wife for her patience and encouragement.

I am thankful to all the faculty and staff members of the Department of Building, Civil, and Environmental Engineering. I would also like to thank my colleagues for their sincere support throughout this research

Finally, I would like to extend my gratitude to the Quebec funding agency Fonds québécois de la recherche sur la nature et les technologies (FQRNT) for its financial support for the current research.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE AND ABREVATIONS

AADT	Average Annual Daily Traffic
AIP	Average Invalidity Percentage
ANN	Artificial neural Networks
ANOVA	Analysis of Variance
ASCE	American Society of Civil Engineers
ASTM	American Society of Testing Materials
AVP	Average Validity Percentage
BNQ	Bureau de normalisation du Québec
BRE	Building Research Establishment
CATT	Centre of Advanced Trenchless Technologies
CCI	Combined Condition Index
CCTV	Closed Circuit Television
CERIU	Centre for Expertise and Research on Infrastructures in Urban Areas
CGSB	Canadian General Standard Board
GIS	Geographic Information System
C_i	Actual Value
E_i	Estimated or Predicted Value
fi	Fitness Function
H_0	Null Hypothesis
H_1	Alternate Hypothesis
MAE	Mean Absolute Error
MS	Microsoft

MSE	Error Mean Square
MSLF	Lack of Fit Mean Square
MSPE	Pure Error Mean Square
MSR	Regression Mean Square
NAAPI	North American Association of Pipeline Inspectors
NASSCO	National Association of Sewer Service Companies
NRC	National Research Council
NZWWA	New Zealand Water and Wastes Association
OOP	Objected Oriented Programming
OPSD	Ontario Provincial Standard Drawings
PE	Polyethylene
PVC	Polyvinyl Chloride
RMS	Root Mean Square Error
SOM	Self-Organizing Maps
SRM	Sewerage Rehabilitation Manual
SSE	Error Sum of Squares
SSET	Sewer Scanner and Evaluation Technology
SSLF	Lack of Fit Sum of Squares
SSPE	Pure Error Sum of Squares
SSR	Regression Sum of Squares
SSTO	Total Sum of Squares
WRc	Water Research Centre
*.xls	Microsoft Excel File Format

Chapter 1

INTRODUCTION

1.1 Overview

With an aging underground infrastructure, the burden on municipal agencies to prioritize and maintain the rapidly deteriorating underground utilities is increasing (Najafi et al 2005). This deterioration of underground infrastructure, such as sewers, poses a serious problem to most developed urban centers (Moselhi et al 2000). Sewer is a conduit (pipe or tunnel) that collects and transports waste and storm water. If a sewer's function is to transport waste water only, it is called sanitary sewer. Sanitary sewer collection systems are an extensive and vital part of any nation's infrastructure (Kulandaivel, 2004).

In Canada, the average age of sanitary sewer system is reported to be 42 years. In particular, about 20 percent of existing trunk sewers (large diameter sewers) have been in service for more than 50 years (Zhao, 1998). Moreover, a report released by the Federation of Canadian Municipalities (FCM) suggests that nearly 55% of Canada's water and wastewater infrastructure can be classified as not meeting current standards (Allouche et al 2002).

Although major part of infrastructure deterioration is attributed to aging, age related deterioration of sewer is unclear (Fenner, 2000). Influence factors such as surrounding soil conditions, hydraulic overloading, corrosion, etc, could accelerate the rate of deterioration of sewers. A major problem in assessing the condition of sewers is the lack of detailed knowledge about pipeline degradation process. Being covered with soil, the condition of buried pipelines cannot be directly and easily monitored (Kathula, 2000). Thus, deteriorating sewers leave communities vulnerable to unexpected catastrophic failures that disrupt not only sewer service but also above ground activities. These failures are difficult to avoid if the cities are unaware of their network's condition (Hahn et al 2000). As a consequence, many municipalities all around the world have been spending a huge amount of their budget on emergency repairs of sewers. The need for emergency repairs of buried pipes can be significantly reduced if critical sections could be identified and repaired before a catastrophic failure occurs. As a result, the utilization of funds can be optimized to dramatically reduce the overall cost of maintenance (MacLeod et al 2000).

In short, effective functioning of sewer network system is extremely important for municipal agencies (Grigg et al 1994). This requires an appropriate maintenance and management system (Hasegawa et al 1999), and valid condition assessment is the key for developing this system (Fenner, 2000).

1.2 Current Approaches

Different sewer management approaches have been adapted by different municipal agencies all across the world. The basic theme of all these approaches is to develop an ideal asset management plan which can prioritize maintenance and rehabilitation of pipeline networks. This involves (Kulandaival, 2004):

- Routine and systematic inspections for sewer structural and hydraulic condition assessment
- Establishment of a standard condition rating system
- Developing and updating prediction models for sewer conditions

The importance of assessing the condition of sewer pipes led to the development of new techniques for inspection of the buried infrastructure. A closed-circuit television (CCTV) camera was first introduced in the 1960s (Reyna et al 1994). Later on, new technologies like laser based scanning and ultrasound inspection systems were introduced (Wirahaikusumah et al 2001). Nevertheless, CCTV inspection remains to be the most commonly used technique (Makar, 1999).

The second aspect of sewer management plan is the establishment of standard condition rating protocols or system. Generally, these protocols consist of some weighted factors which are used to grade the severity of a sewer's condition. These weighted factors have been developed by many institutions all across the world. The first sewer condition assessment protocols were developed by Water Research Centre (WRc), UK in 1978 (Thornhill et al 2005).

The third part of sewer management approach consists of developing and updating condition prediction tools for prioritizing detailed inspection. Many tools have been introduced to assist municipal engineers for optimizing decision regarding infrastructure inspection and condition assessment. Current decision tools are largely in the form of general guidelines where distress indicators observed in the asset are translated into asset condition state (Kleiner, 2001).

1.3 Problem Statement

The current sewer management practices have some limitations. Therefore, there is an urgent need to address these issues for standardizing the current sewer condition assessment process. The limitations associated to current practices are described below:

1.3.1 Limitations of Inspection Techniques for Condition Assessment

There are three major limitations associated to the inspection techniques for sewer condition assessment:

- Random inspections of sewers for condition assessment are expensive (Zhao, 1998). As a consequence, the adoption of current sewer inspection methods has become practically impossible for a majority of Canadian municipalities, and only 22% of the Canadian municipalities have a complete condition assessment program (Rahman et al 2004).
- The inspection process is highly dependent upon the skills of the inspector and the quality of the equipment used. A small error in observation could generate erroneous results. Therefore, the process may lead to diagnostic errors due to lack of concentration and/or experience of inspector (Shehab-Eldeen, 2001).
- The current inspection practices are very time consuming, and a very huge amount of data is involved for the mangers for analysis. Therefore, the process is considered tedious by most engineers and practitioners (Shehab-Eldeen, 2001).

The limitations mentioned above can be minimized by adapting an alternate approach to random inspection of sewers. This approach should suggest solutions for prioritizing sewer inspections to critical sewers. Thus, the undesirable use of resources could be minimized.

1.3.2 Lack of Unification of Sewer Condition Assessment Protocols

Although WRc protocols are in wide usage in many municipalities across Canada, many sewer condition assessment protocols have been developed. As can be seen, it is difficult for a municipality or utility to select amongst the available protocols. Therefore, there is an urgent need of developing a unified sewer condition assessment system for Canadian municipalities and utilities (Rahman et al 2004).

1.3.3 Limitations of Condition Prediction Tools

There are numerous documented studies that focus on various aspects of drainage systems including different methodological approaches to predict the condition of sewers (Ruwanpura et al. 2004). However, there is an urgent need of an integrated and easy to use approach towards condition prediction. Moreover, current research mainly emphasizes upon structural condition prediction; nevertheless, an existing sewer's performance also depends upon its hydraulic, or commonly known as operational, condition.

1.4 Research Objectives

The objective of this research is to develop a methodology that could facilitate the standardizing of the current sewer pipeline condition assessment process. The objective can be divided into the following sub-objectives:

- Develop structural condition prediction model for sewers
- Develop operational condition prediction model for sewers
- Build deterioration curves for sewers
- Propose a technique for unification of sewer condition assessment protocols
- Design a combined condition index (CCI) through integration of structural and operational conditions of sewers; thus, help decision makers in visualizing a complete picture of a sewer's condition

- Develop an automated web-based sewer condition prediction tool by utilizing the findings of this research to facilitate municipal personnel for planning future sewer inspection needs

1.5 Research Methodology

The methodology of current research consists of different steps. The detailed description of the research methodology will be presented in chapter 3. The main steps include:

- Literature review
- Comparison of different sewer condition assessment protocols
- Data collection and pre-processing
- Development of sewer structural and hydraulic condition prediction models
- Development of structural and operational deterioration curves
- Development of integrated sewer condition assessment system protocols
- Design a combined condition index (CCI) for sewers which should integrate structural and hydraulic condition rating of sewers
- Development of web-based automated condition prediction tool

1.6 Thesis Organization

Chapter 2 presents literature review of different types of sewer pipes, major factors which affect a sewer's structural and operational condition, different sewer inspection techniques for condition assessment, historical background of sewer condition assessment protocol development, an overview of two major sewer condition assessment protocols, different adapted research techniques for sewer management system, and an overview of the two techniques which have been adapted in this research.

Chapter 3 provides an overview of the proposed research methodology. This includes a brief introduction and layouts for building sewer condition prediction models, development of combined condition index (CCI) for sewers, and web-based condition rating program.

Chapter 4 describes the data collection and pre-preparation procedures in detail. It illustrates an overview of different assumptions which were made during data pre-processing. The descriptive statistics and histograms of collected data are also presented.

Chapter 5 illustrates the sewer condition assessment regression models design and validation processes. The different statistical checks and diagnostics applied during the process have been discussed in detail. The summary of results is also tabulated.

Chapter 6 proposes a methodology to convert CERIU (Centre for Expertise and Research on Infrastructures in Urban Areas) sewer condition assessment protocols into WRc and vice versa. The methodology verification process is also presented. The proposed methodology of development of combined condition index (CCI) for sewers is presented. The clustering process of structural and operational condition grades through unsupervised neural network is presented in detail.

Chapter 7 describes a methodology of developing a web-based decision support system for condition prediction of existing sewers. The step by step process of the web application is illustrated.

Chapter 8 presents conclusions, limitations of the research, research contributions, and recommendations for the future research work.

Chapter 2

LITERATURE REVIEW

2.1 Overview

The literature review of this research covers six main sections as shown in Figure 2-1:

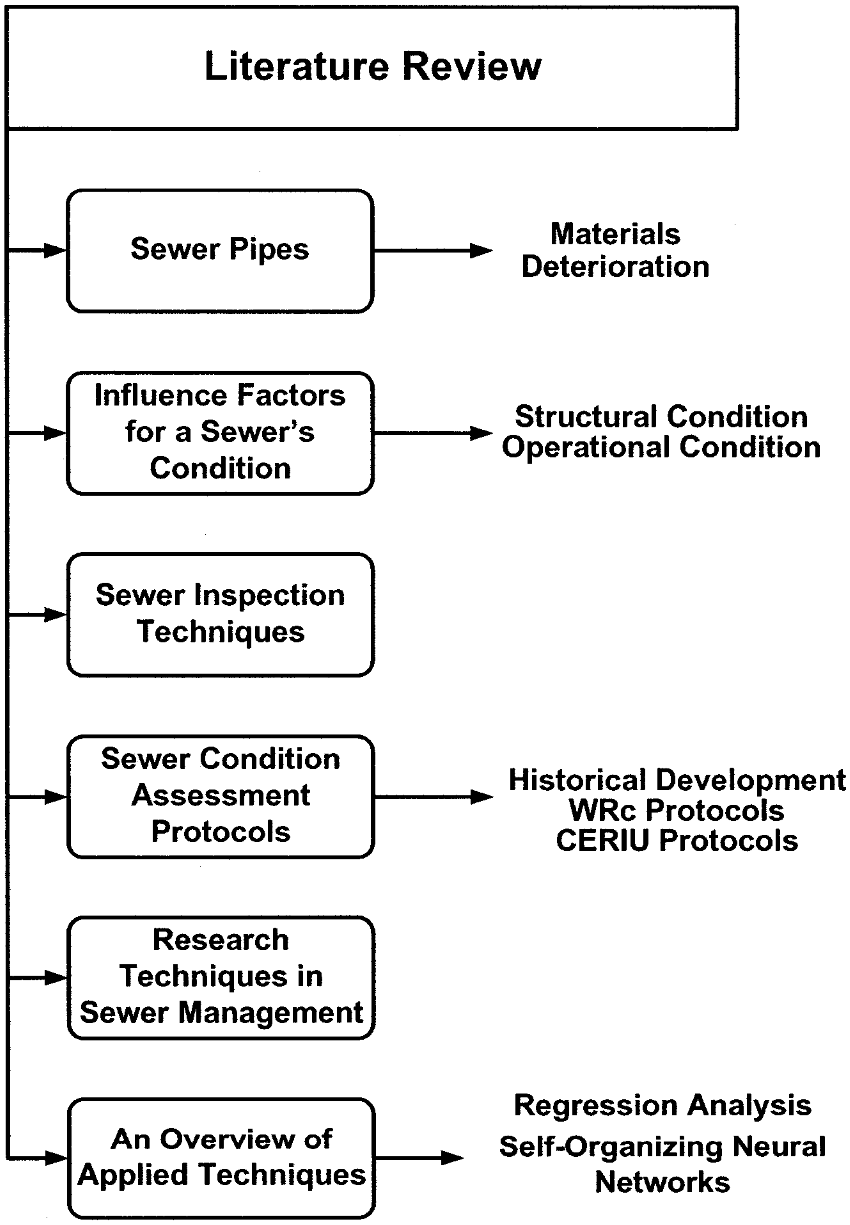


Figure 2-1: Chapter 2 Content Diagram

As illustrated in Figure 2-1, a detailed review of all aspects of sewer pipeline networks has been presented. The first section covers a review of different types of sewer pipe materials, their characteristics, and their deterioration phenomena. The second section describes the major factors which affect a sewer's structural and operational condition. The third section presents an overview of different sewer inspection techniques for condition assessment. Fourth section illustrates a history of sewer condition assessment protocols development. This section further presents an overview of two major sewer condition assessment protocols which have been adapted in Canada.

The fifth and sixth sections deal with the different adapted research techniques for sewer management system. The fifth section represents a review of all major techniques which have been utilized in the past by researchers for finding solutions to different sewer management problems. The sixth and last section presents an overview of the two techniques which have been adapted in this research.

2.2 Sewer Pipes

Sanitary sewer system can be divided into two main categories: trunk and regular sewers. Trunk sewers are typically used to intercept regular sewers, and receive and transport sewage to a few central places, such as treatment plants or discharge points on river banks. These sewers vary in size, and are installed along or deeper than other buried utilities. The minimum recommended diameter of these sewers has been classified differently by different protocols and is dependent upon local administrative requirements. NASSCO (National Association of Sewer Service Companies, USA) defines sewers with diameter 686mm (27 inch) or more as trunk sewers (Zhao, 1998). The other category of sewers, regular sewers, is usually designed for collection of sewage from end users (residential, commercial, and industrial). The minimum recommended size by many Canadian municipalities for this type of sewer is 200mm.

2.2.1 Pipe Materials

There are different pipe materials available for sewer system, each with unique characteristics and is used in different conditions. Until 1850, sewers were generally constructed by bricks. Although some sewer systems still contain brick sewers, very few are left (Kulandaivel, 2004). Now there are several types of sewer pipe material available. The choice of sewer pipe material depends upon several factors. The main factors include (ACPA, 1980):

- Physical strength
- Cost of material and availability in required sizes
- Ease of handling, installation, maintenance, and repair
- Flow characteristics (Friction Coefficient)
- Resistance to abrasion, corrosion, acids, alkalis, gases, and solvents

In general, sewer pipe materials can be divided in two categories depending upon their behaviour towards load carrying capacity (Peggs, 1985), and are shown in Figure 2-2. Furthermore, Figure 2-3 represents the distribution of different pipe materials in Canada. A brief description of these categories is as follows:

(i) Rigid Pipes: Pipe materials in this classification derive a substantial part of their basic earth load carrying capacity from the structural strength inherent in the rigid pipe wall. Commonly specified rigid pipe material includes (ASCE, 1992) asbestos cement, concrete, cast iron, and vitrified clay.

Asbestos cement pipe composed of a mixture of Portland cement and asbestos fibre and is manufactured in such a way that a very strong bond exists between cement and asbestos fibres (McGhee, 1991). The concrete pipes can be further classified into two sub categories: plain and reinforced concrete pipes. Reinforced and pre-stressed concrete pipes are used for pressure and gravity sewers (ASCE, 1992).

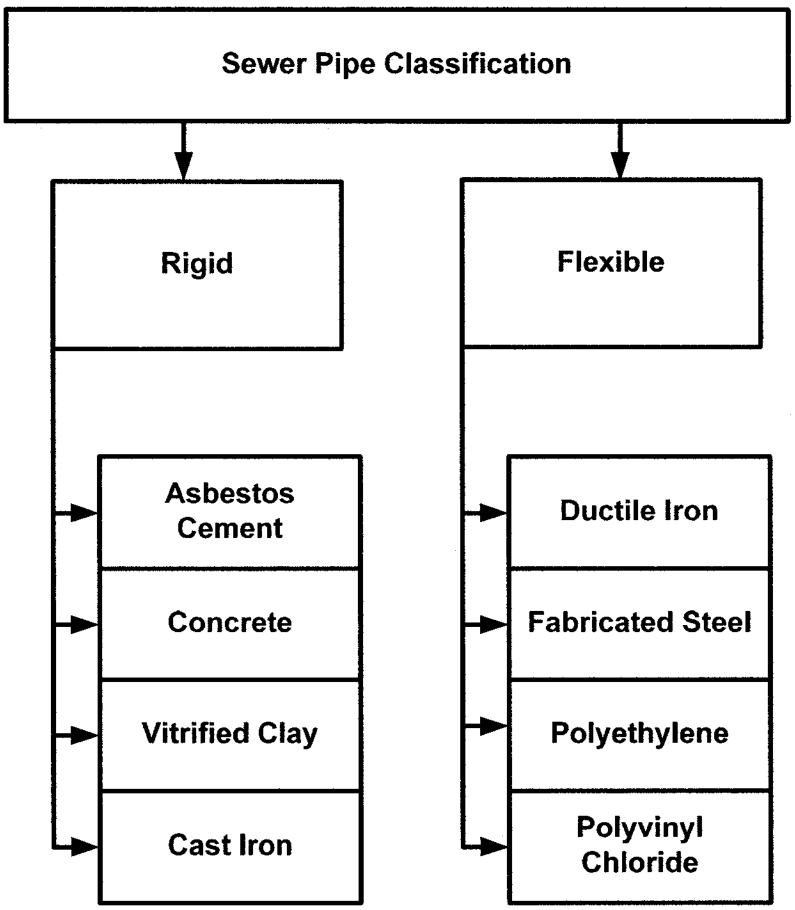


Figure 2-2: Commonly Used Sewer Pipe Material Classification

Table 2-1: Comparison of Different Sewer Pipe Materials

Type	Sub-Type	Advantages	Disadvantages
Rigid Pipes	**Asbestos Cement**	o Long laying lengths o Wide range of strength classification o Wide ranges of fittings available	▪ Subject to corrosion where acid is present ▪ Subject to shear ▪ Low beam strength
	Plain & Reinforced Concrete	o Wide ranges of structural strengths o Wide ranges of nominal diameters o Resistant to abrasion o Moderately low friction losses	▪ Heavy Weight ▪ Subject to corrosion where acids are present ▪ Requires careful installation for avoiding cracks
	Cast Iron	o Withstands higher external loads o Corrosion resistant o Available long laying lengths	▪ Subject to corrosion where acids are present ▪ Brittle
	Clay (Vitrified)	o Resistant to chemical corrosion o Low frictional losses o High resistance to abrasion	▪ Heavy weight ▪ Subject to shear and beam breakage ▪ Limited ranges of sizes available
Flexible Pipes	**Ductile Iron**	o High impact strength o High load bearing capacity o High beam strength o Low frictional losses	▪ Subject to corrosion when acids are present ▪ Subject to chemical attack in corrosive soils
	Fabricated Steel	o Light weight o Flexibility o Wide range of coating available	▪ Relatively poor hydraulic coefficient ▪ Subject to corrosion in aggressive environments
	PVC and PE	o Light weight o High impact strength o Ease in installation o Long laying lengths	▪ Subject to excessive deflection ▪ Limited range of size available ▪ Subject to attack by certain chemicals

Cast iron has been used for both gravity and pressure sewers, although recently ductile iron pipe has been specified in its place (Butler et al. 2000). Vitrified clay pipes are composed of crushed and blended clay that are formed into pipes. These pipes are then dried and fried in a succession of temperatures. These pipes have been used for hundreds of years (Kulandaivel, 2004).The main advantages and disadvantages of all these pipe materials are described in Table 2-1 (Adapted from ASCE 1970 & 1992).

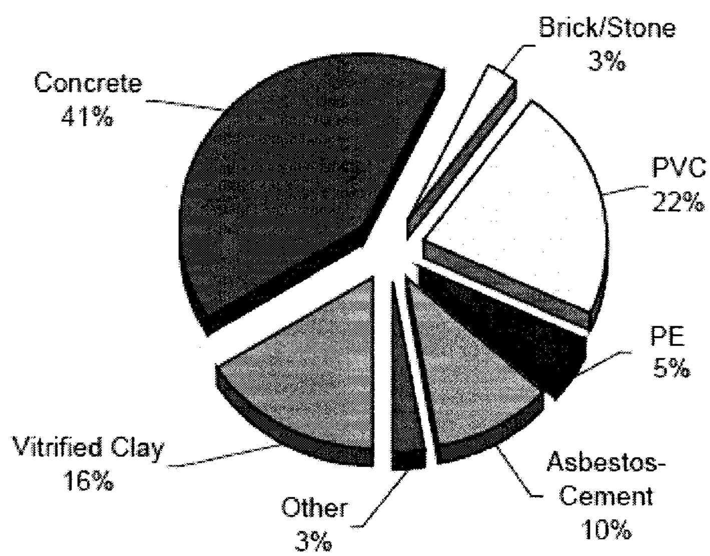


Figure 2-3: Distribution of Sewer Pipe Martials in Canada (Allouche et al 2002)

(ii) Flexible Pipes: Pipe materials in this classification drive their load carrying capacity from a combination of the inherent strength of the pipe and through the side support of the surrounding soil (Butler et al 2000). Commonly used flexible sewers include ductile iron, fabricated steel, Polyethylene (PE) and Polyvinyl Chloride (PVC).

As mentioned above, ductile iron pipes are considered to be a replacement for cast iron pipes. Not only ductile iron pipe has all the cast iron pipe's properties, but also it has an ability to deform without cracking. Fabricated steel pipes are available in different varieties; such as, corrugated steel pipes, arches and galvanized corrugated arch pipes

(ASCE, 1992). Polyvinyl chloride (PVC) and polyethylene (PE) are forms of plastics, and are obtained through the process of polymerization. These pipes are used both for gravity and pressure sewers. The advantages and disadvantages of all the flexible pipes are also shown in Table 2-1.

2.2.2 Sewer Pipe Deterioration

Pipe deterioration is a very complex process and related to various pipe characteristics (Yan et al 2003).

Table 2-2: Sewer Pipe Deterioration Factors Considered by Some Researchers

Country/Area	Research Reference	Sewer Deterioration Factors
Canada	McDonald et al 2001	Location, Soil Type, Pipe Depth, Size, Type of Waste, Seismic Zone
	Ariaratnam et al 2001	Age, Diameter, Material, Depth
	Ruwanpura et al 2004	Age, Material, Length
USA	Abraham et al 1998	Location, Traffic Loads, Pipe Size
	Yan et al 2003	Age, Diameter, Material, Depth
	Najafi et al 2005	Length, Diameter, Material, Age, Depth, Gradient
European Union	Rutsch et al 2000	Age, Material, Length, Diameter, Type of Soil, Ground Water Level
Germany	Baur et al 2002	Age, Material, Location, Diameter, Gradient, Type of Waste
Norway	Rostum et al 1999	Material, Diameter, Soil Condition, History of Failure
Japan	Hassegawa et al 1999	Age, Material, Depth, Ground Water Level, Soil Properties, Proximity of Other Underground Installations

Although sewers are designed for a particular lifespan under standard operating condition, their deterioration never follows a set pattern (Najafi et al 2005). Researchers all around the world have been working on analyzing the effects of different sewer attributes on its deterioration, and different factors have been considered as major factors. Some of the important research history regarding factors influencing sewer deterioration is shown in Table 2-2. Table 2-2 shows that pipe attributes, such as age, size, material, length, size, and environmental attributes, such as type of surrounding soil, water table and proximity of other underground installations are important factors.

2.3 Factors Affecting Sewer Pipeline Condition

It is clear from the previous discussion that there are several factors which should be taken into account for the assessment of sewer pipeline condition. In general, sewer collapses are caused by structural and hydraulic failures (Abraham et al 1998). The structural failures depend upon the existing structural condition of sewers. The hydraulic failures depend upon a sewer's operational condition; which describes the capability of a sewer pipe to meet its service requirements and indicates the loss of capacity, potential of blockage and water tightness.

2.3.1 Factors Affecting Sewer's Structural Condition

Three types of factors, physical, functional and environmental, can influence the structural condition of sewer pipes. The physical factors comprise of the general pipe characteristics such as its length, diameter etc. While the functional factors deal with the adapted operational and maintenance strategies for network management. The third category is concerned to certain environmental factors directly influencing a pipe's criticality and deterioration. These factors include type of surrounding soil, traffic volume above pipe etc.

Table 2-3: Factors Affecting Existing Sewer's Structural Condition

Factors		**Explanation**
Physical	*Pipe Length*	Pipe in longer length and having greater length to diameter ratio are more likely to suffer from bending stresses
	Pipe Diameter	Small diameter pipes are more susceptible to beam failure
	Pipe Material	Pipes manufactured with different materials show different failure patrons.
	Age	More probability of collapse for aged pipes
	Average Depth	If the depth is very low, the pipe is susceptible to surface live load. If depth is high, the pipe is susceptible to overburden. Moderate depths increase the life of sewers
	Pipe Gradient	Steeper slopes of pipe cause high flow velocity which increases erosion phenomena
Functional	*M & R Strategies*	Good maintenance and repair strategies increase the service life of sewers
Environmental	*Type of waste*	Different types of waste react with different pipe materials in a different manner causing pipe erosion
	Ground Water	Groundwater can cause infiltration, which washes soil particles and reduces the soil support along the pipe
	Type of Soil	Different types of soils provide side supports to pipes according to their own physical and chemical properties
	Bedding Conditions	The chance of pipe failure increases with improper bedding condition of pipes
	Frost Factor	The load on buried sewers increases due to additional frost load in winter
	Other Utilities	Proximity of other underground installations increases the criticality of a sewer
	Traffic Volume	The bending stresses in the pipe increase with the increase in live load above pipe

Table 2-3 (Adapted from Barqawi, 2006 & Kulandaivel, 2004) shows these factors in detail. Based on the literature review the table also explains how these factors contribute in pipeline structural deterioration phenomena.

2.3.2 Factors Affecting Sewer's Operational Condition

There are several factors which could deteriorate the over all operational condition of sewers causing over flows. These factors can be divided into two categories (May et al. 1998): non-hydraulic and hydraulic.

Non-hydraulic problems are generally defined as those deficiencies in sewer performance which are not due to lack of flow capacity within the sewer system. As shown in Figure 2-3 (May et. al 1998), these problems are uncertain; for example, random blockage of flow due to some object or pumping station failure etc. However, some factors such as structure condition of a pipe are more likely predictable. The structural condition of a sewer affects directly on its flow capacity. Older pipes have rougher inner surface, more structural cracks, and deformations. These results in more debris and reduction in diameter due to deformation, as well as problems related to the infiltration phenomenon. A pipe's structural condition further depends upon many factors (Table 2-3). Therefore, all these factors directly or indirectly have an influence on the operational condition of sewers. Further, non-hydraulic problems are greatly dependent upon the operational and maintenance strategies and history. Routine maintenance and repair programs could increase the service lives of sewers.

Hydraulic problems occur when a sewer is not adequate enough to sustain high volume of flow. The causes of these problems could be faulty design for pipe size and its gradient. Pipe size includes its diameter and length. Larger diameter pipes can accommodate larger volume of flow. Similarly, longer lengths of pipe mean less bends to accumulate debris creating blockage (Kulandaivel, 2004). Another major hydraulic factor is infiltration (Abraham et al 1998) and inflow. Infiltration occurs when groundwater enters a sewer system through broken pipes, defective pipe joints, or illegal connections of foundation drains; while inflow is surface runoff that enters a sewer system through manhole covers and exposed broken pipe.

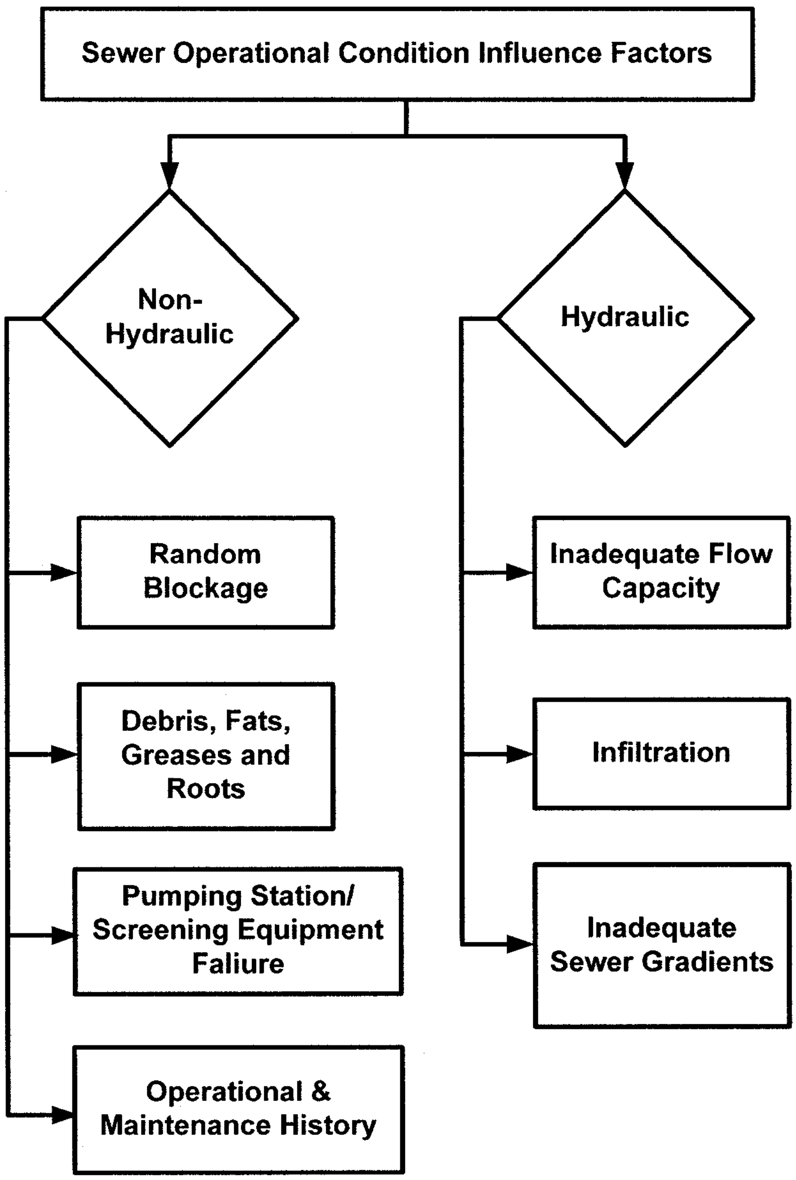


Figure 2-4: Major Factors Affecting Sewer Pipeline Operational Condition

2.4 Inspection and Evaluation of Existing Sewers

Sewer inspection and evaluation are routine tools for the municipal agencies. The first method of inspecting pipelines was developed during the aftermath of Second World War (Allouche et al 2002). Since then, several inspection techniques have been developed to evaluate existing condition of sewers.

2.4.1 Inspection Objectives

There are three main objectives of inspecting sewers (Adapted from Butler et al 2000):

1) Periodic inspection to assess the condition of existing sewers
2) Crisis inspection to investigate emergency conditions or the cause of repeated problems along a particular sewer length
3) Inspection of workmanship and structural condition of new sewers before adoption

2.4.2 Sewer Inspection techniques

There are various sewer inspection techniques that are used for condition assessment of sewers and can be classified into three different groups (Makar, 1999). An overview of this classification is illustrated in Figure 2-4 (Adapted from Makar, 1999). The first group consists of techniques that determine the internal condition of a sewer. The commonly used examples for this group are closed circuit television (CCTV) inspections and sewer scanner and evaluation technology (SSET). The second group examines the overall condition of sewers and the surrounding soil. The last group detects specific problems within or behind a sewer wall (Kulandaivel, 2004). Among all the inspection techniques, CCTV inspection remains to be the mostly used technique (Makar, 1999). Table 2-4

(Adapted from Makar, 1999 and Kulandaivel, 2004) presents a brief summary of commonly used inspection techniques for assessing the existing condition of sewers. IT suggests the most appropriate usage of different techniques, and illustrates the important advantages as well as disadvantages of these techniques.

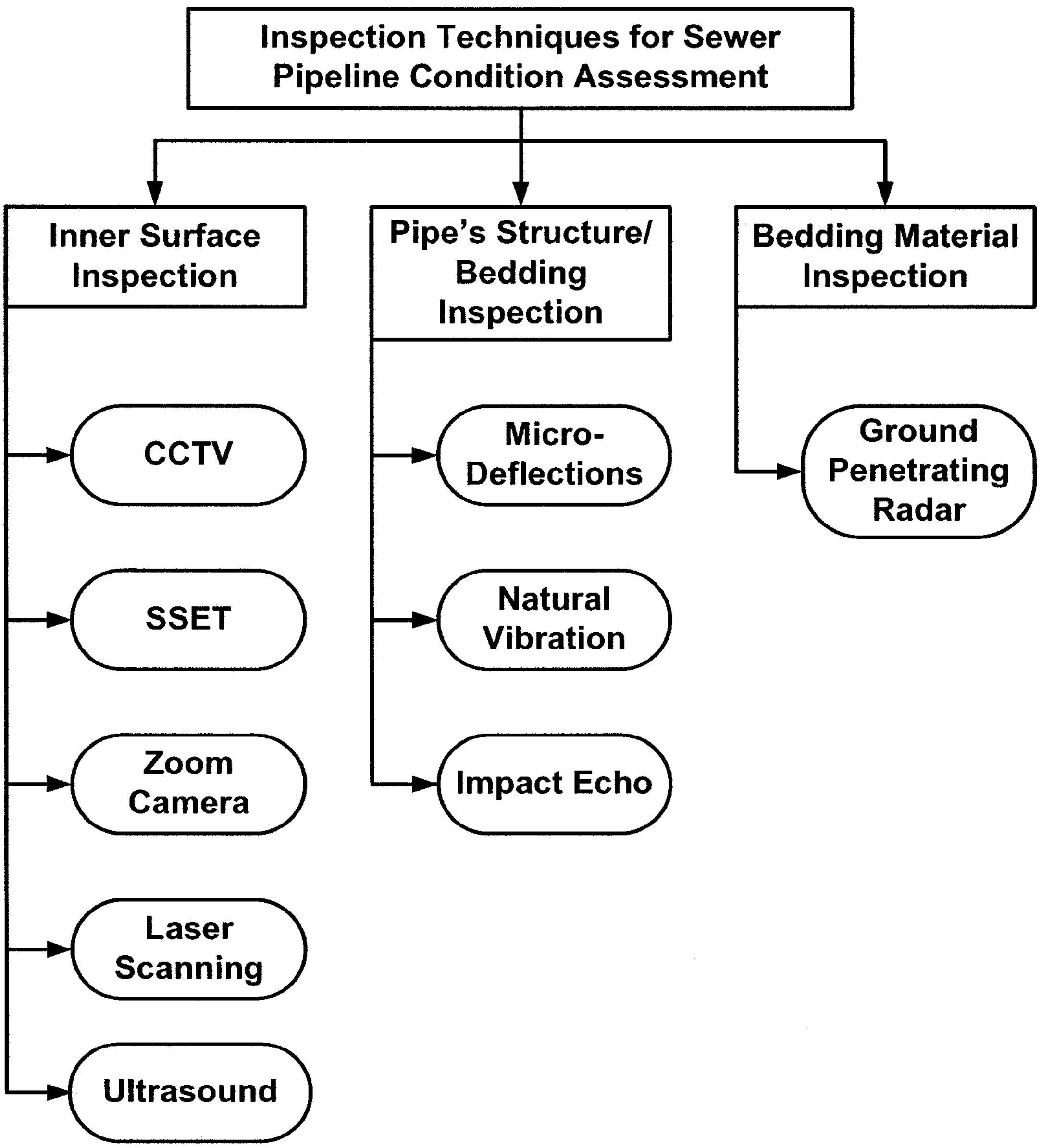


Figure 2-5: Inspection Techniques for Sewer Pipeline Condition Assessment

Table 2-4: A Comparison of Commonly Used Sewer Inspection Technique

Inspection Group	Technique	Usage	Common Detections	Advantages	Disadvantages
Inner surface	CCTV	Empty and partially filled pipes	Surface cracks, deformations, erosion, infiltration, blockage	Standard technique, relatively cheap, easily available, evaluates entire length	May miss hidden defects, interpretation of results
	SSET	Pipes of diameter from 8 to 24 inches	As CCTV but with higher accuracy	Higher accuracy, good judgment of pipe deformation	Expensive
	Zoom Camera	Empty and partially filled pipes	As CCTV but with lower accuracy	No cleaning/flushing is required before inspection, cheaper than CCTV	Cannot cover whole length of a pipe, results are less accurate than CCTV
	Laser Scanning	Empty and partially filled pipes	Surface cracks, erosion, missing bricks, deformations	Computer based analysis, accurate defect and geometry measurement	Only works above water lines
	Ultrasound	Empty to flooded pipes	Deformation, erosion, brick damage	Measures defect above and below water line, computer based analysis	Extensive cleaning of sewer necessary, more expansive
Pipe Structure/ Bedding	Micro-deflections	Rigid pipes	Overall mechanical strength	Not affected by bedding condition, directly measures pipe's structural integrity	Rigid pipes only, not locates individual defects
	Natural Vibrations	Empty sewers	Combined soil and pipe condition, crack region, exfiltration	Directly measures pipe's structural integrity without traveling entire length	Effect of bedding condition unknown, requires more sewer cleaning
	Impact Echo	Large diameter sewers	Combined soil and pipe condition, crack region, exfiltration	Detects voids behind sewers, good for brick and concrete sewers	Manually operated equipment, does not locate individual defects
Bedding	Ground Penetrating Radar	Empty and partially filled pipes	Voids around pipe, water content in bedding	Detection of exfiltration, voids, rocks and other objects in bedding	Interpretation of results, more expensive

2.5 Sewer Pipeline Condition Assessment Protocols

2.5.1 Introduction

Sewer defect coding has become of paramount importance for the worldwide sewer rehabilitation industry to ascertain critical information regarding the underground infrastructure (Thornhill et al 2005). The historical background of the development of sewer defect codes or condition assessment protocols goes back to 1977; when for the first time, sewer defect codes were developed by Water Research Centre (WRc) UK. The first manual for sewer condition classification was published by WRc in 1980. On the basis of the guidelines provided by WRc, several condition assessment protocols were developed through out the world during the past twenty five years. These developments are illustrated in Figure 2-5 (Developed from Thornhill et al 2005).

The WRc sewer condition assessment protocols are accepted world wide, including Canada. In Canada, North American Association of Pipeline Inspectors (NAAPI), Centre of Advanced Trenchless Technologies (CATT), and many municipal agencies have adapted the WRc sewer defect coding techniques. However, National Research Council (NRC) of Canada has introduced its own coding and grading system. Nevertheless, NRC Coding system is strongly based upon WRc theory. In the Province of Quebec, CERIU (Centre for Expertise and Research on Infrastructures in Urban Areas) with the help of BNQ (Bureau de normalisation du Québec) developed their own sewer defect codes in 1997. The CERIU codes have been adapted by most of the municipalities in the Province. As a consequence, WRc and CERIU protocols are the two basic sewer condition assessment codes which have been adapted by most of municipal agencies in Canada.

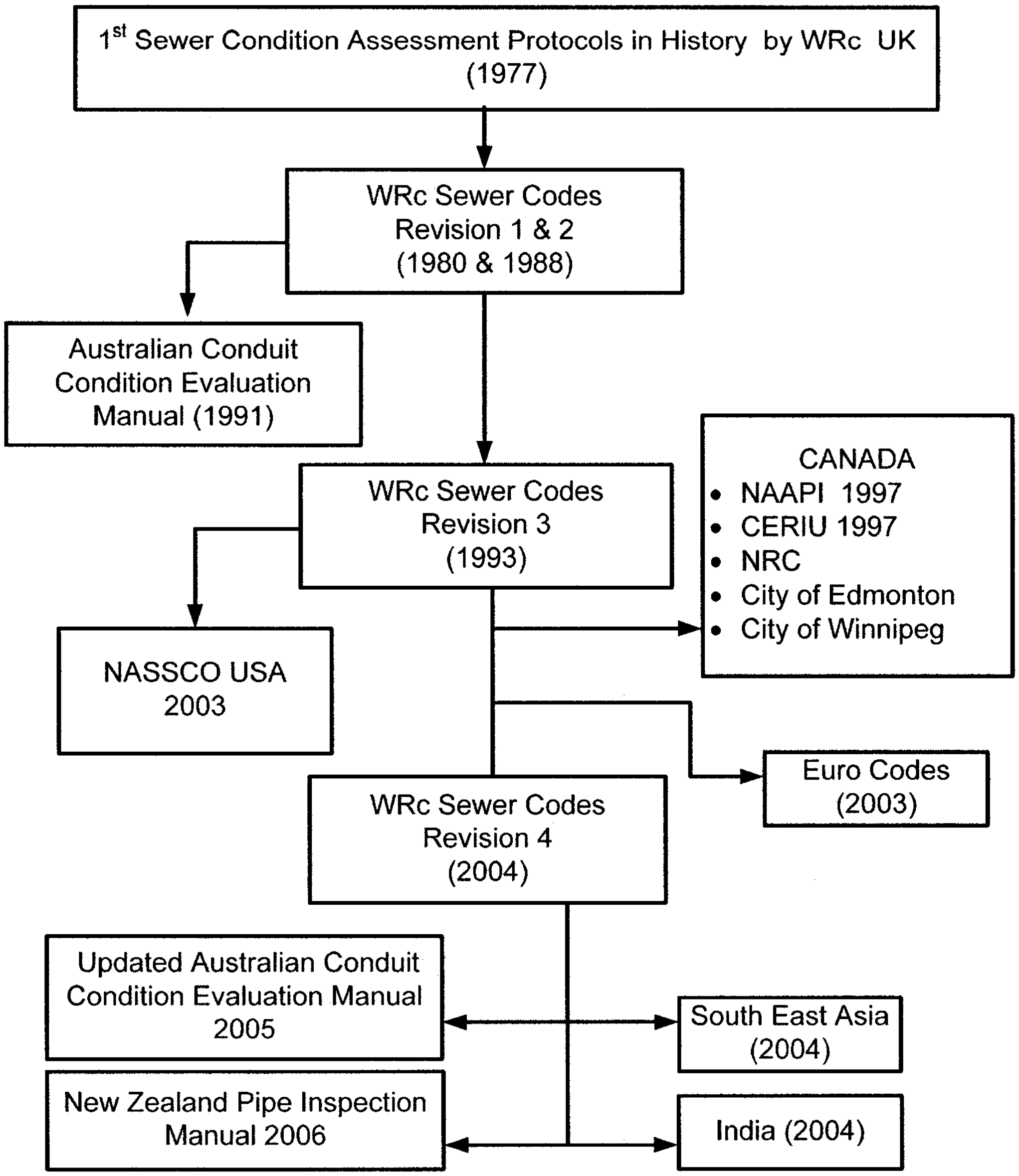


Figure 2-6: An Overview of the Historical Background of Sewer Defect Codes Development

2.5.2 WRc Condition Assessment Protocols:

According to SRM (Sewerage Rehabilitation Manual) 4th Edition (2004) of WRc, sewer pipeline defect codes have been divided into two major categories: structural and operational defects. The structural and service condition rating is generated from the number and severity of defects for each pipeline. Severity or criticality scores have been allocated in accordance with the general principles outlined in Table 2-5 (Developed from New Zealand Water and Waste Water Association (NZWWA, 2006).

Table 2-5: WRc Defect Scores General Guideline

Criticality Code	Description
L	**Light Defects:** Defects which should not cause any problem in the near future (10 years or more). (e.g. for Structural Grading; Defect Score is < 10)
M	**Medium Defects:** Minimal short term failure risk; however no urgent action is required (e. g. for Structural Grading: Defect Score is from 10 to 25).
S	**Severe Defects:** Immediate risk of failure or severe loss (e. g. for Structural Grading: Defect Score is > 30).

An overall sewer condition grade for the whole pipe segment is identified by a number from 1 to 5 (WRc, 2004); the overall condition grade is also called as Internal Condition Grade (ICG).These numbers determines the probability of collapse and are illustrated in table 2-6:

Table 2-6: The Severity of WRc Condition Grades (Developed from WRc, 2004)

Condition Grade	Description
1	Acceptable Condition
2	Minimal Collapse Risk but Potential for Further Deterioration
3	Collapse unlikely but Further deterioration likely
4	Collapse Likely in Near Future
5	Collapse Imminent or Collapsed

WRc also defines Internal Condition Grades (ICG) in terms of criticality and rehabilitation requirements which are tabulated in Table 2-7.

Table 2-7: WRc Condition Classification as per Criticality and Rehabilitation Priority

Condition Grades	Criticality Code	Rehabilitation Priority
1	L	Not Required
2	L	Low
3	M	Medium
4	S	High
5	S	Immediate

Internal Condition Grade (ICG) is determined by a defect score calculation that is based on various defects in a pipe segment. The value, or weight, for each defect is assigned, and the impact of the defect on the service life and performance of the sewer pipe segment is determined. The total score represents the summations of all deduct values while the peak score represents the highest deduct value in the pipe segment. The mean score reflects the overall condition of a pipeline and is the average of defect scores per meter of pipeline (NZWWA, 2006). The mean score can be determined by the formula in equation 2-1.

$$Mean\ Score = \frac{\sum(Deduct\ Values)}{Length\ of\ Pipe} = \frac{Total\ Score}{Length\ of\ Pipe} \qquad \text{Equation 2-1}$$

The peak score reflects the magnitude of the worst defects in each pipeline. The peak score is the maximum defect score for any one meter length of pipe within a pipeline (NZWWA, 2006).

$$Peak\ Score = Maximum\ Deduct\ Value \qquad \text{Equation 2-2}$$

(i) Structural Defects

These defects describe the physical condition of a pipe. The defect score assigned to structural defects depends upon severity of defects and type of pipe material. Table 2-8 shows some defect scores for some common defects in concrete pipes.

Table 2-8: WRc Structural Defect Scores for Some Common Defects (WRc, 2004)

Defect	Detail	Score	Unit
Joint Opening	Slight	0.1	Per Joint
	Medium	0.5	Per Joint
	Large	2	Per Joint
Joint Displacement	Slight	0.1	Per Joint
	Medium	0.2	Per Joint
	Large	5	Per Joint
Crack	Circumferential	1	Per Crack
	Longitudinal	2	Per Crack
	Multiple	5	Each
Fracture	Circumferential	8	Per Crack
	Longitudinal	15	Per Crack
	Multiple	40	Each
Deformation	5%	10	Each
	10%	30	Each
	15%	60	Each
	20%	90	Each
	25%	125	Each
	30% or More	165	Each
Hole	<1/4 Circumferential	80	Each
	>1/4 Circumferential	165	Each
Broken Pipe		80	Each
Collapsed Pipe		165	Each

The defect scores are calculated and on the bases of peak defect score or deduct value, a single condition grade for the structural condition of pipe is assigned according to Table 2-9.

