

## **NOTE TO USERS**

**This reproduction is the best copy available.**

UMI<sup>®</sup>



# **Quality of Spatial Information for Municipal Infrastructure Management**

**Bao Jin Hu**

A Thesis

in

The Concordia Institute for Information Systems Engineering

Presented in Partial Fulfillment of the Requirements  
For the Degree of Master of Applied Science (Quality Systems Engineering)  
at  
Concordia University  
Montreal, Quebec, Canada

March 2009

©Bao Jin Hu, 2009



Library and Archives  
Canada

Published Heritage  
Branch

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

Bibliothèque et  
Archives Canada

Direction du  
Patrimoine de l'édition

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file* *Votre référence*  
ISBN: 978-0-494-63258-1  
*Our file* *Notre référence*  
ISBN: 978-0-494-63258-1

**NOTICE:**

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

---

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

**AVIS:**

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

---

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

  
**Canada**

## **ABSTRACT**

### **Quality of Spatial Information for Municipal Infrastructure Management**

**Bao Jin Hu**

Management of municipal infrastructure involves many processes such as planning, construction, operation, and maintenance of various assets. Municipal infrastructure management systems require gathering and combining a large amount of data from different sources. These data consist of spatial and non-spatial data for describing the process information about each facility. In recent years, Geographic Information Systems (GIS) are widely used in municipal infrastructure management to spatially locate the elements of roads, sewers and water networks. However, the data used in these systems are collected from different sources using different methods with little information about the quality of the data. Problems regarding spatial data quality can affect all fields that use geographic data. Furthermore, the aging of municipal infrastructure assets combined with limited maintenance budgets presents unprecedented challenges to municipalities and public work agencies. A new integrated Municipal Infrastructure Management System (MIMS) is required to perform better quality performance for optimizing maintenance, repair and replacement activities. In this research, first a literature review is conducted about the existing MIMS software solutions, various spatial technologies are introduced, and data standards and quality concepts are discussed. Then, a new framework for MIMS spatial data quality assurance process is proposed. This framework is developed to cover all aspects of data quality and several practical methods for achieving spatial data quality assurance. One case study with four implementations is used to demonstrate the applications of the proposed approach.

## ACKOWNLEGEMENTS

It is definitely a great pleasure for me to express my sincere gratitude to my respectable supervisor Dr. Amin Hammad who gave me the wonderful opportunity to pursue my Master's degree and provided the initial concept for this interesting topic. His profound knowledge and thoughtful instructions have always shed some light on my way to pursue this thesis work. I cannot forget that he encouraged me to participate in various conferences, to communicate with specialists, and to present my work in public seminars. His advice, guidance and confidence were an invaluable resource that always inspired me.

Also, I would like to thank Mr. Peter Ghelerter, Mr. Michel Carignan, and Mr. Ghayath Haïdar from the City of Montreal in the Applications Division, GIS section for their fruitful cooperation and great help in providing me with knowledge, data, and manuals. I would like to thank Mr. Sylvain Marcoux and his assistant in the Infrastructure Division of Pierrefonds Borough for providing Pierrefonds sewer and water systems data, maps and reports. Moreover, I would like to thank Mr. Jacques Crawford from ESRI Canada for providing support in using ArcGIS Server applications.

During these two years of my study, the support of our Infra-group was of extreme importance. I would like to thank Danial Daoud, Andrea Judge, and Hui Wang, for their help in the part of collecting data and preparing the case study.

Most of all, I would like to thank my parents and all my relatives for their continuous encouragement and endless love. You are my spiritual support that encouraged me to keep on moving forward. To them, I dedicate this thesis.

# TABLE OF CONTENTS

LIST OF FIGURES .....	ix
LIST OF TABLES .....	xiii
LIST OF ABBREVIATIONS .....	xiv
CHAPTER 1 INTRODUCTION .....	1
1.1    GENERAL BACKGROUND.....	1
1.2    RESEARCH OBJECTIVES .....	2
1.3    THESIS ORGANIZATION.....	3
CHAPTER 2 LITERATURE REVIEW .....	4
2.1    INTRUCTION .....	4
2.2    REVIEW OF MUNICIPAL INFRASTRUCTURE MANAGMENT.....	5
2.2.1    Definition of Municipal Infrastructure Management.....	5
2.2.2    Software Applications in Municipal Infrastructure Management .....	7
2.2.2.1    Synergen System.....	8
2.2.2.2    CityWorks.....	9
2.2.2.3    Municipal Infrastructure Management System in Alberta.....	11
2.2.2.4    Real-time Asset Valuation Analysis (RIVA).....	11
2.2.2.5    Hansen Information Technologies.....	13
2.2.2.6    Infrastructure 2000.....	14
2.2.2.7    Harfan Infrastructure Management Approach (HIMA).....	15
2.2.3    Summary of MIMS Software Applications .....	16
2.3    SPATIAL TECHNOLOGIES.....	19
2.3.1    Geographic Information Systems (GISs).....	19