Table 2-9: WRc Guidelines for Calculating the Overall Structural Condition Grade of a Sewer Pipe Based upon Peak Score (Developed from NAAPI, 2002)

Overall Structural Condition Grade of a Pipe Segment	**Peak Structural defect Score Found in Segment**
1	< 10
2	10-39
3	40-79
4	80-164
5	165 & above

(ii) Operational Defects

Operational defects describe the capability of a sewer pipe to meet its service requirements and indicate the loss of capacity, potential of blockage and water tightness due to the main factors described in Table 2-10. The individual defect scores for these most common operational defects are also tabulated. General guide lines for evaluating over all operational conditions of pipes are same as described for structural conditions. As mentioned above, WRc suggests peak scores in determining internal condition grade (Rahman et al 2004); nevertheless, NAAPI takes into account both mean and peak scores for evaluating the operational condition grade (Table 2-11).

With the help of all the codes mentioned above, structural and operational condition of a pipe can be assessed separately. However, WRc does not suggest any condition grading index for combined structural and hydraulic condition assessment of pipes.

Table 2-10: WRc Operational Defect Scores for Some Common Defects (WRC, 2004)

Defect	Detail	Score
Roots	Fine	1
	Tap	5
	Mass < 5%	2
	5% - 20%	4
	20% +	10
Encrustation	Light	1
	Medium	2
	Heavy	5
Debris (Including Silt & Grease)	5%	1
	5% - 20%	2
	20% - 50%	5
	50% - 75%	8
	>75%	10
Obstruction		10
Infiltration	Water Seeping Through Cracks	Sipper
	Water Dripping Through Cracks	Dripper
	Water Gushing Through Cracks	Gusher

Table 2-11: WRc Operational Condition Grade for a Sewer Pipe (NAAPI, 2002)

Overall Operational Condition Grade for a Pipe Segment	Peak Operational defect Score Found in Segment	Mean Operational Defect Score of the Segment
1	< 1	< 0.5
2	1 – 1.9	0.5 – 0.9
3	2 – 4.9	1 – 2.4
4	5 – 9.9	2.5 – 4.9
5	> 10	> 5

(iii) WRc Construction Codes

In addition to the structural and operational defects, WRc addresses some additional features which are classified as construction codes. These codes deal with different features associated with the installation of sewers. These codes are used to identify encountered, pre-existing construction features; such as, connections, manholes, linings, etc. These codes are also important, especially for analyzing defective and protruding service connections affecting the structural and operational condition of sewers.

2.5.3 CERIU Condition Assessment Protocols

CERIU sewer condition classification protocols were developed in 1997 and were revised through "Manuel de standardization des observations" 2nd Edition, 2004. The CERIU condition classification codes have been adapted by municipalities all across the Province of Quebec.

Table 2-12: The Severity of CERIU Condition Grades

CERIU Condition Grade	Description
1	No Intervention, Action Required
2	Action Required but not Major
3	Action Required but Not Urgent
4	Action Required and Urgent
5	Immediate Action Required

CERIU does not suggest overall structural or hydraulic internal condition grade (ICG) for a whole sewer pipe segment like WRc codes do. On the contrary, CERIU suggests 5 different classes for each structural or hydraulic defect. These numbers consider the intervention or rehabilitation requirements as the key factor. Table 2-12 better explains these condition grades.

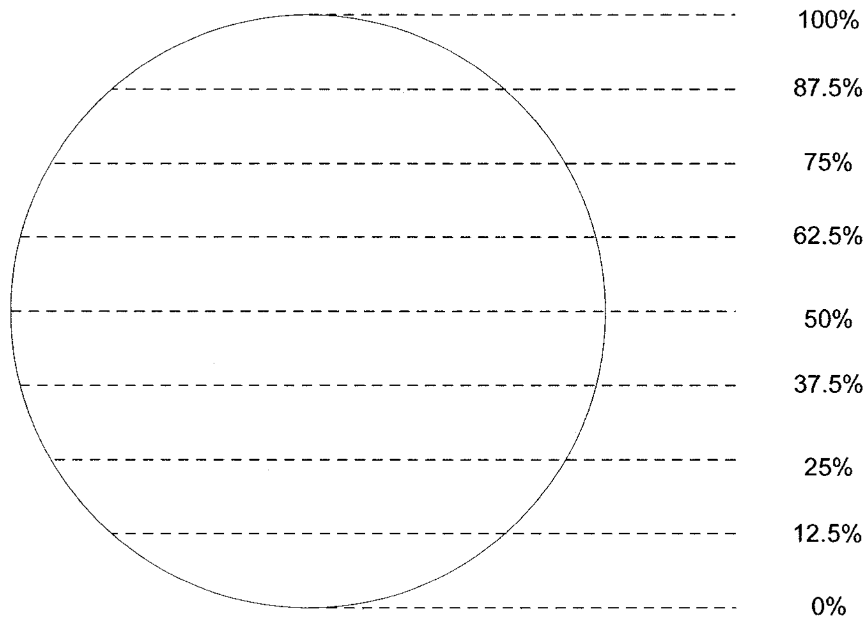


Figure 2-7: CERIU Division of Sewer Pipe's Cross-Sectional Area (CERIU, 2004)

In order to judge the severity of the defect, CERIU suggests the methodology of referencing defects of a sewer pipe by dividing a pipe's cross-sectional area into eight equal parts (1/8, 1/4, 3/8,..., 1) and calculating the percentage as 6, 12.5, 25, 37.5%, etc; thus slightly differs from WRc's methodology which considers the percentage of cross-sectional area of pipe directly (5%, 10%, 20%, 25%, etc.). Figure 2-6 illustrates the CERIU approach more clearly. CERIU addresses the issue of sewer pipeline condition assessment in four different scenarios; structural defects, hydraulic defects, infiltration, and junction/connection condition. All of them are discussed below:

(i) Structural Defects

CERIU describes the physical condition of a pipe by classifying each defect in the pipe individually. The condition class assigned to structural defects depends upon severity of the defect and type of pipe. Table 2-13 shows some defect classes for some of the most common defects in rigid pipes:

Table 2-13: CERIU Structural Defect Condition Classes for Some Common Defects

Defect	Class	Description
Joint Opening	1	Opening ≥ 12 mm and ≤ 25 mm
	2	Opening ≥ 25 mm and ≤ 50 mm
	3	Opening ≥ 50 mm and ≤ 100 mm
	4	Opening > 100 mm
Joint Displacement	2	Displacement ≤ 6% of Pipe Diameter
	3	Displacement > 6% and ≤ 12% of Pipe Diameter
	4	Displacement > 12% and ≤ 25% of Pipe Diameter
	5	Displacement > 25%
Circumferential Crack	1	Crack Visible without Any Visible Opening
	2	Visible Opening ≤ 1.5mm
	3	Visible Opening > 1.5mm and ≤ 8 mm
	4	Visible Opening > 8 mm and ≤ 16 mm
	5	Visible Opening > 16 mm
Longitudinal Crack	1	Crack Visible without Any Visible Opening
	2	Visible Opening ≤ 1.5mm
	3	Visible Opening > 1.5mm and ≤ 5 mm
	4	Visible Opening > 5 mm and ≤ 10 mm
	5	Visible Opening > 10 mm
Multiple Crack	1	Multiple Crack Visible without Any Visible Opening
	2	Visible Opening ≤ 1.5mm
	3	Visible Opening > 1.5mm and ≤ 5 mm
	4	Visible Opening > 5 mm and ≤ 10 mm
	5	Visible Opening > 10 mm
Deformation	4	Rigid Pipe Deformation > 6% to ≤ 12.5% of Diameter
	5	Rigid Pipe Deformation > 12.5% of Diameter
Hole	3	Hole Diameter ≤ 25 mm
	4	Hole Diameter > 25 mm and ≤ 100 mm & Hole Diameter < 25% of the Pipe's Circumference
	5	Hole Diameter > 100 mm and > 25% of the Pipe's Circumference

(ii) Hydraulic Defects

CERIU describes the capability of a sewer pipe to meet its hydraulic capacity by the main factors described in Table 2-14. The individual condition classes for these most common hydraulic defects are also tabulated.

Table 2-14: CERIU Hydraulic Defect Condition Classes for Some Common Defects

Defect	Class	Description
Roots	1	Roots Formation on Rim ≤ 12.5% of Pipe Diameter
	2	Roots Formation on Sill ≤ 12.5% of Pipe Diameter
	3	Roots Occupying > 12.5% and ≤ 25% of Pipe Diameter
	4	Roots Occupying > 25% and ≤ 37.5% of Pipe Diameter
	5	Roots Occupying > 37.5% of Pipe Diameter
Deposits (Soft/Hard/ Lime Stone)	1	Deposits ≤ 6% of Pipe Diameter
	2	Deposits > 6% and ≤ 12.5% of Pipe Diameter
	3	Deposits > 12.5% and ≤ 25% of Pipe Diameter
	4	Deposits > 25% and ≤ 37.5% of Pipe Diameter
	5	Deposits > 37.5% of Pipe Diameter
Grease	1	Grease ≤ 6% of Pipe Diameter
	2	Grease > 6% and ≤ 12.5% of Pipe Diameter
	3	Grease > 12.5% and ≤ 25% of Pipe Diameter
	4	Grease > 25% and ≤ 37.5% of Pipe Diameter
	5	Grease > 37.5% of Pipe Diameter
Visible Material	2	Apparent Trimming < 10% of Pipe Diameter without collapsing
	3	Apparent Trimmed Obstruction on Upper Half Periphery of Pipe
	4	Apparent Trimmed Obstruction Make the Most Vertical or Slanting Part of Pipe
	5	Apparent Trimmed Obstruction on Lower Half Periphery of Pipe
Obstructing Object	1	Object Little Harmful to Flow
	3	Hydraulic Capacity of Pipe Reduced Due to Obstruction
	5	Hydraulic Capacity of Pipe almost Reduced to Zero Due to Obstruction
Water Level	1	Water Depth ≤ 6% of Pipe Diameter
	2	Water Depth > 6% and ≤ 12.5% of Pipe Diameter
	3	Water Depth > 12.5% and ≤ 25% of Pipe Diameter
	4	Water Depth > 25% and ≤ 37.5% of Pipe Diameter
	5	Water Depth > 37.5% of Pipe Diameter

(iii) Service Connections

CERIU recognizes three major and most common defects for service connections which may cause obstruction to flow or loss of hydraulic capacity of whole system. These defects along with their condition classes are explained in Table 2-15.

Table 2-15: CERIU Service Condition Classes for Sewers

Connection Defects	Class	Description
Connection Penetration	1	Penetration < 6% of Pipe Diameter
	2	Penetration > 6% and ≤ 12.5% of Pipe Diameter
	3	Penetration > 12.5% and ≤ 25% of Pipe Diameter
	4	Penetration > 25% and ≤ 37.5% of Pipe Diameter
	5	Penetration > 37.5% of Pipe Diameter
Clogged or Choked Connection	2	Obstruction ≥ 6% and ≤ 12.5% of Connection's Diameter
	3	Obstruction > 12.5% and ≤ 25% of Connection's Diameter
	4	Obstruction > 25% and ≤ 37.5% of Connection's Diameter
	5	Obstruction > 37.5% of Connection's Diameter
Flow from Connection	2	Connection's Flow ≤ 60 drops/min (≤ 6 ml/min or 1 Drinking Glass/Hour)
	3	Connection's Flow > 6 ml/min and ≤ 500ml/min
	4	Connection's Flow > 500ml/min ≤ 10 L/min
	5	Connection's Flow > 10 L/min

(iv) Infiltration

CERIU code deals with infiltration phenomena separately and does not consider it as a part of hydraulic defects.

Table 2-16: CERIU Infiltration Condition Classes for Sewers

Infiltration Class	Description
1	Trace of Infiltration
2	Infiltration ≤ 60 drops/min (≤ 6 ml/min or 1 Drinking Glass/Hour)
3	Infiltration > 60 drops/min (> 6 ml/min and ≤ 500ml/min)
4	Infiltration > 500 ml/min and ≤ 10L/min
5	Infiltration > 10L/min

Infiltration through service connections and through main pipes is usually dealt separately by assigning them appropriate condition classes as defined by the code. The condition classes for infiltration have been defined in Table 2-16.

2.6 Research Techniques in Sewer Management

In order to plan sewer management and maintenance activities more effectively, tools are required which can prioritize any proactive work by predicting sewer condition and performance. This requires the development of suitable analytical techniques to analyze past performance in an attempt to better direct future proactive maintenance activities (Fenner, 2000). Therefore, different research techniques have been adapted for the assessment of condition, performance, and deterioration of sewer pipes all around the world. The theory behind the adoption of these techniques is inspired by the assessment and deterioration models for pavements and bridges which have been the primary focus for research in infrastructure systems (Abraham et al 1998). Following are the important research techniques for sewer management:

2.6.1 Regression Analysis

Regression analysis is a statistical tool for the investigation of relationships between variables. Usually, the investigator tries to find the effect or some kind of relationship of one variable upon another/others. The investigator also typically assesses the degree of confidence that the true relationship is close to the estimated relationship (Sykes, 1986). Applications of regression analysis exist in almost every field. The common aspect of the applications is that the dependent variable is a quantitative measure of some condition or behaviour (Andreu et al 1987).Prediction models currently used for predicting the performance of infrastructure system includes regression techniques (Butt, 1994). However, regression techniques are valid only if the predictive variables can be found that are related to sewer condition deterioration (Abraham et al 1998).

2.6.2 Deterministic Models

Deterministic modelling procedure consists of mathematical modeling, numerical simulation, clustering the solutions, and sensitivity analysis (Sumida et al 2001).In deterministic model development, parameters under consideration are classified according to their similarities. When historical condition data are not readily available, the use of deterministic dynamic programming together with a heuristic prediction model based on expert opinion is suggested (Abraham et al 1998). In short, deterministic methods may provide general guidelines, but they are inadequate for establishing accurate water and sewer management decisions. An alternative and more favourable approach is to be used that essentially simulate a range of potential outcomes (Stengel, 1994).

2.6.3 Logistic Models

Logistic models are a special form of regression models. It is a multiple regression with a categorical response variable (dependent variable) and explanatory variable(s) which can be either continuous or categorical (Varela, 2005). Logistic models can be developed for assessing the condition of existing sewer pipeline network as well as understanding the deterioration phenomena of the network. Ariaratnam et al (2001) used logistic models to predict the like hood that a particular infrastructure system could be in a deficient state. The methodology was illustrated in a case study involving modeling of different attributes related to sewer pipes for providing decision makers with a means of evaluating sewer sections for the planning of future scheduled inspections. The logistic model for pipe deficiency probability (π) was developed through pipe parameter: pipe age, pipe diameter, and type of waste. A sensitivity analysis was also performed to validate the proposed model. However, it was concluded that the model results would only be as good as the quality of data collected.

2.6.4 Probabilistic Models

Probabilistic models estimate a distribution of values for future condition. By applying probabilistic model analysis, judgment can be built into estimates (Maze, 2005). The concept of probabilistic models can be used to determine the probability of a pipe to enter in a stage; thus calculating the economic life of a pipe (Andreou, 1987). Fenner et al (2000) developed a probability based model to predict the like hood of sewer failure in a grid square. The model was based on the analysis of pipe data contained in a series of grid squares defined by GIS software. A consequence factor was allocated to each grid square which was based on global and local matrices; which affect the community and individual customers respectively. These likelihood and consequence values were combined in a two dimensional risk plot which enabled the identification of "Critical Grid Squares"

Another probabilistic approach is known as the cohort survival model for urban infrastructure network which was developed by Herz (1996). Cohorts are defined as a set of elements installed in the same year with a particular failure probability. The state survival function of a sewer system specifies the probability of a transition between the various classes in the form of a Herz distribution curves. Frangopol et al (2004) presented a review of different probabilistic models developed for maintaining and optimizing the life cycle performance of the deteriorating structures. After analyzing different modeling approaches, it was found that none of the probabilistic approaches was generally applicable, and the use of each model was limited to certain extent.

2.6.5 Markovian Models

The basic theory of Markovian modeling is that probabilities involving how the process will evolve in the future depends upon the present state of the process; therefore, independent of the processes in the past. Abraham et al (1998) explained the concept of adaptation of Markovian chain process in deterioration modeling for sewers. To model a sewer's deterioration, Markov probability transition matrix could be developed. The transition matrix P would be a square matrix of order $m \times m$, where m would be the number of possible states. Therefore, if there were five categories in sewer condition, then five possible states would be involved in the transition matrix of order 5×5. The components of P, p_{ij} would represent the probabilities of being in state i at the time 0 and transitioning to state j over a given period Δt.

Markovian models provide a reliable mechanism for development of prediction models, and Markov chain can be employed to model stochastic processes (Hillier et al 1995). However, the model development requires sufficient statistical data for establishing sound transition probability matrices (Kulandaivel, 2004).

2.6.6 Fuzzy Logic Based Approach

Fuzzy provides a remarkably simple way to draw definite conclusions from vague, ambiguous or imprecise information (Mills et al 1996). One of the key applications of this approach is the transformation of non – mathematical linguistic variables into fuzzy variables. The main idea is to have an almost direct correspondence between qualitative data of the rough models and fuzzy values of fuzzy variables. The naming the values of fuzzy variables using suitable words in our everyday language was developed by Zadeh in 1975; who created this idea of calling fuzzy variables as linguistic variables (Bandemer et al 1995).

Yan et al (2003) proposed a method for assessing the condition of pipes by applying Fuzzy set theory to convert linguistic descriptions of pipe condition indicator into numerical format. These numerical values of linguistic variables were used to develop a model that ranked pipes in order of their condition. The main linguistic variables which were considered in developing the condition assessment model were traffic density and environmental conditions in the surrounding area of pipe. It was concluded that the use of Fuzzy theory in pipeline condition assessment would be very helpful for decision makers.

2.6.7 Artificial Neural Networks

ANN (Artificial Neural Networks) is an information processing model that is inspired by the human or biological nervous systems. The key element of this model is the structure of information processing system. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process (Stergiou et al 2005). ANN are applicable in virtually every situation in which a relationship between predictor and response (independent and dependent) variables exists, even when that relationship is very complex (Smith, 2003).

Kulandaivel (2004) developed an Artificial Neural Network (ANN) Model for predicting the condition of sewer pipes based on the historic condition assessment data. The model development in the study included training, testing and validation. The author was in opinion that using neural network methodology for predicting the condition of sewer pipes was the best possible option available to him. Shehab-Eldeen (2001) developed an automated system for detection, classification and rehabilitation of sewer pipes using neural networks. The developed model was based upon image analysis techniques, artificial neural networks and visual basic programming.

Combining two research techniques like ANN and Fuzzy logic has been gaining popularity in the sewer management operations. Chae et al (2001) proposed a Neuro-Fuzzy approach for accurately analyzing and interpreting data regarding the condition of sanitary sewer pipelines. The involvement of Fuzzy estimation techniques in ANN enabled the modeling of uncertainty associated with the input data.

2.6.8 Simulation

Simulation is a mathematical exercise in which a model of a system is established, and then the model's variables are altered to determine the effects on other variables. Simulation is an efficient and cost-effective tool for decision making and analyzing real systems (Ruwanpura et al 2004). Many simulation tools have been developed around the world for strategic decision making planning for sewer maintenance.

Stein et al 2005 developed a model for analyzing the environmental impacts caused by the defects in sewer network system using Monte Carlo simulation. Results of the simulation established the links between the local ancillary conditions and the specific sewer defect characteristics. Denys et al 2004 developed a simulation model for factors influencing sewer system properties to give sewer managers an idea of all elements in their decision making process for sewer network maintenance. The simulator contained some "moduli" which computed the level of performance of the system from the data actually available or when there was a deficiency. The risks were also evaluated through statistical analysis of available data.

A rule based simulation condition assessment model was developed by Ruwanpura et al (2004). The procedure of the simulation involved CCTV inspection data analysis. The developed simulation model predicted the condition rating of pipes based on four inputs: type of material, age class, pipe length, and the APE value. The actual probability of existence (APE) curve was developed through simulation of the collected data. Allouche et al (2003) introduced educational based simulation software "SIMSEWER" for the enhancement of sewer management skills. The main objective of the software was to introduce the basic concepts of infrastructure management, engineering economics and various techniques of sewer rehabilitation. The program was based on real case history and enabled the user to develop management strategies for a simulated trunk sewer network.

2.6.9 Bayesian Belief Networks

Bayesian belief networks (BBN) are compact networks of probabilities that capture probabilistic relationship between variables, and historical information about their relationships. Bayesian belief networks are very effective for modeling situations where some information is already known and incoming data is uncertain or partially unavailable (Russel et al 2003). The probability of any node in the Bayesian belief network being in one state or another without current evidence is described using a conditional probability table. Probabilities on some nodes are affected by the state of other nodes, depending on causality. Prior information about the relationships among nodes may indicate that the likelihood that a node is in one state is dependent on another node's state (Grzymala, 1991).

Hahn et al (2000) proposed a sewer management system for prioritizing sewer line inspections on the basis of Bayesian belief networks. The belief network logic system was based on information provided by the sewer utilities of ten medium sized cities. The concepts of Bayesian belief networks were used for creating an inference engine that catalogued information for both inspected and non-inspected sewer lines. The system developed was found helpful in prioritizing the inspection of sewer lines.

2.6.10 Dynamic Programming

Dynamic programming is a way of decomposing certain hard to solve problems into equivalent formats that are more amenable to solution. Basically, the dynamic programming approach is to solve a multi-variable problem by solving a series of single variable problems. This is achieved by tandem projection onto the space of each of the variables (Benli, 1999). Dynamic Programming Technique starts with a small subset of the original problem which is called sub-problem. The second step of this technique is to find the optimal solution of the sub problem. Consequently, it gradually enlarges the sub problem and finds the current optimal solution from the preceding sub problem. This step is repeated until the entire problem is solved (Trick, 1997).

Dynamic programming in conjunction with Markov chain modeling has been successfully used in highway management systems (Abraham et al 2003). In sewer management systems, decision regarding maintenance and rehabilitation needs to be supported by a sound decision making procedure. Dynamic programming has significant potential for developing solutions in this regard. Therefore, this method is viable for life cycle cost analysis for sewer management system because it provides a quick method of finding optimal maintenance and rehabilitation options for each analysis period or stage through out the life cycle of a sewer (Wirahadikusumah, 1999).

2.7 An Overview of Applied Techniques

Regression analysis and unsupervised neural network are applied in this research. The basic ideology of both the techniques is discussed below:

2.7.1 Regression Analysis

Regression analysis is a process used to estimate the parameter values of a function, in which the function predicts the value of a response variable (Y) in terms of the values of other variables (X). In its simplest form the model can be stated as follows (Neter et al 1996):

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad \text{(Equation 2.3)}$$

where, Y_i is value of response variable in the i^{th} trial, β_0 & β_1 are regression parameters, X_i is the value of predictor variable in the i^{th} trial, and ε_i is random error. In multiple regression models, more than one variable are used to predict the behaviour of response variable. Therefore Equation 2.3 can be transformed into Equation 2.4 for p-1 predictors:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots\dots\dots\dots\dots\dots + \beta_{p-1} X_{ip-1} + \varepsilon_i \qquad \text{(Equation 2.4)}$$

To estimate the regression parameters, the method of least square can be applied. According to the method, the values b_0 and b_1 are estimated for the regression parameters β_0 & β_1 (equation 2.3) respectively to minimize the sum of square deviation for the given sample observations (X_1, Y_1), (X_2, Y_2),…………, (X_n, Y_n). Therefore, b_0 and b_1 are called point estimators of β_0 and β_1 respectively, and their values can be estimated from the following equations (Kutner et al 2005):

$$b_1 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} \qquad \text{(Equation 2.5)}$$

$$b_0 = \frac{1}{n}(\sum Y_i - b_1 \sum X_i) = \bar{Y} - b_1\bar{X} \qquad \text{(Equation 2.6)}$$

where $\bar{X}$ and $\bar{Y}$ are the mean of X_i and Y_i observations respectively. Consequently, b_1 is the slope of regression line and b_0 is the Y intercept for the line. In the method of least square, the variation in regression relation is measured in the form of total sum of squares (SSTO); which is the measure of variation of the Y_i values around their mean $\bar{Y}$.

$$SSTO = \sum_{i=1}^{n}(Y_i - \bar{Y})^2 \qquad \text{(Equation 2.7)}$$

SSTO is further divided into two categories: regression sum of squares (SSR) and error / residual sum of squares (SSE) (Levine et al 2002). *SSR* is equal to the sum of squared difference between each predicted and average value of response variable.

$$SSR = \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2 \qquad \text{(Equation 2.8)}$$

SSE is equal to the sum of squared difference between each observed and predicted value of the response variable

$$SSE = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \qquad \text{(Equation 2.9)}$$

For the development of regression model, some assumptions are necessary for to check the validity of any conclusion reached. Three important assumptions of regression (Levine et al 2002) are listed below:

1) Error around a regression line should be normally distributed at each value of X

2) Variation around a regression line be constant for all values of X.

3) Errors around a regression line be independent for each value of X.

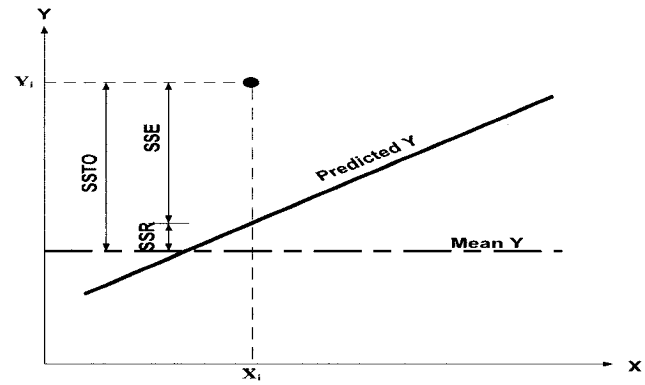


Figure 2-8: Measures of Variance in Regression (Developed from Levine et al 2002)

Following are some important tests or statistics which should be taken into account for checking the appropriateness of a regression relationship:

(i) Coefficient of Multiple Determination

The coefficient of multiple determination, denoted by R^2, is equal to the regression sum of squares divided by the total sum of squares (Levine et al 2002). In other way it can be defined as follows (Kutner et al 2005):

$$R^2 = \frac{SSR}{SSTO} = 1 - \frac{SSE}{SSTO} \qquad \text{(Equation 2.10)}$$

The value of R^2 varies from 0 to 1. R^2 will be 1 when all Y observations fall directly on the fitted regression surface that is when $Y_i = \hat{Y}_i$, for all i. Similarly R^2 will be 0 when there would be no regression fitted surface for Y observations. R^2 usually increases with the addition of more predictive variables in the model. This value could be misleading; therefore, a modified measure is suggested for the adjustment of the value R^2 in the model. The adjusted coefficient of multiple determination, denoted by R_a^2, adjusts R^2 by dividing each sum of squares by its associated degrees of freedom (Kutner et al 2005). The term degrees of freedom (df) is a measure of the number of independent pieces of information on which a parameter estimate is based (Fox, 1997). If more parameters are estimated, the value of "df" decreases. Therefore, R_a^2 can be calculated by the following equation;

$$R_a^2 = 1 - \frac{\frac{SSE}{n-p}}{\frac{SSTO}{n-1}} = 1 - \left(\frac{n-1}{n-p}\right)\frac{SSE}{SSTO} \qquad \text{(Equation 2.11)}$$

(ii) "F" & "t" Tests

To test whether there is a regression relation between the response variable Y and the set of X variables $X_1, X_2 \ldots X_{p-1}$, following alternatives is considered:

$$H_0: \beta_1 = \beta_2 = \ldots\ldots\ldots\ldots = \beta_{p-1} = 0$$
$$H_a: \text{not all } \beta_k \ (k=1 \ \ldots . p-1) \text{ equal to zero} \qquad \text{(Equation 2.12)}$$

The test statistics would be:

$$F^* = \frac{MSR}{MSE} \qquad \text{(Equation 2.13)}$$

where MSR and MSE are Regression Mean Squares and Error or Residual Mean Squares respectively, which can be obtained by dividing the respective sum of squares by the degrees of freedom. The decision rule to control the Type 1 error at α is:

$$\text{If } F^* \leq F\ (1-\alpha;\ p-1,\ n-p), \text{ Conclude } H_0$$
$$\text{If } F^* > F\ (1-\alpha;\ p-1,\ n-p), \text{ Conclude } H_a \qquad \text{(Equation 2.14)}$$

In order to check the significance of any predictor variable X_k in the regression relationship, the alternatives are defined as: H_0: $\beta_k = 0$ and Ha: $\beta_k \neq 0$. The test statistics would be:

$$t^* = \frac{b_k}{s\{b_k\}} \qquad \text{(Equation 2.15)}$$

where, b_k is the estimated co-efficient of the predictor variable X_k, and $s\{b_k\}$ is the standard error associated with it (Schleifer et al, 1995).

Therefore, the alternatives for the t test are:

If $|t^*| \leq t(1-\alpha/2;\ \text{n-p})$, conclude H_0, otherwise, conclude H_a (Equation 2.16)

(iii) Residual Analysis

Residual e_i is the difference between the observed value Y_i and the fitted value $\hat{Y}_i$.

Table 2-16: Residual Analysis Diagnostics (Menzefricke, 1995)

Assumptions	Diagnostic Checks
Linearity	Standardized Residual vs. Fitted Plot Standardized Residual vs. Predictor Variable Plot
Constant Standard Deviation	Standardized Residual vs. Fitted Plot Standardized Residual vs. Predictor Variable Plot
Randomness	Sequence Plots; Control Tests; Run Charts
Normality	Normal Probability Plot

Residual may be regarded as the observed error, in distinction to the unknown random error ε_i in a regression model. For an appropriate regression relationship, error terms should justify the assumptions of linearity, constant standard deviation, randomness and normality. Table 2-17 shows the different diagnostics for the assumptions and their explanation.

(iv) Lack of Fit Test

Lack of fit test determines whether a specific type of regression function adequately fits the data. The test requires repeat observations of Y at one or more levels of X. These observations are called replicates. In the test, error sum of squares SSE is decomposed into pure error and lack of fit. The pure error sum of squares SSPE is obtained by first calculating for each replicate group the sum of squared deviations of the Y observations around the group mean, where a replicate has the same values for each of the X variables.

$$SSLF = SSE - SSPE \qquad \text{(Equation 2.17)}$$

The F test would be

$$F^* = \frac{SSLF}{c-p} \div \frac{SSPE}{n-c} = \frac{MSLF}{MSPE} \qquad \text{(Equation 2.18)}$$

The appropriate decision rule would be;

$$\begin{aligned} &If F^* \le F(1-\alpha; c-p, n-c),\ conclude\ H_0 \\ &If F^* > F(1-\alpha; c-p, n-c),\ conclude\ H_1 \end{aligned} \qquad \text{(Equation 2.19)}$$

Minitab, statistical software, performs an approximate lack of fit test, data subsetting lack of fit test, when replicate observations are not available (Anderson et al 2005).

2.7.2 Unsupervised Neural Networks (Self-Organizing Maps)

Self-Organizing Maps (SOM) belong to a general class of unsupervised neural network methods, which are non-linear regression techniques that can be trained to learn or find relationships between inputs and outputs or to organize data so as to disclose so far unknown patterns or structure.

In unsupervised neural network learning, there is no performance evaluation available (Gallant, 1993). Therefore, there is no information which could be used to improve network's behaviour. Without any specific knowledge of what constitutes a correct or incorrect answer, unsupervised models constructs groups of similar input patterns. This phenomenon of constructing similar pattern groups is known as clustering. Cluster boundaries can be found based on the large and representative sample of inputs as shown in Figure 2-8 (Developed from Hrycej, 1992).

SOM are also known as competitive-learning neural networks that uses an unsupervised training algorithm (Deboeck et al 1998). In competitive learning, cells receive identical input information. By means of lateral interactions, they compete in their activities. Each cell or a group of cells is sensitized to a different domain of vectorial input signal values and act as a decoder of that domain (Kohonen, 1997). Moreover, the process of self-organization means that the network becomes oriented and adaptively assumes a form by which it best describes the input in an ordered and structural fashion (Kohonen, 1993).

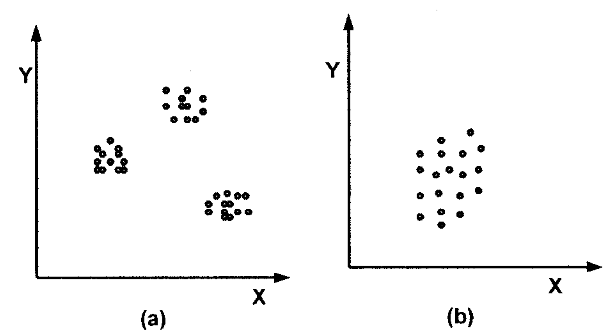


Figure 2-9: Two dimensional Patterns (a) Clustered and (b) No Apparent Cluster

It is important to state that SOM algorithm is not a clustering algorithm. However, each node on the map can be theoretically considered as a cluster centroid. The key concept of SOM algorithm is the neighbourhood function that is adjusting not only the best matching unit but also its surrounding units (Schatzmann et al 2003). Nevertheless, SOM algorithm was compared with hierarchical clustering methods, and it was found that SOM is superior in both robustness and accuracy (Wang et al 2002). Furthermore, data division through SOM have three important advantages (Shahin et al 2004):

- There is no need to decide which proportion of the available data to use for training, testing and validation
- The statistical properties of the resulting training, testing, and validation is sufficiently small, provided that intra-cluster variation is sufficiently small.
- Information is provided about whether "outliers" exist in the data or not

(i) Historical Development and Applications

The SOM is a fairly well known neural network and one of the most popular unsupervised learning algorithms. It was invented by Finnish Professor Teuvo Kohonen in 1980s. Since the invention of SOM, more than 4000 research articles have been published on the algorithm, its visualization and applications (Schatzmann et al 2003). The application of SOM in the field of civil engineering includes various data classification problems. Mukherjee (1997) employed the SOM algorithm to predict the natural mode shapes of building frames with a varying number of stories. Lingras (1995) used the concept of SOM algorithm in classifying a large number of traffic patterns. Shahin et al 2004 developed four different ANN models using four different data division methods for the study of settlement prediction of shallow foundations on granular soils. It was concluded that SOM was more suitable approach for data division. Lee et al 2006 used SOM for the assessment of alternative methods of analyzing water quality performance indicators for constructed treatment wetlands. It was concluded that Self-Organizing Maps (SOM) had better potential to visualize the relationship between complex biochemical variables and to search for optimal map.

(ii) SOM Structure

A self organized map consists of neurons organized on a regular low dimensional grid (Figure 2-9). Each neuron is a "d" dimensional weight vector, where "d" is equal to the dimension of input vector. Neurons are connected to the adjacent neurons by a neighbourhood relation, which indicates the topology or structure of the map (Vesanto et al 1999).

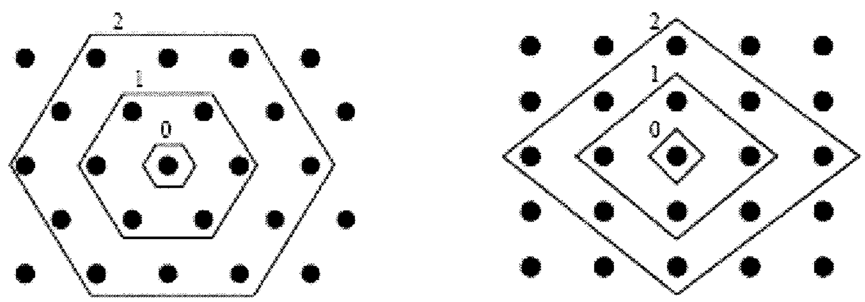


Figure 2-10: Neighbourhoods (0, 1 and 2) of the Winning Neuron (The Centremost Cell): Hexagonal on the Left and Rectangular on the Right (Vesanto et al 1999)

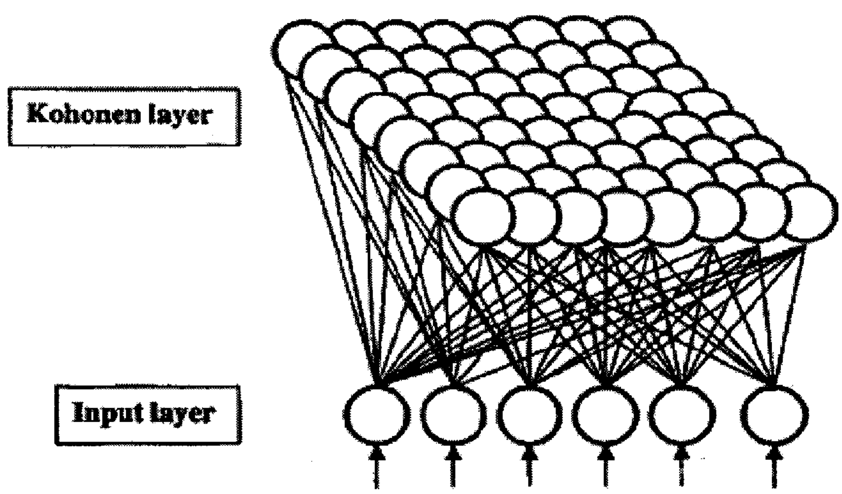


Figure 2-11: SOM Input and Output (Kohonen) Layers (Shahin et al 2004)

The self-organization of this out layer is dependent on the input layer patterns. The input layer consists of a set of "n" input patterns (x_i for i = 1, 2, 3,….., n). This layer is fully connected to the output layer or self-organized layer (Kohonen Layer) through connection weight (w_{ji}, where j = 1, 2, 3, ..., m). The connections between input and output layers are illustrated in Figure 2-10.

(iii) Training of SOM Algorithm

The training of SOM algorithm consist of four basics steps (Kohonen, 1990):

i. Calculation of similarities between the input pattern and the weights arriving at each node or neuron
ii. Finding the most similar neuron which is usually called as "winner"
iii. Selection of a set of output nodes which are located close to the winner in the output grid or neighbourhood
iv. Updating the weights of all nodes in the neighbourhood in such a way that their weights are moved closer to the input pattern

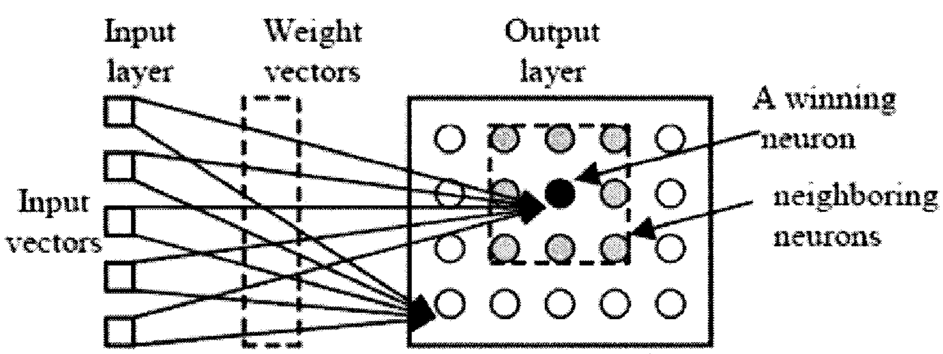


Figure 2-12: A typical Architecture of a Winning Neuron (Tangsripairoj et al 2005)

The weights are initially assigned randomly. At each node in the output layer, the input (x_i) is presented without providing a desired output. In this context, a matching value is calculated for each output node. This value is the Euclidean distance or Straight line distance (D_j) between the weights of each node and the corresponding input values. The value of D_j is calculated by the equation below (Shahin et al 2004):

$$D_j = \sum_{i=1}^{n} (x_i - w_{ji})^2 \text{, where j = 1, 2,......, m} \qquad \text{(Equation 2.20)}$$

The node that has the minimum Euclidean value is considered as a winning neuron (Figure 2-11). This process is repeated until desire groups are obtained.

2.8 Summary

The deterioration phenomenon of sewers is very complex and is dependent upon many factors. There are different types of sewer inspection techniques available, and proper selection of these techniques depends upon local requirements. Similarly, different municipalities have adapted different sewer condition assessment protocols according to their respective needs.

Many analysis techniques have been utilized by different researchers for proposing solution for different sewer management problems. Among these techniques, regression analysis and artificial neural networks are important for data analysis and modelling purposes.

Chapter 3
RESEARCH METHODOLOGY

3.1 Overview

The methodology of current research is presented. Current research consists of different steps: literature review, data collection and pre-processing, sewer condition prediction models development, development of deterioration curves for sewers, comparison of different sewer condition assessment protocols, integration of sewer condition assessment protocols, development of combined condition index (CCI) for sewers, web-based condition rating program, and conclusions and recommendations.

3.2 Description

The research methodology is presented in Figure 3-1, and is described below:

3.2.1 Problem Statement and Literature Review

The first step of this research is to clearly describe the problem statement and research objectives. The study objectives have been defined in chapter 1. The relevant literature has been reviewed in chapter 2. That includes: types and failure of sewer pipes, factors contributes to sewer deterioration, sewer inspection techniques for condition assessment, adapted sewer condition assessment protocols, research techniques in sewer management, and regression analysis and unsupervised neural network learning processes.

After the literature review, the research is to be carried out in two parallel paths: development of sewer condition prediction models and combined condition index (CCI). Both the methods have been briefly described below.

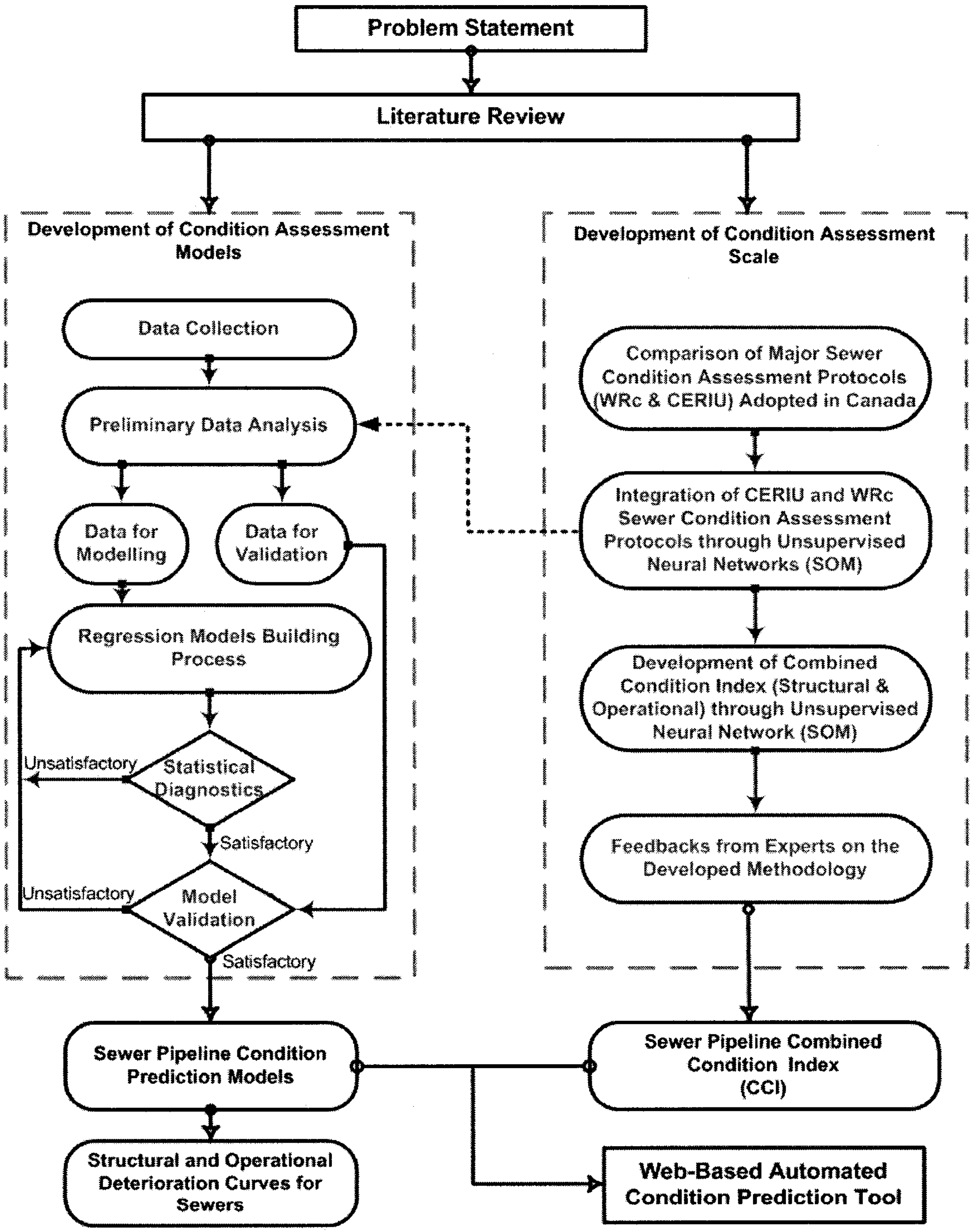


Figure 3-1: Research Methodology

3.2.2 Sewer Condition Prediction Regression Models

The main objective of this part of research is to develop sewer pipeline structural and operational condition prediction models. The main steps of the process are described below:

(i) Data Collection and Preliminary Analysis

The first step in model building process is to collect historical data of sewer condition assessment. Two data sets are collected from municipalities of Niagara Falls, Ontario and Pierrefonds, Quebec. The condition assessment data collected from both the municipalities are according to two different condition assessment protocols: WRc (Water Research Centre) UK protocols and CERIU (Centre for Expertise and Research on Infrastructures in Urban Areas), Canada. Therefore, a methodology is developed to convert CERIU sewer condition classification data into WRc protocols, the most widely used protocols in the world, through unsupervised neural network clustering process. This methodology will be discussed later.

Preliminary analysis of the data is carried out; which consists of sorting and ordering of data obtained from sources, comparing data, defining assumptions, and finalizing input parameters. In order to eliminate the categorical variable "pipe material" from structural condition prediction models, the prepared data is transformed into three groups with respect to their pipe material (concrete, asbestos cement, and PVC). Separate structural condition prediction models are to be built for the three groups. For the operational condition model the categorical variable "pipe material" is transformed according to its Manning's coefficient of roughness values for each material. Therefore, total four data sets are prepared. The detail of this process has been presented in chapter 4.

(ii) Regression Model Development Methodology

The regression model building methodology is presented in Figure 3-2.

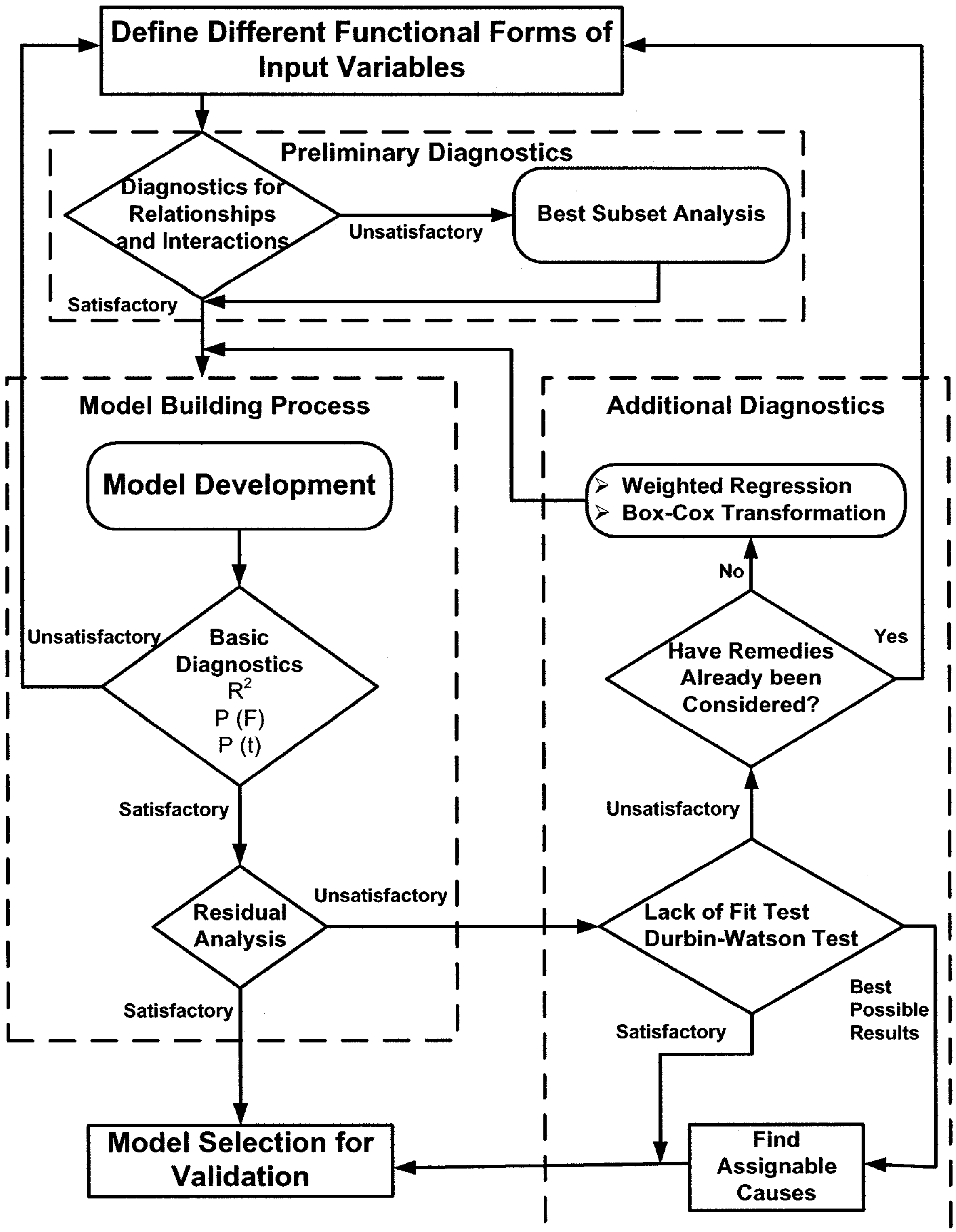


Figure 3-2: Regression Model Building Methodology

From all the four data groups approximately 25% of data is picked randomly for model validation process, and the rest of the data is used in developing regression models. The regression models are built using Minitab ® statistical software. Figure 3-2 shows that regression model building process included four main steps: preliminary diagnostics for interactions, model building, statistical diagnostics of the built model, and remedial measures. The detail of all these diagnostics and remedies will be presented in chapter 5.

(iii) Regression Model Validation Methodology

The validation data are embedded into their respective developed most appropriate regression models for comparing results with the actual results using Microsoft Excel spread sheet procedures.

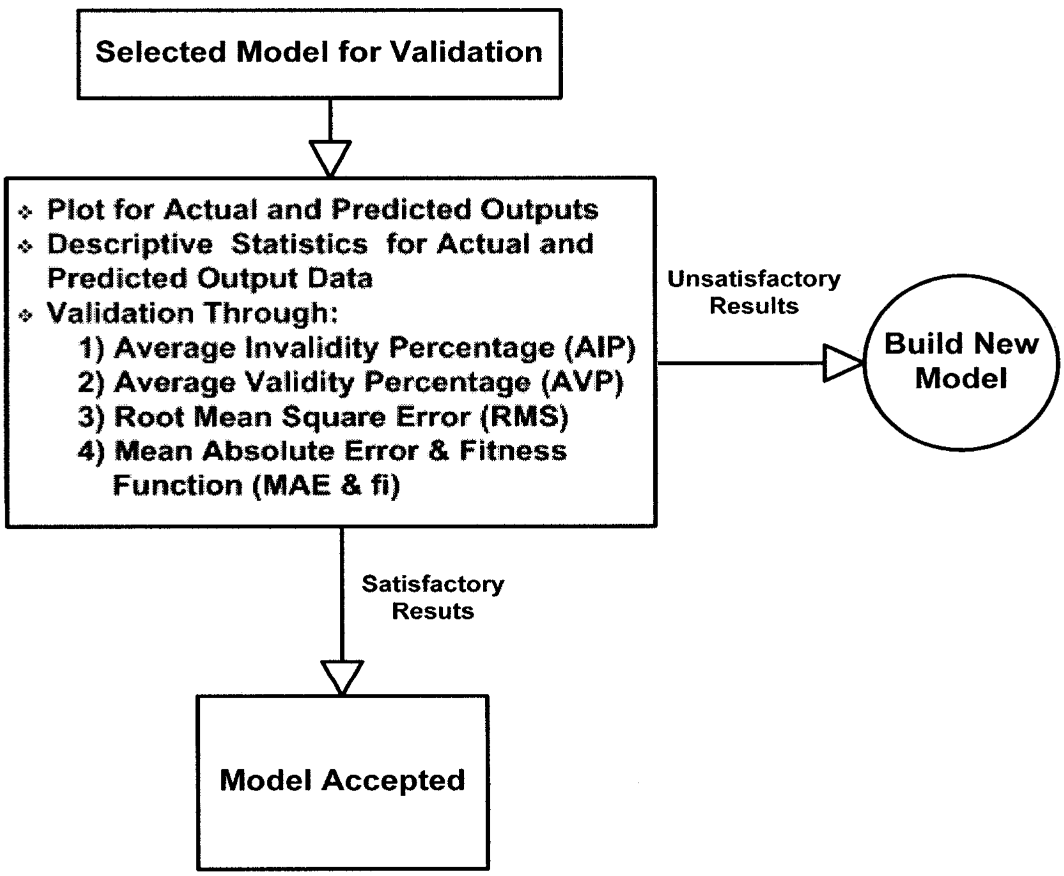


Figure 3-3: Regression Model Validation Methodology

All the models are validated on four basic criteria: average invalidity percentage (AIP), average validity percentage (AVP), root mean square error (RMS), and mean absolute error and fitness function (MAE & fi). An overview of model validation methodology has been presented in Figure 3-3. The detail description of the validation methodology will be presented in chapter 5.

(iv) Deterioration Curves for Sewers

Current research develops structural and operational deterioration curves for sewers. The developed regression models are used to build these curves for sewers, which will also be presented in chapter 5. The curves are built to develop and understand the relationship between a sewer's structural and hydraulic condition to its age. The curves are intended to assist decision makers in prioritizing inspections, maintenance, and rehabilitation programs.

3.2.3 Combined Condition Index (CCI) for Sewers

Figure 3-1 illustrates an overview of the methodology for developing a combined condition index (CCI) for sewers. In order to achieve the objectives, the first step is to compare the adapted sewer condition assessment protocols, so that an integrated sewer condition classification system could be proposed. In this context, the two major protocols adapted by many municipal agencies in Canada: WRc and CERIU; are compared. As WRc protocols are accepted world wide and also called "Embryo Codes" (Thornhill et al 2005), the protocols are considered as standard in this research. Consequently, the other major protocol adapted in Canada, CERIU, is converted into WRc for unification and integration of sewer condition assessment procedures.

(i) CERIU Protocols Conversion Methodology

An overview of the methodology for conversion of CERIU sewer condition assessment protocols into WRc is shown in Figure 3-4. The detailed methodology will be presented in chapter 6. The key difference between the two protocols is that WRc assigns different deduct values to each defect in a sewer pipe, and based on these deduct values it calculates an overall condition class for the whole pipe; while CERIU directly assigns a condition class to each defect in a sewer pipe, and it does not calculate any condition class for the whole pipe segment.

Based on WRc severity ranking for different sewer defects, transformed deduct values for CERIU classifications are generated. The generated values are clustered into five groups. The methodology of unsupervised neural network clustering (Kohonen's self-organizing maps) is adapted with the help of Neuroshell ® software. The principal objective of clustering deduct values is to develop five different condition classes for CERIU protocols, compatible to WRc protocols, for structural and operational conditions of sewers. This methodology is verified through feedbacks from CERIU sub-committee for the development of a unified condition assessment protocol (2006). The details will be discussed in chapter 6.

(ii) Combined Condition Index (CCI) for Sewers

After the conversion of CERIU protocols into WRc, the next step is to propose a combined condition index (CCI) for sewers. Usually, an existing sewer's condition is considered in two different scenarios: structural and operational. An integrated approach to condition assessment is proposed by combining the effects of structural and hydraulic condition ratings of sewers.

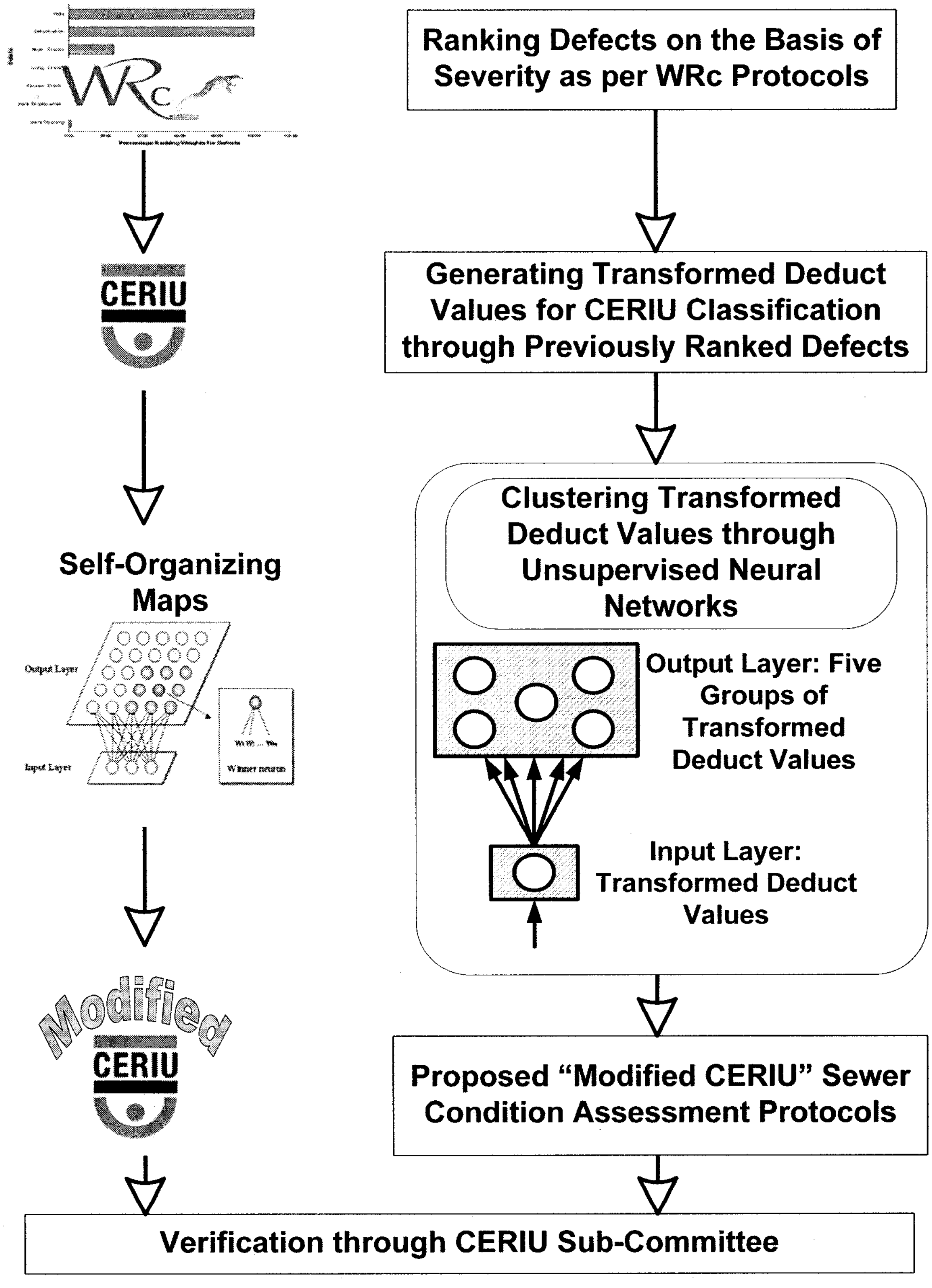


Figure 3-4: Proposed Modification Methodology in CERIU Protocols

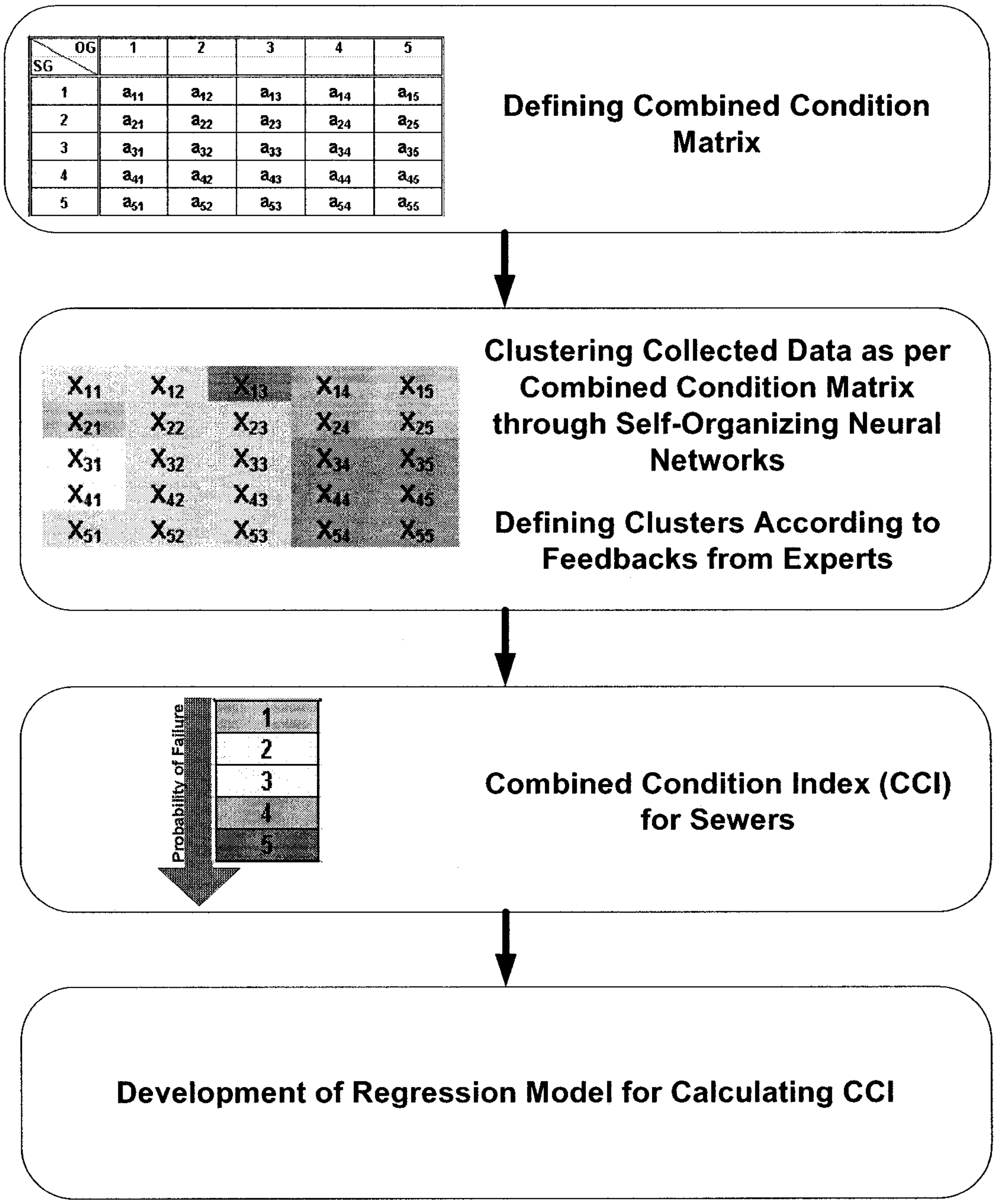


Figure 3-5: Combined Condition Index (CCI) Development Methodology

As a result, a methodology is adapted for development of combined condition index (CCI) for sewers. Figure 3-5 presents an overview of the methodology adapted for the development of combined condition index (CCI) for sewers. A combined condition matrix is defined by taking account of all possible scenarios for a sewer's condition. The matrix is clustered into five classes by using the collected data and adapting self-organizing neural networks methodology. The clustering is done with the help of Neuroshell ® software. The developed clusters are examined through feedbacks from experts, and a final combined condition index (CCI) is proposed. CCI varies fro 1 to 5; where 1 represents an acceptable combined (structural plus operational) condition of a sewer, 5 represents critical combined condition. Further, a regression model is developed to directly convert structural and operational condition ratings into CCI. The detailed methodology of the development of CCI will be presented in chapter 6.

3.2.4 Web-Based Automated Condition Prediction Tool

After the development of condition prediction models and combined condition index (CCI) for sewers, a web-based automated tool is designed using Java programming language. The web-based program will assist municipal engineers to predict the existing structural and operational condition of sewers for prioritizing detailed sewer inspection. Thus, the tool will assist the decision makers to minimize the current, costly practises of random inspections of sewers.

The program has been designed in a simple, user friendly environment. After login, a user can import relevant data in *.xls file format with unknown pipe condition. After processing, the program displays the results which show the most likely existing condition ratings of sewers (structural condition, operational condition, and CCI). The program utilizes all the developed regression models during this research for data processing. A flowchart for the web-based model process is shown in figure 3-6. The detailed process of the program will be discussed in chapter 7.

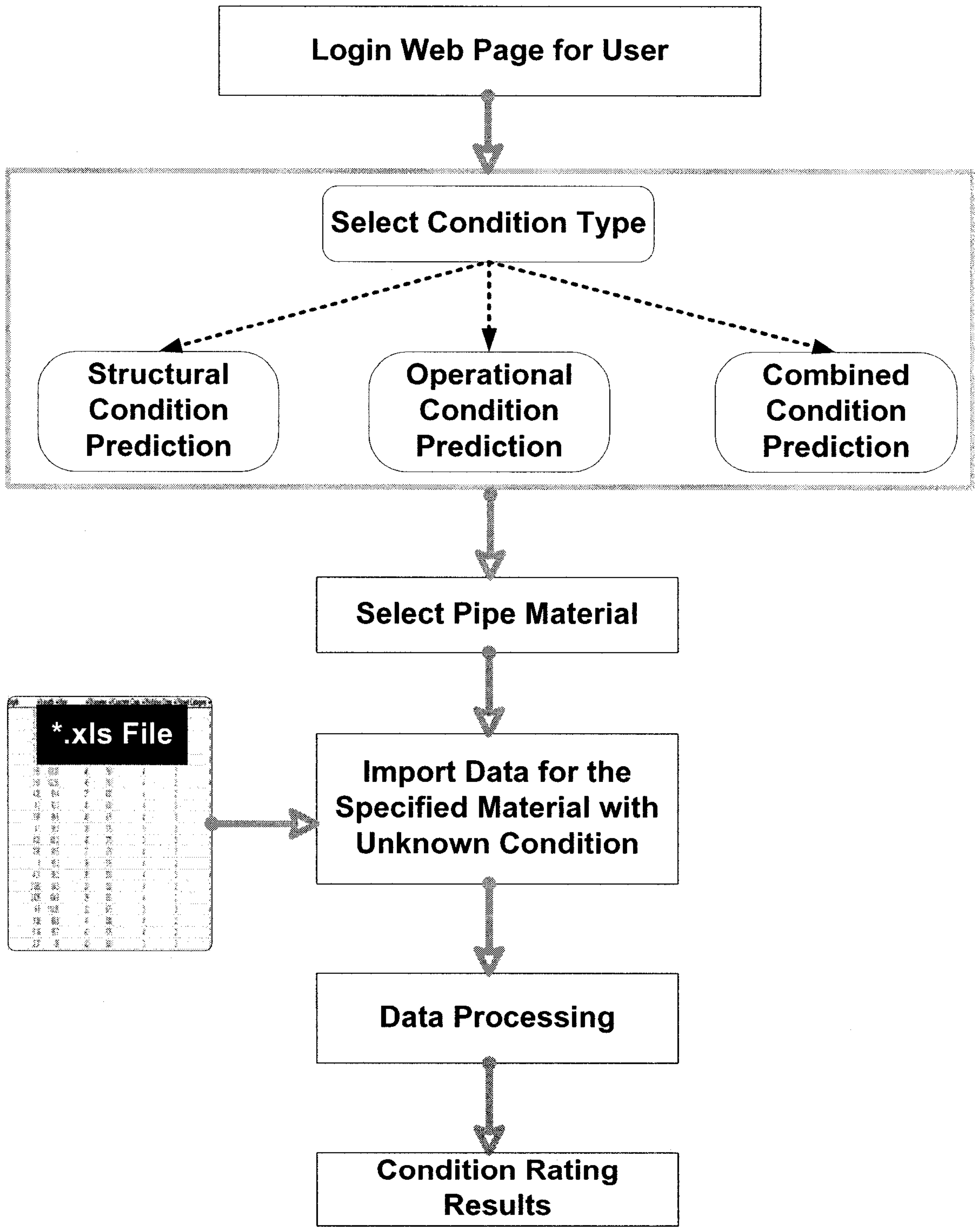


Figure 3-6: Flowchart for the Web-Based Model

3.3 Summary:

The methodology of current research is presented. The methodology includes literature review, data collection and preliminary analysis, condition prediction regression models design and validation, development of structural and operational deterioration curves for sewers, comparison and conversion of sewer condition assessment protocols, development of combined condition index (CCI) for sewers, and development of an automated web-based sewer pipeline condition prediction tool.

Chapter 4

DATA ACQUISITION AND PREPARATION

4.1 Introduction

One of the main objectives of this research is to design model for sewer pipeline condition prediction that can prioritize the sewer inspection; consequently, could reduce the cost due to random sewer inspection.

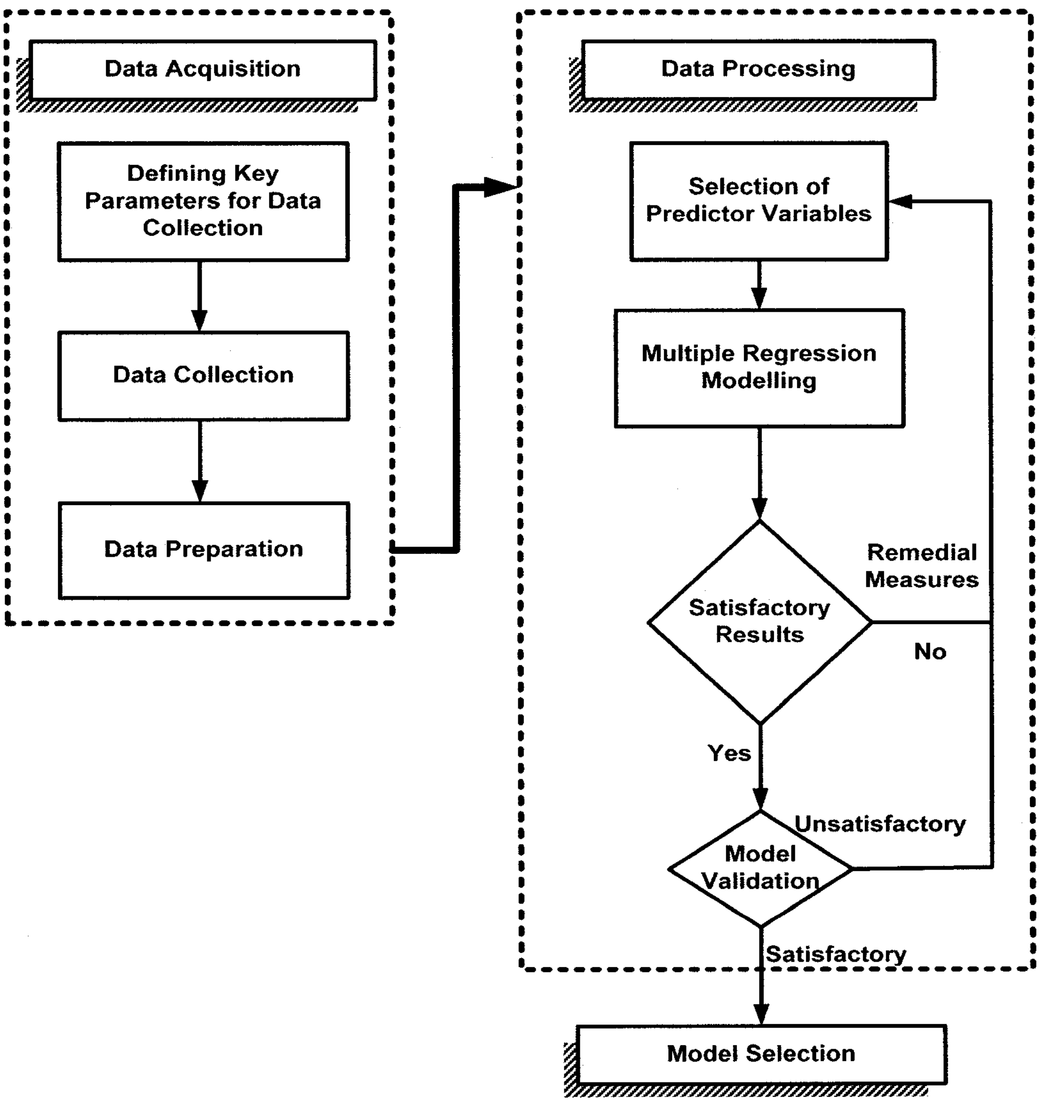


Figure 4-1: Overview of Condition Prediction Model Development Process

Model building procedure requires a detailed analysis of historical data. In this context, data acquisition and processing is performed according to the requirements. Figure 4-1 shows an overview of this process. As clear from the figure, the process can be divided into two main steps: data acquisition and preparation, and data processing. This chapter discusses the first step, data acquisition, in detail. Data collection and pre-preparation procedures are described in detail. The chapter also illustrates an overview of different assumptions which are made during data pre-processing. The descriptive statistics and histograms of collected data are also presented in the chapter.

4.2 Data Acquisition

Many municipalities across Canada were contacted for data collection purposes. As a result, two data sets were received from municipalities: Niagara Falls, Ontario and Pierrefonds, Quebec. The information received regarding sewer pipeline networks is shown in Figure 4-2.

A major difference between both the data sets is that the municipality of Niagara Falls, Ontario has adopted the WRc condition grading systems for their sewers; whereas, the municipality of Pierrefonds, Quebec has adopted CERIU classification system. Moreover, there are some other minor differences in both the data sets. Missing information from the data collected from Pierrefonds is traffic volume data. In order to develop an integrated approach for model development and its applications, interpolation and preparation of data in a generalize manner is performed which is described in the next section.

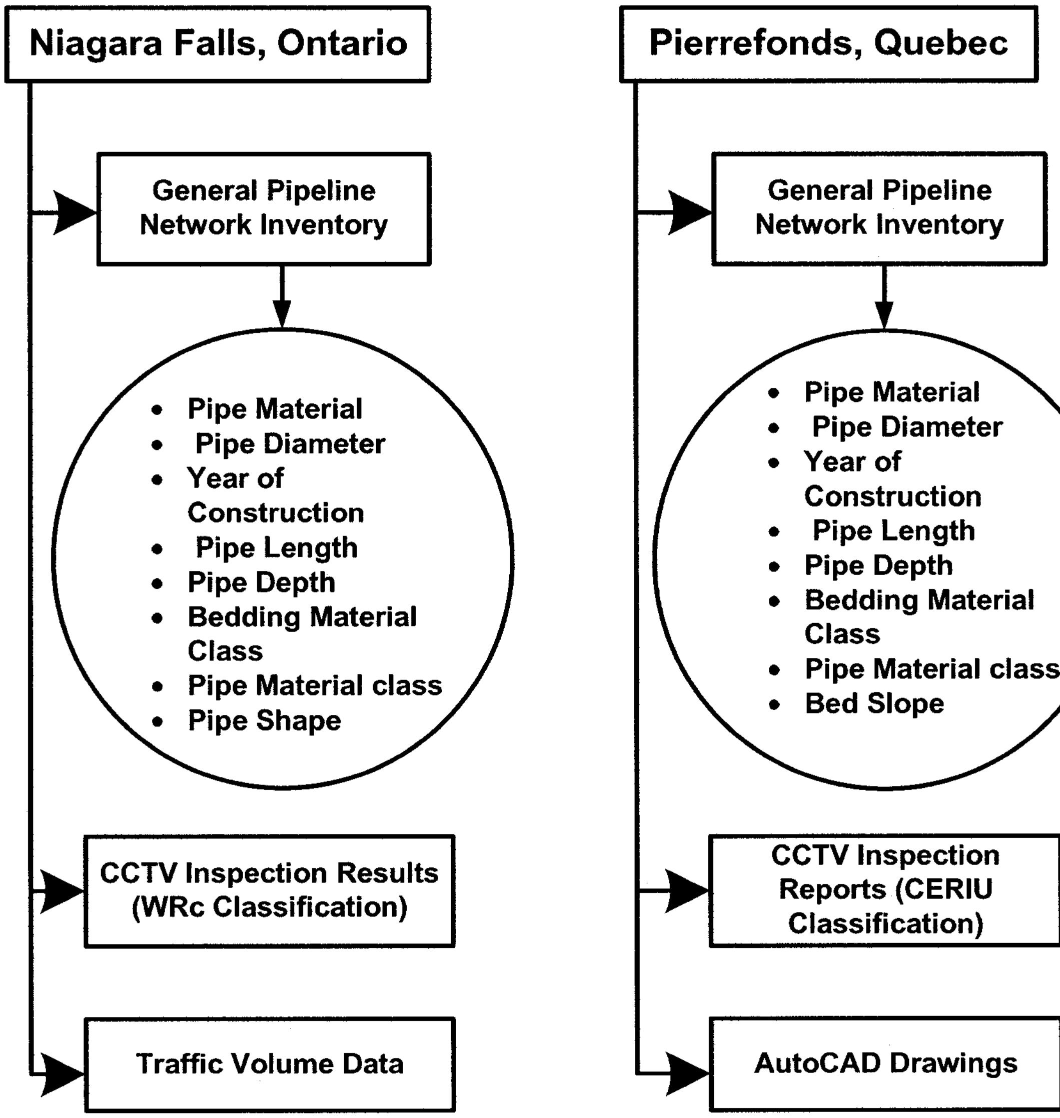


Figure 4-2: Summary of Data Received from Municipalities

4.3 Data Preparation

As mentioned above that the collected data sets have some differences and can not be simultaneously modeled into a single pattern for model development. Therefore, a complete preliminary analysis is performed for both the data sets. Consequently, a generalized approach for analyzing the data is developed and adapted for both systems.

Figure 4-3 presents the adapted methodology for data preparation. It clearly shows the four key steps for data preparation. The detail of theses steps are describe below.

4.3.1 Data Sorting

The first step is to sort the data in some order. Figure 4-2 shows the systematic order of data; however, it is found out that there are less data set points available which have all the complete required details. For example, if a pipe segment has a record of CCTV inspection; it may not have record of other parameters.

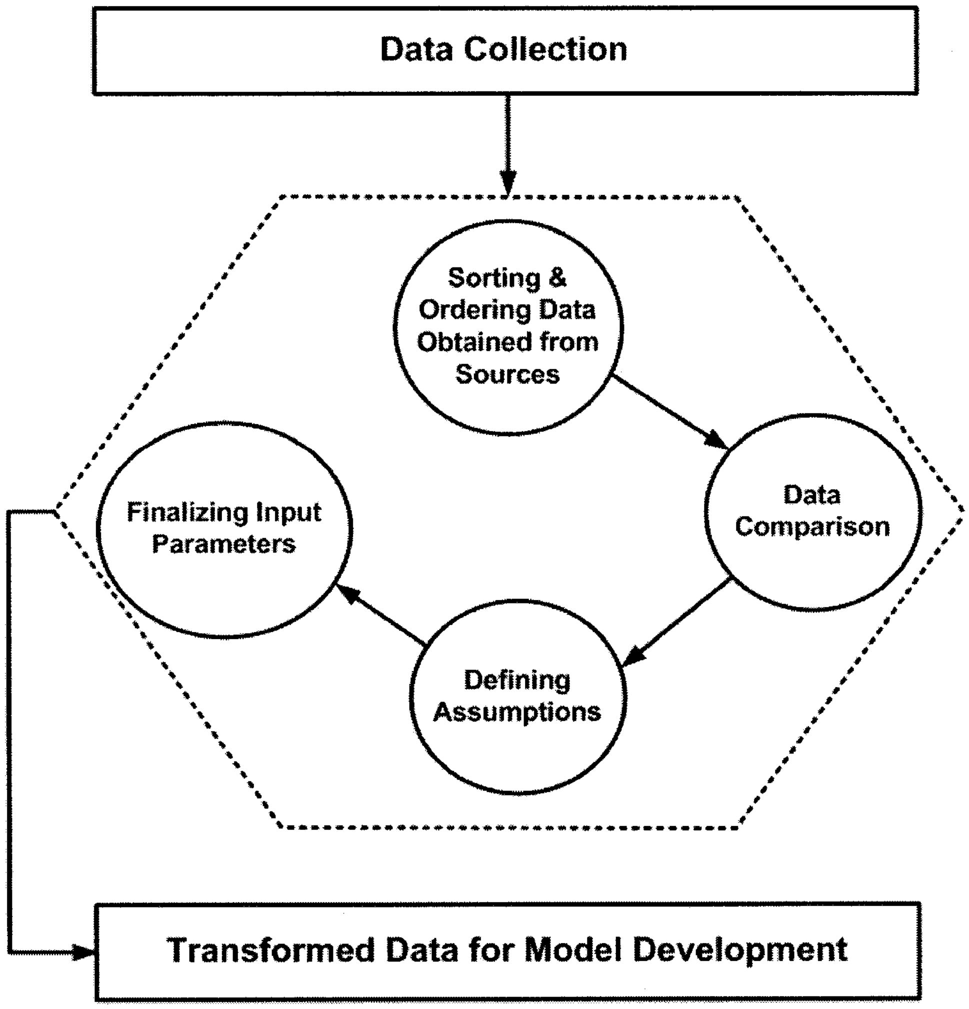


Figure 4-3: Data Preparation Process

Therefore, all the data set points from both the available databases which have some missing information are removed from the prepared database.

4.3.2 Data Comparison

After sorting the data, the next step is to determine the differences in the two data sets, and to find some solution for generalizing the input data. The major difference, as mentioned above, was the adapted different sewer condition classification systems by the municipalities. The utilization of both data sets simultaneously is not possible because there is no standard method available for comparison and conversion of these two different classification systems. Therefore, in order to compare one sewer condition classification system with the other, a complete analysis of both the condition classification systems is carried out, and a methodology is adapted to suggest conversion factors for data conversion. This methodology and analysis will be presented in chapter 6. Condition assessment data obtained from Pierrefonds is converted into WRc classification system by using the suggested methodology for further processing.

4.3.3 Defining Assumptions

The next step is to define some of the parameters in detail. The different assumptions and interpretation made for different parameters are discussed below.

(i) Pipe Material

The data received from the municipalities consist of three different categories for pipe material: concrete, asbestos cement, and polyvinyl chloride (PVC). The structural condition grading models are to be developed separately for each pipe material for simplifying the multiple regression inputs.

In case of developing operational condition assessment model, the care is given to the fact that operational condition depends upon flow velocity inside a pipe and some other factors. In order to assess flow velocity, a formula should be selected which must include a suitable co-efficient of friction for the internal surface of the pipe (Perkins, 1974). Consequently, the categorical pipe material variable is redefined by a suitable co-efficient of friction for each type of pipe material; thus transformed into quantitative variable.

In this context, Manning's roughness coefficient is considered which is widely applied in partly filled conduits (Casey, 1992). The Manning's roughness coefficient can be found out from the Manning's formula which is expressed as:

$$V = \frac{R^{0.67}\sqrt{S}}{n} \qquad \text{(Equation 4.1)}$$

where, V is mean velocity of flow (m/s), n is Manning's co-efficient of roughness, S is channel slope (m/m), and R is hydraulic radius (m) (cross-sectional area/wetted perimeter). There are wide disagreements between researchers on the value of n and extensive research on its determination is on going (Bilgil, 2003). However, the general ranges of n for various pipe materials have been defined (Gribbin, 2001) and the input values of n are taken as a model parameter according to Table 4-1.

Table 4-1: Manning's Roughness Co-efficient Values (Adapted from Gribbin, 2001)

Pipe Material	**"n" Value Range**	**Input Value of "n" for Model**
Concrete	0.011 to 0.015	0.011
Asbestos Cement	0.011 to 0.015	0.011
PVC	0.009 to 0.011	0.009

Table 4-1 shows that for simplicity, the value of n is considered to remain constant through out a pipe's life. However, the value of n increases as the pipe age increases.

(ii) Pipe Material Class

Pipe material class is very important in determining the existing condition of pipe. A high strength material can with stand more crushing load as well as bending stresses. As mentioned, the data collected from both the municipalities consist of three materials: concrete, asbestos cement and PVC.

Table 4-2: Description of Reinforced Concrete Pipe Classes (ASTM C-76-02)

Pipe Class	Dead Load to Produce 0.01 inch or 0.3mm Crack		Ultimate Load	
	Lbs/foot/foot of Diameter	N/m/mm of Diameter	Lbs/foot/foot of Diameter	N/m/mm of Diameter
1	800	40	1200	60
2	1000	50	1500	75
3	1350	65	2000	100
4	2000	100	3000	150
5	3000	140	3750	175

In case of concrete pipes, both municipalities have been using the ASTM (American Society of Testing Materials) standards (ASTM C-76). ASTM divides concrete pipe into five classes according to the ultimate crushing strength of the material. Table 4-2 illustrates the ASTM classes. The data sets received from municipality of Pierrefonds have some missing information regarding the type of concrete (reinforced or plain) and concrete class. It is found out through Pierrefonds municipality's official that if a pipe's depth is greater or equal to 4m, and the municipality does not have proper record regarding the concrete class of the pipe; it considers the pipe to be a class 5 reinforced concrete pipe. Therefore, the missing information regarding concrete class for depth greater than 4m is interpolated accordingly. The same input weights are assigned to each class for model development; for example, the input weight for class 5 would be 5.

Table 4-3: Canadian Standards for AC Sewer Pipe Classes (CAN/CGSB-34.9-94)

Pipe Class	Minimum Crushing Load		Diameter Ranges
	Lb/foot	KN/m	
1500	1500	22	200mm to 400mm
2400	2400	35	200mm to 600mm
3300	3300	48	200mm to 750mm
4000	4000	58	250mm to 1055mm
5000	5000	73	250mm to 1055mm
6000	6000	88	900mm to 1055mm
7000	7000	102	900mm to 1055mm

In case of asbestos cement pipe, the data of municipality of Niagara Falls is interpolated as per Canadian General Standards for asbestos cement sewer pipes. Canadian General Standard Board (CGSB) defines 7 different asbestos cement classes which are in accordance with ASTM standard ASTM C-428. These classes are shown in Table 4-3. The pipe diameter ranges (Table 4-3) show that first three classes are most appropriate for pipes up to 750mm. The acquired data from municipality of Niagara Falls for asbestos cement pipe is consisted of diameter range below than 750mm mark.

Consequently, the pipe classes defined by municipality of Niagara Falls are redefined as per CGSB standards and influence factors or input weights are assigned to them for condition rating model's input as shown in Table 4-4.

Table 4-4: Approximate Equivalent CGSB Classes for Niagara Falls Data

Municipality's Defined Class	Approximate Equivalent CGSB Class	Input Weights for Model
A	3300	3
B	2400	2
C	1500	1

As far as the municipality of Pierrefonds data is concerned, there is very little information available for asbestos cement pipes; therefore, they are eliminated from model input data. Furthermore, no data is available for PVC pipe classes; therefore, PVC material classes are not considered during the condition prediction model development for PVC pipes.

(iii) Bedding Material Class

Pipe stresses also depend upon bedding material properties: material class, material thickness and physical properties. Five different types of bedding material have been considered by Building Research Establishment (BRE) UK, which are also acceptable in USA (Perkins, 1975).

Table 4-5: Bedding Material Classes as per BRE and OPSD Standards

Bedding Class	Description	Bedding Factor B_f	
		BRE Classification	OPSD Classification
A	Reinforced Concrete Cradle or Arch	3.4	2.8
	Plain Concrete Cradle or Arch	2.6	
B	Well Compacted Granular Material	1.9	1.9
C	Well Compacted Backfill	1.5	1.5
D	Flat Sub Grade	1.1	--
Others	Cement Stabilized Material	2.6 to 3.4	--

In Ontario, Canada, OPSD (Ontario Provincial Standard Drawings) defines four classes of bedding material (Zhao et al 2001). Table 4-5 (Adapted from Perkins, 1974 and Zhao et al 2001) shows the summary of different bedding material classes. The bedding factor B_f can be defined in equation 4.2 (Zhao et al, 2001):

$$B_f = \frac{W}{S_{eb}} \qquad \text{(Equation 4.2)}$$

Where, W is the calculated external load and S_{eb} is the three edge bearing strength

Table 4-6: Transformed Values for Table 4-5 for Model Inputs

Bedding Class	**Description**	**Input Weight for Model**
A	Reinforced or Plain Concrete Cradle or Arch	4
B	Well Compacted Granular Material	3
C	Well Compacted Backfill	2
D	Flat Sub Grade	1
Others	Cement Stabilized Material	Not Considered

The data obtained from municipalities consist of some sub classes of bedding material e.g. class "B-B", ¾ inch crushed stone etc. These sub classes are sorted out according to the specification presented in Table 4-6. For example, class "B-B" is simply considered as class "B' etc. Four basic classes are defined which are generally considered in Canada. The classes are assigned with there respective weights to convert the categories into numbers for input data. The description of these classes and their assigned weights are presented in Table 4-6.

(iv) Traffic Volume and Street Categories

The next step is to define the influence of average daily traffic above a pipe on its condition in a generalized manner. It is difficult to collect average traffic volume data for each street in a municipal area. Therefore, a methodology is adapted by categorizing streets and assigning them weights according to street sizes. This methodology is developed by considering the ASCE (American Society of Civil Engineers) classification for residential streets. Figure 4-4 illustrates the basic ASCE classification.

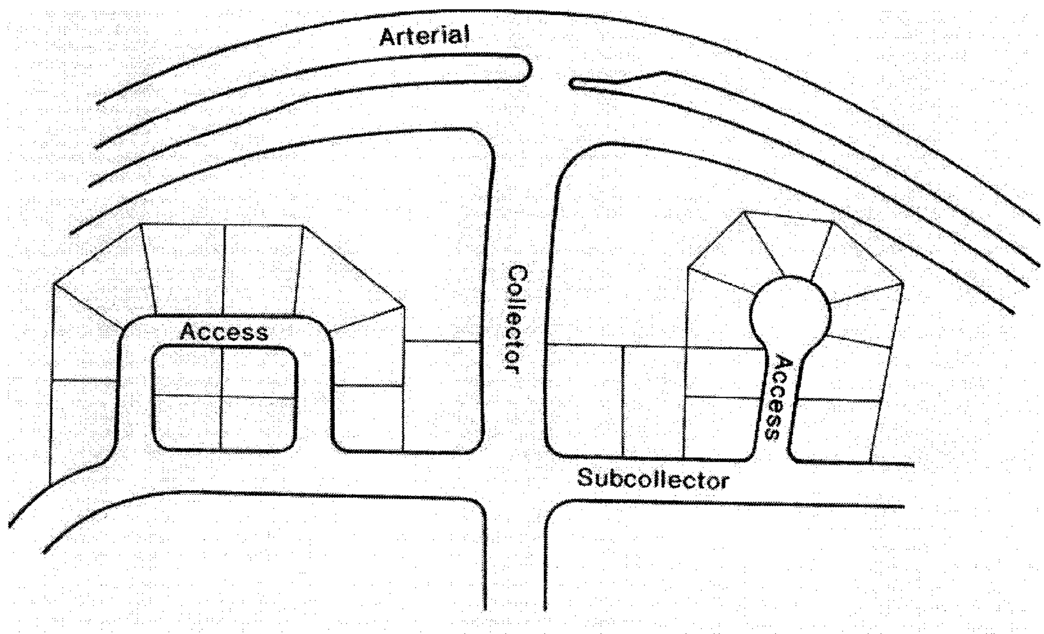


Figure 4-4: Basic Hierarchy of Urban Streets (ASCE, 1990)

ASCE defines the four different categories of streets (Figure 4-4) as follows:

***a) Arterial Street*:** It is a high volume street which should not have any residences on it. Its function is to conduct traffic between communities and to connect the communities with major highways.

b) Collector: It conveys the traffic from arterial streets to lower order streets. The collector carries relatively higher traffic volume then lower order streets.

c) Sub-Collector: It provides passage to access streets and conveys traffic to collectors. It provides frontage and access to residential lots but also carries some through traffic to lower order streets. The sub – collector is a relatively low volume street.

d) Access Streets: It is designed to conduct traffic between dwelling units and higher order streets. It is a very low traffic volume street.

The obtained data from municipality of Pierrefonds contains AutoCAD drawings. As the average annual daily traffic (AADT) volume data is not obtained, the streets are categorized according to ASCE classification by analyzing drawings. Corresponding weights are assigned to streets according to their respective classes as an influence factor for sewer pipeline condition. These classes and their respective weights are shown in Table 4-7.

In case of municipality of Niagara Falls, the data contain average annual dailly traffic (AADT) volume record along with the closest intersection location information. The municipality categorizes its roads as per average annual daily traffic (AADT) volume into five categories. After thorough analysis of traffic volume data with respect to its location information, it is found out that the fifth category consisted of highways or major roads of the area. For simplicity and comparison with the other data set, this category is ignored. Therefore, all of the categorical traffic volume data obtained from Niagara Falls is transformed according to ASCE urban street classification.

Table 4-7 clearly shows that the traffic volume above a pipe is categorized only on the basis of urban streets classification and on the effect of extremely high traffic volume; which might be the case for highways and expressways, is not considered in the model development.

Table 4-7: Transformation of Data into ASCE Urban Street Classification System

ASCE Street Classification	Description	AADT Count		Input Weight for Model
		Niagara Falls	Pierrefonds	
1	Arterial	10,000 to 12,500	Data Not Available	4
2	Collector	7,500 to10,000		3
3	Sub-collector	5,000 to7,500		2
4	Access	< 5,000		1
Not Specified		> 12,500		Not Considered

4.3.4 Finalized Transformed Data

The data prepared and described in the previous section was finalized for the regression model development.