2.3.2	GIS Technology in MIMS .....	20
2.3.3	Web-based GISs.....	23
2.3.4	Global Positioning System (GPS).....	26
2.3.5	Differential and Real-time Kinematics GPS.....	29
2.3.6	Other Spatial Technologies.....	29
2.4	SPATIAL DATA INTEROPERABILITY STANDARDS .....	30
2.4.1	Interoperability of GISs .....	30
2.4.2	XML Technology.....	32
2.4.3	Geography Markup Language (GML).....	33
2.4.4	SVG (Scalable Vector Graphics).....	34
2.4.5	Associated Technologies .....	35
2.4.5.1	CityGML.....	36
2.4.5.2	Product Models .....	38
2.5	QUALITY ISSUES IN MUNICIPAL INFRASTRUCTURE.....	39
2.5.1	Uncertainty of Spatial Data Quality.....	39
2.5.2	Background of Six Sigma Principle.....	40
2.6	SUMMARY AND CONCLUSIONS .....	42
CHAPTER 3: RESEARCH METHODOLOGY .....		44
3.1	INTRODUCTION .....	44
3.2	FRAMEWORK FOR MIMS DATA QUALITY PROCESS.....	47
3.3	DATA COLLECTION .....	51
3.3.1	Sources of Data Errors.....	51
3.3.2	Data Collection Methods .....	52



3.3.3	Data Availability.....	53
3.3.4	Data Conversion.....	54
3.4	DATA ACCURACY ANALYSIS .....	55
3.4.1	Position Measurement.....	55
3.4.2	Network Analysis.....	57
3.4.3	Accuracy Measurement .....	58
3.5	DATA IMPROVEMENT AND SHARING.....	59
3.5.1	Data Sharing and Access .....	59
3.5.2	Coordinate Systems and Map Projections .....	60
3.5.3	WebGIS Application.....	61
3.5.4	WebGIS Architectures.....	62
3.6	DATA MAINTENANCE .....	65
3.6.1	Surveying Standards .....	65
3.6.2	Data Standardization.....	66
3.6.3	Metadata.....	68
3.6.4	Scheduled validation.....	69
3.7	SUMMARY AND CONCLUSIONS .....	70
CHAPTER 4: CASE STUDY.....		71
4.1	INTRODUCTION .....	71
4.2	SELECITION OF DEVELOPMENT TOOLS.....	71
4.3	BACKGOURND OF CASE STUDY.....	72
4.4	CASE STUDY IMPLEMENTATIONS.....	73
CHAPTER 5 SUMMARY, CONCLUSIONS, AND FUTURE WORK .....		106

5.1	SUMMARY.....	106
5.2	CONCLUSIONS AND CONTRIBUTIONS.....	107
5.3	FUTURE WORK.....	108
	REFERENCES .....	110
	APPENDIX A: GPS Configuration.....	119
	APPENDIX B: Manhole’s Coordinates Collecting Procedure by Using GPS.....	126
	APPENDIX C: Calculating the X, Y Coordinates in Attribute Table.....	130
	APPENDIX D: Pierrefonds Sewer Pipe Network Metadata .....	134
	APPENDIX E: Macro 1 to Insert X, Y Coordinates to Each Manhole .....	137
	APPENDIX F: Macro 2 to Create a Drawing File by Using X, Y Coordinates.....	138
	APPENDIX G: Pierrefonds Borough Sewer Network Different Layers Explanation ...	139
	APPENDIX H: Commercial GIS Software Vendors.....	141
	APPENDIX I: ESRI Supported OGC Specifications (ESRI 2008b).....	142
	APPENDIX J: Other Spatial Technologies .....	144
(1)	Radio Frequency Identification (RFID).....	144
(2)	Ground Penetrating Radar (GPR).....	144
(3)	Closed-Circuit Television (CCTV).....	147
	APPENDIX K: Survey Report Example .....	150
	APPENDIX L: Cumulative Standard Normal Distribution (Montgomery 2005) .....	151
	APPENDIX M: WebGIS Application .....	152
	APPENDIX N: List of Publications .....	154