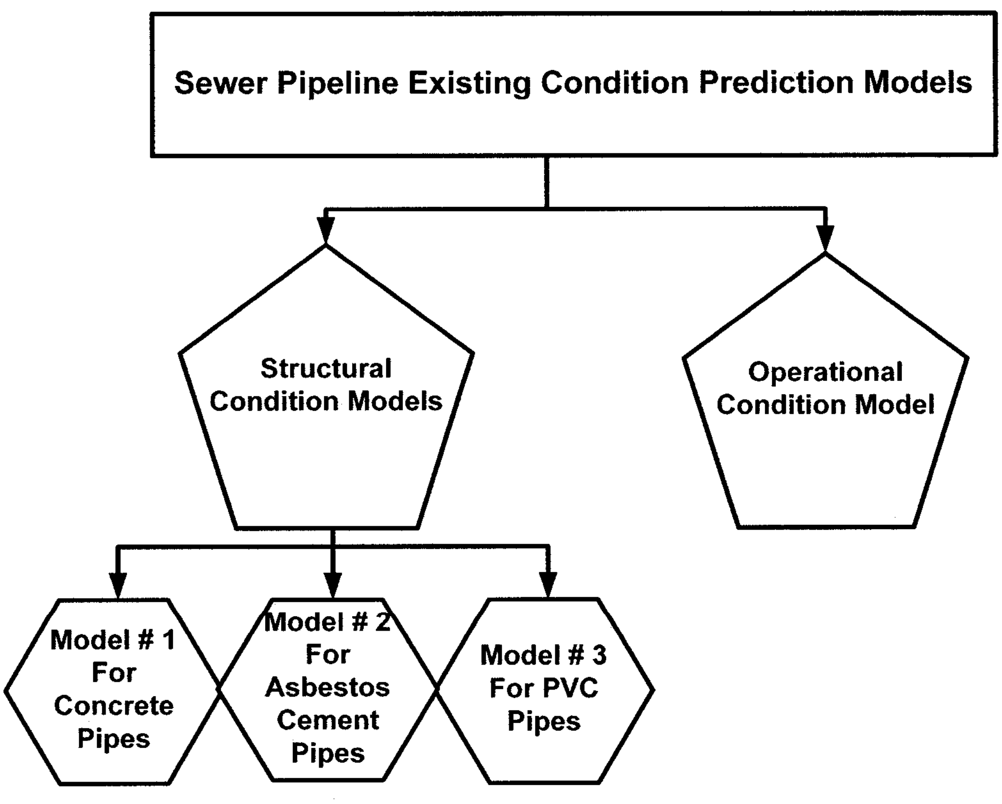


Figure 4-5: Types of Sewer Pipeline Condition Assessment / Prediction Models

As illustrated in the previous section the structural condition prediction models are to be developed for each pipe material separately; nevertheless, for operational condition grading models, the categorical parameter "pipe material" is defined in terms of Manning's roughness co-efficient. As clear from Figure 4-5 that data is prepared and transformed into four categories for the development of four different statistical models. All the four data sets are further divided into two parts; data for model development and data for model validation. Approximately, 25% of data set points from the above mentioned four prepared groups of data are randomly picked for model validation.

A brief summary of the four prepared data groups is illustrated below:

(i) Group Number 1

The first group of data is prepared for the concrete pipe structural condition assessment model.

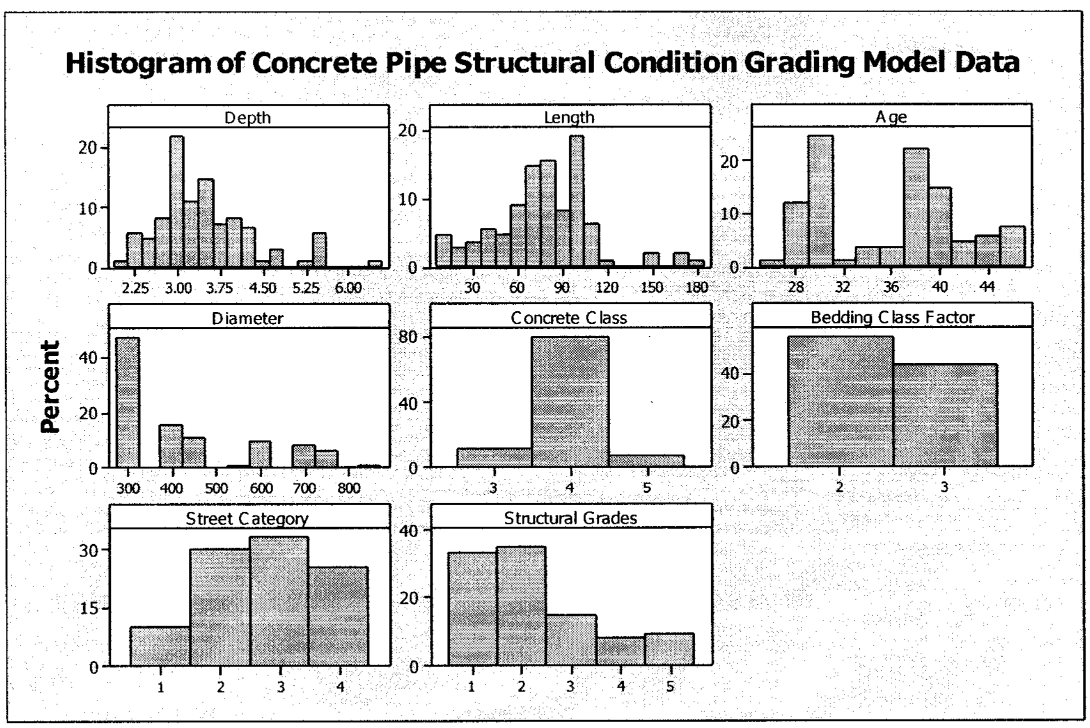


Figure 4-6: Histogram of Data for Group Number 1

Figure 4-6 presents the histogram for the transformed data for the different parameters. Further, the descriptive statistics for the data is also presented in Table 4-8.

Table 4-8: Descriptive Statistics for the First Group Data

Parameter	Mean	Standard Deviation	Minimum	Median	Maximum
Depth (m)	3.48	0.83	2.03	3.3	6.6
Length (m)	77.21	33.29	6.1	79.1	179.83
Age (Years)	35.14	5.92	26	37	47
Diameter (mm)	422.5	156.3	300	375	825
Concrete Class	3.95	0.44	3	3	5
Bedding Factor	2.44	0.5	2	2	3
Street Category	2.75	0.95	1	2	4
WRc Structural Rating	2.26	1.26	1	1	5

(ii) Group Number 2

The second group consists of data for asbestos cement pipe structural condition assessment model.

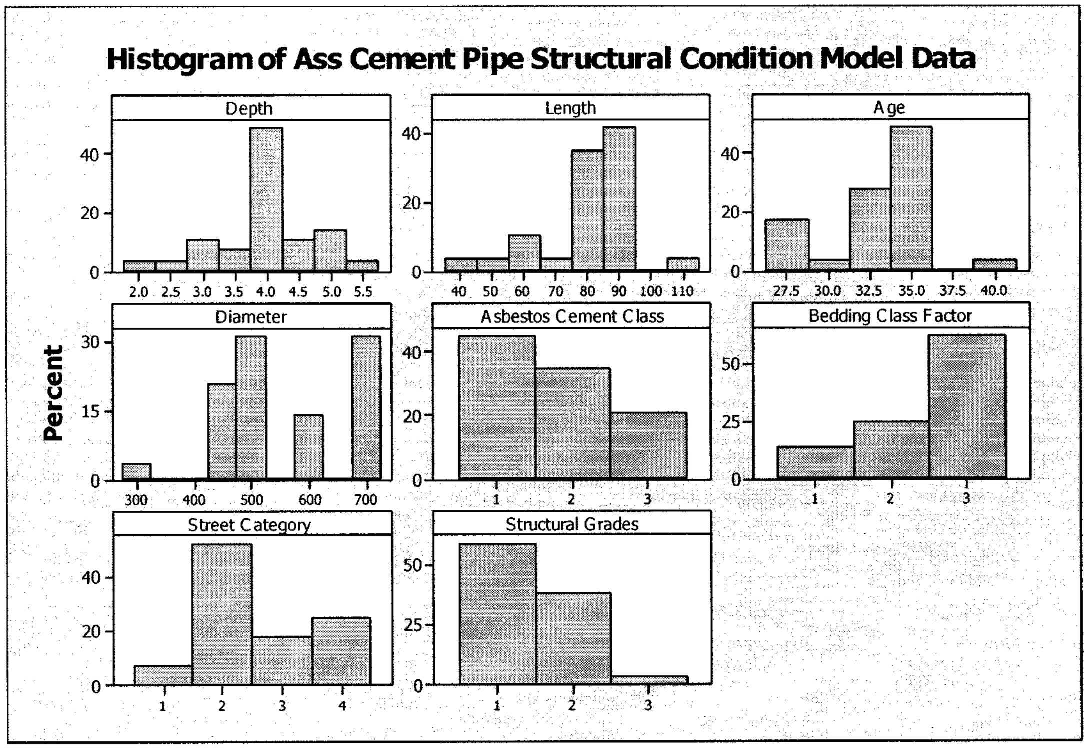


Figure 4 -7: Histogram of Data for Group Number 2

Figure 4-7 presents the histogram for the transformed data for the different parameters. Further, the descriptive statistics for the data is also presented in Table 4-9.

Table 4-9: Descriptive Statistics for the Second Group Data

Parameter	Mean	Standard Deviation	Minimum	Median	Maximum
Depth (m)	4.03	0.72	2.21	4	5.34
Length (m)	80.46	14.81	35.66	82.3	107.59
Age (Years)	32.55	3.26	27	34	41
Diameter (mm)	550.9	102.7	300	500	675
A. Cement Class	1.76	0.79	1	2	3
Bedding Factor	2.48	0.74	1	3	3
Street Category	2.59	0.95	1	2	4
WRc Structural Rating	1.45	0.57	1	1	3

(iii) Group Number 3

The third group consists of data for PVC pipe structural condition assessment model.

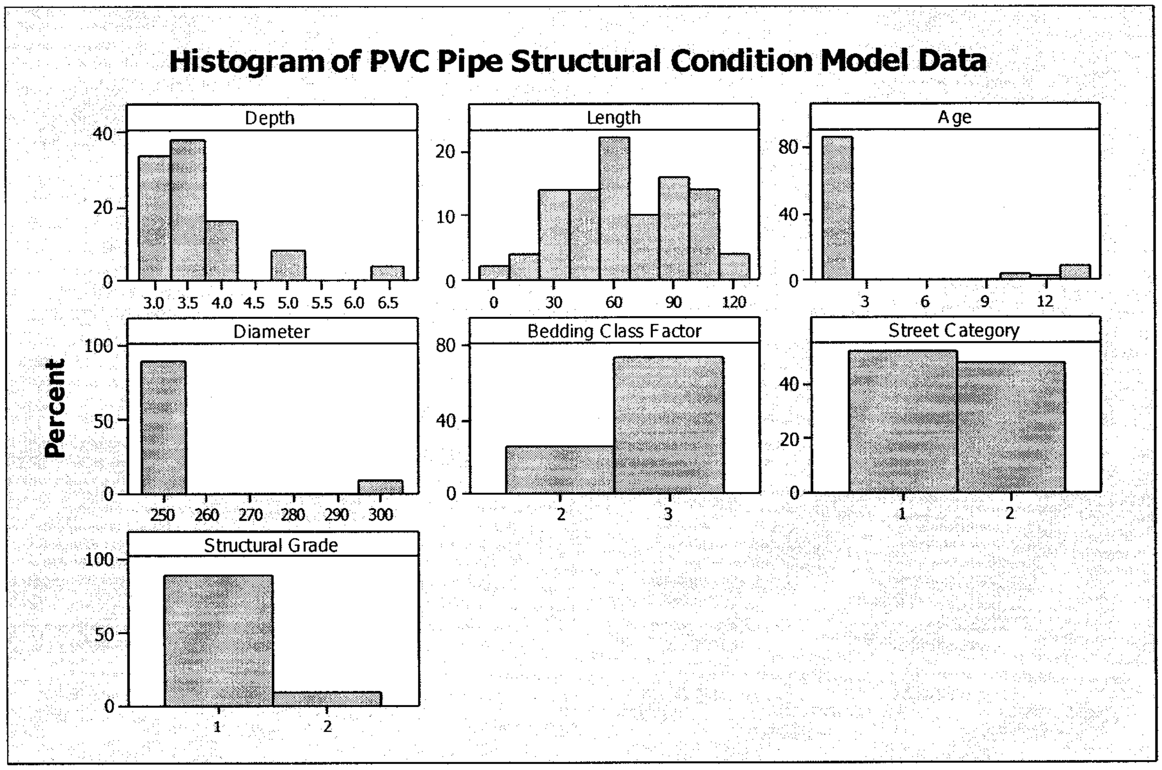


Figure 4-8: Histogram for Data of Group Number 3

Figure 4-8 presents the histogram for the transformed data for the different parameters. Further, the descriptive statistics for the data is also presented in Table 4-10.

Table 4-10: Descriptive Statistics for the Third Group Data

Parameter	Mean	Standard Deviation	Minimum	Median	Maximum
Depth (m)	3.66	0.8	2.93	3.36	6.48
Length (m)	66.03	29.26	5	62.9	126
Age (Years)	2.64	4.10	1	1	14
Diameter (mm)	255	15.15	250	250	300
Bedding Factor	2.74	0.44	2	3	3
Street Category	1.48	0.5	1	1	2
WRC Structural Rating	1.1	0.303	1	1	2

4.3.4.4 Group Number 4

The fourth and last group consists of data for the development of operational condition assessment model. Figure 4-9 presents the histogram for the available data for the different parameters.

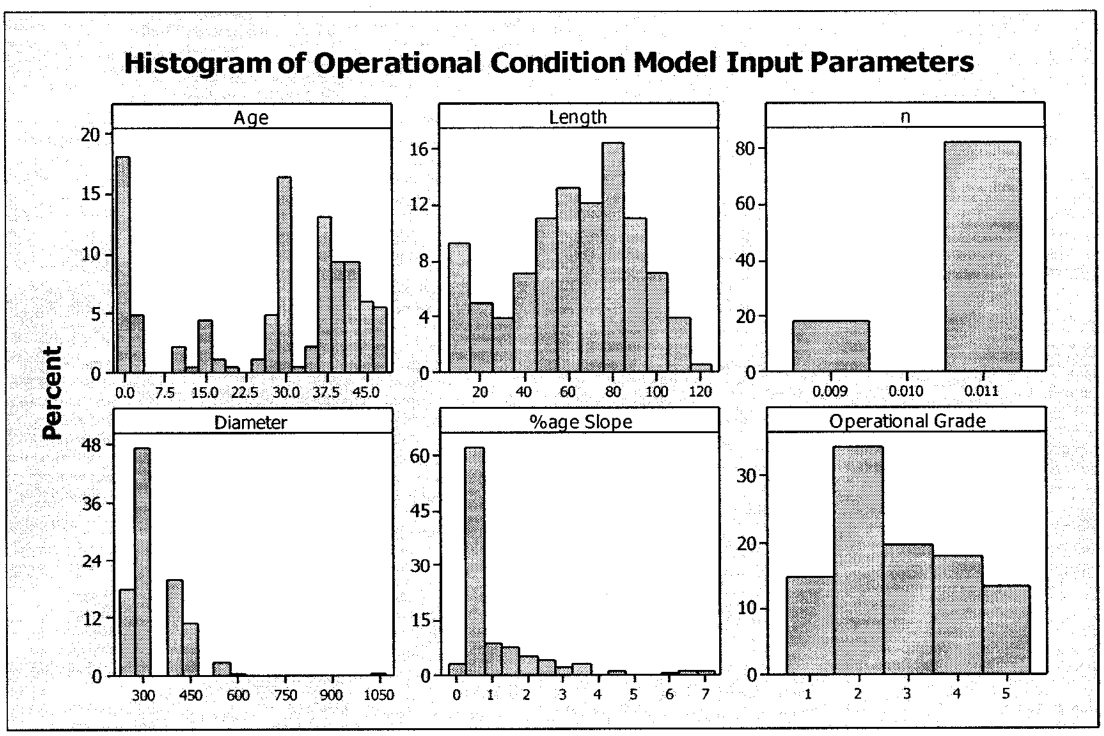


Figure 4-9: Histogram of Data Group Number 4

In Figure 4-9, "n" stands for the Manning's co-efficient for roughness; which indicates the different type of materials. The descriptive statistics for the data is also presented in Table 4-11.

Table 4-11: Descriptive Statistics for the Fourth Group Data

Parameter	Mean	Standard Deviation	Minimum	Median	Maximum
Age (Years)	26.69	16.26	1	31	47
Length (m)	62.23	27.97	6.1	65.5	121.2
Manning's Co-efficient	0.11	0.0007	0.009	0.011	0.011
Diameter (mm)	334.02	88.61	250	300	1050
%age Slope	1.06	1.29	0.11	0.5	6.82
WRc Operational Rating	2.8	1.27	1	3	5

4.4 Summary

A detailed discussion regarding data collection and preparation is presented in this chapter. The raw database from two different sources is transformed into a standardized format, which is ready to use for regression model development. The available parameters for model development are identified, and their preliminary statistical analysis is performed. The descriptive statistics and histogram presented in the chapter are beneficial in identifying outliers, data variability issues, and data ranges for analyzing and defining the limitations of developed models in the later stages.

Chapter 5

SEWER CONDITION PRIDICTION MODELS DEVELOPMENT

5.1 Overview

In the previous chapter, data acquisition, preparation and preliminary analysis procedures are described. This chapter deals with the regression model building and validating process. The application of regression analysis in this chapter is concentrated on building the most appropriate models for condition assessment or prediction of sewer pipelines. In Figure 4-1, a general methodology regarding data processing or regression model building has been presented. An overview of model building and validating methodologies has already been presented in Figures 3-2 and 3-3 respectively. The detailed description of these methodologies is described in this chapter. Furthermore, the chapter presents the designed structural and operational deterioration curves.

5.2 Model Building Process

Regression is a data oriented technique because it deals directly with the collected data without considering the process behind it (Zayed et al 2005). As a result the original data has been carefully pre-processed according to the adapted procedures described in the previous chapter. The original data is stored and pre processed in Microsoft Excel because of its versatility of spreadsheet analysis. For data processing or model building part, Minitab ® statistical software package is selected. Minitab ® is one of the most powerful, flexible, and easy to use (Kulandaivel, 2004) statistical software package. Due to broad spectrum influence of predictor variables upon response variable "pipe condition", many regression models are designed by defining different functional forms of explanatory or predictor variables. The necessary statistical diagnostics are applied to each model for further processing and decision making as shown in Figure 3-2.

As illustrated in Figure 4-5, four separate models are to be designed to describe the best possible fit for the prepared four data sets. The step by step methodology for designing the four models is same. Therefore, in this section, the model building process is described in a general manner with specific examples. The final outcome of all the models is summarized after the detailed explanation of the methodology. Following is the step by step explanation of the model building process:

5.2.1 Selection of Variables and Their Functional Forms

The selection of proper variables for model development depends upon the data in hand regarding the explanatory and response variables. Furthermore, a variety of diagnostics should be employed to identify the functional forms in which the explanatory variables should enter a regression model, and important interactions that should be included in the model (Neter et al 1996). The predictor selected for developing regression models are based upon the literature review of the pipe condition influence factors, and the actual data information in hand regarding these factors.

The second step is to define functional forms for the input variables. Linear regression model include not only first – order models in predictor variables but also more complex models. Therefore, models with transformed variables or with different interaction terms will be considered as linear regression models due to their respective linear parameters. For example, the following two equations present linear regression models (Kutner et al 2005):

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i2} + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + \varepsilon_i \qquad \text{(Equation 5.1)}$$

$$\log_{10} Y_i = \beta_0 + \beta_1 \sqrt{X_{i1}} + \beta_2 \exp(X_{i2}) + \varepsilon_i \qquad \text{(Equation 5.2)}$$

In this context, different functional forms of variables are defined for different possible scenarios. Each combination of variables is tested according to the methodology described in Figure 3-2 for finding the best fit.

Furthermore, the standardised function is also used to standardise the data. The function assigns a normalized value from a distribution characterized by mean and standard deviation. The normalized value for i^{th} observation of a specific distribution is obtained through the equation 5.3.

$$Z_i = \frac{X_i - \mu}{\sigma} \qquad \text{(Equation 5.3)}$$

where, Z_i is normalised or standardised value of the i^{th} observation, X_i is the i^{th} observation, μ is the mean of distribution, and σ is the standard deviation of the distribution.

5.2.2 Preliminary Diagnostics for Relationships and Interactions

The next step is to find out if any multi-colinearity or possible interactions existed in the variables. The matrix scatter plot for all the model input variable is obtained for detection and remedial measures. This plot is useful in detecting any existing bivariate relationships between predictors and response variables as well as among predictor variables. In matrix plots, Y variable for any scatter plot is the name found in its row and the X variable is the name found in its column. Figures 5-1 and 5-2 describe it in more detail.

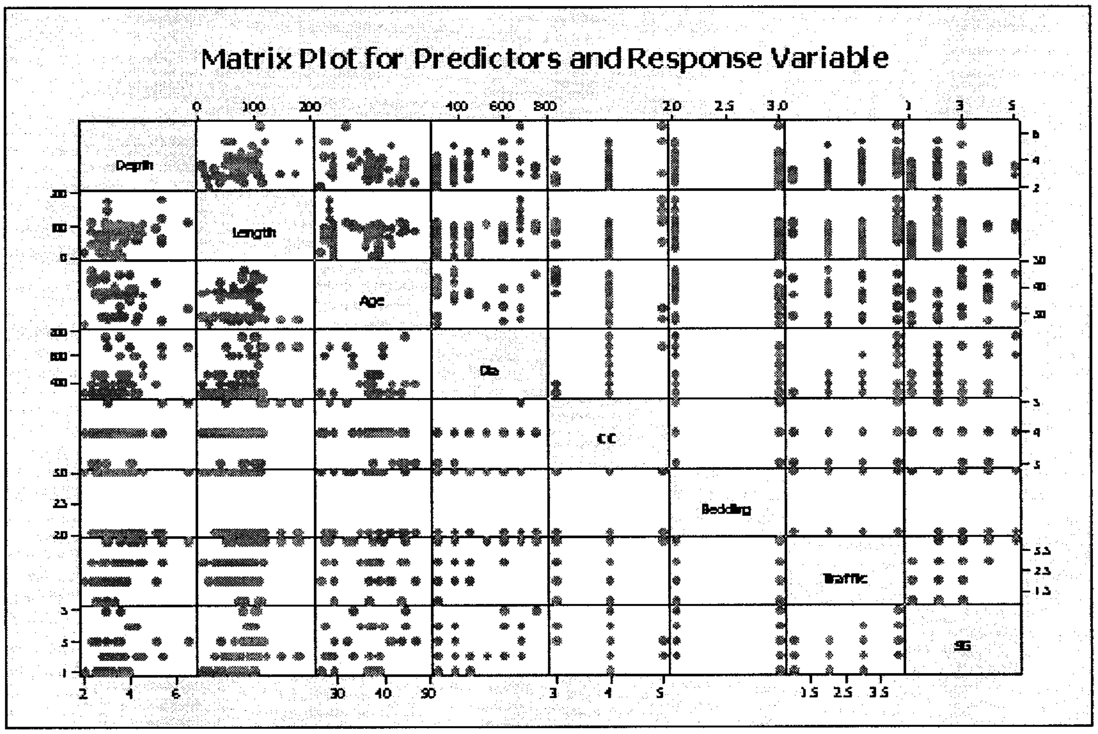


Figure 5-1: Matrix Plot for Concrete Pipe Structural Condition Assessment Model Variables before Defining Forms

Figure 5-1 shows the matrix plot for model variables prior to defining any forms. It also shows replicates in case of integer variables; nevertheless, the relationships between decimal variables are quite satisfactory. In case of integer variables with very less range (e.g. from 1 to 5 in most of cases under consideration), the replications are expected. Therefore, the plot is considered satisfactory.

Figure 5-2 presents a matrix plot for transformed variables. All the variables have been predefined by assigning some forms as described in step number 1. Figure 5-2 shows existing relationships and interactions between some variables. This fact is due to the defined functional forms and interactions; therefore, are expected in this kind of situation. Therefore, the results of plot are considered satisfactory.

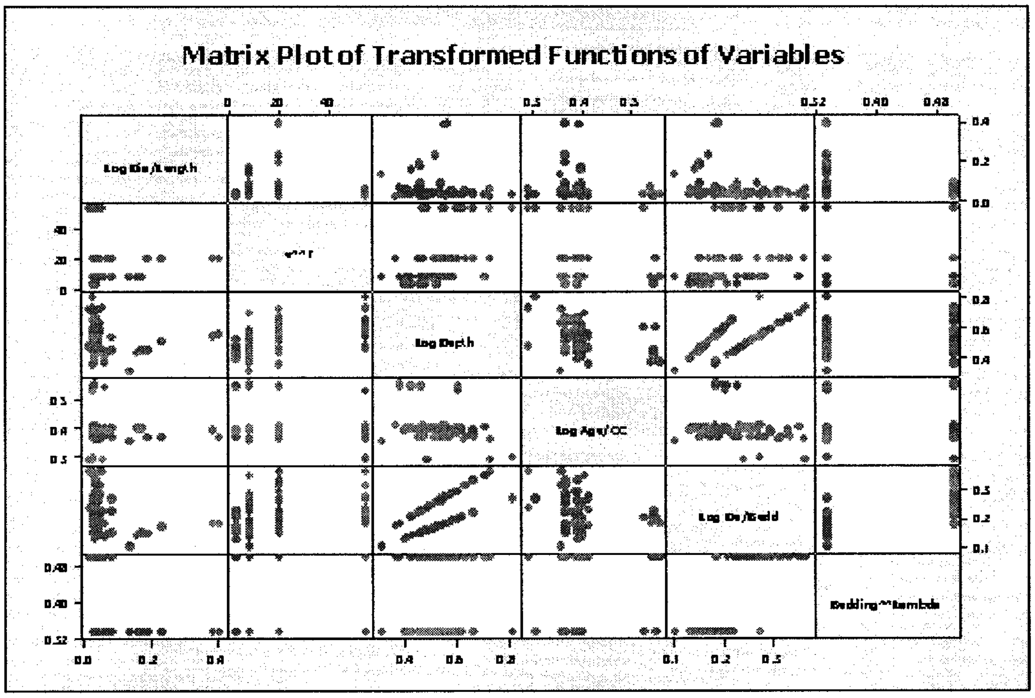


Figure 5-2: Matrix Plot for Concrete Pipe Structural Condition Assessment Model Variables after Defining Forms

5.2.3 Best Subset Analysis

If there are some unexpected interactions or bivariate relationships found in a matrix plot, the combination of variables is redefined through best subsets regression analysis. The best subsets analysis determines the best possible combination of variables with regards to lowest error and variation and the highest R^2 (adjusted) value. Therefore, best subset analysis identifies the best fit regression model that can be constructed with the specified number of variables.

Figure 5-3 shows an example of Minitab out put for best subset analysis. Each line of the output represents a different model. The first column represents the number of variables or predictors in the model. R^2 and R^2 (adjusted) have been converted into percentages.

Predictors that are present in the condition assessment model are indicated by a symbol "x".

Vars	R-Sq	R-Sq(adj)	Mallows C-p	S	Depth	Length	Age	Dia	CC	Bedding	Traffic
1	22.1	21.1	65.4	1.1332				X			
1	19.6	18.6	69.8	1.1512							X
2	40.6	39.0	34.4	0.99645				X	X		
2	34.2	32.4	45.8	1.0487					X		X
3	50.8	48.8	18.2	0.91332				X	X	X	
3	49.3	47.2	20.8	0.92672		X		X	X		
4	56.2	53.8	10.4	0.86734		X		X	X	X	
4	55.6	53.2	11.4	0.87287		X			X	X	X
5	60.1	57.3	5.4	0.83326		X		X	X	X	X
5	58.0	55.0	9.3	0.85546		X	X		X	X	X
6	60.9	57.6	6.0	0.83087		X	X	X	X	X	X
6	60.2	56.8	7.3	0.83842	X	X		X	X	X	X
7	60.9	57.0	8.0	0.83676	X	X	X	X	X	X	X

Figure 5-3: Minitab Output for Best Subset Analysis for Concrete Pipe Structural Condition Assessment Model Trial

It is clear from Figure 5-3 that the obtained model with higher R^2 (adjusted), lower C_p and lowest S value is the most appropriate model. Therefore, the variable "pipe depth" should be excluded from the model (C_p = 6.0 & S = 0.83087). Where, S is the standard deviation of residuals and C_p is described as follows:

$$C_p = \frac{SSE_p}{MSE(X_1 \ldots\ldots X_{p-1})} - (n - 2p) \qquad \text{(Equation 5.4)}$$

where, SSE_p is error sum of squares for the fitted subset regression model with p parameters (p-1 predictors), $MSE(X_1 \ldots\ldots X_{p-1})$ is unbiased estimate of variance, and n is the number of observations.

5.2.4 Model Building

After all the above mentioned preliminary diagnostics the next step is to build a multiple regression model for further analysis. The corresponding data for the deformed variables is put in Minitab for regression analysis. The Minitab output consisted of a certain regression equation with an estimate of regression coefficients "β_k" for the specified data and some other results for further analysis.

5.2.5 Preliminary Tests for Model Adequacy

The preliminary tests include; coefficient of multiple determination, F test for regression relation ant t test for each regression parameter "β_k". Figure 5-4 shows the Minitab output for these tests. R^2 and R^2 (adjusted) values are 84.9% and 83.1% respectively. The R^2 value indicates that the predictors explain 84.9% of the variance in "structural grade" (response variable). The R^2 (adjusted) accounts the number of predictors in the model. Both values indicate that the model fits the data well.

The next test is the F test for regression. To determine P(F) for the whole model, a hypothesis test is carried out. The null hypothesis (H_0) assumes that all regression coefficients, β_0, $\beta_{1...}$ βp_{-1} are zero i.e. $\beta_0 = \beta_1 = \beta p_{-1} = 0$. The alternate hypothesis (H_a) assumes that not all of them equal to zero. In Figure 5-4, the p-value (statistical significance) in analysis of variance table is 0.000. That means that null hypothesis is rejected. This shows that the estimated model is significant at α - level of 0.05. Therefore, at least one coefficient in the estimated regression equation is not zero.

The next step is to test that all predictors are significantly related to the response variable or not. To determine the validity of regression coefficient individually, "t-tests" are performed separately for β_0, $\beta_{1...}$ βp_{-1} in a similar fashion. In case of β_0, the null hypothesis (H_0) of t-test assumes that $\beta_0 = 0$; while alternative hypothesis (H_a) assumes that $\beta_0 \neq 0$. Similarly, the other null hypothesis assumes that $\beta_1 = 0$ and vise versa. The results of these tests are shown in Figure 5-4.

Predictor	Coef	SE Coef	T	P
Constant	-0.056508	0.004799	-11.77	0.000
Log Depth	-0.009911	0.005630	-1.76	0.080
Depth/Length	-0.000011	0.000230	-1.36	0.180
Length	0.00001311	0.00001375	0.95	0.342
Age	0.00011606	0.00005646	2.06	0.041
Diameter	-0.00000082	0.00000395	-0.21	0.837
Concrete Class	-0.000218	0.001013	-0.22	0.830
Bedding Factor	0.007641	0.001496	5.11	0.000
Street Class	0.0020209	0.0005337	3.79	0.000

S = 0.00634790 R-Sq = 84.9% R-Sq (adj) = 83.1%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	8	0.631275	0.078909	1958.25	0.000
Residual Error	182	0.007334	0.000040		
Total	190	0.638609			

Figure 5-4: Minitab Output for Preliminary Test Results of a Concrete Pipe Structural Condition Assessment Trial Model

Figure 5-4 shows that the p-value for the estimated coefficients for predictors "Bedding Factor" and "Street Class" is 0.000. Similarly, the p-value for predictor "Age" is 0.041. As a result, alternative hypothesis is accepted. This indicates that the predictors are significantly related with the response variable "Structural Grade" at α - level of 0.05. However, the case is different for other predictors. For example, p–value for the estimated coefficient for predictor "diameter" is 0.837; which is closer to 1 and greater than α - level of 0.05. This shows that predictor is not significantly related to response variable.

Therefore, due to the unsatisfactory results, the model is rejected and some other forms of variables should be introduced for the next trial. This process is repeated, and some models with better results are selected for further diagnostics. One of them is shown in Figure 5-5. Figure 5-5 shows that the model has less R^2 value as compare to the model shown in Figure 5-4; however, its t tests for β_k are better. Therefore, on this criterion the model should be selected for further diagnostics

Predictor	Coef	SE Coef	T	P
Constant	-0.8491	0.7147	-1.19	0.239
Log Dia/Length	0.5921	0.3151	1.88	0.064
e**T	-0.006807	0.001196	-5.69	0.000
Log Depth	-3.217	1.213	-2.65	0.010
Log Age/CC	-1.6044	0.4161	-3.86	0.000
Log De/Bedd	6.919	2.754	2.51	0.014
Bedding	0.9582	0.2511	3.82	0.000

S = 1.02281   R-Sq = 72.7%   R-Sq(adj) = 70.5%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	6	200.979	33.497	32.02	0.000
Residual Error	72	75.322	1.046		
Total	78	276.301			

Figure 5-5: Minitab Output for Preliminary Test Results for one of the Trial Models

5.2.6 Residual Analysis

After obtaining satisfactory results from the previous step, the next step is to analyze the residuals and their patterns. These diagnostic checks are essential to verify the linear regression assumptions. These diagnostic are described below:

(i) Normality of Error

Consider the Minitab output for normal probability and frequency plots for residuals for the selected model (Figure 5-6) from the previous step. The normal probability plot shows that error terms are nearly normal. As small departures from normality do not create any serious problems (Kutner et al 2005); the results could be considered as satisfactory. However, there could be a possibility of outliers. The possibility of outliers is also clear from the histogram of residuals plot. The bar on the far right and two bars on far left indicate that the data correspond to these values are not fit with the model.

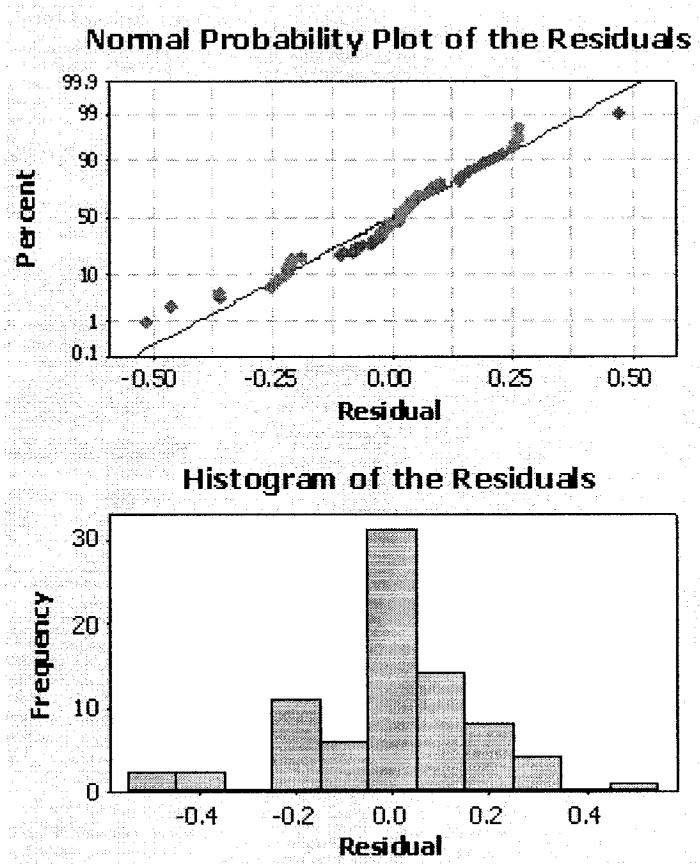


Figure 5-6: Normal Probability and Histogram of Residual Plots for Concrete Pipe Structural Condition Assessment Chosen Model

In order to check the possibility of outliers and errors in normal probability plots, Minitab output for unusual observations are analyzed. Figure 5-7 shows observations with large standardized residuals, and some with large influence on the model characteristics. These observations are affecting the normal probability plot of residuals. A question arises that either should these observations be removed by considering them outliers or should be considered important to be included in the model? After a careful examination of the unusual observations, it is found out that the pipeline deterioration phenomenon is sometimes extremely uncertain. For example, one of the unusual observations shows that the pipe condition is in class 1 (excellent) while the pipe's age is 40 years.

```
Unusual Observations

Obs  Log Dia/Length   1/SG    Fit   SE Fit  Residual  St Resid
  1            0.027  0.200  0.558   0.046    -0.358     -2.09R
  2            0.027  0.200  0.558   0.046    -0.358     -2.09R
 18            0.026  0.333  0.439   0.107    -0.106     -0.73 X
 35            0.035  0.500  0.965   0.042    -0.465     -2.51R
 36            0.023  0.500  1.013   0.045    -0.513     -2.78R
 56            0.406  1.000  1.027   0.111    -0.027     -0.17 X
 57            0.031  1.000  0.525   0.047     0.475      2.58R
 69            0.390  1.000  0.980   0.106     0.020      0.15 X

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large influence.
```

Figure 5-7: Program Output for Unusual Observations

The program considers these types of observations as possible outliers. Eliminating these unusual observations from the model would give better results in terms of R^2 values and other statistical parameters; however, the model could not be considered as the best representation of the real world data in hand. Figure 5-8 shows the normal probability plot of residuals after eliminating all the unusual observations. The results show that there is minimum possibility of outliers.

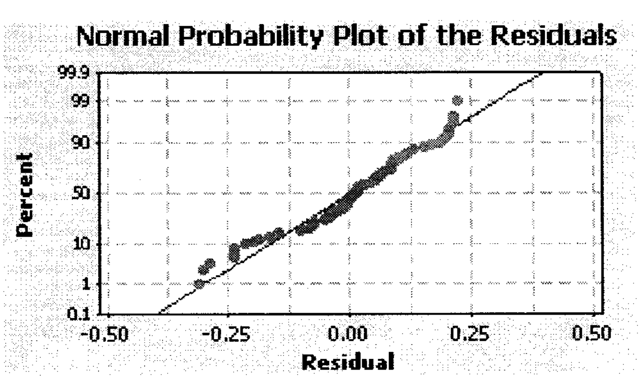


Figure 5-8: Normal Probability Plot of Residuals for the Chosen Model after Eliminating Outliers

Moreover, Figure 5-9 shows that the value of R^2 has also been improved by eliminating the possible outliers. Furthermore, there seems to be an improving trend in p-value for β_0. However, p-value for β_k of predictor "Log Diameter/Depth" is 0.920. Previously, when the outliers were considered, it was 0.064 (Figure 5-5). Therefore, two major predictors under consideration, pipe diameter and pipe depth, have almost been eliminated from the model; which could lead to pitfall. Consequently, these possible outliers are considered important to be represented in the model.

```
Predictor                Coef     SE Coef      T      P
Constant               4.0572      0.6107   6.64  0.000
Log Dia/Length         0.0488      0.4848   0.10  0.920
e**T               -0.0062616   0.0009923  -6.31  0.000
Log Depth              -2.710       1.197  -2.26  0.027
Log Age/CC            -2.0221      0.3447  -5.87  0.000
Log De/Bedd             5.245       2.630   1.99  0.050
Bedding**Lambda        -5.302       1.423  -3.73  0.000


S = 0.798235   R-Sq = 83.5%   R-Sq(adj) = 82.0%
```

Figure 5-9: Minitab Output for Preliminary Test Results for the Chosen Model after Eliminating Outliers

(ii) Homoscedasticity

The second assumption that the variation around a regression line be constant for all values of X can be verified through the residuals vs. the fitted value plot. Figure 5-10 shows the fitted value plot for the model under consideration. In ideal scenario, constant data would be distributed evenly across the plot. That would show the consistent variance across the fitted value range. However, Figure 5-10 shows diagonal bands across the centre line. The reasons for these types of results could be due to:

- Important variable(s) might be omitted from the model (Neter et al 1996)

➤ Data variability issues: data composed of integer variables (Anderson et al 2005)

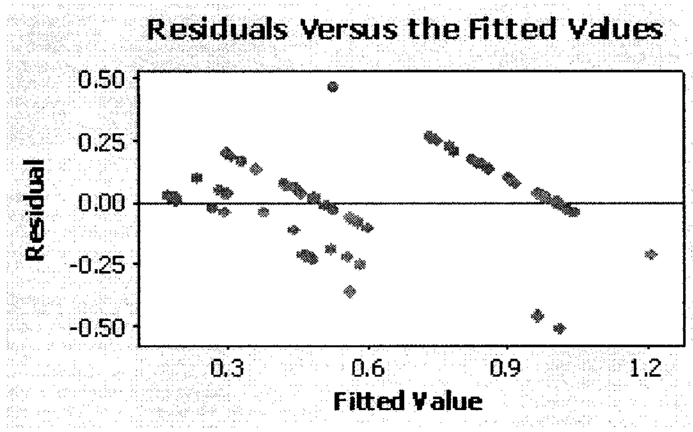


Figure 5-10: Residual vs. Fitted Values Plot for Concrete Pipe Structural Condition Assessment Chosen Model from Step 4

The careful examination for the data in hand shows that both above mentioned possibilities exist in this case. Some of the important variables which could have a strong effect on existing pipe conditions could be missing. For example, type of soil, maintenance and repair history, infiltration etc. are important parameters which affects existing pipe conditions directly. As information regarding these kinds of parameters was not available; the study of the effect of these parameters on the pipe condition is recommended for future research.

The second possibility of diagonal bands could be due to existence of integer predictors causing data variability. The model under consideration has integer predictors and response: concrete class, bedding class factor, street categories, and pipe condition. The discrete values of these variables could cause the problem of unequal variance. The remedial to the unequal variance is weighted least square regression (Kutner et al 2005). If some important parameters are not omitted, the weighted least square regression will produce better results. The weighted least square regression will be discussed later.

(iii) Independence of Error

Errors around a regression line should be independent for each value of predictors. Figure 5-11 shows the residuals vs. the order of data plot for the model under consideration. The results shows positive residuals at outer bands of X values, and the inner bands largely consist of negative residuals.

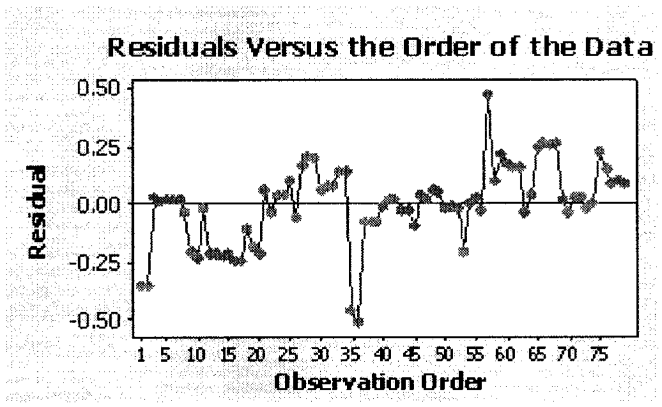


Figure 5-11: Residual vs. Order of Data Plot for Concrete Pipe Structural Condition Assessment Chosen Model

The results could lead to a conclusion that there could be a pronounced shift in the regression equation in the outer bands of the series; resulting in lack of fit of regression function. Therefore, some additional diagnostics are necessary to evaluate the model.

5.2.7 Additional Diagnostics

In the previous section it has been seen that residual plots are one of the most necessary tools in identifying a regression model's adequacy. If the required results are not achieved, the adequacy of a regression equation becomes objectionable. Therefore, some additional statistical test should be performed before reaching on a final conclusion. There are many statistical diagnostics available to check the adequacy of a model; however, the two important diagnostics are described below:

(i) Lack of Fit Test

The first diagnostic is the lack of fit test. As mentioned in literature review, Minitab performs this test in two ways;

- If enough replicates are available → pure error test
- If not enough replicates are available → data subsetting test

Lack of fit test is performed on the model under considerations and it is found out that the program cannot perform pure error test due to lack of replicates. Therefore, data subsetting test results are displaced as, "Overall lack of fit test is significant at p = 0.000". The results also show some possible interaction in the predictor "Bedding Factor" at p = 0.000. Therefore, the model can not be presented as fit for the data in hand.

(ii) Durbin-Watson Test

This test is important in identifying the possibility of auto-correlation among the predictors. The Minitab out put for the model under consideration for Durbin-Watson Statistics is 0.804345. The Durbin-Watson test bound tables provide the values for test statistics up to five predictors (Neter, 1996). In the case under consideration, the predictors are six.

$d_L = 1.49$ ($n = 80$, $p-1 = 5$)
$d_U = 1.77$ ($n = 80$, $p-1 = 5$)
H_0: $\rho = 0$ (Error terms are independent)
H_1: $\rho > 1$ (Error terms are positively correlated)
If
$D > d_U$, conclude H_0
$D < d_L$, Conclude H_1
$d_L \leq D \leq d_U$, the test is inconclusive
Result: $0.804345 < 1.49$
$D < d_L$, Error terms are positively correlated

Figure 5-12: Durbin-Watson Test Statistics for the Model under Consideration

Therefore, the table is interpolated accordingly, and the results are shown in Figure 5-12. The results show that error terms are positively correlated. This shows that the error terms are not independent and remedial are required in this regard.

5.2.8 Remedial Measures

In the previous two steps, some necessary diagnostics for checking the adequacy of regression models have been discussed. This section deals with some remedies of the problems detected through these diagnostics in the regression models.

(i) Weighted Least Square Regression

One of the remedies for unequal error term variances is weighted least square regression. The model under consideration is tested for different weighted trials by assigning weights in variables one by one. Minitab performs weighted regression analysis using weights in one variable at a time. In order to obtain better results, all variables in the model under consideration were assigned weights one by one, and obtained results were analysed on hit and trial basis. Two of the examples are explained in Figure 5-13.

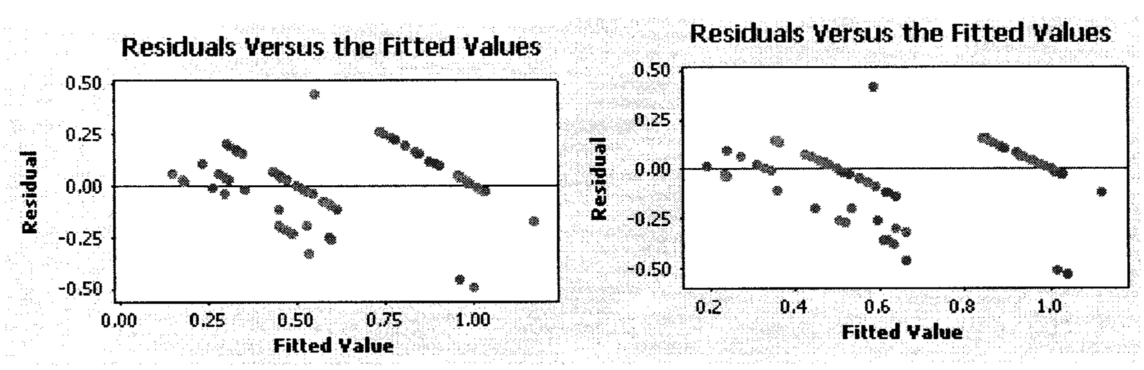


Figure 5-13: Residual vs. Fitted Values Plot using weights in Predictor 'bedding Factor' (left) and in response 'SG' (Right) for Concrete Pipe Structural Condition Assessment Chosen Model

Figure 5-13 shows the results for residual vs. fitted value plots for two different types of weighted regression. This analysis is performed by assigning weights in each variable to judge any improvement in the inconsistency of error term. The purpose of weighted regression is the reduction of sequential error terms variance; consequently, to minimize the possibility of variability issues or noise in data. However, Figure 5-13 shows almost the same results as already been presented in Figure 5-10. Therefore, it is concluded that these diagonal bands of residual vs. fitted values are due to some omitted important variable which could have an influence on a pipe's structural condition.

(ii) Box-Cox Transformation

For remedies of lack of fit and inconsistent error terms, the response variable for the model under consideration is transformed according to the Box-Cox procedure, and is shown in Figure 5-14.