## LIST OF FIGURES

Figure 2.1 Synergen asset module interface.....	9
Figure 2.2 CityWorks add-on ArcGIS interface.....	10
Figure 2.3 GIS integration capabilities of RIVA interface.....	13
Figure 2.4 Hansen demo screen property browser from public works module.....	14
Figure 2.5 RoadManager 2000 pavement condition map.....	15
Figure 2.6 HIMA integrated evaluation interface.....	16
Figure 2.7 ArcGIS Server system architecture .....	24
Figure 2.8 Satellites used for GPS.....	26
Figure 2.9 Ionosphere and troposphere in the vertical structure of the Earth's Atmosphere.....	27
Figure 2.10 GML three elements .....	33
Figure 2.11 A partial example of GML (using a Unified Modeling Language) static structure diagram as applied to represent a pipe .....	34
Figure 2.12 GML map making process.....	35
Figure 2.13 OGC, Building Smart Alliance.....	37
Figure 3.1 Example of data conflict caused by different data sources.....	46
Figure 3.2 Proposed GIS data quality assurance process.....	48
Figure 3.3 GIS-based spatial data quality management flowchart for MIMS .....	56
Figure 3.4 Polyline's disconnection when applying network analysis .....	57
Figure 3.5 Positional error measurements.....	58
Figure 3.6 WebGIS Development Cycle.....	62
Figure 3.7 Web services framework.....	63

Figure 3.8 Thin client architecture applications .....	64
Figure 3.9 Thick client architecture applications.....	64
Figure 3.10 Surveying standards and data standardization across departments.....	65
Figure 4.1 Example of errors: sewer lengths in CAD map and Excel file do not match..	73
Figure 4.2 Scanned design drawing of sewers network for Pierrefonds Borough.....	74
Figure 4.3 CAD drawing of sewer network in Pierrefonds .....	75
Figure 4.4 CAD drawing of water network in Pierrefonds.....	75
Figure 4.5 Pierrefonds sewer pipes and manholes converted from CAD to GIS.....	76
Figure 4.6 Pierrefonds sewer pipes network attribute table .....	77
Figure 4.7 Sewer pipe slop calculation .....	78
Figure 4.8 Pierrefonds sewer and manhole layer presented in ArcMAP.....	80
Figure 4.9 Process to retrieve pipe’s starting and ending manhole by using “buffer” function.....	82
Figure 4.10 IDs of starting and ending manholes of pipes.....	83
Figure 4.11 Append manhole attribute data with graphic file by using “join” function in ArcMap.....	83
Figure 4.12 Pierrefonds sewer network in AutoCAD.....	84
Figure 4.13 Detailed CAD drawings for partial sewer network.....	85
Figure 4.14 Codes of the manholes in Excel data sheet.....	86
Figure 4.15 Error automated inspection process.....	87
Figure 4.16 Attribute data exaction from CAD to EXCEL.....	88
Figure 4.17 Adding coordinates of manholes from CAD data .....	90
Figure 4.18 Example of ID errors by using wrong punctuation.....	92

Figure 4.19 Example of ID errors by typing wrong section code in CAD file.....	92
Figure 4.20 Example of processing error in logical inconsistency .....	93
Figure 4.21 Example of processing error in wrong direction .....	94
Figure 4.22 Examples of duplication of manhole ID.....	95
Figure 4.23 Independent dataset collections.....	98
Figure 4.24 Create metadata for Pierrefonds sewer pipe network by using Datalogr....	101
Figure 4.25 Pierrefonds sewer network WebGIS architecture .....	103
Figure 4.26 Pierrefonds sewer system network with Montreal roads presented in WebGIS.....	104
Figure A.1 Trimble GPS 5700 receiver physical connection .....	119
Figure A.2 Trimble GPS 5700 receiver top panel .....	120
Figure A.3 Trimble GPS 5700 receiver configuration setting.....	121
Figure A.4 ArcPad communication parameters setting.....	122
Figure A.5 Port settings.....	122
Figure A.6 Connecting to GPS.....	123
Figure B.1 Selecting “WGS84” as track data projection.....	126
Figure B.2 Montreal mainroads and TrackLog layers are apart before projection .....	127
Figure B.3 Open layers properties.....	128
Figure B.4 Create new Transverse Mercator .....	128
Figure B.5 North American 1983 HARN as “Coordinate System”.....	129
Figure B.6 Montreal mainroad and Tracklog data are in the same layer.....	129
Figure C.1 Add field in attribute table.....	130
Figure C.2 Insert two fields in attribute table for X and Y coordinates’ value.....	131