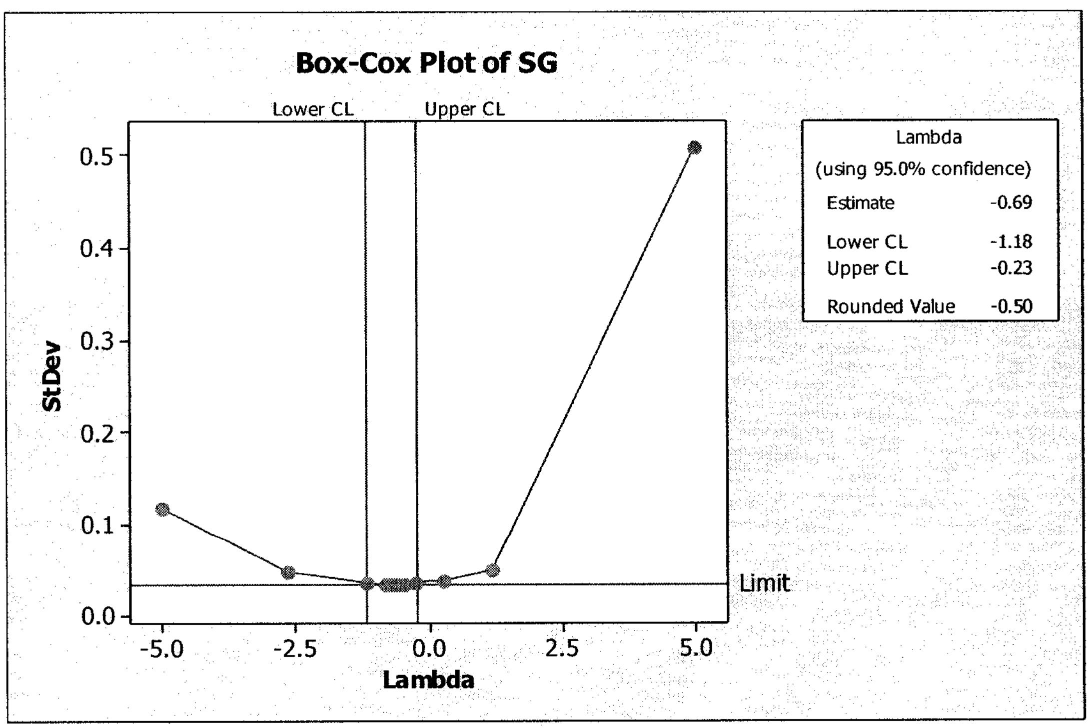


Figure 5-14: Box-Cox Plot for Response Variable

Figure 5-14 shows the Box-Cox plot of response variable "Structural Grade". It is clear that if power transformation on Y (λ) is near to -0.5, the standard deviation in the data is minimum. Therefore, the response variable is transformed accordingly for the transformed model development. As described earlier, the lack of fit test displayed that there could be possible interaction in the predictor "bedding Factor". Therefore, the predictor is also transformed accordingly and is shown in the Figure 5-15. Figure 5-15 shows that the rounded value of λ in this case is -1.00. Consequently, the predictor is transformed according to the obtained Box-Cox power transformation results.

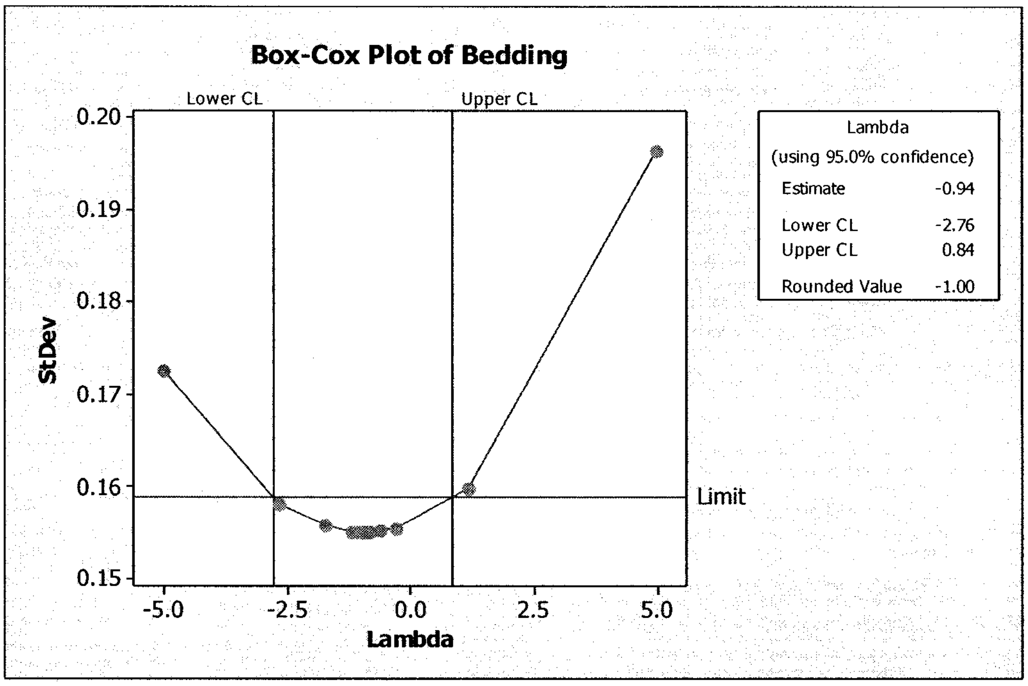


Figure 5-15: Box-Cox Plot for Predictor "Bedding Factor"

The transformed variables are again regressed, and preliminary results of the obtained model are shown in Figure 5-16. As far as the R^2 and F test for regression is concerned, there is no difference between the values obtained previously. However, there is a considerable improvement in the t test for β_0. The p – value in this case is 0.000; where the p - value in the previous case was 0.239. Therefore, the intercept of regression plane is better represented in the transformed model.

```
Predictor                 Coef    SE Coef      T      P
Constant                3.9420     0.6225   6.33  0.000
Log Dia/Length          0.5921     0.3151   1.88  0.064
e**T                 -0.006807   0.001196  -5.69  0.000
Log Depth               -3.217      1.213  -2.65  0.010
Log Age/CC             -1.6044     0.4161  -3.86  0.000
Log De/Bedd              6.919      2.754   2.51  0.014
Bedding**Lambda         -5.749      1.507  -3.82  0.000


S = 1.02281   R-Sq = 72.7%   R-Sq(adj) = 70.5%


Analysis of Variance

Source          DF       SS      MS      F      P
Regression       6  200.979  33.497  32.02  0.000
Residual Error  72   75.322   1.046
Total           78  276.301
```

Figure 5-16: Minitab Output for Preliminary Test Results for the Concrete Pipe Structural Condition Assessment Transformed Model

In case of residual analysis, the results are almost the same except the diagonal band slope of the residual vs. fitted value plot are inversed. This is due to the negative power transformation of the response variable after Box-Cox procedure. In case of Lack of Fit test, the results show some improvement, as the p-value of data subsetting test is increased from 0.000 to 0.049. Furthermore, the Durbin-Watson statistics value is increased form 0.804 to 0.95; nevertheless, is still less than d_L, showing auto correlation.

In short, the transformed model shows better results for statistical diagnostic than the previous model; therefore, it is selected for the model validation process. Table 5-1 shows the summary of results for both the models: model before Box-Cox transformation and model after Box-Cox transformation.

Table 5-1: Summary of Statistical Results for Concrete Pipes Structural Condition Grading Models under Consideration

Model	R^2	R^2_{adj}	P (F)	P (t)							Durbin-Watson Test	P (F) Lack of Fit	
				β_0	β_1	β_2	β_3	β_4	β_5	β_6		Pure Error	Data Sub-setting
Before Box-Cox Transformation	72.7	70.5	0.000	0.239	0.064	0.000	0.010	0.000	0.014	0.000	0.804 $D<d_L$	-	0.000
After Box-Cox Transformation	72.7	70.5	0.000	0.000	0.064	0.000	0.010	0.000	0.014	0.000	0.95 $D<d_L$	-	0.049

The regression model building process and decision making criterion for selection or rejection of the built models have been explained above through one example i.e. concrete pipe structural condition assessment model. The other models for the remaining three data groups are built according to the same methodology. All of these models are than validated through their respective validation data.

5.3 Model Validation Process

It is noted that all the statistical diagnostics which have been described above are not enough to check the adequacy of a regression model. Therefore, a comprehensive model validation procedure is applied to all selected models for validating them. An overview of this methodology has already been shown in Figure 3-3.

The validation data for all the four groups is embedded into the regression model for comparing its results with the actual results using Microsoft Excel spread sheet procedures. Furthermore, descriptive statistics and histograms for the actual and predicted output data is obtained through Minitab statistical software package. All of these steps are explained below through examples in an order of sequence.

5.3.1 Actual vs. Predicted Output Plot

The first step is to compare the actual observation with the predicted values for the validation data for each group. Figure 5-17 shows actual vs. predicted output plot for the concrete pipe structural condition model as an example. The figure shows that the predicted values are well in acceptable limits and are scattered around the actual values for response variable. Therefore, the first validation test's results are satisfactory.

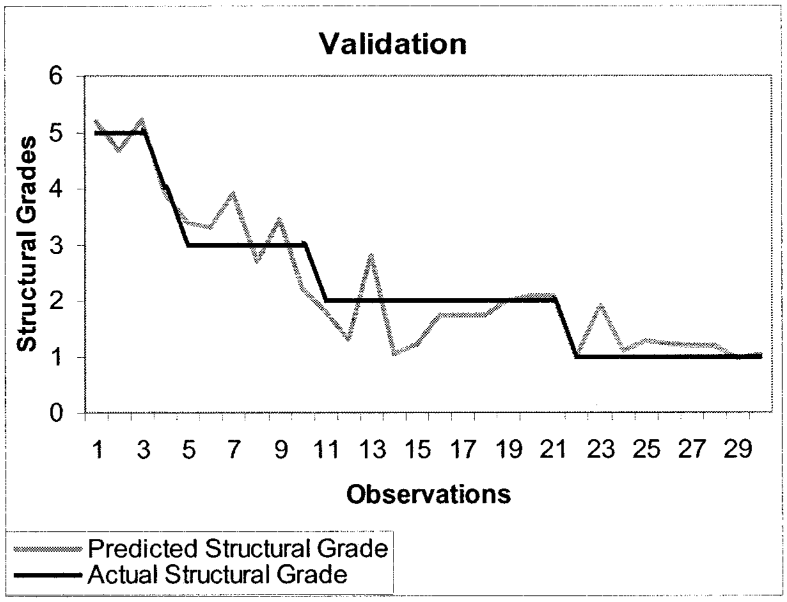


Figure 5-17: Model Validation Plot as per Order of Observation for the Selected Concrete Pipe Structural Condition Prediction Model

5.3.2 Descriptive Statistics

The second step is to check descriptive statistics of the actual and predictive output observations.

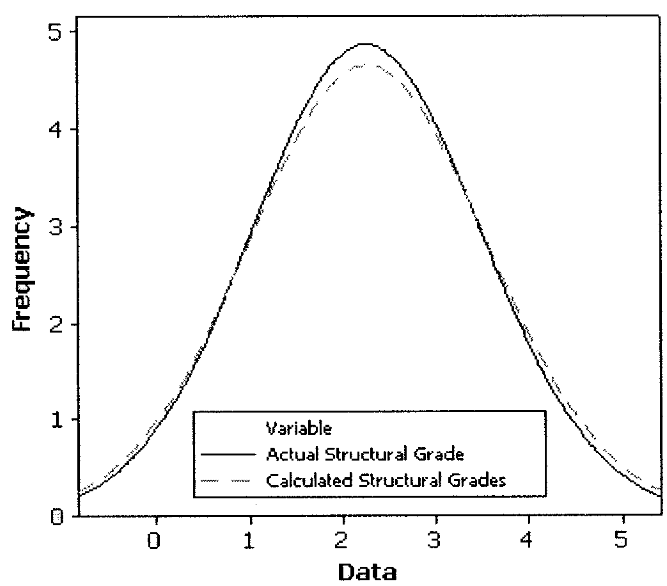


Figure 5-18: Minitab Output for Histogram of Actual and Predicted Values

Table 5-2: Descriptive Statistics for Actual and Predicted Outputs for Validation Data

Descriptive Statistics	**Mean**		**St Deviation**	
Observations	Actual	Predicted	Actual	Predicted
Value	2.267	2.283	1.23	1.287

Figure 5-18 and Table 5-2 show descriptive statistics for the concrete pipe structural condition prediction model validation data. The results show that the mean and standard deviation of actual and predicted outputs are quite closer to each other. The predicted outputs have slightly more values of mean and standard deviation; however, the results are satisfactory.

5.3.3 Mathematical Validation Diagnostics

The validation data is checked for all commonly used mathematical parameters for model validation. All these mathematical parameters are discussed below:

(i) Average Invalidity and Validity Percent (AIP & AVP)

Average Invalidity and Validity Percent can be calculated for validation data by the following formulae (Zayed et al 2005):

$$AIP = \frac{\sum_{i=1}^{n} \left| 1 - \left(\frac{E_i}{C_i} \right) \right|}{n} \qquad \text{(Equation 5.5)}$$

and

$$AVP = 1 - AIP \qquad \text{(Equation 5.6)}$$

where, AIP is Average Invalidity Percent, AVP is Average Validity Percent, E_i is Estimated or Predicted Value, C_i is Actual Value, and n is the number of observations.

AIP value varies from 0 to 1. If the value AIP is closer to 0; the model is fit for its validation data. AIP value closer to 1 shows that the model is not appropriate for its validation data. On the contrary, for a satisfactory validation, AVP value should be closer to 1. For the model under consideration, the values are listed below:

- AIP = 0.1808
- AVP = 0.8192

The value show that predicted outputs are almost 82% accurate. The results can be considered as more than satisfactory because the model which is being validated has R^2 value of about 72%. That means it explains 72% of variation in model building data; further, it explains about 82% of variation in validation data.

(ii) Root Mean Square Error (RMSE)

Root mean square error (RMSE) can be estimated by the formula (Barqawi, 2006):

$$RMSE = \frac{\sqrt{\sum_{i=1}^{n}(C_i - E_i)^2}}{n} \qquad \text{(Equation 5.7)}$$

where, $RMSE$ = Root Mean Square Error, E_i = Estimated or Predicted Value, C_i = Actual Value, and n = No of Observations. The value of $RMSE$ close to 0 shows that the model is fit for its validation data. For the model under consideration, the value of RMSE is 0.0827.

5.3.3.3 Mean Absolute Error and Fitness Function (MAE & f_i)

The mean absolute error (MAE) is defined as (Barqawi, 2006):

$$MAE = \frac{\sum_{i=1}^{n} |C_i - E_i|}{n} \qquad \text{(Equation 5.8)}$$

where, MAE is Mean Absolute Error, E_i is Estimated or Predicted Value, C_i is Actual Value, n is the number of observations.

MAE value varies from 0 to infinite. However, the value of mean absolute error should be close to zero for the validity of a model. Further, the mean absolute error value is used to define the fitness function f_i for a model validation. The fitness function can be calculated as (Dikmen et al 2005):

$$f_i = \frac{1000}{1 + MAE} \qquad \text{(Equation 5.9)}$$

where, MAE is Mean Absolute Error, and f_i is Fitness Function. The equation for fitness function shows that if the value of fitness function is closer to 1000 for a model, the model is fit for validation data. f_i value closer to 0 indicates that the model is inappropriate for the representation of validation data. The values for MAE and f_i for the model under consideration are:

- MAE = 0.3465
- f_i = 742.66

The above results show that the model's accuracy is about 74.26%.

All the above validation checks show that the model is fit for the validation data. Therefore, the model was selected as the best fit model for the representation of data in hand; includes model building as well as validation data.

5.4 Summary Developed Models

All the above mentioned methodology for model building and validating processes has been illustrated by giving examples from concrete pipe structural condition assessment model. As already been mentioned, four data groups are prepared to develop four different models. Consequently, all the models are built and tested according to the same adapted methodology. Total number of predictors in a developed model may differ with other models. This is due to available input data and results of different statistical tests e.g. best subset analysis etc. The regression equations for the developed models are listed below. Further, Table 5-3 presents the results for the different statistical diagnostics applied to all the developed models

5.4.1 Structural Condition Assessment Models

Three different models for asbestos cement, concrete, and PVC pipes have been designed. The final outcome of the models is described below:

(i) Asbestos Cement Pipe Structural Condition Assessment Model

The developed regression equation is:

$$Structural\_Grade = \sqrt{\begin{array}{l} 20.9 + 542\dfrac{Log_{10}Depth}{Length} + 0.207Age - 0.742Aabestos\_Cement\_Class \\ -14.8Diameter^{0.1} \end{array}}$$

(Equation 5.10)

The units of all variables are same as described in Table 4-9.

(ii) Concrete Pipe Structural Condition Prediction Model

$$Structural\_Grade = \begin{bmatrix} 3.94 + 0.592\dfrac{Log_{10}Diameter}{Length} - 0.00681e^{Street\_Category} \\ -3.22Log_{10}Depth - 1.6\dfrac{Log_{10}Age}{Concrete\_Class} \\ +6.92\dfrac{Log_{10}Depth}{Bedding\_Factor} - 5.75\dfrac{1}{Bedding\_Factor} \end{bmatrix}^{-1}$$

(Equation 5.11)

The units of all variables are same as described in Table 4-8.

(iii) PVC Pipe Structural Condition Prediction Model

$$Structural\_Grade = -Log\begin{bmatrix} 2.25 - 0.00642Age - 1.89Length^{0.01} \\ -0.0302Bedding\_Factor - 0.0405Street\_Category \\ -0.0000(Diameter)^{0.3}(Depth)^{4} \end{bmatrix}$$

(Equation 5.12)

The units of all variables are same as described in Table 4-10.

5.4.2 Operational condition Prediction Model

As mentioned above, operation condition assessment model was developed for all the three pipe materials and is given in the following equation:

$$Operational\_Grade = \left[\frac{0.308 + 0.567\left(\dfrac{Age}{Diameter^{n}}\right)(Length)^{Slope}}{Age}\right]^{\frac{1}{0.63}} \qquad \text{(Equation 5.13)}$$

Where, n is the Manning's roughness coefficient for a specific pipe material. The units of all variables are same as described in Table 4-11.

Table 5-3: Summary of Statistical Results for Condition Prediction Models

Model		R^2 (%)	R^2 (Adj) (%)	P (F)	P (t)							Durbin-Watson Statistics	P (F) Lack of Fit	
					β_0	β_1	β_2	β_3	β_4	β_5	β_6		Pure Error	Data Sub-setting
Structural Condition Prediction	Concrete Pipes	72.7	70.5	0.000	0.000	0.064	0.000	0.010	0.000	0.014	0.000	0.95 $D<d_L$	-	0.049
	Asbestos Cement Pipes	82.4	78.3	0.000	0.041	0.001	0.034	0.093	0.085	--	--	1.43 $d_L \le D \le d_U$	0.874	≥ 0.1
	PVC Pipes	81.8	78.6	0.000	0.003	0.000	0.008	0.021	0.000	0.000	--	1.82 $D > d_U$	--	0.056
Operational Condition Prediction		87.9	87.8	0.000	0.007	0.000	--	--	--	--	--	0.465 $D < d_L$	--	0.073

Table 5-4: Validation Summary Results for Built Regression Models

Condition Rating Model		AIP	AVP	RMS	MAE	f_i	Descriptive Statistics for Outputs			
							Mean		Standard Deviation	
							Actual	Predicted	Actual	Predicted
Structural	**Concrete**	0.181	0.819	0.083	0.346	742.94	2.267	2.283	1.23	1.287
	Asbestos Cement	0.139	0.860	0.074	0.144	873.88	1.286	1.42	0.488	0.42
	PVC	0.142	0.857	0.043	0.152	867.99	1.063	0.971	0.25	0.215
Operational		0.164	0.836	0.094	0.413	707.91	2.718	2.529	1.191	1.227

The summary of results of all developed models is shown in Tables 5-3 and 5-4.

5.4.3 Limitation of Models

The developed regression models are limited to a certain range of input data. These input data ranges have been described in chapter 4 (Tables 4-8 to 4-11).

5.4.4 Validation Summary and Parameter Comparison

Table 5-4 shows the summary of results for all concerned validation checks for the selected models. All of the validation checks shown in Table 5-4 are in satisfactory range with respect to data in hand and R^2 values for the respective models. However, some of the factors are more sensitive to variation or noise in the validation data than others. Figure 5-19 shows a comparison between two of the validation parameters; average validity percent (AVP) and fitness function (*fi*). As the fitness function is dependent upon MAE value, it is more sensitive than average validity percent. Nevertheless, the values of both the parameters are compatible to their respective R^2 values.

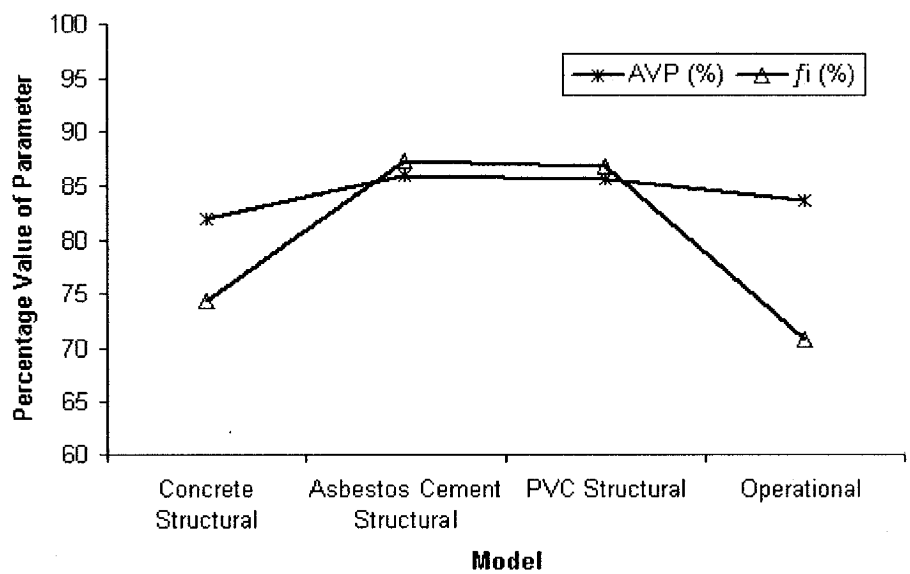


Figure 5-19: Sensitivity Comparison between AVP and Fitness Function

5.5 Deterioration curves for Sewers

Based on developed model, structural and operational deterioration curves for sewers are built. Due to limitation of input data ranges for asbestos cement and PVC structural condition models, the structural deterioration curves are only developed for concrete pipes. Nevertheless, the developed operational deterioration curves represent all pipe materials under consideration.

5.5.1 Structural Deterioration Curves

Equation 5.11 shows that deterioration of a concrete pipe is a complex phenomenon, which involves many sewer attributes. The structural deterioration curves are built by varying one or two attributes at a time; while other attributes are kept constant.

Figures 5-20 and 5-21 represent structural deterioration of concrete sewers with respect to their road categories. The curves have been drawn by considering average values of other attributes in concrete pipe regression model. For example, average values of depth and length are taken as 3.5m and 80m respectively. Figure 5-20 represent collector sewers with the diameter ranges from 200mm to 525mm, and Figure 5-21 represent trunk sewers with the diameter ranges from 600mm to 825mm. It is observed that pipes buried under arterial streets will reach the critical condition class in 50 years when all the other values will be kept constant. It is also observed that this deterioration trend is slightly more in case of collector sewers. This could be due to greater length to diameter ratios for collectors, which could increase bending stresses.

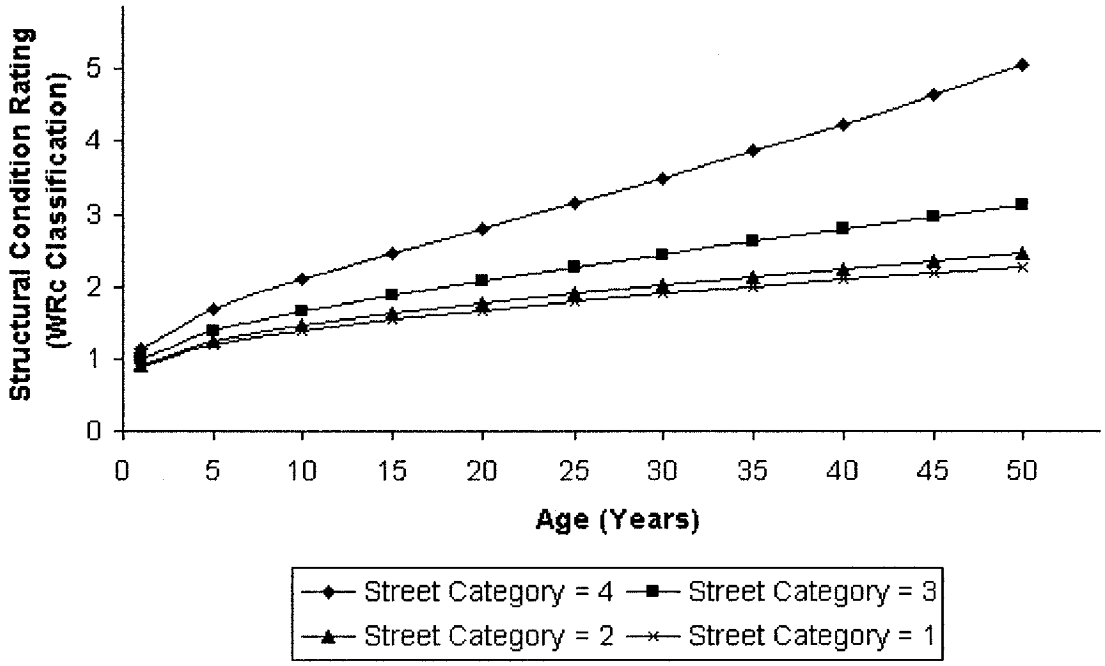


Figure 5-20: Structural Deterioration Curves for Concrete Pipes for Average Depth, Length, Bedding and Concrete Classes (Diameter 200mm to 525mm)

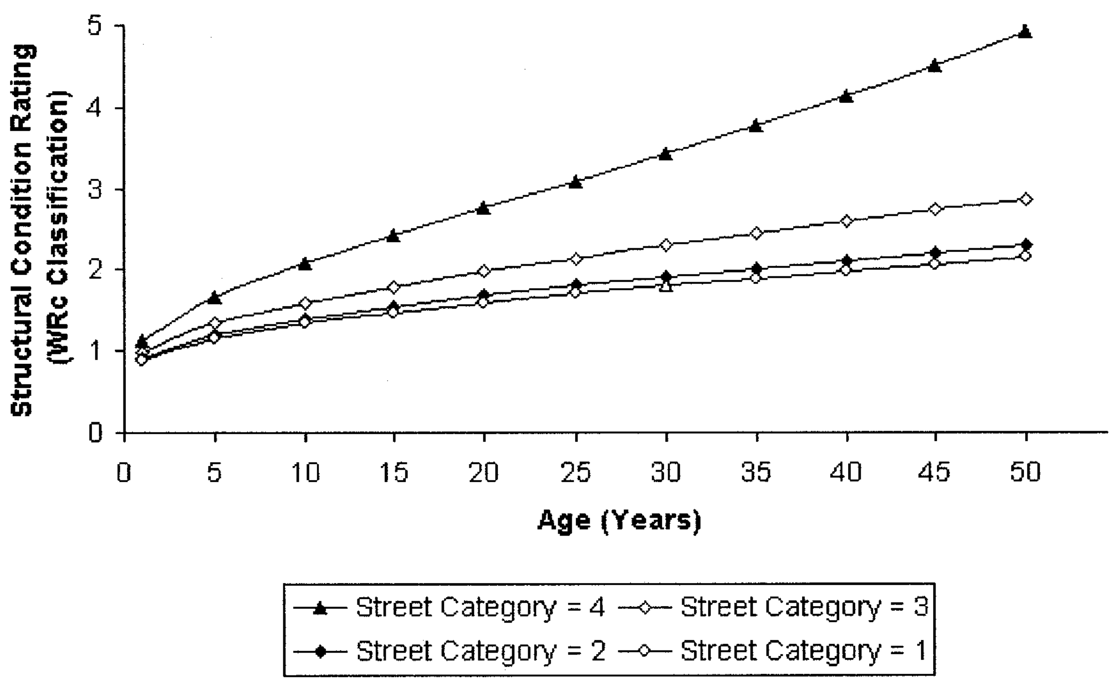


Figure 5-21: Structural Deterioration Curves for Concrete Pipes for Average Depth, Length, Bedding and Concrete Classes (Diameter 600mm to 825mm)

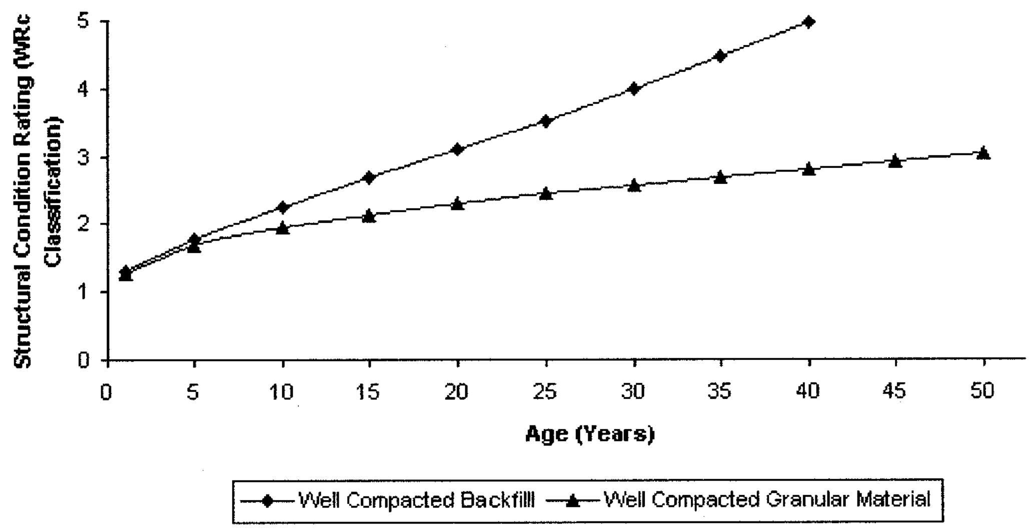


Figure 5-22: Structural Deterioration Curves for Concrete Pipes for Average Depth, Length, Diameter and Concrete Classes (Collector Street, Bedding Material Class B & C)

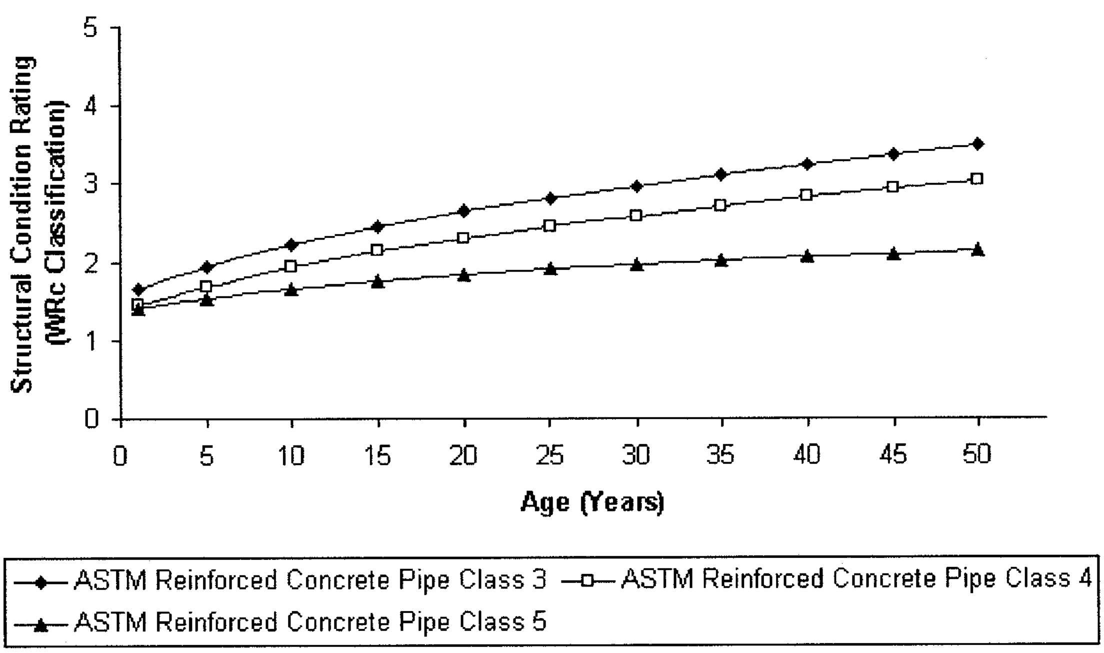


Figure 5-23: Structural Deterioration Curves for Concrete Pipes for Average Depth, Length, Diameter and Bedding Material Classes (Collector Street, Pipe Classes 3 to 5)

Figure 5-22 represents structural deterioration of concrete sewers with respect to their bedding material classes. It is observed that for average conditions, deterioration rate for concrete sewers will almost remain the same for pipes placed on class B or C bedding material during the first five years. However, after five years there is a significant difference in the rate of deterioration. This could be due to more vulnerability of displacements in weaker bedding materials. Figure 5-23 represents structural deterioration of concrete sewers with respect to their concrete classes. It is clear that the rate of deterioration is significantly less in case of high strength classes.

5.5.2 Operational Deterioration Curves

Operational deterioration curves are drawn with the help of equation 5.12. For simplicity, Manning's roughness coefficient is considered constant for all the pipe materials under consideration: concrete, asbestos cement, and PVC. The operational deterioration curves are built by varying one or two attributes at a time; while other attributes are kept constant.

Figure 5-24 represents operational deterioration of a sewer with respect to its bed slopes (length to diameter ratio ranges between 200 and 300). It is observed that operational deterioration rate is more for steeper bed slopes. This means that performance of sewers is best for a certain range of bed slopes, which can be called as optimum or critical bed slopes. These slopes are associated with the designed self cleansing velocity of sewers. Steeper slopes may cause super critical flows resulting in erosion of pipe; thus, increasing pipe's roughness. Therefore, inspection priority should be given to those sewers which have extremely steeper bed slopes for average length to diameter ratios.

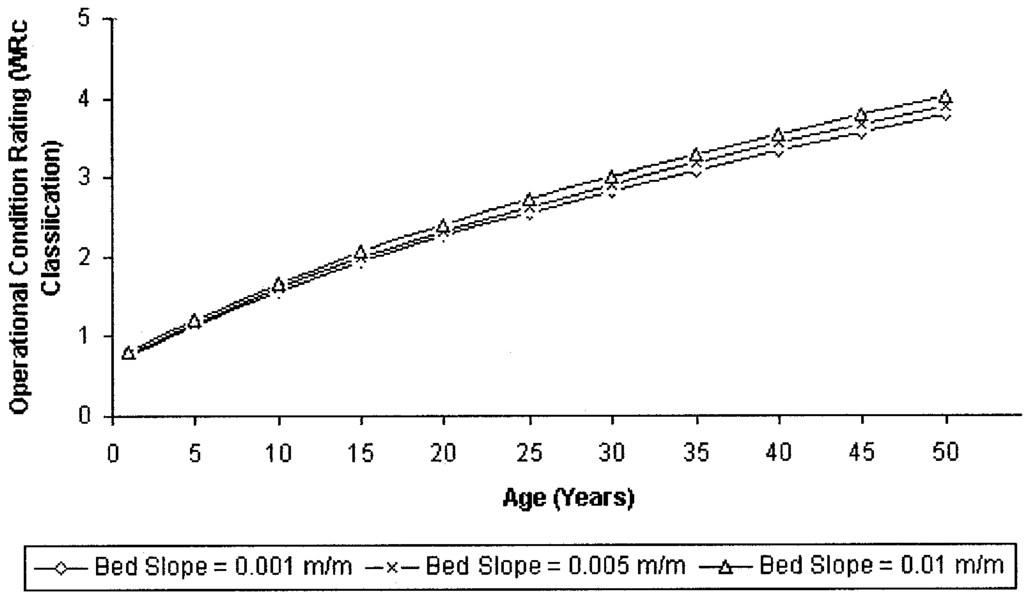


Figure 5-24: Operational Deterioration Curves for Concrete, Asbestos Cement, and PVC Pipes for Different Bed Slopes (L/D Ratio between 200 and 300)

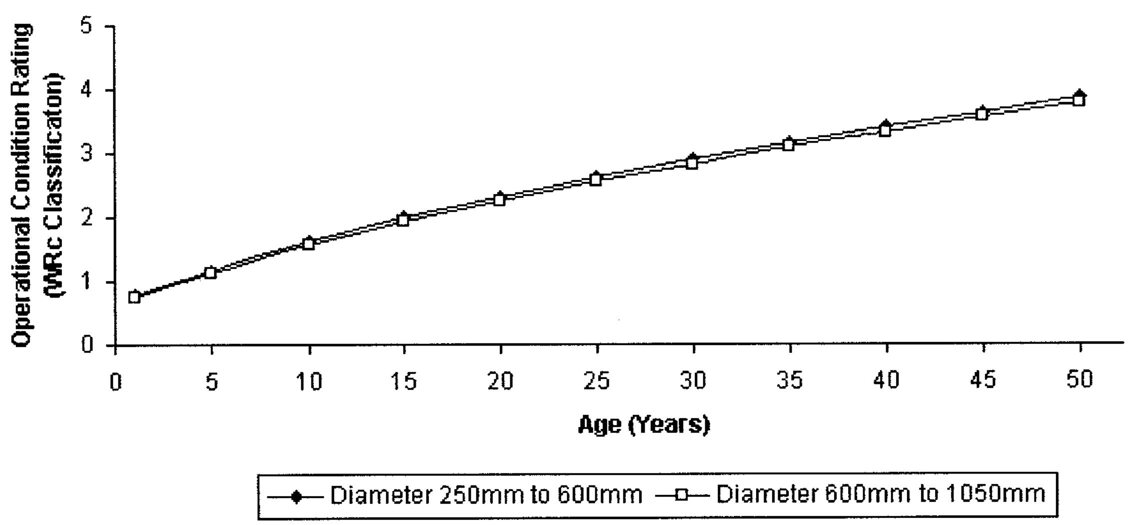


Figure 5-25: Operational Deterioration Curves for Concrete, Asbestos Cement, and PVC Pipes for Average length and Bed Slope

Figure 5-25 compares operational deterioration of trunk sewers with collectors. All the other attributes are considered to be in average condition. It is observed that the rate of operational deterioration is slightly less than that of collectors. However, the difference is not significant. As many factors have not been considered during the design of operational condition assessment model, the detailed review of operational deterioration is recommended for future research.

5.5 Summary

Regression Model building and validating methodology has been presented. The different statistical checks and diagnostics applied during the process have been discussed in detail. Three different structural condition grading models are built for different pipe materials: concrete, asbestos cement, and PVC. However, only one operational condition grading model is built for all the three pipe materials. The summary of results have been tabulated and found satisfactory. Based on the developed models, deterioration curves are drawn to better understand the structural and operational deterioration phenomena.

Chapter 6

INTEGRATION OF SEWER CONDITION ASSESSMENT PROTOCOLS

6.1 Overview

As mentioned in chapter 2, there are many condition assessment protocols which have been adopted by Canadian municipal agencies. Among them, WRc and CERIU protocols are more popular. Nevertheless, there is an urgent need of developing and adapting an integrated and unified approach towards condition assessment. As a first step towards achieving this objective, this chapter compares and analyzes both the protocols. Moreover, the chapter proposes a methodology to convert CERIU protocols into WRc and vice versa. Conversion factors have been developed through unsupervised neural network clustering in this regard. Based on the self-organizing methodology for the integration of protocols, a combined condition index (CCI) is proposed through clustering of combined condition matrix. The combined condition matrix has been developed based on WRc's approach. The methodology has been verified through CERIU's sub-committee for development of a unified condition assessment protocol and other experts.

6.2 Protocol Comparison

6.2.1 General

WRc and CERIU sewer condition assessment protocols consider the same concept of judging a defect's severity and its impact on service life by assigning a certain numerical value on a specified scale. The key difference is that the CERIU protocols are not tied to condition assessment but rather to the standardisation of the terms or language used to describe what one can see when doing an inspection. The CERIU's manual for standardization of observations was and still is a manual for sewer condition classification, near equivalent to WRc manual of sewer condition classification and does not stand to address the equivalent of the WRc's SRM (Sewerage Rehabilitation Manual, 2004).

For calculating a pipe's condition, sewer defects in the pipe need to be ranked in some order of severity. In this context, WRc SRM assigns different deduct values to each defect in a pipe, and on the basis of these deduct values it assigns a condition class to the whole pipe. On the contrary, CERIU directly assigns a condition class to each defect in a pipe, and it does not calculate any condition class for the whole pipe segment. Therefore, the results obtained by the application of CERIU codes are complex and need more time for analysis. Nevertheless, CERIU divides cross-sectional area of a sewer pipe in a more convenient way (Figure 2-6) resulting in easy judgment of flow depth and other observations inside a pipe. Moreover, CERIU addresses the problems caused by infiltration and service connections in more detail than WRc.

The main drawback of CERIU codes, could be minimized by comparing condition classes developed by CERIU with the deduct values developed by WRc for each sewer defect. Therefore, the first step should be to develop tables for comparison between CERIU condition assessment classes for sewer defects and their corresponding WRc deduct values. This methodology is illustrated below and deals with structural and operational defects separately.

6.2.2 Structural Condition Protocols

Table 6-1 shows a comparison between the codes for structural defects. It is developed by taking the three criticality conditions (light, medium and severe) into account.

Table 6-1: Comparison of WRc Structural Deduct Values with CERIU Condition Classes

Defects	Criticality Level	Unit	CERIU Condition Class	WRc Deduct Values
Joint Opening	Light	Per Joint	1	0.1
	Medium	Per Joint	2 to 3	0.2
	Severe	Per Joint	4	2
Joint Displacement	Light	Per Joint	2	0.1
	Medium	Per Joint	3 to 4	0.5
	Severe	Per Joint	5	2
Circumferential Crack	Light	Per Crack	1 to 2	1
	Medium	Per Crack	3	-
	Severe	Per Crack	4 to 5	8
Longitudinal Cracks	Light	Per Crack	1 to 2	2
	Medium	Per Crack	3	-
	Severe	Per Crack	4 to 5	15
Multiple Crack	Light	Each	1 to 2	5
	Medium	Each	3	-
	Severe	Each	4 to 5	40
Deformation	Light	Each	-	10
	Medium	Each	4	75
	Severe	Each	5	165
Hole	Light	Each	3	80
	Medium	Each	4	-
	Severe	Each	5	165

Table 6-1 shows that sometimes both the codes do not define either a certain condition class or a deduct value for a particular defect's criticality level. However, both codes usually define a condition class or a deduct value for every defect when the criticality level is severe. Therefore, it is essential to take into account these scores for critical defects when comparing the two grading systems.

6.2.3 Operational Condition Protocols

In the same context, a comparison between the two codes for some common operational defects can be done.

Table 6-2: Comparison between WRc Operational Deduct Values with CERIU Classes

Operational Defect		Criticality Level	CERIU Condition Class	WRc Deduct Values
CERIU Pronunciation	**WRc Pronunciation**			
Roots	**Roots**	Light	1 to 2	2
		Medium	3	4
		Severe	4 to 5	10
Deposits	**Encrustation**	Light	1 to 2	1
		Medium	3	2
		Severe	4 to 5	5
Grease/ Visible Material	**Debris**	Light	1 to 2	1
		Medium	3	5
		Severe	4 to 5	10
Obstructing Object	**Obstruction**	Light	1	-
		Medium	3	-
		Severe	5	10
Infiltration	**Infiltration**	Light	1 to 2	-
		Medium	3	-
		Severe	4 to 5	-

Similarly, Table 6-2 has been developed by taken into account the three conditions of criticality i.e. light, medium and severe, which shows that sometimes both the codes do not define either a certain condition class or a deduct value for a particular defect's criticality level. However, both codes usually define a condition class or a deduct value for every defect when the criticality level is severe.

6.3 Proposed Modification Methodology

As mentioned, the main drawback of CERIU protocols is that CERIU directly assigns a condition class to each defect in a sewer pipe, and it does not calculate any condition class for the whole pipe segment. A methodology for the solution to this problem was adapted and has already been overviewed in Figure 3-4. The detailed stepwise analysis is performed by adapting the methodology and is described below.

6.3.1 Defect Ranking

WRc assigns different peak deduct values for different defects. That means some defects have more weights than others for determining the overall condition of a pipe. For example, longitudinal crack has a maximum deduct value of 15 per crack as compare to 40 for multiple crack. As a consequence, it can be said that a multiple crack affects the overall condition of a pipe 2.67 times more than a longitudinal crack. In this context, all defects can be ranked on the basis of their contribution towards the overall condition of a pipe. This methodology is illustrated below for structural and operational defects.

(i) Structural Defects

For structural condition assessment, WRc assigns a maximum deduct value of 165 for a defect. Therefore, all structural defects can be ranked according to the above mentioned value and are shown in the table 6-3. The percentage ranking weights obtained from Table 6-3 for common structural defects have been plotted in Figure 6-1 for comparison of severity.

Table 6-3: WRc Ranking Weights for Common Structural Defects

Structural Defect	WRc Maximum Deduct Value	Ranking Weights (Col 2/165)	%age Ranking Weights (Col 3*100)
Joint Opening	2	0.01	1.21
Joint Displacement	2	0.01	1.21
Circum. Cracks	8	0.05	4.85
Long. Cracks	15	0.09	9.09
Multi. Cracks	40	0.24	24.24
Deformation	165	1.00	100.00
Hole	165	1	100.00

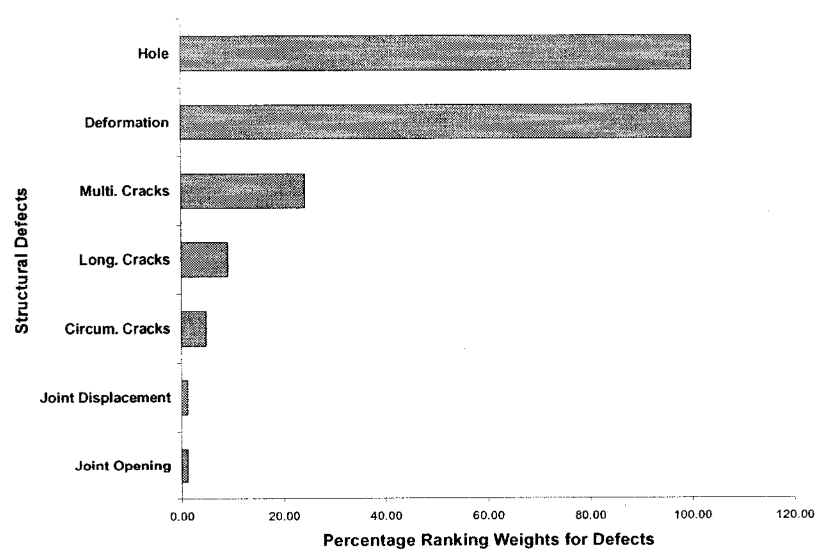


Figure 6-1: Severity Contribution in Overall Structural Condition of a Pipe by Common Structural Defects (WRc Approach)

(iii) Operational Defects:

For operational condition assessment, WRc assigns a maximum deduct value of 10 for a defect. Therefore, operational defects can be ranked accordingly and are shown in Table 6-4.

Table 6-4: WRc Ranking Weights for Common Operational Defects

Operational Defect	WRc Maximum Deduct Value	Ranking Weights (Col 2/10)	%age Ranking Weights (Col 3*100)
Roots	10	1.00	100.00
Encrustation	5	0.50	50.00
Debris	10	1.00	100.00
Obstruction	10	1.00	100.00

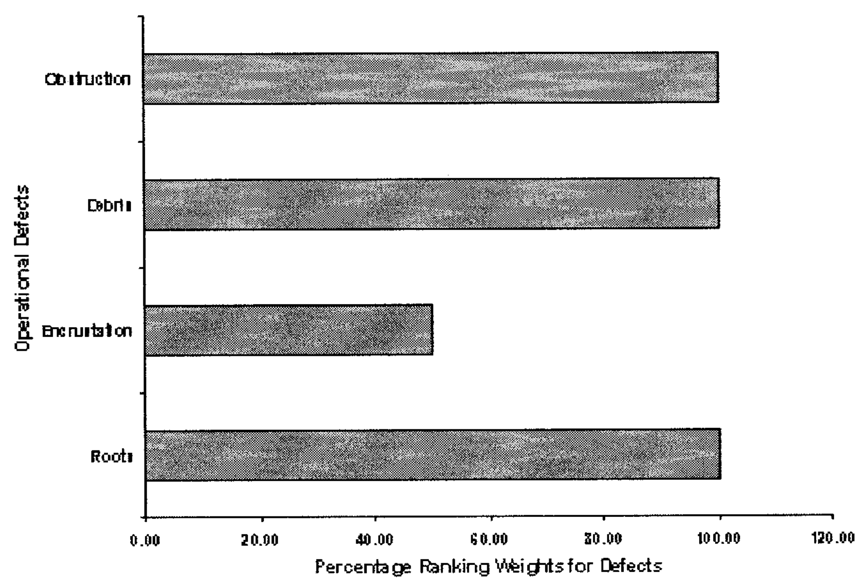


Figure 6-2: Severity Contribution in Overall Operational Condition of a Pipe by Common Operational Defects (WRc Approach)

Similarly, the percentage ranking weights obtained from Table 6-4 for some common operational defects have been plotted in Figure 6-2 for comparison of severity.