Figure C.3 Calculate x, y coordinates' value.....	132
Figure C.4 The result of x, y coordinates' value.....	133
Figure J.1 3D image of underground pipes can be drawn by GPR data.....	146
Figure J.2 Examples of GPR applications .....	147
Figure J.3 Pipe inspection's image in CCTV video.....	148
Figure J.4 Examples of CCTV application .....	149
Figure M.1 Publishing GIS resources .....	152
Figure M.2 Creating Web Application.....	153

## LIST OF TABLES

Table 2.1 Review of municipal infrastructure systems review.....	18
Table 2.2 Errors of the GPS system .....	28
Table 2.3 Approaches of GIS interoperability .....	31
Table 4.1 Positional accuracy statistic from manhole RD.2.0100 to RD.2.0168.....	100

## LIST OF ABBREVIATIONS

<b>Abbreviation</b>	<b>Description</b>
2D	Two-dimensional
3D	Three-dimensional
AEC/FM	Architecture Engineering and Construction/ Facilities Management
API	Application Programming Interface
APISE	Quebec Alliance for the Protection of Underground Facilities
ASP	Active Server Pages
AVL	Automated Vehicle Location
CAD	Computer-Aided Design
CCTV	Closed - Circuit Television
CGA	Common Ground Alliance
CGI	Common Gateway Interface
CII	Construction Industry Institute
CWWA	Canadian Water and Wastewater Association
DEM	Digital Elevation Model
DGPS	Differential GPS
DMADV	Define Measure Analyze Design Verify
DMAIC	Define Measure Analyze Improve Control



DTP	Document Type Definition
FGDC	Federal Geographic Data Committee
GPR	Ground Penetrating Radar
GPS	Global Positioning System
HIMA	Harfan Infrastructure Management Approach
HTML	Hyper Text Markup Language
ID	Identification Number
MIMS	Municipal Infrastructure Management System
IMS	Internet Map Server
JIT	Just-In-Time
JSP	Java Server Pages
M&R	Maintenance & Rehabilitation
MIDS	Municipal Infrastructure Data Standard
MIIP	Municipal Infrastructure Investment Planning
MTM	Modified Transverse Mercator
NSSDA	National Standard for Spatial Data Accuracy
OGC	Open Geospatial Consortium
OOP	Object-Oriented Programming
PPS	Pulse per second

PRN	Pseudorandom noise
QA	Quality Assurance
QC	Quality Control
RDBMS	Relational Database Management System
RFID	Radio Frequency Identification
RIVA	Real-time Asset Valuation Analysis
RMSE	Root Mean Squared Error
RSL	Remaining Service Life
RTK	Real-time Kinematics
RWIS	Road Weather Information Systems
SCADA	Supervisory Control and Automatic Data Acquisition
SDTS	Spatial Data Transfer Standard
STEP	Standard for the Exchange of Product model data
SVG	Scalable Vector Graphics
TDS	Tripod Data System
UML	Unified Modeling Language
URL	Unified Resource Locator
UTM	Universal Transverse Mercator
VBA	Visual Basic for Application

VML	Vector Markup Language
W3C	World Wide Web Consortium
WGS	World Geodetic System
XML	Extensible Markup Language

# CHAPTER 1 INTRODUCTION

## 1.1 GENERAL BACKGROUND

The existence of reliable data on the state of Canadian infrastructure is an important requirement for evidence-based infrastructure policy. Without reliable and comprehensive data, the capacity of policy-makers at all levels to plan, prioritize and evaluate investments in public infrastructure is hindered, as is their ability to develop new policy approaches (Infrastructure 2004).

Although many studies define municipal infrastructure in different ways, generally the definition of infrastructure comprises transportation (e.g. roadways, bridges, sidewalks, and transit), protection (e.g. fire, police, EMS facilities and equipment), community (e.g. parks, recreation, cultural and community services and amenities), general government (e.g. civic buildings, information technology, municipal fleet), and accommodation utilities (e.g. gas, electricity, storm drainage, water and sewer). These infrastructures often share the transportation right-of-way with the roadway, yet accommodation facilities such as sewers and water pipes are complex because many facilities are buried underground and difficult to locate during planning and construction activities.