6.3.2 Assigning Transformed Deduct Values for CERIU Classifications

The deduct values, or weights, for defects are assigned according to the condition assessment protocol used and they determine the impact of defects on the service life and performance of a sewer pipe segment. Deduct values for defects for any one protocol should be assigned in a consistent manner (Rahaman et al 2004). Therefore, care should be taken into account while proposing deduct values for CERIU classification, so that the deduct values should be consistent and compatible with other codes. The proposed methodology assigns deduct values for CERIU classification by multiplying WRc severity ranking weight for a particular defect with the specified CERIU class for the same defect. This methodology can be applied to structural and operational defects as follows:

(i) Structural Defects

Table 6-5 presents the obtained deduct values for CERIU classification for some common structural defects. These values have been transformed from their respective ranking weights obtained from WRc classification. For some defects, CERIU does not specify all five condition classes. For simplicity, these classes have been ignored while transforming deduct values. Therefore, the missing values will be considered as zero during the self-organizing neural network clustering process.

Table 6-5: Transformed Deduct Values for CERIU Condition Classes for Common Structural Defects

Structural Defect	WRc Ranking Weights	Transformed Deduct Values for CERIU Condition Classes (Ranking Weight * CERIU Condition Class)				
		Class 1	Class 2	Class 3	Class 4	Class 5
Joint Opening	0.01	0.01	0.02	0.03	0.04	–
Joint Displacement	0.01	–	0.02	0.03	0.04	0.05
Circum. Crack	0.05	0.05	0.1	0.15	0.2	0.25
Long. Crack	0.09	0.09	0.18	0.27	0.36	0.45
Multi. Crack	0.24	0.24	0.48	0.72	0.96	1.2
Deformation	1	–	–	–	4	5
Hole	1	–	–	3	4	5

Table 6-6: Transformed Deduct Values for CERIU Condition Classes for Common Operational Defects

Operational Defect	WRc Ranking Weights	Transformed Deduct Values for CERIU Condition Classes (Ranking Weight * CERIU Condition Class)				
		Class 1	Class 2	Class 3	Class 4	Class 5
Roots	1	1	2	3	4	5
Deposits	0.5	0.5	1	1.5	2	2.5
Grease	1	1	2	3	4	5
Visible Material	1	–	2	3	4	5
Obstruction	1	1	–	3	–	5

(ii) Operational Defects

Similarly, deduct values for CERIU classification for some common operational defects are shown in Table 6-6, and the values have been transformed from their respective ranking weights obtained from WRc classification. Again for easy transformation, deduct values for missing classes have been ignored.

6.3.3 Development of Self-Organizing Maps

Overall condition class of a pipe can be calculated on peak or mean deduct values methods; where peak score represents the highest deduct value in a pipe segment and mean score represents an average of the deduct values for a particular pipe segment. For simplicity, the method of peak deduct values has been adapted for developing modified CERIU classification system.

In order to develop an overall structural or operational condition grading system for CERIU classification, the obtained transformed deduct values need to be grouped or clustered into five categories of condition classes. For this purpose, self-organizing maps are developed for grouping structural and operational transformed deduct values through unsupervised neural network applications.

The clustering or groupings of deduct values for structural and operational grades are done separately. Computer software Neuroshell ® is used for this purpose. The key topology of SOM is shown in Figure 6-3. The input layer consists of transformed deduct values, and the output layer represents the topology of five groups for these values. These groups are obtained separately for structural and operational deduct values which are described below.

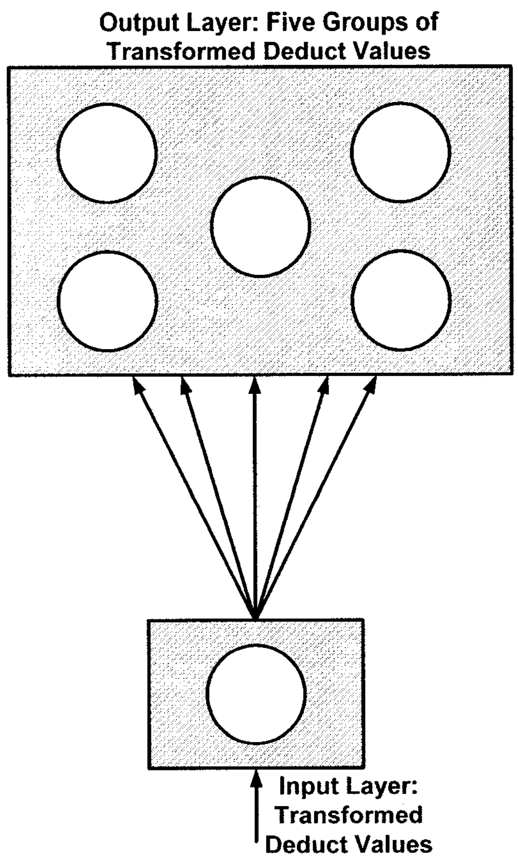


Figure 6-3: Topology of Applied Kohonen's Self-Organizing Map

(ii) Structural Defects

The transformed deduct values for structural defects (Table 6-5) are taken as the input values for the development of self-organizing map. The total number of deduct values are 28. Initially, these 28 values are entered in the program. The input layer is trained from 500 to 500,000 epochs for generating the desired five category output. The initial learning rate was 0.5 and neighbourhood size was taken as 4. During the process of training, the learning rate and neighbourhood size eventually decreased to 0.000001 and 0 respectively (Figure 6-4). However, the five categorical desired outputs are not achieved. This process is repeated for different learning epochs; nevertheless, the final results are the same i.e. 4 categorical outputs and one unused output category (Figure 6-4).

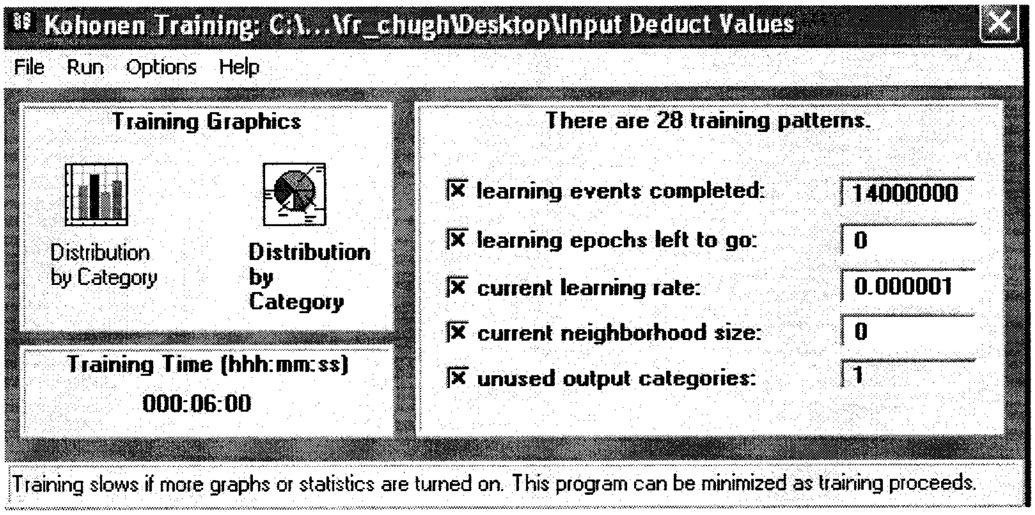


Figure 6-4: Initial Results after Training Showing Four Categorical Outputs (One Unused Output)

As the neural network models are always data hungry; the input data set points are gradually increased to get the desired five categorical outputs. Therefore, in the second step, the 28 input deduct values are entered twice, resulting in total 56 input dataset points. Further, all the possible scenarios are adapted regarding pattern selection, distance, learning epochs, neighbourhood size, etc. The input data set points are increased by adding one more batch of deduct values (28 values) every time. This process is repeated until consistent results of five categorical outputs are achieved. It was found out for optimum results, deduct values are entered in the program 14 times simultaneously resulting in total 392 input data set points. The deduct values are entered in the program 15, 16 and 17 times to check any change in the categorical outputs, but no significant change is observed in this regard. Therefore, the results obtained with 392 data set points were considered as satisfactory, and are shown in the Figure 6-5. It also shows the class boundaries of obtained clusters. The five clusters (groups) of deduct values obtained through Kohonen self-organizing maps are shown in Table 6-7.

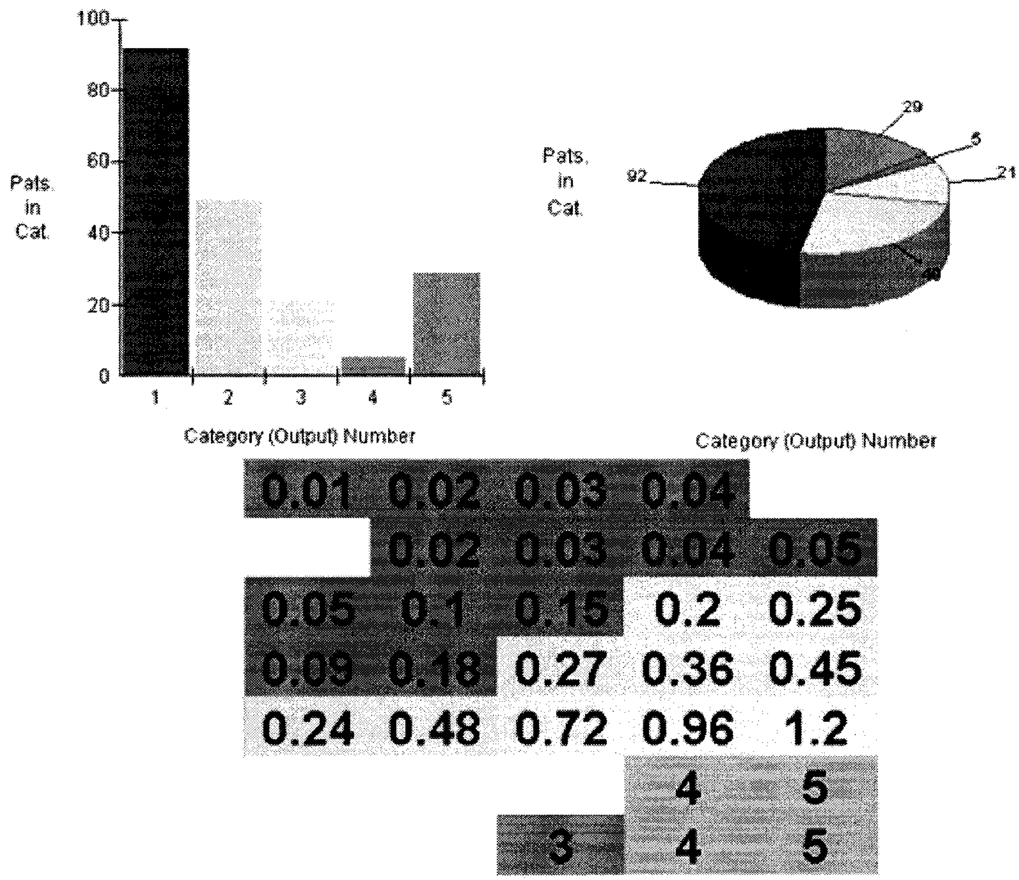


Figure 6-5: Final Categorical Output for CERIU Transformed Deduct Values for Structural Defects

Table 6-7: Group Divisions for CERIU Transformed Deduct Values for Structural Defects Obtained through Unsupervised Neural Network (Kohonen) Learning Process

SOM Groups	**Transformed Deduct Values**
Group # 1	0.01, 0.02, 0.03, 0.04, 0.05, 0.09, 0.1, 0.15, 0.18
Group # 2	0.2, 0.24, 0.25, 0.27, 0.36, 0.45, 0.48
Group # 3	0.72, 0.96, 1.2
Group # 4	3
Group # 5	4, 5

(ii) Operational Defects

The same methodology for grouping operational deduct values is adapted. Total numbers of dataset points as obtained from Table 6-6 are 22. For obtaining consistent results, the total number of input values are increased and decreased in the similar fashion. It is found out that when the deduct values are entered 9 times simultaneously, resulting in 198 input data points, the results are more consistent with well defined output categories. These results are chosen as the best possible scenario (Figure 6-6). Figure 6-6 shows the class boundaries of obtained clusters. The obtained categorical divisions are illustrated in Table 6-8.

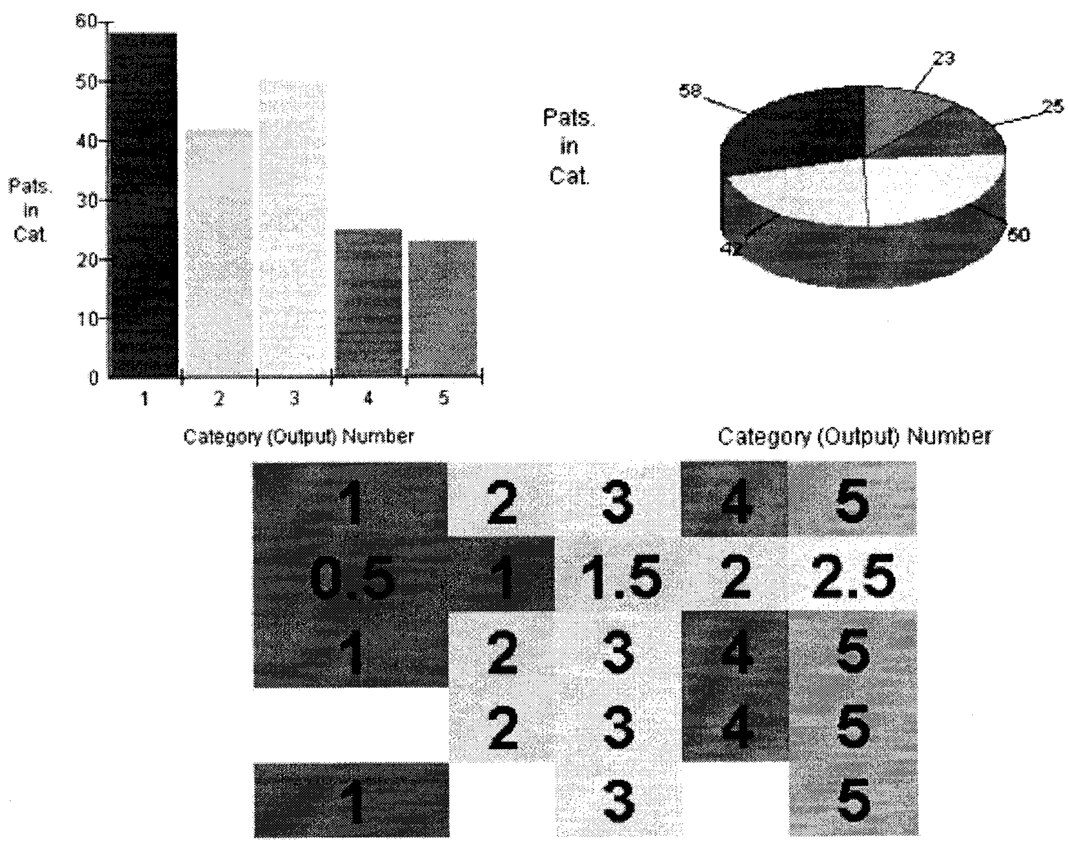


Figure 6-6: Final Categorical Output for CERIU Transformed Deduct Values for Operational Defects through Kohonen Learning Process

Table 6-8: Group Divisions for CERIU Transformed Deduct Values for Operational Defects

SOM Groups	Transformed Deduct Values
Group # 1	0.5, 1
Group # 2	1.5, 2
Group # 3	2.5, 3
Group # 4	4
Group # 5	5

6.3.4 Proposed Modification in CERIU Protocols

The class boundaries for each group or cluster can be easily defined from the self-organizing map's results which are tabulated in Tables 6-7 and 6-8. For example (Table 6-7), group no 1 has a minimum deduct value of 0.01 and maximum of 0.18. Therefore, the peak value for this group is 0.18. Further, this maximum value is less than all values of group 2 and so on. This shows a holistic picture of developed self-organized condition classes for both structural and operational defects. These condition classes are tabulated separately for structural and operational conditions in Table 6-9.

Table 6-9: Holistic CERIU Structural and Operational Condition Classes for Sewers

Proposed Overall CERIU Structural and Operational Condition Class	Peak Structural Transformed Deduct Values	Peak Operational Transformed Deduct Values
1	≤ 0.18	≤ 1.00
2	$0.19 - 0.48$	$1.1 - 2$
3	$0.49 - 1.2$	$2.1 - 3$
4	$1.21 - 3$	$3.1 - 4$
5	> 3	> 4

The obtained peak transformed deduct values for CERIU classifications for structural and operational defects are compared with their corresponding WRc deduct values for comparison. The results are shown in Tables 6-10and 6-11 respectively.

Table 6-10: Comparison between Obtained Modified CERIU and WRc Structural Condition Classes According to Peak deduct Values

Overall WRC/CERIU Structural Condition Class for a Pipe	Modified CERIU Peak Structural Deduct Value Found in a Pipe Segment	WRc Peak Structural Deduct Value Found in a Pipe Segment
1	$\leq$ 0.18	< 10
2	0.19 – 0.48	10-39
3	0.49 – 1.2	40-79
4	1.21 – 3	80-164
5	> 3	165 & above

Table 6-11: Comparison between Obtained Modified CERIU and WRc Operational Condition Classes According to Peak deduct Values

Overall WRC/CERIU Operational Condition Class for a Pipe	Modified CERIU Peak Operational Deduct Value Found in a Pipe Segment	WRc Peak Operational Deduct Value Found in a Pipe Segment
1	$\leq$ 1.00	< 1
2	1.1 – 2	1 – 1.9
3	2.1 – 3	2 – 4.9
4	3.1 – 4	5 – 9.9
5	> 4	> 10

6.5 Feedbacks from Experts on the Developed Methodology

CERIU has setup a sub-committee which is going to prepare a round-table of all the stakeholders and other people interested in the development of a unified condition assessment protocol. In short, there seems to be at three camps, one pro-CERIU, one pro-WRc and some on the fringe who support the approach of integrating both approaches. Before proceeding with any approach, the sub-committee at large must give recommendations and then only will CERIU be able to trace a clear development path (Bergeron, 2006).

The proposed methodology was presented to the CERIU sub-committee during its meeting in its meeting in October 2006. The main comments given by the committee on the research (Bergeron, 2006) are described below:

- The sub-committee admired the work and considered it as very interesting and promising
- The sub-committee acknowledged that the question of CERIU protocols had been lingering for a long time now within the community, and there was an urgent need to address it.
- The sub-committee agreed that the proposed conversion factors would be helpful in providing a documented link between CERIU protocols and WRc condition assessment system.
- Some of the committee's members expressed their concerns about the inner working of the unsupervised neural networks.

- Another question was raised by a member that how the network addressed the various observations noted within CERIU's manual that were not even considered by WRc.
- After consultation, the sub-committee recommended that future research should be carried out for finding solutions to both the raised question

6.6 Combined Condition Index (CCI) for Sewers

In previous sections, discussion is mostly concerned with the comparison and conversion of sewer condition assessment protocols into one another for the integration of condition assessment protocols. The very next step is to develop a combined condition rating system for sewers which takes into account both structural and hydraulic conditions simultaneously. This section presents a methodology of clustering structural and operational condition grades through unsupervised neural network learning into five well defined categories. A combined condition matrix has been developed based on WRc protocols, and has been clustered through Kohonen's self-organizing procedure for developing a combined condition index (CCI) for sewers.

6.6.1 Combined Condition Matrix

As described, a sewer's existing condition is usually defined in two ways: structural condition and operational condition. Generally, a condition rating scale varies from 1 to 5, where 1 represents the good condition and 5 represents the worst case scenario. A question arises over here for municipal managers that what would be the condition of a sewer which would not be following the path of balanced deterioration? For example, if a pipe has structural condition rating 1 and operational condition rating 5 according to a certain code, what would be the criteria of judging the overall condition of that pipe?

In order to better understand it let us consider Figure 6-7. It shows a matrix of all possible combinations of structural and operational conditions for a sewer as per WRc specification; therefore, this matrix can be called as combined condition matrix. The matrix a_{ij} is a square matrix of order 5. Where, i and j represent the possible structural and operational condition ratings of a pipe respectively. It can be noticed from Figure 7-1 that

- If i = j, balanced deterioration of a pipe
- If i > j, pipe more structurally deteriorated
- If i < j, pipe more operationally deteriorated

SG \ OG	1	2	3	4	5
1	a_{11}	a_{12}	a_{13}	a_{14}	a_{15}
2	a_{21}	a_{22}	a_{23}	a_{24}	a_{25}
3	a_{31}	a_{32}	a_{33}	a_{34}	a_{35}
4	a_{41}	a_{42}	a_{43}	a_{44}	a_{45}
5	a_{51}	a_{52}	a_{53}	a_{54}	a_{55}

Legends	
OG	WRc Operational/Service Grades
SG	WRc Internal Structural Condition Grades
	Balanced Deterioration of pipe
	Pipe Structurlarliy Deteriorated
	Pipe Hydraulically Deteriorated

Figure 6-7: Sewer Pipeline Combined Condition Matrix

The matrix also shows that there are 25 possible scenarios for assigning a combined condition of a sewer. As a consequence, it shows a very complex picture to municipal managers for taking any decision.

6.6.2 Clustering of Combined Condition Matrix

The idea of clustering the combined condition matrix through unsupervised neural network is introduced in the same fashion as have been carried out for conversion of

CERIU protocols. The main objective is to generate five well defined clusters out of the 25 possible scenarios for defining overall condition classes for sewer pipes; consequently, developing combined condition index (CCI) for sewers.

The procedure adapted for generating the required clusters is almost the same as is adapted for clustering the transformed deduct values for CERIU classification. The same software Neuroshell ® is used in this regard, and the same topology of Kohonen's self-organizing map is applied to this problem as shown in the Figure 6-3.

Data obtained from the municipality of Niagara Falls is chosen for clustering purposes. Total 966 data set points are available, which show the required description of a pipe's structural and operational condition rating. All these values are taken as the input values for the development of self-organizing map.

Initially, objectives are defined to divide these 966 data set points into 5 clusters. The input layer is trained from 500 to 500,000 epochs for generating the desired five category output.

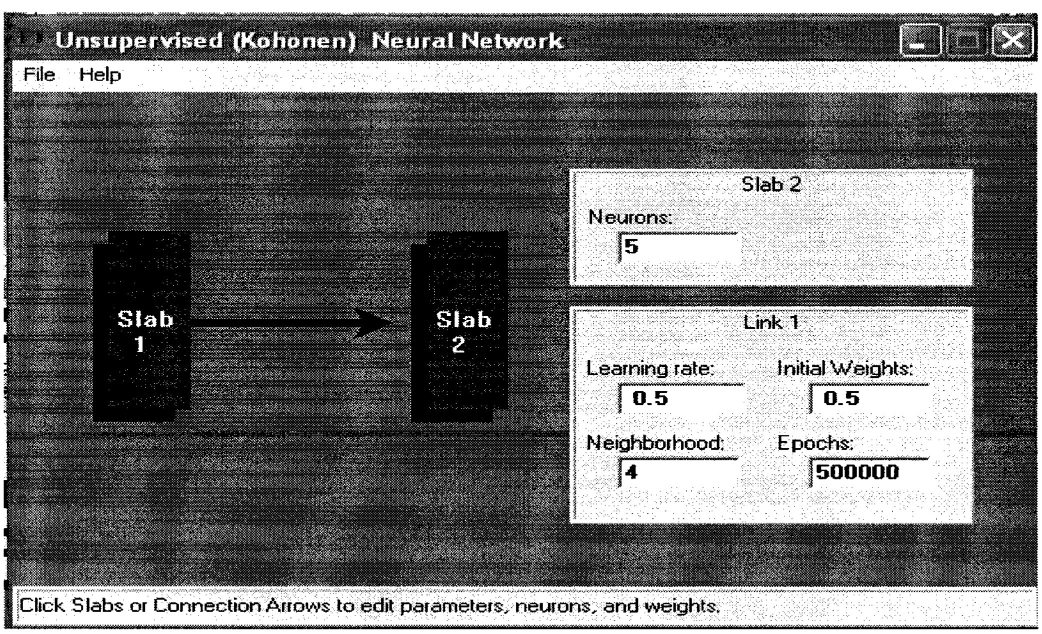


Figure 6-8: Output Layer Design for Neurons and Neighbourhood Size

The initial learning rate is 0.5 and neighbourhood size is taken as 4 (Figure 6-8). The output layer design for neurons is set at 5 neurons as five clusters are desired. Furthermore, the pattern selection criterion for clusters is set at random and the Euclidean distance is used to measure the distance between the clusters and is shown in Figure 6-9.

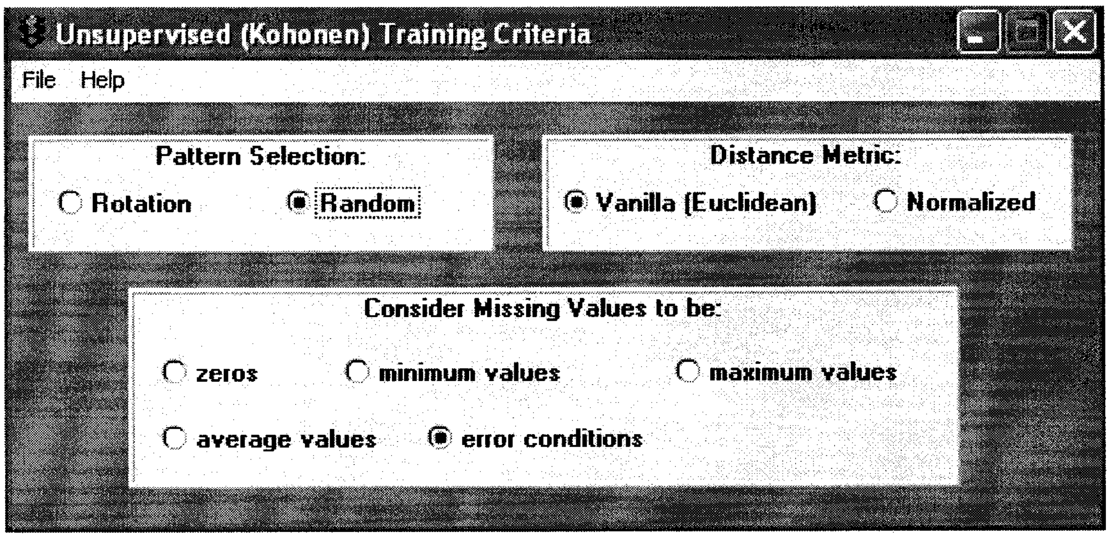


Figure 6-9: Pattern and Distance selection Parameters for Clusters

During the process of training, the learning rate and neighbourhood size eventually decrease to minimum possible value i.e. 0.000001 and 0 respectively. However, the groups or clusters obtained are not according to basic logics.

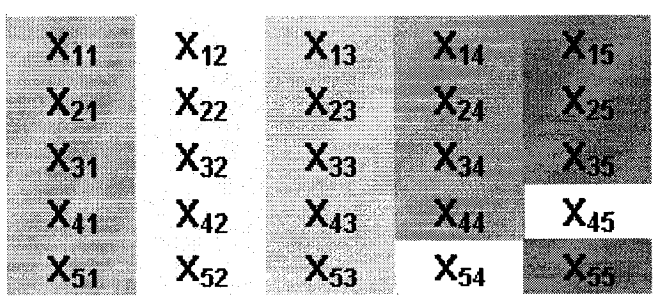


Figure 6-10: Achieved Five Clusters through Niagara Falls Data

Figure 6-10 shows the 5 categorical outputs for the input data in terms of combined condition matrix. The entries without any colour show the missing possible scenarios from the collected data; that means the collected data have 23 combinations of different structural and operational conditions out of 25 possibilities.

Table 6-12: Class Boundaries of Obtained Clusters

Cluster Number	Structural Condition Rating	Operational Condition Rating
1	1 to 5	1
2	1 to 5	2
3	1 to 5	3
4	1 to 5	4
5	1 to 5	5

Table 6-12 presents the obtained class boundaries for the desired five clusters which have been shown in Figure 6-10. It is clear that all the five clusters have been produced by considering more weights in operational condition ratings and no importance has been given to structural condition. Therefore, the obtained clusters should be rejected. Consequently, for achieving desired results, the same procedure is adapted to increase the data set points as have been done in deduct value clustering. However, instead of increasing data set points, desired outputs are gradually increased to cluster the data into maximum possible groups. As real world data is used for clustering purposes, it is not recommended to change the collected set of observations.

Therefore, in the second attempt, the networks are set to generate 6 output clusters. Similar procedures are adapted as described above, and the obtained results are analyzed. This procedure is repeated with increasing the desired output categories. It is found out that the maximum groups in which the collected data could be clustered are 9.

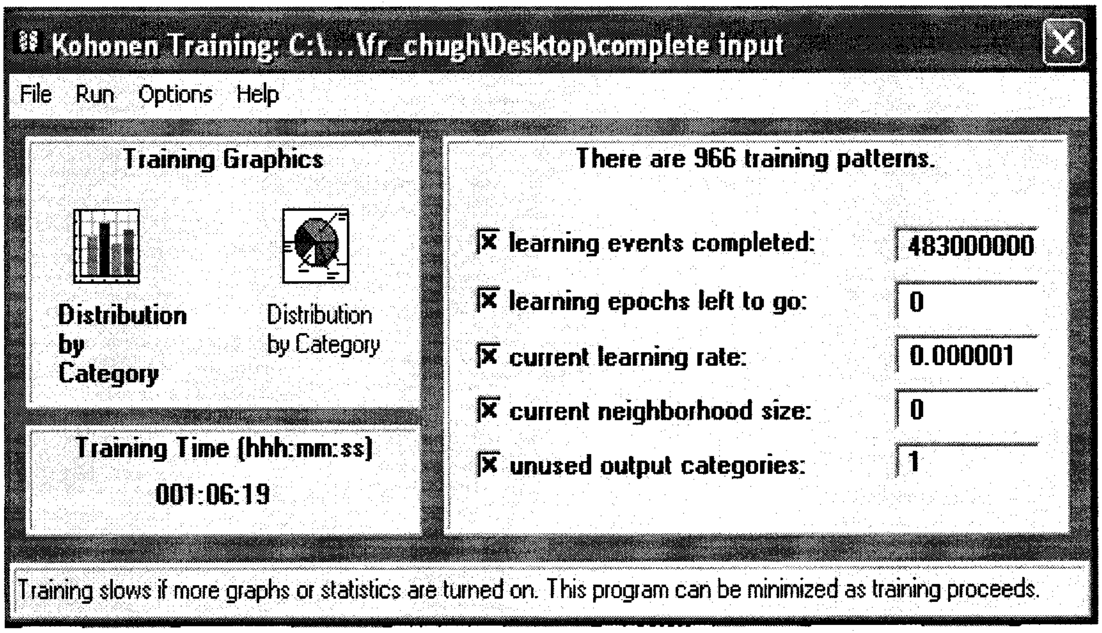


Figure 6-11: Final Outcome of Neighbourhood size and Learning Rate

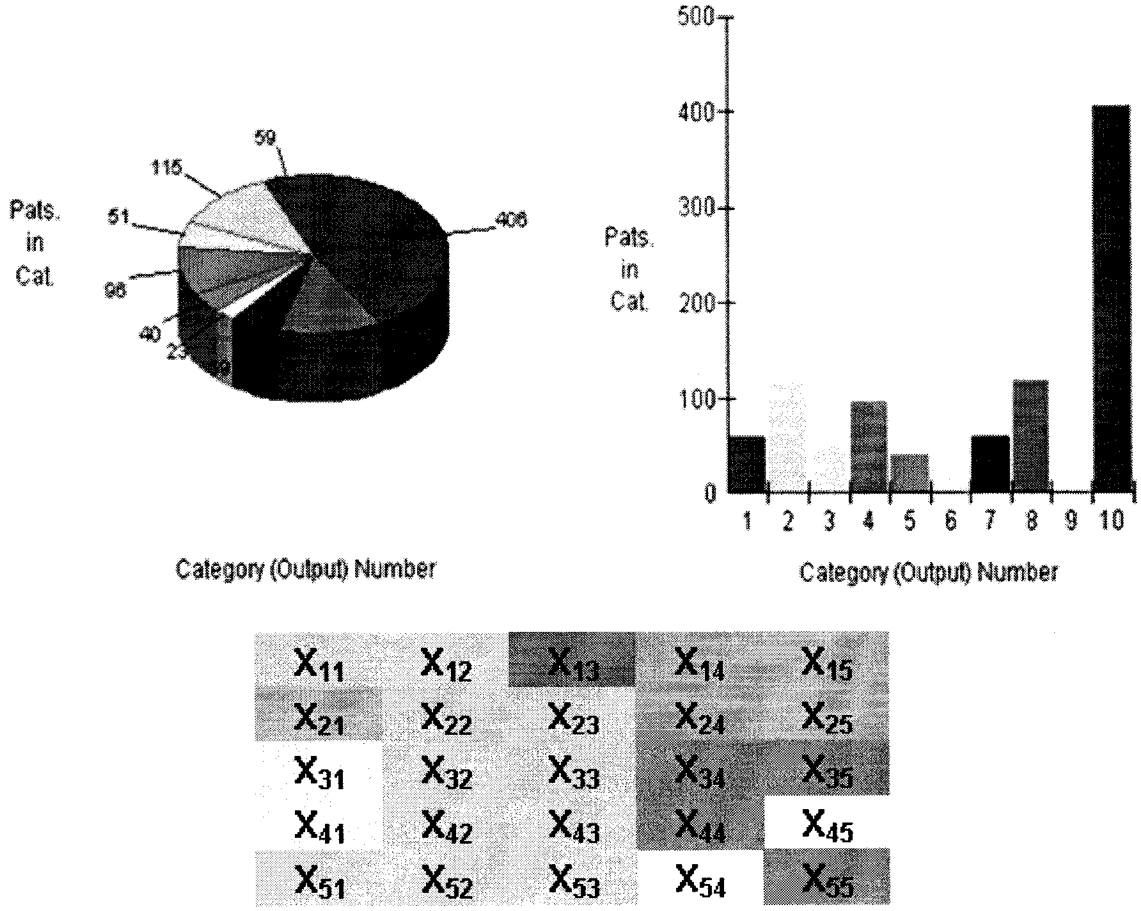


Figure 6-12: Generated Groups or Clusters: Categorical Output Summary

Figure 6-11 shows the final outcome of neighbourhood size and learning rate for 10 desired output categories. It shows that the network is not able to produce 10 desired clusters for the input data; therefore, the number of unused output categories is 1. The learning epochs completed for this particular trial are 500,000, and the achieved learning rate is almost 0. As a result, it is concluded that the input data could be conveniently transformed only into maximum 9 clusters. The categorical output patterns for the selected simulation are shown in Figure 6-12. The figure also illustrates the generated boundaries for groups or clusters for the combined condition matrix.

Table 6-13: Class Boundaries of Obtained 9 Clusters (Total Number of Entries = 966)

<table>
<tr><th>Cluster Number</th><th>Number of Entries</th><th>Structural Condition Rating</th><th>Operational Condition Rating</th></tr>
<tr><td>1</td><td>406</td><td>1</td><td>1 to 2</td></tr>
<tr><td>2</td><td>96</td><td>1</td><td>3</td></tr>
<tr><td>3</td><td>117</td><td>2</td><td>1</td></tr>
<tr><td>4</td><td>51</td><td>1 to 2</td><td>4</td></tr>
<tr><td>5</td><td>115</td><td>1 to 2</td><td>5</td></tr>
<tr><td rowspan="2">6</td><td rowspan="2">40</td><td>2</td><td>2 to 3</td></tr>
<tr><td>3</td><td>3</td></tr>
<tr><td>7</td><td>59</td><td>3 to 4</td><td>1</td></tr>
<tr><td rowspan="2">8</td><td rowspan="2">23</td><td>3 to 4</td><td>2 to 3</td></tr>
<tr><td>3 to 5</td><td>1 to 3</td></tr>
<tr><td>9</td><td>59</td><td>3 to 5</td><td>4 to 5</td></tr>
</table>

Table 6-13 shows class boundaries for the obtained clusters. The missing values have been assumed to be a part of the 9th cluster. This assumption has been made by extending the 9th cluster's boundaries for the missing elements: X_{45} and X_{54}. This extension is described in Figure 6-13.

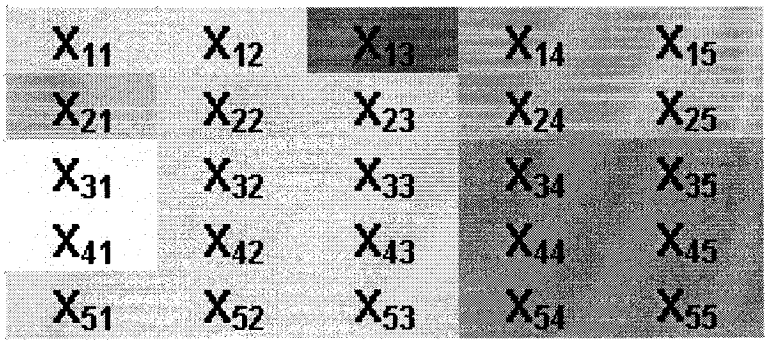


Figure 6-13: Logical Extension of 9th Cluster for Missing Elements

Table 6-14 shows that for the first 5 clusters, structural condition rating varies from 1 to 2, and operational condition varies from 1 to 5. Therefore, after combining these 5 clusters with the 6th cluster and rearranging, five clusters with definite class boundaries can be developed (Table 6-14).

Table 6-14: Integration of Obtained 9 Clusters into 5 Categories

Cluster Number	Structural Condition Rating	Operational Condition Rating
1	1 to 2	1 to 3
2	1 to 2	4 to 5
3	3 to 4	1
4	3 to 4	2 to 3
	3 to 5	1 to 3
5	3 to 5	4 to 5

Table 6-14 clearly indicates that the obtained clusters have been transformed into five well defined categories. These clusters have been sorted according to the criticality of structural condition ratings and then according to the criticality of operational condition ratings.

6.6.3 Proposed Combined Condition Index for Sewers

This research suggests that the obtained five groups (Table 6-14) through self-organizing neural networks should be considered as five different combined condition classes or indexes for sewers. Therefore, based on these clusters, a combined condition index (CCI) for sewers is proposed. Table 6-16 shows the description of the proposed combined condition index (CCI) for sewers. Each class or index has well defined boundaries for its respective structural and operational condition classes. This index is divided into 5 categories, ranging from "1" to "5", and linguistically, from "Acceptable" to "Critical". The index has been proposed by giving more weights to a sewer's structural condition for defining the rehabilitation or action requirements. In addition, criteria for assessment risk of collapse and flooding is defined for each class. The proposed remedial actions depend upon the developed risk criteria for collapse, over flow and basement flooding problems, as well as impact assessment factor. These criteria have been developed through the general guidelines provided by experts, which will be discussed in the next section. The premise for an impact factor follows that not all pipe segments have the same likelihood of failure or the same consequences of failure. This impact factor for a sewer can be calculated with the following formula (McDonald et al 2001):

$$I_W = (0.2)f_l + (0.16)f_s + (0.16)f_z + (0.16)f_d + (0.16)f_f + (0.16)f_q \qquad \text{(Equation 6.1)}$$

Where, f_l is location factor, f_s is embedment soil factor, f_z is size factor, f_d is burial depth factor, f_f is sewer function factor, and f_q is seismic factor.

Table 6-15: Impact Factors for a Sewer's Criticality (Developed from McDonald et al 2001)

Impact Factor Subtype	Description	Severity	Input Value for the Formula
Location Factor	Pipe under commercial area, 6 lanes street, or an airport etc	High	3
	Moderate pipe locations	Medium	1.5
	Pipe under a park, 2 lanes street, etc	Low	1
Soil Factor	Silty and sandy supporting soil	High	3
	Other soils	Medium	1.5
	Medium to high plastic clays	Low	1
Size Factor	Pipe diameter greater than 1800mm	High	3
	Pipe diameter between 900mm and 1800mm	Medium	1.5
	Pipe diameter less than 900mm	Low	1
Depth Factor	Burial depth greater than 10m	High	3
	Burial depth between 3m and 10m	Medium	1.5
	Burial depth less than 3 m	Low	1
Functionality Factor	Pipe entering/exiting a treatment plant	High	3
	Moderate functionality	Medium	1.5
	Collector pipes	Low	1
Seismic Factor	High seismic zone	High	3
	Medium seismic zone	Medium	1.5
	Low seismic zone	Low	1

Table 6-16: Proposed Combined Condition Index (CCI) for Sewers

Combined Condition Index (CCI)		Equivalent WRc (UK) Internal Condition Grades (ICG)		Description	Impact Rating (I_w)	Action Required
Numeric Scale	**Linguistic Scale**	**Structural Grading**	**Operational Grading**			
1	Acceptable	1 to 2	1 to 3	Acceptable overall condition	1 to 5	Routine Monitoring
2	Adequate	1 to 2	4 to 5	Overflow problems	1 to 5	Cleaning and Flushing
3	Moderate	3 to 4	1	➢ Collapse risk with no overflows (Light to medium cracks / deformation) ➢ Light to moderate service connection / construction defects	1 to 3	Low Rehabilitation Priority
					4 to 5	Medium Rehabilitation Priority
4	Poor	3 to 4, and 5	2 to 3, and 1 to 3	➢ Greater collapse risk with minimal overflow problems (Medium to severe cracks/ deformation) ➢ Medium to severe service connection / construction defects ➢ Risk of basement flooding and end user complaints	1 to 4	Medium Rehabilitation Priority
					5	High Rehabilitation Priority
5	Critical	3 to 5	4 to 5	➢ Extreme collapse risk with overflow problems (Medium to severe cracks / deformation) ➢ Severe service connections / construction defects ➢ Extreme Overflows with basement flooding and loss of property	1 to 5	Immediate Rehabilitation Priority

The impact rating is helpful in defining criticality level of a sewer. Table 6-15 shows the method of calculating the impact of some important factors for defining criticality level of a sewer. The calculated impact rating for a sewer (Equation 6.1) is integrated with other scenarios in Table 6-16 for proposing an appropriate rehabilitation or action plan. The integration of all scenarios in defining a specific class of the proposed CCI will be helpful in understanding the overall condition of sewers. For example, if the CCI is "1" for a certain sewer, it has acceptable overall condition; therefore, no particular action except routine monitoring is required. On the contrary, if a pipe has CCI value of "5", immediate rehabilitation action is proposed for that pipe. In this context, the proposed combined condition index is intended to provide a framework for municipal engineers to decide and plan maintenance and rehabilitation actions for sewer networks.

6.6.4 Verification of Proposed CCI

In order to verify the proposed combined condition index, a questionnaire is designed and has been sent to different municipal experts and consultants. The questionnaire consists of three basic questions:

1) Is the index adequate according to maintenance and rehabilitation requirements?
2) Is the index requires some revisions/reassessments in terms of assigned equivalent WRc structural and operational condition class boundaries?
3) Is the defined criteria for each category of the index is acceptable?

Four comprehensive feedbacks have been received from experts. Three out of four municipal practitioners are agreed on the point that the idea of combining structural and operational condition ratings into a single scale will help municipal engineers in prioritizing detailed inspection, maintenance, and rehabilitation operations. However, one expert suggests that structural and operational conditions should be analyzed separately. The important comments from experts have already been embraced into the description of different classes of CCI (Table 6-16). Some of the points are summarized as follows:

- Pipes collapse occurs for reasons like severe cracking or exposed aggregate due to hydrogen sulphide or chemical attack. Light, moderate or severe cracking should be considered in determine collapse risk
- There are other defect conditions that may cause overflow problems similar to a collapse pipe. These defects may be tree root intrusions, debris or encrustations etc. Depending on the severity, the required action may range from cleaning to immediate rehabilitation
- In all separated sewer systems and some combined systems, collapsed pipes may cause flooded basements instead of overflow problems. The response to flooded basements may require a higher rehabilitation priority than the priority given to overflow
- A good CCI should also cover construction defects such as sags in the pipe, protruding services, and misaligned joints etc
- Pipe rehabilitation is expensive and also depends upon available resources and budget location etc

6.6.5 Automated Conversion of Structural and Operational Ratings into CCI

To facilitate an automated conversion of a sewer's structural and operational condition observations into CCI, a regression model is designed. All the possible scenarios shown in Table 6-16 are taken as input data for the model. The response variable "CCI" is regressed against its corresponding values of predictor variables (structural and operational ratings) using the Minitab ® statistical software. The similar procedures for model development are adapted as shown in Figure 3-2.

Equation 6.2 shows the final outcome of the adapted procedure. The equation clearly indicates that CCI can be found for any sewer if its structural and operational conditions are known. The structural and operational condition ratings are according to WRc classification

$$CCI = \sqrt{\begin{array}{l}0.541 + 0.273(Structural\_Condition\_Rating) \\ + 0.37(Operational\_Condition\_Rating)\end{array}} \qquad \text{(Equation 6.2)}$$

Table 6-17: Important Statistical and Validation Diagnostics for the Regression Model

R^2 (%)	R^2_{adj} (%)	P (F)	P (t)			Validation	
			β_0	β_1	β_2	AIP	AVP
81.2	81.1	0.000	0.000	0.000	0.000	0.208	0.791

As mentioned, the equation 6.2 is verified through all necessary statistical diagnostics as well as validation checks. Some of the important statistical and validation diagnostics are shown in Table 6-16. The Box-Cox power transformation for CCI data was found out to be 0.50. Therefore the response variable, CCI, in the model has been transformed accordingly (Equation 6.2). The fitted response plane for the regression model is shown in Figure 6-14.It shows the variation in response (CCI) with the variations in predictors (structural and operational condition ratings). The developed model will facilitate a user to calculate CCI from given structural and operational condition information of a sewer.

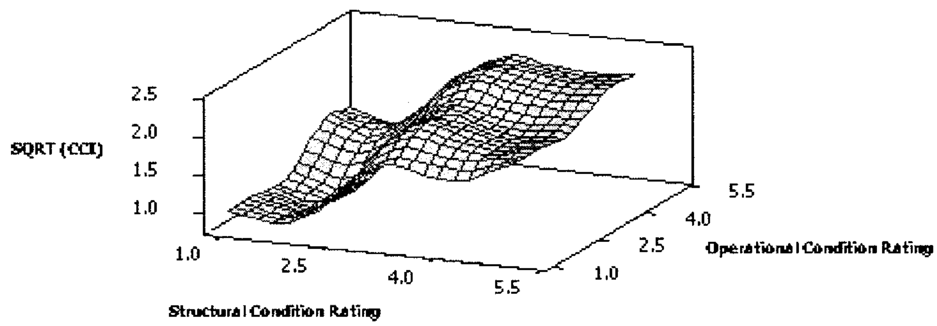


Figure 6-14: Minitab Plot of Fitted Response Plane

6.7 Summary

The proposed modification in CERIU classification system will not change the current CERIU classification; only some additional details have been proposed. Moreover, the proposed system is based on WRc approach; therefore, data conversion from WRc to CERIU and vice versa is possible. The proposed modification is a first step towards integrating sewer condition assessment protocols. Consequently, unified, standardized, and integrated sewer condition assessment protocols could be developed, as compare to using different protocols according to different local requirements.

The proposed combined condition index (CCI) for sewers has been developed through unsupervised neural network modeling, and has been modified through feedbacks from experts. The basic idea of developing CCI is to facilitate municipal agencies in prioritizing maintenance and rehabilitation works by taking into account the combined effects of structural and hydraulic conditions of sewers.

Chapter 7

WEB-BASED AUTOMATED SEWER CONDITION PREDICTION MODEL

7.1 Introduction

Recent developments in internet technologies have provided a wide range of applications which can be shared with other users. This chapter describes a methodology of developing a web-based decision support system for condition prediction of existing sewers. This system is developed to assist municipal engineers in predicting structural and operational condition ratings of sewer pipeline networks. Further, the system can also predict the combined condition index (CCI) for sewers.

7.2 Model Program

The program of the web-based condition prediction model is written in Java (version 5.0) using the JBoss application server (JBoss 4.0.1SP1). Java is a well-known object oriented programming (OOP) language for internet applications created by Sun Microsystems in 1990s. OOP is a methodology that views a program as consisting of objects that interact with each other by means of actions. Small Java applications are called Java applets. Java applets are ideal for running program applications on any computer after downloading them from a server (Savitch, 2006).

The web-based program includes procedures that link different web-pages, import and export of Excel files, calculations and interpretations, and result generation. Finally the program generates and displays the condition grading results.

7.3 Framework and Process

The web-based condition prediction tool utilizes the previously developed regression and unsupervised neural network models in predicting existing condition of sewers. The web-based model uses MS Excel in order to import and export data. The model has been designed in a simple, easy to use format to facilitate users. Figure 3-6 shows an overview of the process flowchart of the model.

The developed system requires data related to all the factors which were considered during regression and neural network modelling process. The model requires the data in a well defined Excel spreadsheet format. A user has to prepare input data as described in chapter 4. For example, the user has to define street categories as per ASCE specifications from category 1 to 4. If a user puts a street category value more than 4, the program will automatically convert it into 4. Similarly, the program processes all input data, and if it finds some data sets exceeding or decreasing from the ranges specified in chapter 4; it takes the maximum or minimum allowable values respectively to calculate the required condition rating. The framework shown in Figure 3-6 is described below:

7.3.1 User Login

The first page of the model will allow a user to register and login as shown in Figure 7-1. This web-page includes a menu bar which enables a user to proceed.

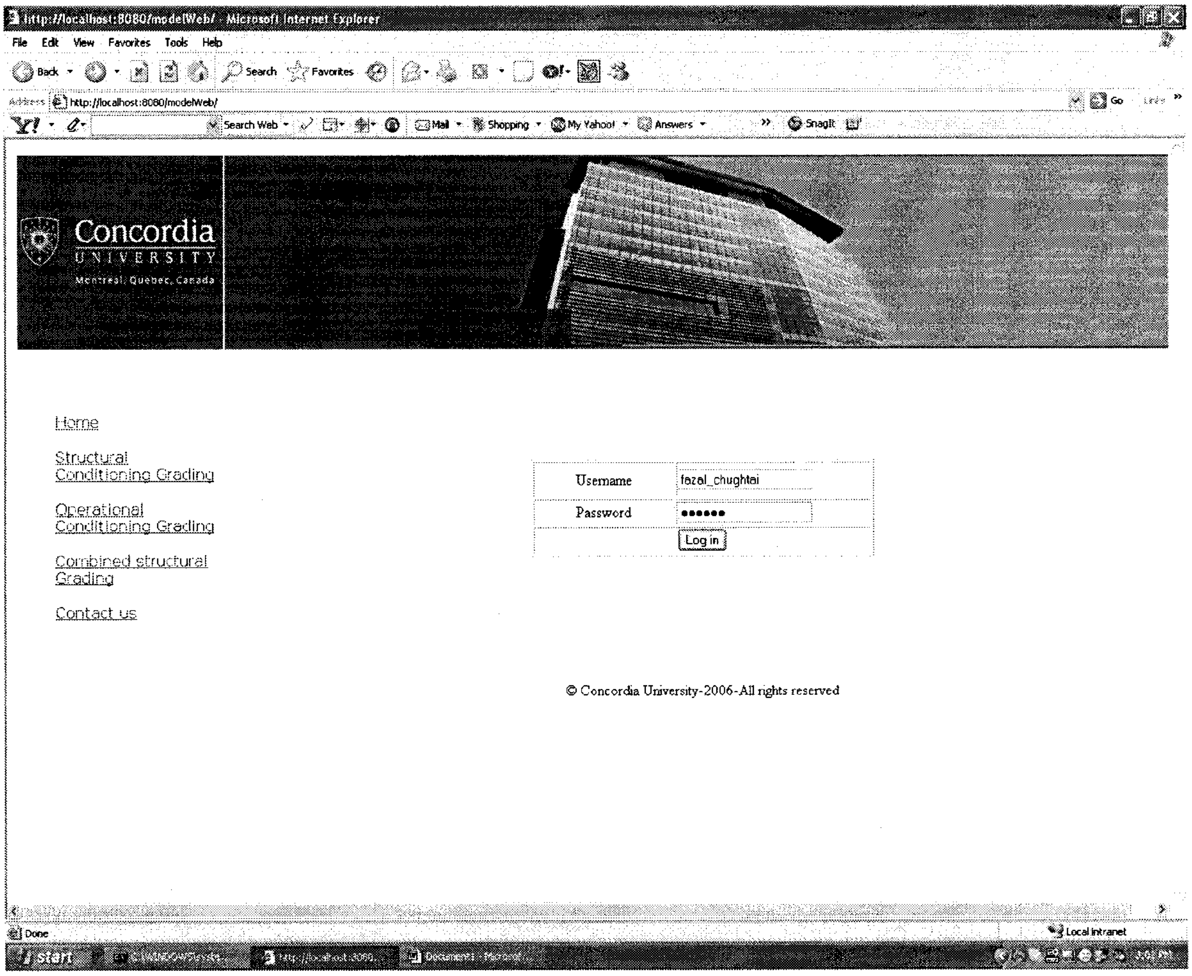


Figure 7-1: Login web-page for users

7.3.2 Selection of Condition Rating and Pipe Material

The next step for a user is to select the type of condition grading system. Three types of condition rating: structural, operational, and combined, can be obtained from this model. Therefore, user has to specify the requirements on the web-page shown in Figure 7-2. The user can only select one option at a time.

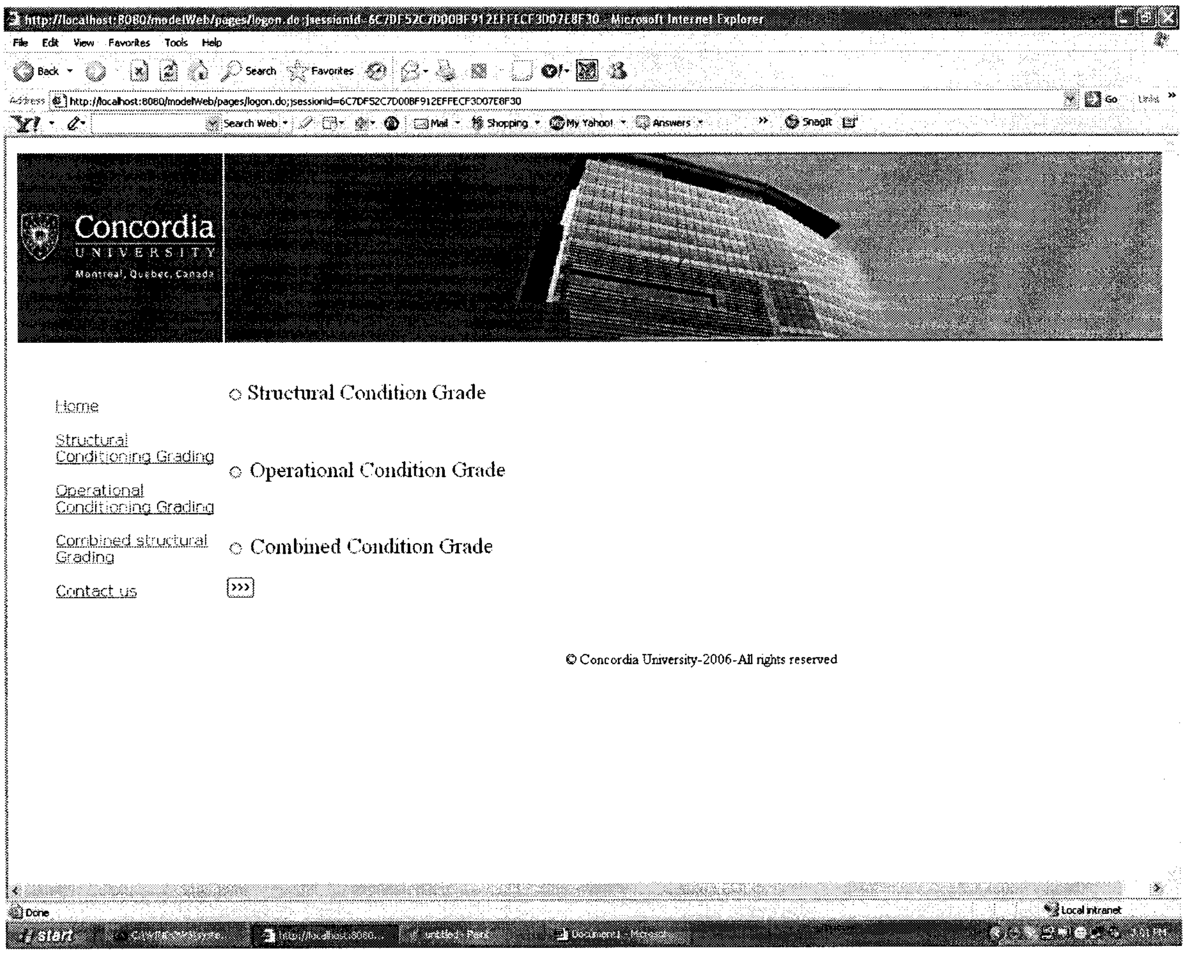


Figure 7-2: Selection of Required Condition Grading

Similarly, in the next step, the program will ask from the user to select a specific pipe material from the three options: concrete, asbestos cement, and PVC. Further, like the previous case, user can select only one pipe material type at a time. The reason for this has been explained in chapter 5 and 6, where different condition prediction models are built for different pipe materials. Therefore, the program has been designed on the similar fashion.

7.3.3 Importing Data

After the selection of pipe material, the program will ask the user to import data. The user can browse through his files and can select an appropriate "*.xls" file according to the pipe material and condition prediction requirements (Figure 7-3).

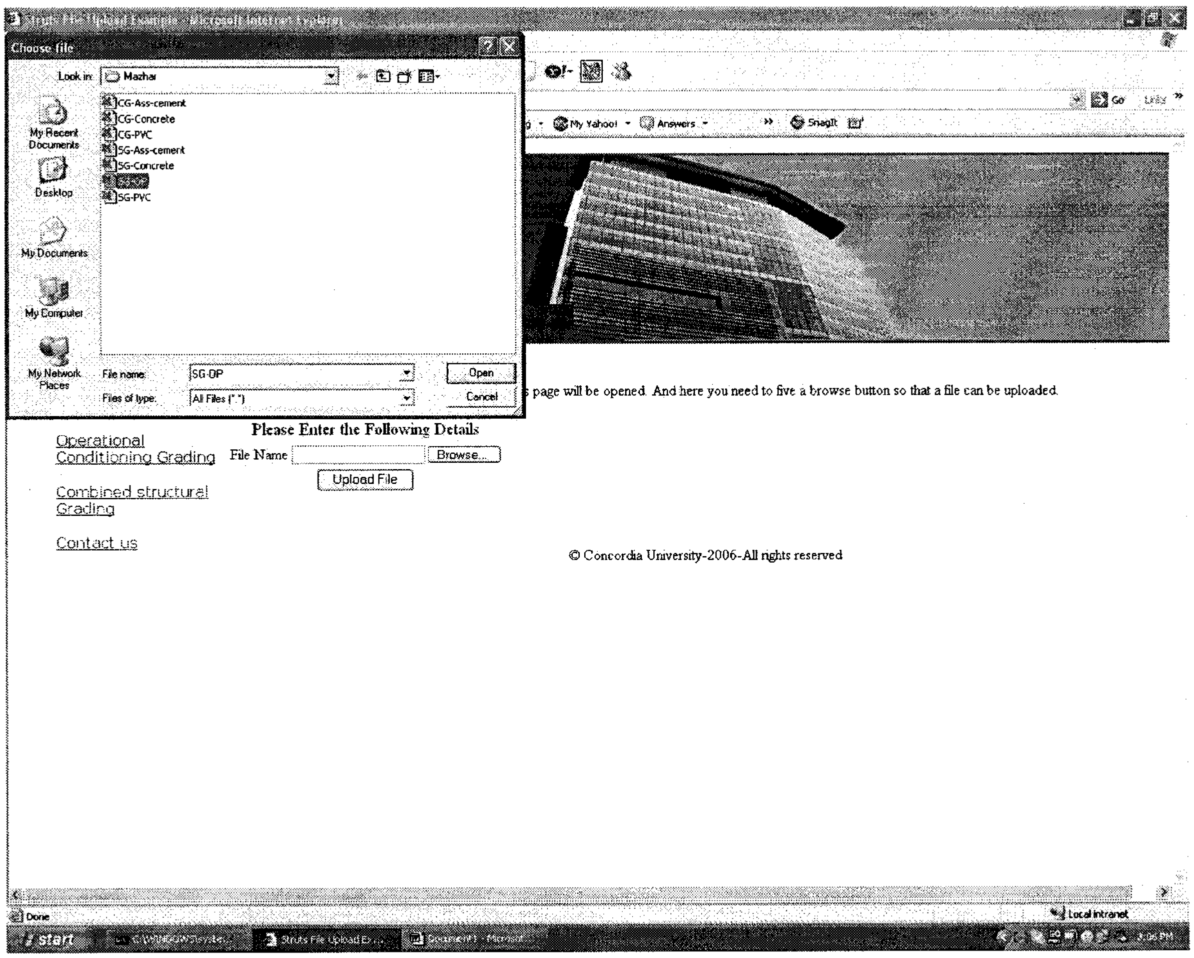
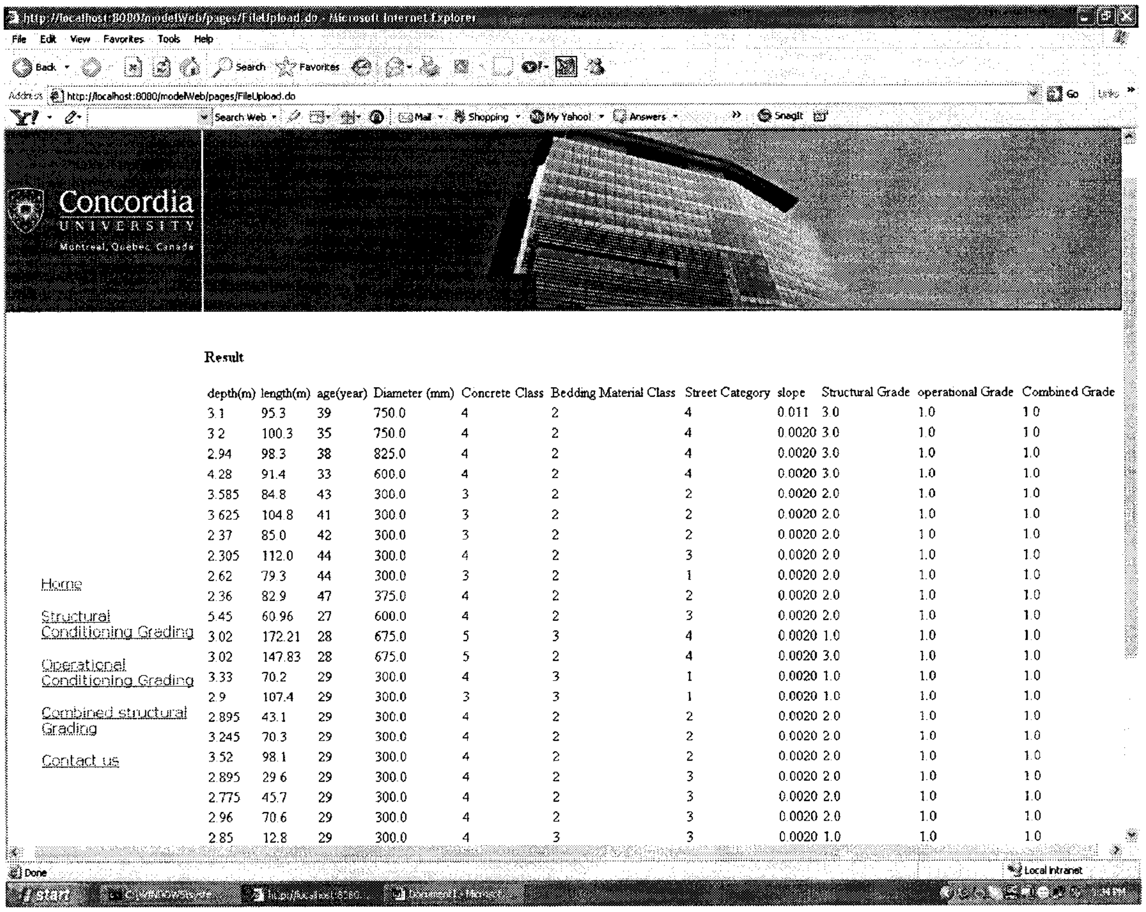


Figure 7-3: Importing Data for Model Input

7.3.4 Data Processing and Results

The program processes the imported file for calculation of required condition rating and displays final results as shown in Figure 7-4. If a file is not appropriate, the program displays an error. As already been explained, input data should be according to a well defined excel spread sheet formats. Some samples of the formats have been embedded in the program to help the user for understanding the proper function of the program.



depth(m)	length(m)	age(year)	Diameter (mm)	Concrete Class	Bedding Material Class	Street Category	slope	Structural Grade	operational Grade	Combined Grade
3.1	95.3	39	750.0	4	2	4	0.011	3.0	1.0	1.0
3.2	100.3	35	750.0	4	2	4	0.0020	3.0	1.0	1.0
2.94	98.3	38	825.0	4	2	4	0.0020	3.0	1.0	1.0
4.28	91.4	33	600.0	4	2	4	0.0020	3.0	1.0	1.0
3.585	84.8	43	300.0	3	2	2	0.0020	2.0	1.0	1.0
3.625	104.8	41	300.0	3	2	2	0.0020	2.0	1.0	1.0
2.37	85.0	42	300.0	3	2	2	0.0020	2.0	1.0	1.0
2.305	112.0	44	300.0	4	2	3	0.0020	2.0	1.0	1.0
2.62	79.3	44	300.0	3	2	1	0.0020	2.0	1.0	1.0
2.36	82.9	47	375.0	4	2	2	0.0020	2.0	1.0	1.0
5.45	60.96	27	600.0	4	2	3	0.0020	2.0	1.0	1.0
3.02	172.21	28	675.0	5	3	4	0.0020	1.0	1.0	1.0
3.02	147.83	28	675.0	5	2	4	0.0020	3.0	1.0	1.0
3.33	70.2	29	300.0	4	3	1	0.0020	1.0	1.0	1.0
2.9	107.4	29	300.0	3	3	1	0.0020	1.0	1.0	1.0
2.895	43.1	29	300.0	4	2	2	0.0020	2.0	1.0	1.0
3.245	70.3	29	300.0	4	2	2	0.0020	2.0	1.0	1.0
3.52	98.1	29	300.0	4	2	2	0.0020	2.0	1.0	1.0
2.895	29.6	29	300.0	4	2	3	0.0020	2.0	1.0	1.0
2.775	45.7	29	300.0	4	2	3	0.0020	2.0	1.0	1.0
2.96	70.6	29	300.0	4	2	3	0.0020	2.0	1.0	1.0
2.85	12.8	29	300.0	4	3	3	0.0020	1.0	1.0	1.0

Figure 7-4: Condition Assessment/Prediction Results

7.4 Summary

The developed web-based decision support tool will help municipal engineers to prioritize inspection and rehabilitation of critical sewers. The tool will be helpful for the decision makers to share their knowledge with others. Therefore, it has a great potential for providing extremely valuable feedbacks of experts from all around the world. Therefore, the program will provide a solid platform for the future expansion of the research.

Chapter 8

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

8.1 Summary and Conclusions

The present research work leads to the development of a combined condition index (CCI) for sewer pipelines. The index has five different categories varying from 1 to 5; where 1 represents acceptable combined (structural plus operational) condition of a sewer, and 5 represents a sewer's critical condition. The proposed index will help municipal engineers in visualizing the combined effects of structural and hydraulic problems on a sewer's existing condition.

A methodology for predicting a sewer's structural and operational condition rating through the use of historical data is proposed. The current research has developed multiple regression models for prediction the most likely structural and operational condition rating of an existing sewer. Different regression models are designed for three different sewer pipe materials: concrete, asbestos cement, and PVC. These models are developed on the basis of identified physical, operational and environmental factors which contribute to a sewer's deterioration. Various functional forms of variables have been experimented during the design procedure for the selection of best possible scenario. The co-efficient of multiple determination (R^2) results show that 72% to 88% of the total variability in structural and operational condition of sewers can be explained through the developed regression models. Similarly, all other necessary statistical diagnostics have been applied to check the adequacy of the regression models.

Furthermore, all regression models have been validated through their respective validation data. Average validity percent of these models is found to be with in range from 82% to 86%. These results are considered as acceptable. Based on the developed models, structural and operational deterioration curves have been generated. These curves represent a relationship between condition rating and age. Consequently, the models are recommended for the future usage of structural and operational condition prediction of sewers.

Moreover, the research has developed a web-based automated tool for a sewer's structural and operational condition prediction. The proposed web-based tool will evaluate the structural, operational, and combined condition (CCI) of sewers on the basis of different physical, environmental and operational factors. The developed web-based decision support tool will help municipal engineers to prioritize inspection and rehabilitation to critical sewers. The tool will assist decision makers to share their knowledge with others. Therefore, it has a great potential of expansion and enhancement through feedbacks of experts from all around the world.

In order to optimize solution for the unification and integration of different developed sewer condition assessment protocols, this research suggests accepting the WRc protocols, the most widely used sewer condition assessment protocols in the world, as standardized condition rating system for sewers. This research further proposes a methodology of integrating other in use protocols with WRc protocols through unsupervised neural network clustering technique (self-organizing maps). As an example, the research proposes modifications in CERIU sewer condition assessment protocols, the protocols adapted in the Province of Quebec, to facilitate its conversion and integration into WRc. The proposed methodology has been examined by CERIU sub-committee for the development of an integrated and unified condition assessment protocol. The sub-committee has agreed that the proposed conversion factors will be helpful in providing a documented link between CERIU and WRc protocols. Furthermore, the sub-committee has recommended some related issues for future research. These issues will be presented in recommendation session. Consequently, the proposed methodology will be helpful for municipal engineers in their research regarding unification, standardization, and integration of sewer condition assessment protocols.

8.2 Research Contributions

Following are the contributions of this research in the current sewer condition assessment process:

1) Design a combined condition index (CCI) for sewers
2) Develop structural condition prediction regression models for sewers
3) Develop an operational condition prediction model for sewers
4) Propose a modification in CERIU sewer condition assessment protocols for its documented conversion to WRc protocols: the first step towards integration of all sewer condition assessment protocols
5) Develop a web-based condition rating model for sewers

The added value of the research contributions is presented in Table 8-1. The added value includes, development of an integrated and easy to understand condition rating scale for sewers, enhancement of the concept of sewer condition prediction for prioritizing sewer inspections, development of a documented link between the two major sewer condition assessment protocols adapted in Canada, and introduction to the web-based applications for sewer pipeline condition assessment.

Table 8-1: Added Value of the Research

Research Contribution	Added Value
CCI	Integration of structural and operational condition ratings of sewers into one easy to understand condition rating scale
Structural Condition Prediction Models	Introducing the effects of bedding material and traffic volume on existing sewer pipeline condition
Operational Condition Prediction Model	Introducing the concept of operational condition assessment for sewer pipelines
Proposed Modification in CERIU	A documented link between WRc and CERIU protocols has been developed
Web-Based Condition Predictor	A platform for municipal engineers to utilize condition assessment tools for sewers is developed

8.3 Research Limitations

The current research introduces combined condition index (CCI) for sewers, structural and operational condition prediction models for sewers, and a methodology for the integration of sewer condition assessment protocols. The research has some limitations which are described below:

- The statistical diagnostics results for sewer condition prediction regression models show that the selected predictors are not enough to entirely explain the variation in existing condition of sewers
- The developed regression models are limited to a certain range of input data. These data ranges have been described in chapter 4 (Tables 4-8 to 4-11)

- The condition assessment models are only designed for three types of sewer pipe materials: concrete, asbestos cement, and PVC
- The operational condition prediction model is built on the assumption that the value of Manning's co-efficient of roughness will not change during the whole life span of a pipe
- The developed models are not appropriate for the condition prediction of sewers buried under highways
- The CERIU protocol conversion methodology does not address the various observations noted within CERIU's manual that are not even considered by WRc

8.4 Future Recommendations

Future recommendations for the extension of this research can be summarized as follows:

- ❖ Current Research Enhancement Areas:
    - More predictors, such as soil conditions, seismic factors, etc, should be incorporated to enhance the developed condition prediction models
    - True validation of models can be performed by acquiring more data from various other municipalities
    - More clustering techniques, such as k-means clusters etc, should be adapted to validate the developed cluster boundaries of transformed deduct values and combined condition matrix
    - The left over observations of CERIU protocols should be taken into account for the enhancement of CERIU protocol modification methodology

- Current Research Extension Areas:
  - Standardization of data acquisition tool for municipalities which should cover all relevant physical, operational and environmental factors
  - Extension of the sewer pipeline condition prediction methodology to other sewer network structures such as manholes, outfalls, pumping stations, etc.
  - Integration of hydraulic performance models for sewers with the developed operational condition prediction model
  - Application of sewer condition prediction methodology to storm water drains
  - Incorporation of the developed web-based tool with GIS
  - Adoption of similar methodologies for conversion of all sewer condition assessment protocols into an integrated condition assessment system for sewers

REFERENCES

Abraham D, Wirahadikusumah R., Short T, and Shahbahrami S, (1998), "*Optimization Modeling for Sewer Network Management*", ASCE Journal of Construction Engineering and Management, Volume 124, Number 5

Abraham D, & Wirahadikusumah R, (2003), "*Application of Dynamic Programming and Simulation for Sewer Management*", Journal of Engineering, Construction and Architectural Management, Volume 10, Nov 2003, Pages 193-208

ACPA, (1980), "*Concrete Pipe Design Manual*", 5th Edition, American Concrete Pipe Association (ACPA), USA

Allouche, N & Freure, P, (2002), "*Management and Maintenance Practice of Storm and Sanitary Sewers in Canadian Municipalities*", Institute of Catastrophic Loss Reduction Research Report, ICLR No 18, Canada

Allouche, N & Wong L, (2003), "*Interactive Simulation-Based Educational Software for Management of Municipal Sewer Network*", Proceedings of the American Society of Civil Engineers (ASCE) International Pipeline Conference, Baltimore, USA

Anderson, M & Davenport, N, (2005), "*A Rural Transit Asset Management System*", UTCA Report 04401, University Transportation Centre for Alabama, USA

Andreou D., Marks D & Clark R, (1987), "*A New Methodology for Modeling Breaks Failure Patterns in Deteriorating Water Systems: Theory*", Journal of Advanced Water Resources, March 1987

Ariaratnam, S, El-Assaly, A & Yang, Y, (2001), "*Assessment of Infrastructure Inspection Needs using Logistic Models*", ASCE Journal of Infrastructure Systems, Volume 7, No 4, December 2001

ASCE & WPCF, (1970), "*Design and Construction of Sanitary and Storm Sewers*", A Joint Committee of the American Society of Civil Engineers and the Water Pollution Control Federation, USA

ASCE, (1990), "*Residential Streets*", 2nd Edition, American Society of Civil Engineers, National Association of Home Builders of the United states, and the Urban Land Institute (ULI) Publication

ASCE, (1992), "*Design and Construction of Urban Stormwater Management Systems*", American Society of Civil Engineers (ASCE) and Water Environment Federation (WEF) Joint Publication, USA

ASTM, (2003), "*Standards for Reinforced Concrete and Asbestos Cement Sewer Pipes*", American Society of Testing Materials Standards (ASTM C-76-02 & ASTM C-428)

Bandemer, H & Gottwald, S, (1995), "*Fuzzy Sets, Logic and Methods with Applications*", John Willy and Sons Inc, USA

Barqawi, H, (2006), "*Condition Rating Models for Underground Infrastructure: Sustainable Water Mains*", Masters Thesis, Concordia University, Montreal

Baur, R& Herz, R, (2002), "*Selective Inspection Planning with Ageing Forecast for Sewer Types*", International Water Association (IWA) Journal of Water Science and Technology, Volume 46, Number 6-7, Pages 389 – 396

Benli O, (1999), "*Industrial Engineering Applications and Practice: Dynamic Programming*", International Journal of Industrial Engineering (Sited in November, 2005) http://benli.bcc.bilkent.edu.tr/~omer/research/dynprog.html

Bergeron, C, (2006), "Internal *Correspondence during the Research*", Director Technical, Centre for Expertise and Research on Infrastructures in Urban areas (CERIU), Montreal, Canada

Bilgil A, (2003), "*Effect of Wall Shear Stress Distribution on Manning Coefficient of Smooth Open Rectangular Channel Flows*", Nidge University Research Report, Turkey

Butler, D & Davis, J, (2000), "*Urban Drainage*", Taylor and Francis Group, New York, USA

Butt A, (1994), "*Application of Markov Process to Pavement Management Systems at Network Level*", Proceedings of 3rd International Conference on Managing Pavements, Transportations Research Council, Washington, USA

CGSB, (1994), "*Asbestos Cement Sewer Pipe standard CAN/CGSB-34.9-94*" Canadian General Standard Board Publication

Casey, T, (1992), "*Water and Wastewater Engineering Hydraulics*", Oxford Science Publications, Oxford, UK

CERIU, (2004), "*Manuel de Standardization des Observations*" 2nd Edition, Centre for Expertise and Research on Infrastructures in Urban Areas (CERIU), Montreal, Canada

Chae, M & Abraham D, (2001), "*Neuro-Fuzzy Approaches for Sanitary Sewer Pipeline Condition Assessment*", ASCE Journal of Computing in Civil Engineering, Volume 15, Number 1, January, 2001

Chung S, (2001), "*Demand Modeling and Analysis for the Management of Underground Infrastructure Systems*", PhD Thesis, Purdue University, USA

Deboeck, G & Kohonen, T, (1998), "*Visual Exploration in Finance with Self Organizing Maps*", Springer Inc

Denys B, Elisio V & Pascal L, (2004), "*Decision Making in Sewer Maintenance Strategies: Simulation as a Practical Tool*", Proceedings of 1st International Forum on Engineering Decision Making (IFED) December 2004, Stoos, Switzerland

Dikmen, I, Birgonul, M & Kiziltas, S, (2005), "*Prediction of Organizational Effectiveness in Construction Companies*", ASCE – Journal of Construction Engineering and Management, volume 131, 2005

Fenner, R, (2000), "*Approaches to sewer Maintenance: A Review*", Urban Water, Volume 2, Page 343 – 356, UK

Fenner, R, Sweeting, L & Marriot, M, (2000), "*A New Approach for Directing Proactive Sewer Maintenance*", Proceedings of Institution of Civil Engineers Water Maritime and Energy Journal, Volume 124(2), Pages 67 – 78

Fox, J, (1997), "*Applied Regression Analysis, Linear Models, and Related Methods*", SAGE Publications Inc, California, USA

Frangopol, D, Kallen, M and Noortwijk, M (2004), "*Probabilistic Models for Life-Cycle Performance of Deteriorating Structures: Review and Future Directions*", Wiley Inter Science (www.interscience.wiley.com)

Gallant, S, (1993), "*Neural Network Learning and Expert Systems*", The MIT press, England

Gribbin, J, (2001), "*Introduction to Hydraulics and Hydrology with Applications for Storm Water Management*", Delmar Thomson Learning, USA

Grigg, N, (1994), "*Maintenance Management Systems for Urban Drainage*", Urban Rehabilitation and Program Techniques, American Society of Civil Engineers (ASCE) Publication, Pages 2 to 10

Grzymala, J., (1991) "*Managing Uncertainty in Expert Systems*" Kulwer Academic Publishers, Dordrecht, Netherlands

Hahn, M, Palmer R, Merrill M and Lukas A, (2000) "*Knowledge Acquisition and Validation of an Expert System for Prioritizing the Inspection of Sewers*", Proceedings of Joint Conference on Water Resource Engineering and Water Resources Planning and Management, Minneapolis, USA

Hasegawa, K, Wada, Y & Miura, H, (1999), "*New Assessment System for Premeditated Management and Maintenance of Sewer Pipe Networks*", Proceedings of 8th International Conference on Urban Storm Drainage, Page 586-593, Sydney, Australia

Herz, R K, (1996) "*Ageing Processes, Rehabilitation Needs and Strategies for Drinking Water Distribution Network*", Journal of Water Supply Research and Technology-Aqua, Volume 45(5), Pages 221-231

Hillier, S. F. and Lieberman G. J. (1995), "*Introduction to Operations Research*" 6th Edition, McGraw – Hill Book Co Inc

Hrycej, T, (1992), "*Modular Learning in Neural Networks*", John Wiley and Sons Inc, USA

Kathula, V, (2000), "*Distress Modeling for Sanitary Sewer Management Systems*", Masters Thesis, Louisiana Tech University, USA

Kleiner, Y, (2001), "*Scheduling Inspection and Renewal of Large Infrastructure Assets*", American Society of Civil Engineers (ASCE) Journal of Infrastructure Systems, Volume 7, Number 4, December 2001

Kohonen T, (1990), "The Self Organizing Map", Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) Conference, Volume 78, No. 9

Kohonen, T, (1993), "*The Things You haven't Hear about the Self Organizing Map*", Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) Conference, Volume 5, 1993

Kohonen, T, (1997), "*Self Organizing Maps*", 2nd Edition, Springer Inc

Kulandaivel, G, (2004), "*Sewer Pipeline Condition Prediction using Neural Network Models*", Masters Thesis, Michigan State University, USA

Kutner, M, Neter, J, Nachtsheim, C & Wasserman, W, (2005), "*Applied Linear Statistical Models*", 5th Edition, McGraw-Hill Companies Inc. USA

Lee, B & Scholz, M, (2006), "*A Comparative Study: Prediction of Constructed Treatment Wetland Performance with K-Nearest Neighbors and Neural Networks*", Springer Inc Scotland, UK

Levine, D, Stephanm D, Krehbiel T, Berenson, M & Bliss, J, (2002), "*Statistics for Mangers – Using Microsoft Excel*", Canadian Edition, Prentice Hall Inc

Lingras, P, (1995), "*Classifying Highways: Hierarchical Grouping versus Kohonen Neural Networks*", Journal of Transportation Engineering, July/August, 1995

MacLeod, C, Ariaratnam, S, and Chua, K, (2000), "*Evaluation of Emergency Repair Attributes for Sewer Infrastructure Management*", Proceedings of CSCE-2000 Conference, Canadian Society for Civil Engineering, Toronto, Canada, Pages 41 – 59

Makar, J, (1999), "*Diagnostic Techniques for Sewer Systems*", ASCE Journal of Infrastructure Systems, Volume 5, Number 2, June 1999, Pages 69 – 78

May, R, Martin, P & Price, N, (1998), "*Low-Cost Options for Prevention of Flooding from Sewers*", CIRIA Inc London, UK

Maze, T, (2005), "*Asset Management Policy Issues*" (Sited in December, 2005) http://www.ctre.iastate.edu/gasb34/policyissues.pdf

McDonald, S & Zhao, J, (2001), "*Condition Assessment and Rehabilitation of Large Sewers*", Proceedings of International Conference on Underground Infrastructure Research, page 361-369, Waterloo, Canada

McGhee, T, (1991), "*Water Supply and Sewerage*", 6th Edition, McGraw-Hill Inc, New York, USA

Menzefricke, U, (1995), "*Statistics for Managers*", Duxbury Press, Wadsworth Publishing Company, USA

Mills R, Patrick, F, & Bowles, J, (1996), "*Fuzzy Logic Enhanced Symmetric Dynamic Programming for Speech Recognition*", Proceedings of the 5th IEEE International Conference on Fuzzy Systems, September 1996, Pages 2013 to 2019

Moselhi, O & Shehab-Eldeen, T, (2000), "*Classification of Defects in Sewer Pipes using Neural Networks*", American Society of Civil Engineers (ASCE) Journal of Infrastructure Systems, Volume 6, Number 3, September, 2000

Mukherjee, A, (1997), "*Self-Organizing Neural Network for Identification of Natural Modes*", Journal of Computing in Civil Engineering, January, 1997

Mutulsky, H, (2002), "*Multicollinearity in Multiple Regression*", GraphPad Software Inc http://www.graphpad.com/ (Sited in September, 2006)

NAAPI/CATT, (2002), "*CCTV Inspection Standards*", North American Association of Pipeline Inspectors (NAAPI) and Centre of Advancement of Trenchless Technologies (CATT) Joint Presentation (Sited in May, 2006), http://www.civil.uwaterloo.ca/catt/index.htm

Najafi, M & Kulandaivel G, (2005), "*Pipeline Condition Prediction Using Neural Network Models*", Proceedings of the American Society of Civil Engineers (ASCE) International Pipeline Conference, USA

Najafi, M & Gokhale, S, (2005), "*Trenchless Technology: Pipeline and Utility Design, Construction and Renewal*", McGraw-Hill Inc USA

Neter, J, Kutner, M, Nachtsheim, C & Wasserman, W, (1996), "*Applied Linear Regression Models*", 3rd Edition, McGraw-Hill Companies Inc. USA

Narendra, K, & Thathachar, M, (1989), "*Learning Automata: An Introduction*", Prentice Hall, Englewood Cliffs, USA

NZWWA, (2006), "*New Zealand Pipe Inspection Manual*", 3rd Edition, New Zealand Water and Waste Water Association, New Zealand

Peggs, L, (1985), "*Underground Piping Handbook*", Robert E. Krieger Publishing Company, Florida, USA

Perkins P, (1974), "*Concrete Pipes and Pipelines for sewerage and Water Supply*", Cement and Concrete association, London, UK

Rahman, S & Vanier D, (2004), "*An Evaluation of Condition Assessment Protocols for Sewer Management*", National Research Council of Canada Research Report Number B-5123.6

Reyana, S, Vanegas, J & Khan, A, (1994), "*Construction Technology for Sewer Rehabilitation*" American Society of Civil Engineers (ASCE) Journal of Construction Engineering and Management, Volume 120, Number, 3, Pages 467 to 487

Rostum, J, Baur, R, Saegrov, S, Horold, S & Schilling, W, (1999), "*Predictive Service-Life Models for Urban Water Infrastructure Management*", Proceedings of the 8th International conference Urban Storm Drainage, Pages 594 – 601, Sydney, Australia

Russell, S & Norvig, P, (2003), "*Artificial Intelligence*", Pearson Education Inc, NJ, USA

Rutch, M & Uibrig, H, (2006), "*Classification System to Estimate the Leakage of Sewers*", (Sited in May, 2006), http://www.insa-lyon.fr/Laboratoires/URGC-HU/apuss/nrl_024_classification_system_leakage.pdf

Ruwanpura J, Ariaratnam, S, and El-Assaly, A, (2004), "*Prediction Models for Sewer Infrastructure Utilizing Rule-Based Simulation*", Journal of Civil Engineering and Environmental Systems, volume 21, No 3, Page 169-185

Savitch, W, (2006), "*Absolute Java*", 2nd Edition, Pearson Education Inc, New York, USA

Schatzmann, J & Ghanem, M, (2003), "*Using Self-Organizing Maps to Visualize Clusters and Trends in Multidimensional Datasets*", Case Project Report, Imperial College, London, UK

Schleifer, A & Bell, D, (1995), "*Data Analysis, Regression and Forecasting*", Course Technology Inc, USA

Shahin, M, Mair, H & Jaksa, B, (2004), "*Data Division for Developing Neural Networks Applied to Geotechnical Engineering*", Journal of Computing in Civil Engineering, April 2004

Shehab-Eldeen, T, (2001), "*An Automated System for Detection, Classification and Rehabilitation of Defects in Sewer Pipes*", PhD Thesis, Concordia University, Montreal, Canada

Smith L, (2001), "*An Introduction to Neural Networks*", (Sited in December, 2005) http://www.cs.stir.ac.uk/~lss/NNIntro/InvSlides.html#what

Sumida Y, Onishi F and Kurata H, (2003), "*Deterministic and Stochastic Models Analyze the Robustness of Circadian Rhythms*" Research Report, Department of Biochemical Engineering and Science Kyushu Institute of Technology, Fukuoka, Japan

Stein, I & Partner, G, (2005), "*European Study of the Performance of Various Pipe Systems*", European Expert Panel Report 6-44801, Bochum, Germany

Stengel, R, (1994), "*Optimal Control and Estimation*", Dover Inc Publications New York, USA

Stergiou, C & Siganos, D, (2005), "*Neural Networks*", (Sited in December, 2005) http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html

Sykes, R, (1986), "*Introduction to Regression Analysis*", (Sited in November, 2005) http://www.law.uchicago.edu/Lawecon/WkngPprs_0125/20.Sykes.Regression.pdf

Tangsripairoj, S & Samadzadeh, M, (2005), "*Organizing and Visualizing Software Repositories using the Growing Hierarchical Self-Organizing Map*", 2005 ACM Symposium of Applied Computing, New Mexico, USA

Thornhill, R & Wildbore, P, (2005), "*Sewer Defect Codes: Origin and Destination*", U-Tech Underground Construction Paper, April, 2005

Trick A, (1997) "*A Tutorial on Dynamic Programming*", (Sited in November, 2005) http://mat.gsia.cmu.edu/classes/dynamic/dynamic.html

Varela G, (2005), "*Logistic Regression Models*", (Sited in December, 2005) http://www.sussex.ac.uk/Units/economics/Logistics/log_index

Vesanto, J, Himberg, J, Alhoniemi, E & Parhankangas, J, (1999), "*Self-Organizing Maps in Matlab*", Proceedings of the Matlab DSP Conference, Espoo, Finland, Pages 35-40

Wang, D, Ressom, H, Musavi, M & Domnisoru, C, (2002), "*Double Self-Organizing Maps to Cluster Gene Expression Data*", European Symposium on Artificial Neural Networks, Bruges, Belgium

Wirahadikusumah, R, (1999), "*Optimization Modeling for Management of Large Combined Sewer Network*" PhD Thesis, Purdue University, USA

Wirahadikusumah, R, Abraham, D, and Iseley, T, (2001), "*Challenging Issues in Modeling Deterioration of Combined Sewers*", American Society of Civil Engineers (ASCE) Journal of Infrastructure Systems, , Vol. 7, No 2, pp 77-84

WRc, (2004), "*Sewerage Rehabilitation Manual*", 4^{th} Edition, Water Research Centre, UK

Yan J & Vairavamoorthy, K, (2003), "*Fuzzy Approach for Pipe Condition Assessment*", Proceedings of the American Society of Civil Engineers (ASCE) International Pipeline Conference, USA

Zayed, T & Halpin, D, (2005), "*Pile Construction Productivity Assessment*", ASCE – Journal of Construction Engineering and Management, volume 131

Zhao, J, Q, (1998), "*Trunk Sewer in Canada*", APWA international Public Works Congress Seminar Series, American Public Works Association, Las Vegas, USA

Zhao, J & Daigle, L, (2001), "*SIDD Pipe bedding and Ontario Provincial Standards*", Proceedings of International Conference on Underground Infrastructure Research, Kitchener, Ontario, pp 143 – 152

Zurada, J, (1992), "*Introduction to Artificial Neural Systems*", West Publishing Company Inc, USA

APPENDICES

APPENDIX – A

SEWER PIPE STRUCTURAL CONDITION PREDICTION MODEL RESULTS

Appendix A- 1: PIPE MATERIAL: ASBESTOS CEMENT

Weighted analysis using weights in Structural Grade

The regression equation is

Structural Grade**2 = 20.9 + 542 Log Depth/Length + 0.207 Age - 0.742 Ass Cmt Class - 14.8 Dia**0.1

Predictor	Coef	SE Coef	T	P
Constant	20.87	20.41	1.92	0.041
Log Depth/Length	541.7	136.9	3.96	0.001
Age	0.20691	0.08946	2.31	0.034
Ass Cmt Class	-0.7416	0.4168	-1.78	0.093
Dia**0.1	-14.83	10.74	-1.99	0.085

S = 1.32426   R-Sq = 82.4%   R-Sq(adj) = 78.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	139.824	34.956	19.93	0.000
Residual Error	17	29.813	1.754		
Lack of Fit	15	22.613	1.508	0.42	0.874
Pure Error	2	7.200	3.600		
Total	21	169.636			

19 rows with no replicates

Source	DF	Seq SS
Log Depth/Length	1	101.251
Age	1	20.509
Ass Cmt Class	1	14.718
Dia**0.1	1	3.346

Unusual Observations

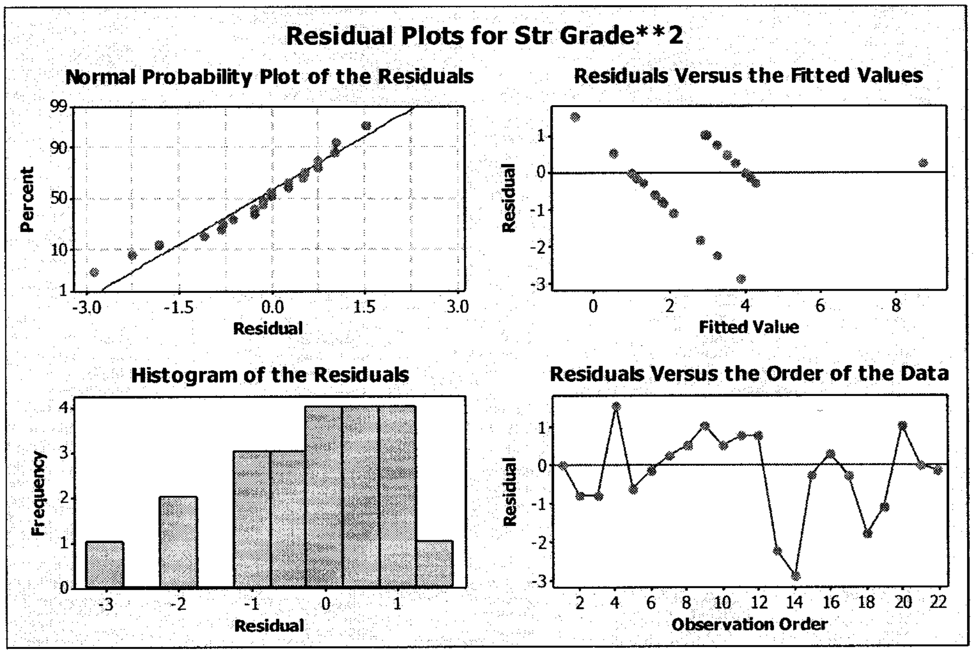
Obs	Log Depth/Length	Str Grade**2	Fit	SE Fit	Residual	St Resid
14	0.0088	1.000	3.883	0.329	-2.883	-2.25R
16	0.0143	9.000	8.731	0.726	0.269	1.12 X

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large influence.