Recently, research on integrated municipal infrastructure driven by the need to find cost-effective approaches for decision makers is envisioned, and Geographic Information Systems (GIS) are widely used in municipal infrastructure management to spatially locate the elements of roads, sewers and water networks. Although many researches associate

the term “quality” with data used in information systems, the data used in these systems are collected from different sources using various methods with little information about the quality of the data. For example, data collected by Global Positioning System (GPS) will be relatively of “better quality” than those digitized from paper maps at a scale of 1:100,000) (Devilleers and Jeansoulin 2006). Indeed, the concept of quality encompasses a much larger spectrum and affects the entire process of the acquisition, management, communication, and use of geographic data. In addition, the integrated municipal infrastructure management approaches require better techniques for correlating the data of different networks to apply spatio-temporal analysis for optimizing maintenance, repair and replacement activities.

## **1.2 RESEARCH OBJECTIVES**

The efforts in this research are directed to propose a realistic, spatial data quality assurance process and relevant quality control approaches for the decision-makers to perform efficient representation of site layouts, to identify and resolve spatial errors during operations, and then facilitate spatial data sharing and maintenance with various systems. The objectives of this research are summarized as follows:

- (1) To investigate different infrastructure data problems, infrastructure GIS data requirements, and techniques for correlating data from different sources in a GIS;
- (2) To develop a new approach to identify and resolve spatial data problems and integrate possible solutions into the proposed method to achieve QA for integrated Municipal Infrastructure Management System (MIMS);
- (3) To provide implementations for testing the proposed approach.

### **1.3 THESIS ORGANIZATION**

This study will be presented as follows:

**Chapter 2 Literature Review:** This chapter reviews municipal infrastructure management systems and the situation in North America. It also reviews relevant spatial technologies used in supporting municipal asset management. Then spatial data interoperability standards and the concepts of quality in municipal infrastructure technologies are introduced. In addition, CAD and GIS technologies are compared.

**Chapter 3 Proposed Approach:** This chapter has two parts. In the first part, a framework for spatial data quality assurance process is developed. In the second part, four main aspects for spatial data quality including data collection, data accuracy analysis, data sharing, and data maintenance are addressed, and the methods to achieve better quality control are presented.

**Chapter 4 Implementation and Case Study:** In this chapter, a case study of Pierrefonds Borough is used to demonstrate the proposed method of using a GIS database to facilitate the MIMS. Several implementations relevant to four aspects in spatial data quality are applied.

**Chapter 5 Summary, Conclusions, Contributions, and Future work:** This chapter summarizes and concludes the present research work, highlights its contributions, and suggests recommendations for future research.

## **CHAPTER 2 LITERATURE REVIEW**

### **2.1 INTRUCTION**

Management of municipal infrastructure involves many processes such as planning, construction, operation, and maintenance of various assets. The Municipal Infrastructure Investment Planning (MIIP) Primer on Municipal Infrastructure Asset Management (Vanier and Rahman 2004) defines municipal infrastructure as: Those assets managed by municipalities. These typically include the classes of assets such as underground utilities, roads, transit systems, bridges, water/sewage treatment plants and parks. These assets often share the transportation right-of-way with the roadway. In particular, sewers and water networks are complex because they are buried underground and difficult to locate during planning and construction activities. Canadian Water and Wastewater Association (CWWA 2004) estimated that upgrading and maintaining the water and wastewater infrastructure in Canada will cost the government about CAD\$88.4 billion over the period of 1997-2012. Statistics Canada (Statistics Canada 2006) stated that 52% of the sewer system had passed their service life for all levels of government in 2003. Therefore, the aging of municipal infrastructure assets combined with limited maintenance budgets presents unprecedented challenges to municipalities and public work agencies.

In the past decades, significant advances have been made in developing management systems for municipal infrastructure. GIS technology has been widely recognized as a useful tool in supporting municipal asset management to query, visualize, explore, and analyze asset data in a spatial context. Several asset management systems have been already integrated with GIS applications. However, the data used in these systems are

collected from different sources using different methods with little information about the quality of the data. Since municipal assets are typically characterized by their long lifecycle, large size, complexity, inter-dependencies, and dynamic nature, one of the primary functions of an integrated MIMS is to maintain the accuracy, consistency, and integrity of the asset data.

This Chapter reviews the state-of-the-art knowledge in the Municipal Infrastructure Management Systems (MIMS) in Canada and gives a comprehensive overview on interoperability and quality issues in this domain. This knowledge was collected by reviewing the literature, current MIMS software packages, and current practice in the MIMS domain.

## **2.2 REVIEW OF MUNICIPAL INFRASTRUCTURE MANAGMENT**

### **2.2.1 Definition of Municipal Infrastructure Management**

At the municipal level, infrastructure is all assets that need to be managed over their entire lifecycle. These diverse assets also compete with each other within a municipality for funding for inspection, maintenance, repair and renewal. Municipal infrastructure management (and the supporting tools and equipment) is part of asset management which is a business process and decision-support framework that: (1) covers the extended service life of an asset; (2) draws from engineering as well as economics; and (3) considers a diverse range of assets (Vanier and Rahman 2004).



A municipal asset management system is used to store and manage asset data, and to support operational and strategic decision making processes (Halfawy et al. 2005). Danylo and Lemer (1998) envisioned the role of an asset management system as “an integrator, a system that can interact with and interpret the output coming from many dissimilar systems.”

The main role of a MIMS is to ensure that utilities such as highways, roads and bridges, as well as systems for water supply, sewer pipe network and storm water in a municipality are able to deliver the essential services which are healthy, prosperous and safe to each community. The scope of a MIMS ranges from a small scale (e.g., a sewer pipe network management) to a large scale (e.g., municipality’s MIMS department).

In practice, MIMS is also defined as the management of services. These services are categorized as hard services and soft services. As an example of hard services, a waste water system should be operating efficiently, reliably, safely and legally. Soft services include ensuring that pipes are cleaned properly and regularly, monitoring the performance of operation, assessing assets condition, etc. (Vanier 2000).

In the field of infrastructure management, many managers and maintenance staff are still performing their work activities using paper-based forms for inspection tasks’ data collection. The inspection records in this case may include some human mistakes or may have some missing data. Therefore, applying the suitable technology and providing

simple-to-use tools to automate the processes can increase productivity, profitability and improve resource utilization.

### **2.2.2 Software Applications in Municipal Infrastructure Management**

With the remarkable development of MIMS, significant advances have been made during the last decade in managing infrastructure systems. Many techniques and software tools have been developed to support activities in various municipal infrastructure domains such as pavement and bridge management, sanitary/storm water sewer management, and water supply management.

A typical municipal asset management system comprises two main components: a relational database containing the asset data, and a set of add-on analysis and decision-support modules. A number of data forms are used for entering, retrieving, and updating asset data stored in the database, and a set of pre-formatted reports can be automatically generated. More modern systems allow interaction with databases through GIS interfaces to graphically represent the asset data in a spatial context. Add-on data analysis and decision-support modules are used to support functions such as assessing current asset condition, predicting future performance, analyzing lifecycle costs, and identifying, prioritizing, and selecting feasible renewal plans. Asset management software can be classified into two broad categories: general-purpose software and asset specific software. General-purpose systems typically offer generic functionality and need to be customized and adapted for specific data and work processes related to specific classes of assets, such as Synergen, CityWorks, Hansen. Asset-specific software solutions such as

MIMS, RIVA and Infrastructure2000 provide a set of built-in data models and processes to support the management of a specific class of municipal assets (e.g. facilities, sewers, roads, bridges, etc.). The rest of this section summarizes the available systems for MIMS and evaluates their main features and limitations (Halfawy et al. 2005).

### **2.2.2.1 Synergen System**

Synergen system, developed by Synergen Inc., is a web-based work management and procurement system that is mainly targeted to large organizations with extensive data and process management requirements. Synergen defines a set of applications organized in a hierarchy of subsystems and modules. The subsystems include: Resources, Maintenance, Purchasing, Inventory, Customer, and Administration (Synergen 2007).

The main feature of this system is using “modules”. A group of applications can be accessed through a set of “Views” or forms to display and edit the data records from a table selected by the user. Each View or form displays a subset of the fields in that table. In addition, each module defines a set of “Actions” that represent data manipulation and analysis functions, or a procedure that a user may need to perform while in a particular module. Moreover, Synergen includes extensive data import/export capabilities, and can interface with external applications such as GIS, email, or Supervisory Control and Automatic Data Acquisition (SCADA) systems. Figure 2.1 shows Synergen asset module in the resources subsystem, and GIS view of selected assets.