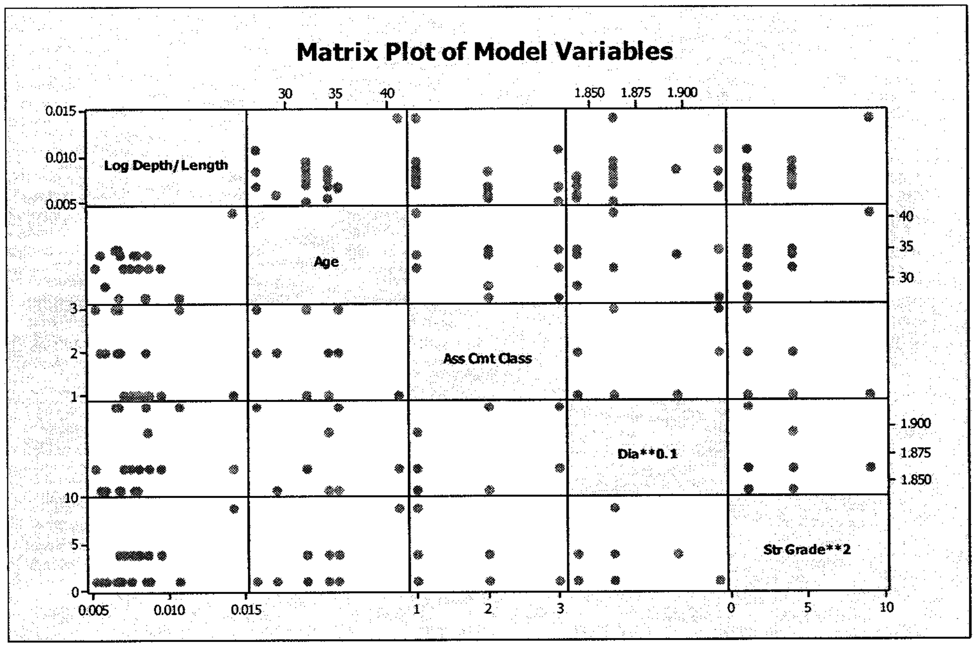
Durbin-Watson statistic D = 1.42776
For the Model
p-1 = 4 (Total No of Predictor variable)
n = 22
From Table B-7 (Neter, 1996)
d_L = 0.96
d_U = 1.80
H_0: $\rho = 0$ (Error terms are independent)
H_1: $\rho > 1$ (Error terms are positively correlated)
If
$D > d_U$, conclude H_0
$D < d_L$, Conclude H_1
$d_L \le D \le d_U$, the test is inconclusive
Result: $d_L \le D \le d_U$, the test is inconclusive

Lack of Fit Test
No evidence of lack of fit (P >= 0.1).

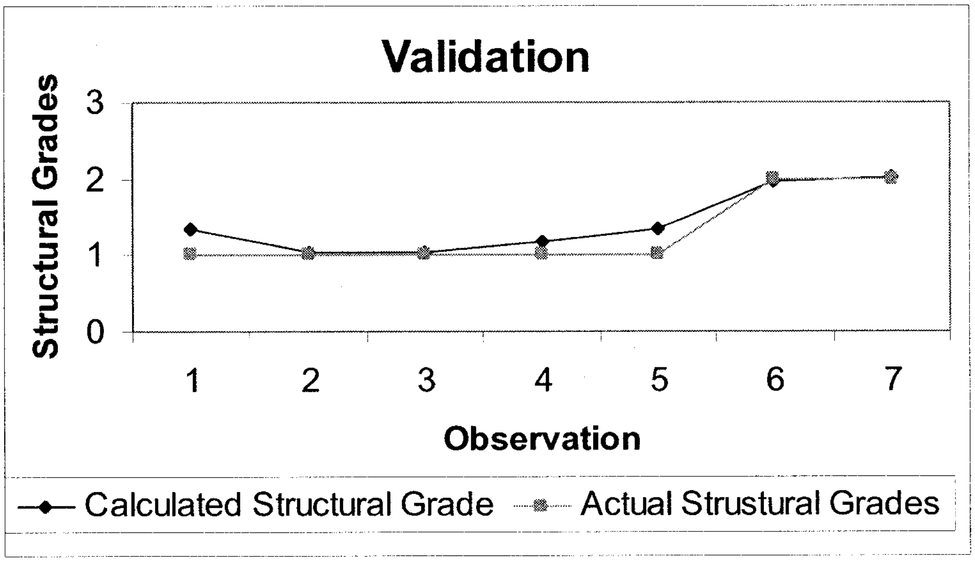
Residual Plots



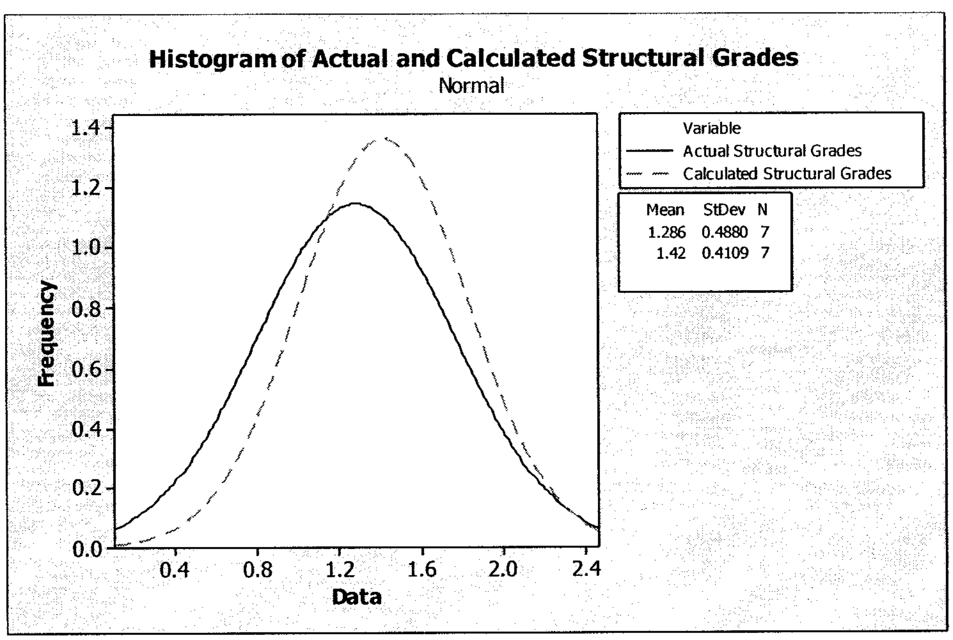
Matrix Plot for Model Variables



Validation Plot for Actual and Predicted Values



Histogram of Actual and Predicted Values



Appendix A-2: PIPE MATERIAL: CONCRETE

```
Weighted analysis using weights in Age

The regression equation is
1/SG = 3.94 + 0.592 Log Dia/Length - 0.00681 e**T - 3.22 Log Depth
       - 1.60 Log Age/CC + 6.92 Log De/Bedd - 5.75 Bedding**Lambda
```

Predictor	Coef	SE Coef	T	P
Constant	3.9420	0.6225	6.33	0.000
Log Dia/Length	0.5921	0.3151	1.88	0.064
e**T	-0.006807	0.001196	-5.69	0.000
Log Depth	-3.217	1.213	-2.65	0.010
Log Age/CC	-1.6044	0.4161	-3.86	0.000
Log De/Bedd	6.919	2.754	2.51	0.014
Bedding**Lambda	-5.749	1.507	-3.82	0.000

S = 1.02281   R-Sq = 72.7%   R-Sq(adj) = 70.5%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	6	200.979	33.497	32.02	0.000
Residual Error	72	75.322	1.046		
Total	78	276.301			

Sum of squares for pure error is (nearly) zero.
Cannot do pure error test

Source	DF	Seq SS
Log Dia/Length	1	41.216
e**T	1	52.236
Log Depth	1	0.408
Log Age/CC	1	28.228
Log De/Bedd	1	63.658
Bedding**Lambda	1	15.233

Unusual Observations

Obs	Log Dia/Length	1/SG	Fit	SE Fit	Residual	St Resid
1	0.027	0.200	0.558	0.046	-0.358	-2.09R
2	0.027	0.200	0.558	0.046	-0.358	-2.09R
18	0.026	0.333	0.439	0.107	-0.106	-0.73 X
35	0.035	0.500	0.965	0.042	-0.465	-2.51R
36	0.023	0.500	1.013	0.045	-0.513	-2.78R
56	0.406	1.000	1.027	0.111	-0.027	-0.17 X
57	0.031	1.000	0.525	0.047	0.475	2.58R
69	0.390	1.000	0.980	0.106	0.020	0.15 X

R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large influence.

Durbin-Watson statistic = 0.804345
For the Model
p-1 = 6 (Total No of Predictor variable)
n = 79
From Table B-7 (Neter, 1996)
d_L = 1.49 (n = 80, p-1 = 5)

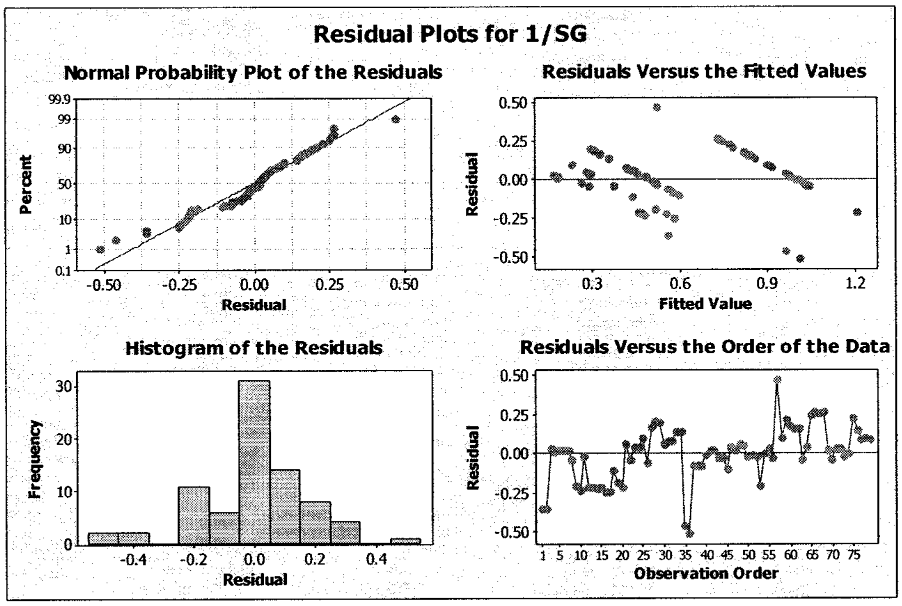
$d_U = 1.77$ (n = 80, p-1 = 5)
H_0: $\rho = 0$ (Error terms are independent)
H_1: $\rho > 1$ (Error terms are positively correlated)

Result: $D < d_L$, Error terms are positively correlated

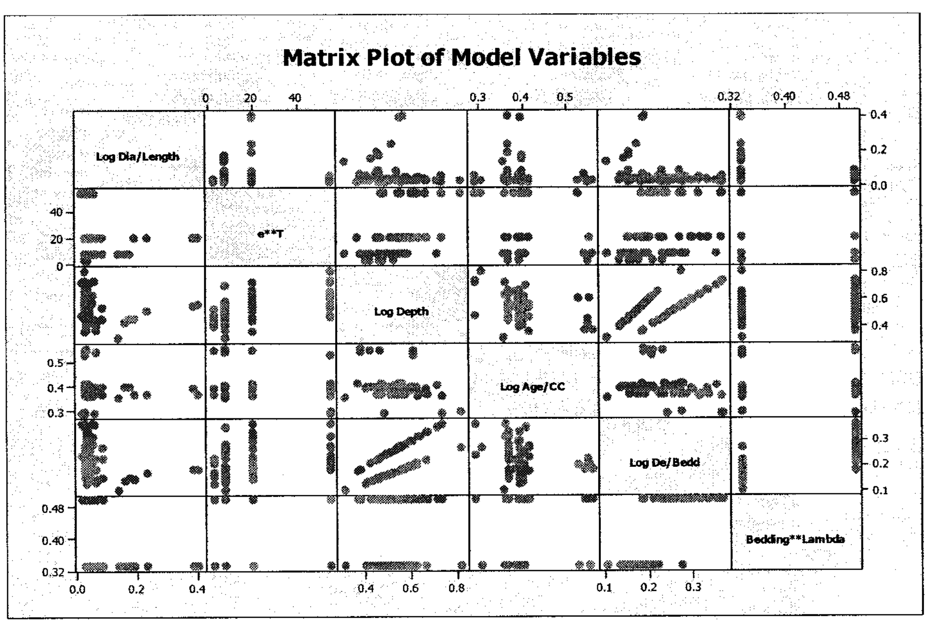
Lack of fit test

Possible curvature in variable Log Age/ (P-Value = 0.065)
Possible interaction in variable Bedding* (P-Value = 0.032)
Overall lack of fit test is significant at P = 0.049

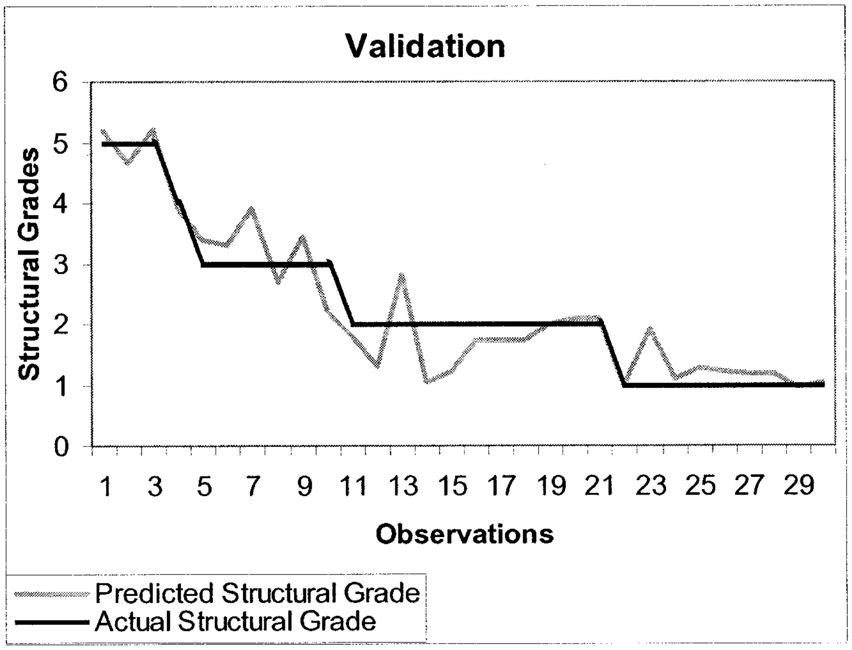
Residual Plots



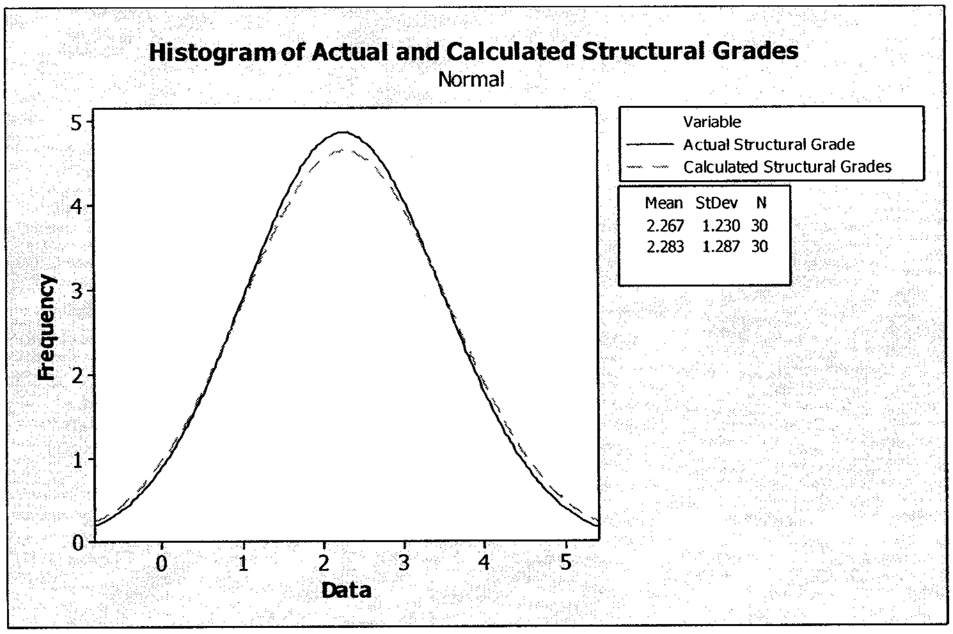
Matrix Plot for Model Variables



Validation Plot for Actual and Predicted Values



Histogram of Actual and Predicted Values



Appendix A-3: PIPE MATERIAL: PVC

```
Weighted analysis using weights in Age

The regression equation is
0.1**SG = 2.25 - 0.00642 Age - 1.89 L**0.01 - 0.0302 Bedding Class
          - 0.0405 Street Class - 0.000013 (Dia**0.3)(Depth**4)
```

Predictor	Coef	SE Coef	T	P
Constant	2.2468	0.6976	3.22	0.003
Age	-0.0064246	0.0009648	-6.66	0.000
L**0.01	-1.8916	0.6645	-2.85	0.008
Bedding Class	-0.03023	0.01241	-2.44	0.021
Street Class	-0.040524	0.009001	-4.50	0.000
(Dia**0.3)(Depth**4)	-0.00001273	0.00000223	-5.71	0.000

S = 0.0368970   R-Sq = 81.8%   R-Sq(adj) = 78.6%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	5	0.171724	0.034345	25.23	0.000
Residual Error	28	0.038119	0.001361		
Total	33	0.209843			

No replicates.
Cannot do pure error test.

Source	DF	Seq SS
Age	1	0.060913
L**0.01	1	0.004925
Bedding Class	1	0.041465
Street Class	1	0.020079
(Dia**0.3)(Depth**4)	1	0.044341

Unusual Observations

Obs	Age	0.1**SG	Fit	SE Fit	Residual	St Resid
1	14.0	0.01000	0.02563	0.00811	-0.01563	-2.79RX
4	1.0	0.10000	0.01411	0.01877	0.08589	2.70R
5	14.0	0.01000	0.02616	0.00820	-0.01616	-2.95RX
20	14.0	0.01000	0.00671	0.00961	0.00329	1.47 X
23	14.0	0.10000	0.07932	0.00600	0.02068	2.64R
34	2.0	0.01000	0.02223	0.02397	-0.01223	-1.19 X

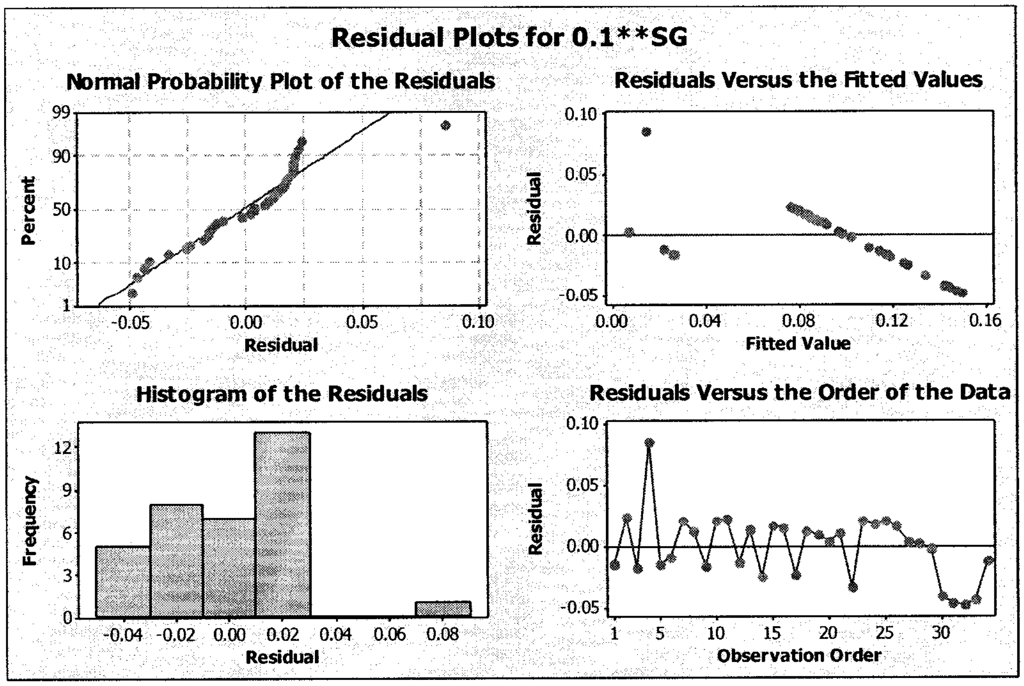
R denotes an observation with a large standardized residual.
X denotes an observation whose X value gives it large influence.

Durbin-Watson statistic = 1.81929 = 1.82
For the Model
p-1 = 5 (Total No of Predictor variable)
n = 34
From Table B-7 (Neter, 1996)
d_L = 1.15
d_U = 1.81
H_0: $\rho = 0$ (Error terms are independent)
H_1: $\rho > 1$ (Error terms are positively correlated)
$D > d_U$, (Error terms are independent)

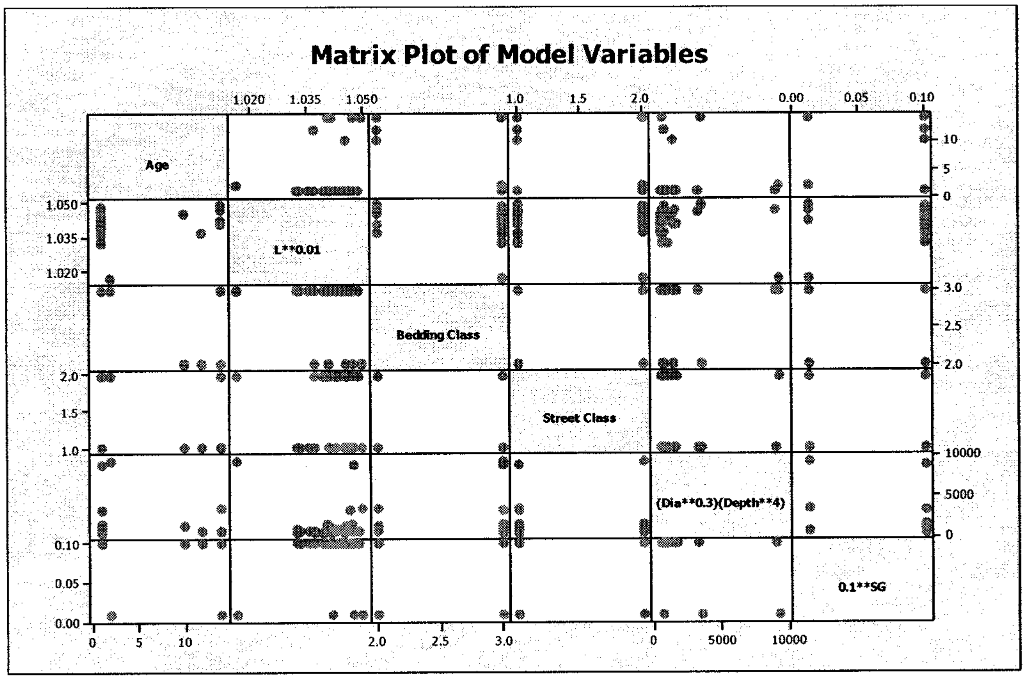
Lack of fit test

Possible interaction in variable Age (P-Value = 0.031)
Possible interaction in variable L**0.01 (P-Value = 0.081)
Overall lack of fit test is significant at P = 0.056

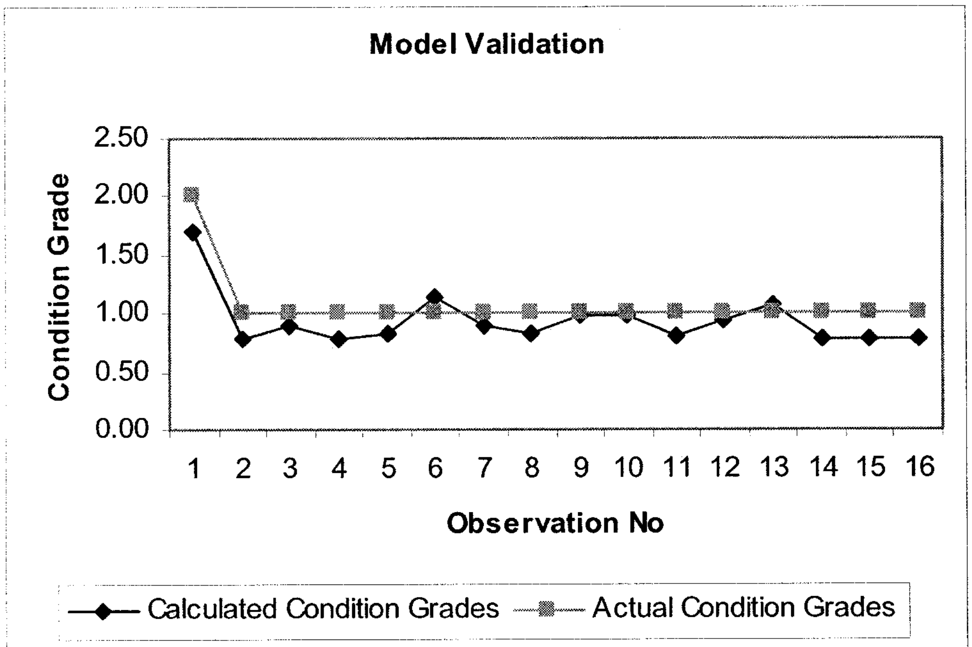
Residual Plots



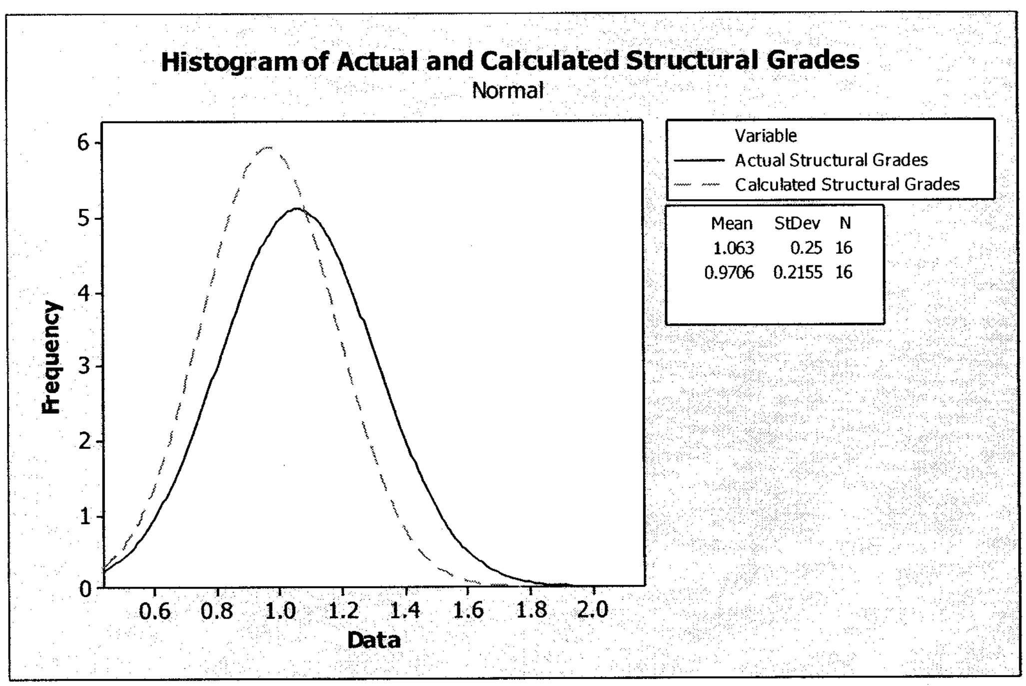
Matrix Plot for Model Variables



Validation Plot for Actual and Predicted Values



Histogram of Actual and Predicted Values



APPENDIX – B

SEWER PIPE OPERATIONAL CONDITION PREDICTION MODEL RESULTS

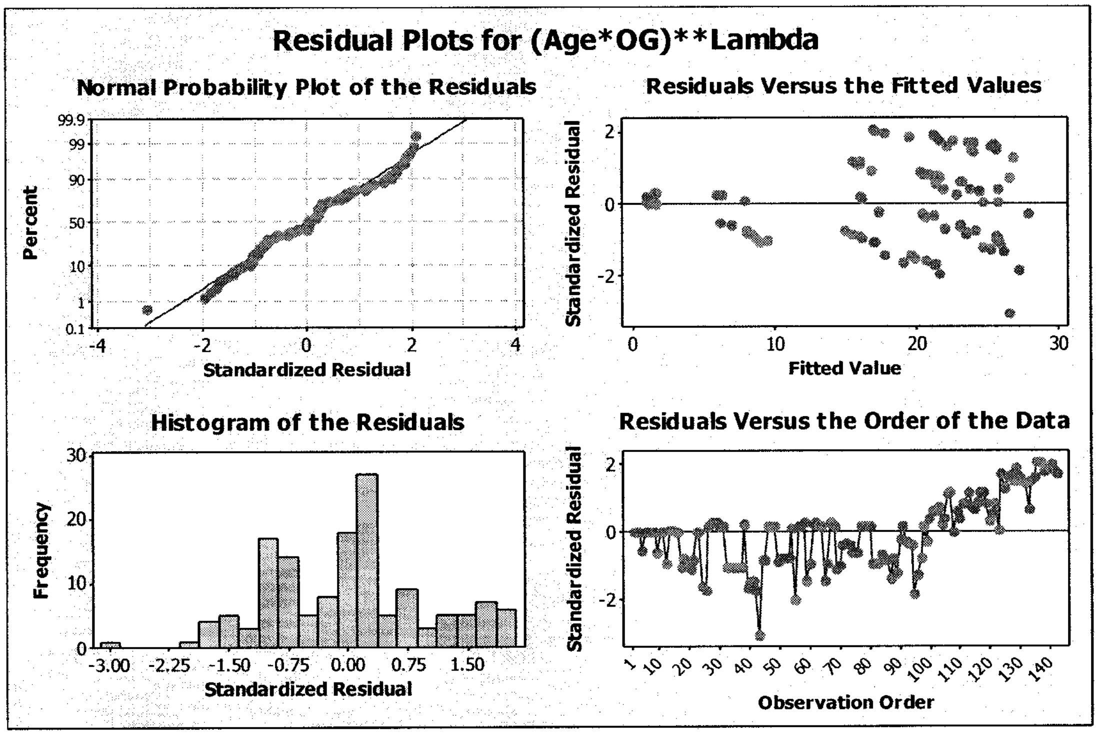
```
The regression equation is
(Age*OG)**Lambda = 0.308 + 0.567 [(Age/Dia**n)(Length**Slope)]

Predictor                            Coef  SE Coef      T      P
Constant                           0.3075   0.5504  -3.05  0.007
[(Age/Dia**n)(Length**Slope)]     0.56703  0.01775  31.94  0.000


S = 3.38404   R-Sq = 87.9%   R-Sq(adj) = 87.8%


Analysis of Variance

Source           DF     SS     MS        F      P
Regression        1  11680  11680  1019.96  0.000
Residual Error  141   1615     11
Total           142  13295


Sum of squares for pure error is (nearly) zero.
Cannot do pure error test.


Unusual Observations

Obs  [(Age/Dia**n)(Length**Slope)]  (Age*OG)**Lambda     Fit  SE Fit  Residual
 43                           46.3            16.262  26.547   0.450   -10.286
136                           29.4            23.912  16.975   0.287     6.937
137                           29.1            23.912  16.830   0.287     7.082

Obs  St Resid
 43    -3.07R
136     2.06R
137     2.10R

R denotes an observation with a large standardized residual.
```

Durbin-Watson statistic D = 0.465104
For the Model
p-1 = 1 (Total No of Predictor variable)
n = 143 > 100 so take n = 100
From Table B-7 (Neter, 1996)
d_L = 1.65
d_U = 1.69
H_0: $\rho = 0$ (Error terms are independent)
H_1: $\rho > 1$ (Error terms are positively correlated)
If
$D > d_U$, conclude H_0
$D < d_L$, Conclude H_1
$d_L \leq D\ d_U \leq$, the test is inconclusive
$D < d_L$
Result: Error terms are positively correlated

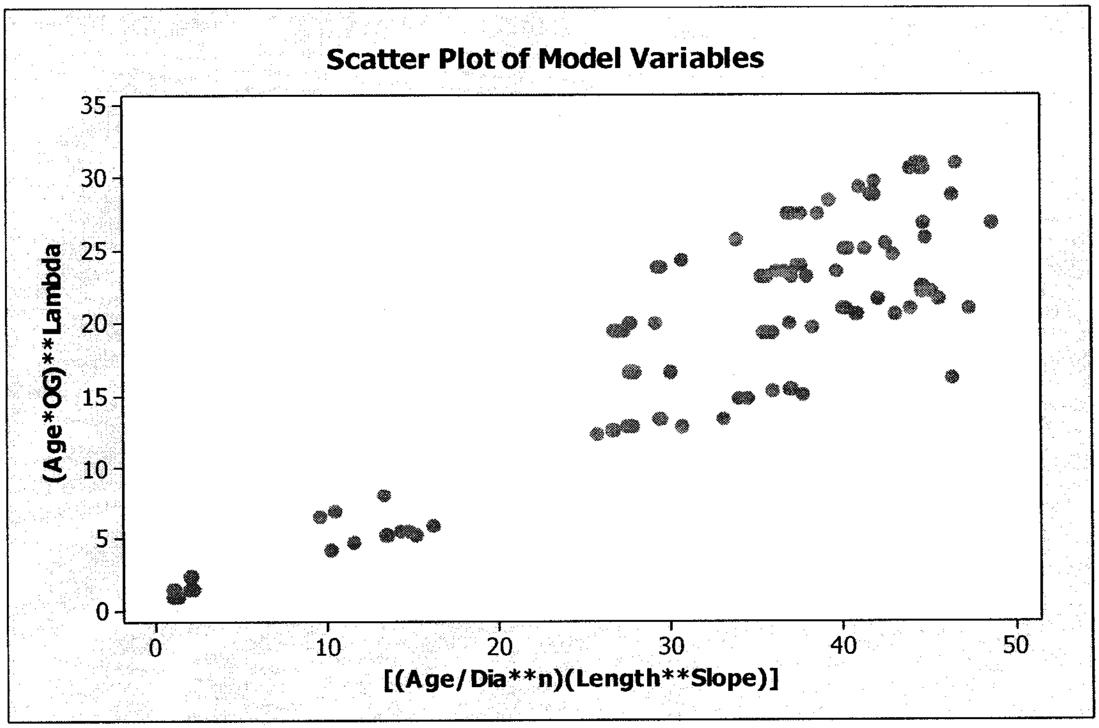
Lack of fit test

```
Possible curvature in variable [(Age/Di (P-Value = 0.073)
Overall lack of fit test is significant at P = 0.073
```

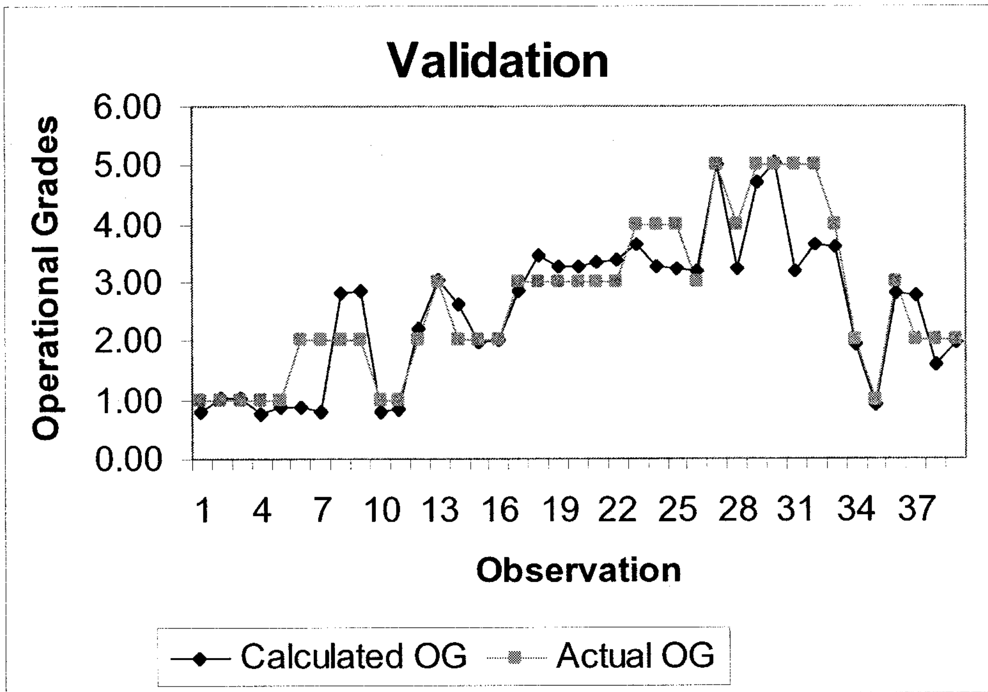
Residual Plots



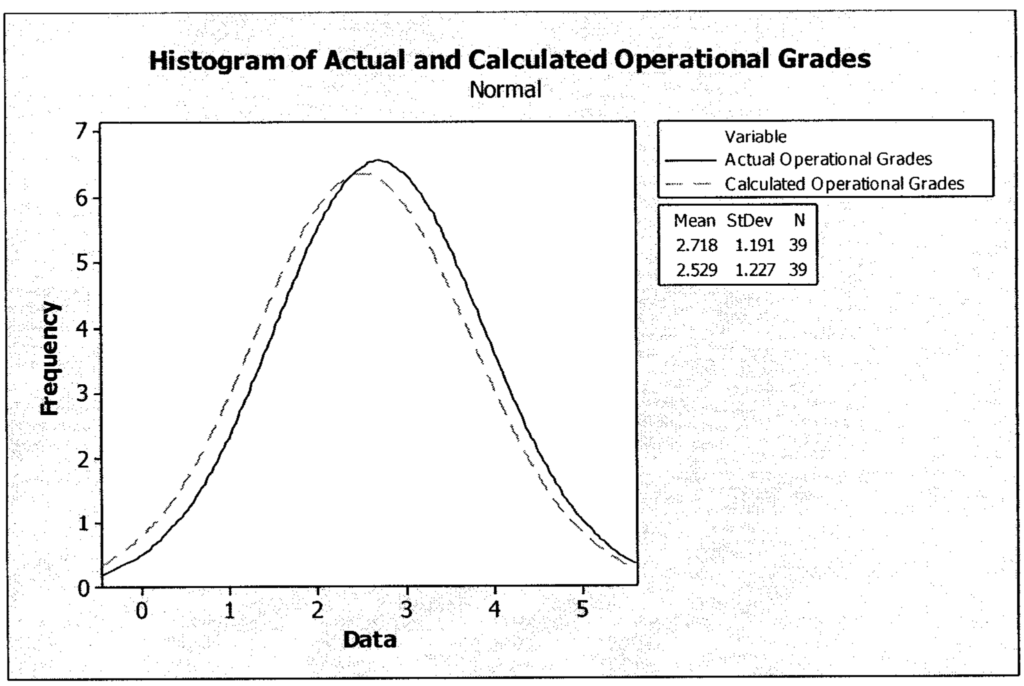
Scatter Plot for Model Variables



Validation Plot for Actual and Predicted Values



Histogram of Actual and Predicted Values



APPENDIX – C

Sample of Collected Data

Appendix C-1: Sample of Data Obtained from Niagara Falls, Ontario

ASSET_NUMBER	DIA	DEPTH	MATERIAL	Pipe Class	Bedding	LENGTH	As Built	SG	OG	AADT	WORK_CATEGORY
808040005	1550	5.82	concrete	4	B	152.1	1960	1	4	5000	MAINTENANCE
816020010	1375	6.43	concrete	4	B	256.34	1973	1	5	20000	MAINTENANCE
816020015	1375	6.43	concrete	4	B	102.11	1973	1	5	20000	MAINTENANCE
406010030	1375	6.12	concrete	3	A	114.3	1977	1	5	20000	MAINTENANCE
816020005	1375	4.25	concrete	4	B	222.81	1973	1	4	20000	MAINTENANCE
809310045	1275	6.04	concrete	4	B	126.9	1960	1	5	10000	MAINTENANCE
402020043	1200	5.29	concrete	5	B	163.2	1973	1	5	5000	MAINTENANCE
206080005	1050	7.65	concrete	5	B	110.12	1960	2	5	12500	MAINTENANCE
206080005	1050	7.65	concrete	5	B	111.3	1960	2	5	12500	MAINTENANCE
206080005	1050	7.65	concrete	5	B	120.3	1960	2	5	12500	MAINTENANCE
206080005	1050	7.65	concrete	5	B	102.63	1960	2	5	12500	MAINTENANCE
206080005	1050	7.65	concrete	5	B	120.5	1960	2	5	12500	MAINTENANCE
811060010	1050	7.56	concrete	5	B	84.12	1975	1	5	10000	MAINTENANCE
1107010070	1050	7.33	concrete	2	B	111.1	1997	1	5	20000	REHABILITATION
1107010080	1050	7.33	concrete	2	B	110.95	1997	1	5	7500	REHABILITATION
1107010080	1050	7.33	concrete	2	B	110.95	1997	1	5	7500	REHABILITATION
1107010090	1050	7.22	concrete	2	B	90.7	1997	1	5	7500	REHABILITATION
1107010040	1050	6.5	concrete	2	B	87.1	1997	1	4	20000	REHABILITATION
1107010110	1050	5.97	concrete	2	B	147.4	1997	1	5	7500	REHABILITATION
1107010110	1050	5.97	concrete	2	B	147.4	1997	1	5	7500	MAINTENANCE
819070015	1050	5.92	concrete	5	B	51.97	1964	1	5	12500	MAINTENANCE
819070025	1050	5.92	concrete	3	B	105.16	1964	1	5	12500	MAINTENANCE
830010020	1050	5.3	concrete	2	C	104.55	1960	1	5	12500	MAINTENANCE
830010020	1050	5.3	concrete	2	C	104.55	1960	1	5	12500	MAINTENANCE
809450013	1050	4.41	concrete	2	B	93.91	1960	1	4	12500	MAINTENANCE
1110120010	900	9.82	concrete	5	A	142.01	1972	1	5	12500	MAINTENANCE
1105040001	900	7.89	concrete	5	B	135.67	1975	1	5	12500	MAINTENANCE
1110120045	900	6.72	concrete	3	A	103.66	1972	1	3	20000	MAINTENANCE
1110120055	900	6.1	concrete	3	A	93.51	1972	1	4	20000	MAINTENANCE

Appendix C-2: Sample of Data Obtained from Pierrefonds, Quebec

IDENTIFICATION DES ÉLÉMENTS DU RÉSEAU D'ÉGOUT DOMESTIQUE											
No. De RUE	DE REGARD No	À REGARD No	Longueur de la section	Diamètre (mm)	Matériaux	Pente	Année de Construction	Radier AMONT	Radier AVAL	Dessus AMONT	Dessus AVAL
A010	RD.3.0504	RD.3.0505	69.0	300	TBA.IV	4.48%	1974	53.74	50.65	56.63	53.46
A010	RD.3.0595	RD.3.0596	64.6	300	TBA.IV	0.99%	1974	50.65	50.01	53.46	52.43
A010	RD.3.0596	RD.3.0597	36.3	300	TBA.IV	3.60%	1974	50.01	40.60	52.43	52.06
A015	RD.3.0178	RD.3.0179	75.0	300	TBA IV	-0.39%	1961	35.22	35.51	38.27	38.20
A015	RD.3.0179	RD.3.0180	75.0	300	TBA IV	-0.44%	1961	35.51	35.84	38.20	38.05
A015	RD.3.0180	RD.3.0181	66.0	300	TBA IV	-0.33%	1961	35.84	36.06	38.05	38.45
A015	RD.3.0181	RD.3.0182	49.5	300	TBA IV	-2.97%	1961	36.06	37.53	38.45	39.81
A015	RD.3.0182	RD.3.0183	92.3	300	TBA IV	2.04%	1961	39.40	37.52	41.55	39.75
A015	RD.3.0183	RD.3.0184	61.8	300	TBA	0.83%	1958	39.23	38.72	41.56	41.59
A015	RD.3.0184	RD.3.0185	97.1	300	TBA	0.55%	1958	39.25	38.72	41.84	41.59
A015	RD.3.0185	RD.3.0186	70.8	300	TBA	0.58%	1958	39.66	39.25	42.40	41.84
A015	RD.3.0186	RD.3.0187	20.0	300	TBA	0.60%	1958	39.78	39.66	42.46	42.40
A015	RD.3.0107	RD.3.0100	70.0	300	TBA IV	0.37%	1961	40.07	39.70	42.92	42.46
A015	RD.3.0188	RD.3.0159	101.7	300	TBA IV	0.54%	1961	40.19	39.64	42.92	42.92
A015	RD.3.0159	RD.3.0189	64.9	300	TBA IV	-0.40%	1961	39.41	39.67	42.67	42.91
A015	RD.3.0189	RD.3.0190	52.4	300	TBA IV	-0.42%	1961	39.67	39.89	42.91	43.10
A015	RD.3.0190	RD.3.0191	16.8	300	TBA IV	-0.42%	1961	39.89	39.96	43.10	43.16
A015	RD.3.0191	RD.3.0192	91.4	300	TBA IV	-1.97%	1962	39.96	41.17	43.16	45.89
A015	RD.3.0192	RD.3.0193	70.1	300	TBA IV	-0.49%	1962	41.12	41.46	45.89	45.75
A015	RD.3.0193	RD.3.0194	74.8	300	TBA IV	-0.49%	1974	41.57	41.94	45.75	44.38
A015	RD.3.0188	RD.3.0195	39.3	300	TBA IV	0.66%	1961	40.33	40.07	43.13	42.92
A015	RD.3.0195	RD.3.0196	20.5	300	TBA IV	0.39%	1961	40.41	40.33	43.14	43.13
A030	RD.3.0481	RD.3.0486	11.0	300	TBA	N.D.	1958	27.66	N.D.	30.63	N.D.
A030	RD.3.0406	RD.3.0407	92.4	300	TBA	N.D.	1950	N.D.	31.37	N.D.	N.D.
A030	RD.3.0487	RD.3.0488	122.7	300	TBA	-1.69%	1958	31.37	33.44	N.D.	N.D.
A030	RD.3.0488	RD.3.0489	27.0	300	TBA	-2.81%	1958	33.44	34.20	N.D.	N.D.
A030	RD.3.0489	RD.3.0490	22.3	300	TBA	-2.15%	1958	34.20	34.68	N.D.	N.D.
A030	RD.3.0490	RD.3.0491	74.7	300	TBA	-1.73%	1959	34.68	35.97	N.D.	N.D.
A030	RD.3.0491	RD.3.0492	72.3	300	TBA	-2.99%	1959	35.97	38.13	N.D.	N.D.
A030	RD.3.0492	RD.3.0493	55.9	300	TBA	-0.34%	1959	38.13	38.32	40.39	40.68
A030	RD.3.0493	RD.3.0494	19.7	300	TBA	-0.41%	1959	38.32	38.40	40.68	41.03
A100	RD.3.0598	RD.3.0599	50.9	300	TBA.IV	0.59%	1976	51.73	51.43	54.40	54.44
A100	RD.3.0599	RD.3.0600	33.0	300	TBA.IV	0.61%	1976	51.43	51.23	54.44	54.22
A100	RD.3.0600	RD.3.0601	89.2	300	TBA.IV	0.61%	1976	51.23	50.69	54.22	53.75
A100	RD.3.0601	RD.3.0602	60.0	300	TBA.IV	-1.21%	1976	50.53	51.35	53.75	54.57

Appendix C-3: Sample of CCTV Inspection Reports Obtained from Pierrefonds, Quebec

From	To	Year of Ins	Class	No of Observations																				
				Raccordements				Structure									Hydraulic							
				OR	RP	Lc	Int	FL	FM	FC	JD	JO	I	Co	De	Int	DM	DD	Cal	Ra	Gr	GA	CL	NL
3-0165	3-0164	2002	0																					
			1		1							8				1								7
			2		2								1											7
3-0164	3-0163	2002	0																					
			1		2	1		11	1	26		15				7			5					3
			2		7								2											3
			3		2	3							1											
3-0163	3-0162	2002	0																					
			1		2			17		21		10				7			15					
			2		5						3											1		2
			3			1																		1
			4								1													
3-0162	3-0161	2002	0																					
			1					1				1							5					3
			2		1												1							3
			3		1								1											
			4								1													
3-0161	3-0160	2002	0																					
			1					2				2				22			6					
			2		3																			1
			3		1	1																		
3-0160	3-0159	2002	0																					
			1					1		1						7			19					
			2		2														14			1	1	
			3			2										4			5					
4-0399	4-0386	2002	0																					
			1		1							5							1					3
			2		4						1													3
			3			4																		1