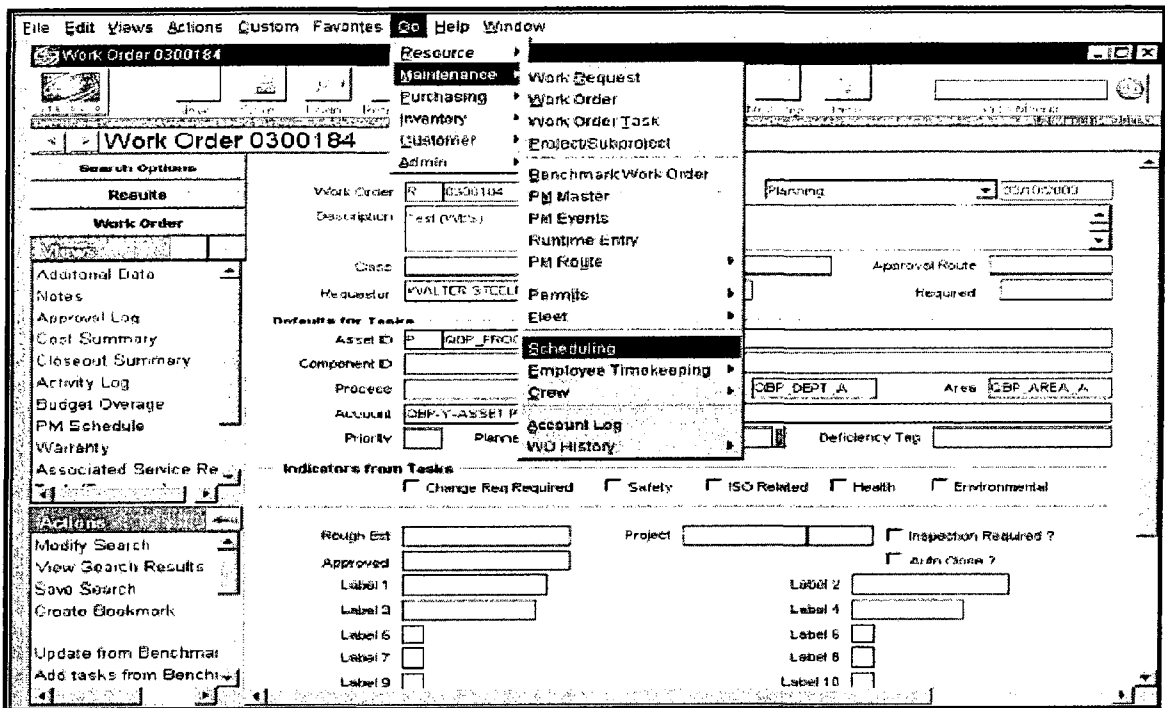


Figure 2.1 Synergen asset module interface (Synergen, 2007)

This system is that this software allows users to define a subset of the system's components relevant to their areas of responsibility, and to organize these components onto a Web portal for easy access. However, the set up and customization process, and the investment required in training and maintaining the system may be significant, especially for small to medium size organizations.

### 2.2.2.2 CityWorks

CityWorks (2007), developed by Azteca Systems Inc., is a GIS-centric and Computerized Maintenance Management System for operational and maintenance management of municipal assets. CityWorks supports functions including asset data management, work order management, recording inspection and condition data, and report generation. It also

provides logging and tracking of service requests using the add-on “Call Center” module, and supports procurement and inventory management operations.

A distinguishing feature of CityWorks is the integration with GIS. CityWorks uses the GIS database (or geodatabase) to maintain and integrate asset data. The advantage of integrating asset spatial and non-spatial data in one database ensures the consistency of asset data, and eliminates the need to duplicate the data in multiple databases or to update different databases to keep the data synchronized. The CityWorks software can also be deployed as an extension to ESRI ArcGIS (as a tool bar) or as a standalone system. Another distinguishing feature of CityWorks is its ability to support field operations by enabling browser-based wireless Internet access (e.g., using a PDA device) to the asset geodatabase, and allowing field staff to access and update work orders and service requests, and to view asset maps. Figure 2.2 presents CityWorks ArcGIS add-on showing a map of water mains and associated work orders.

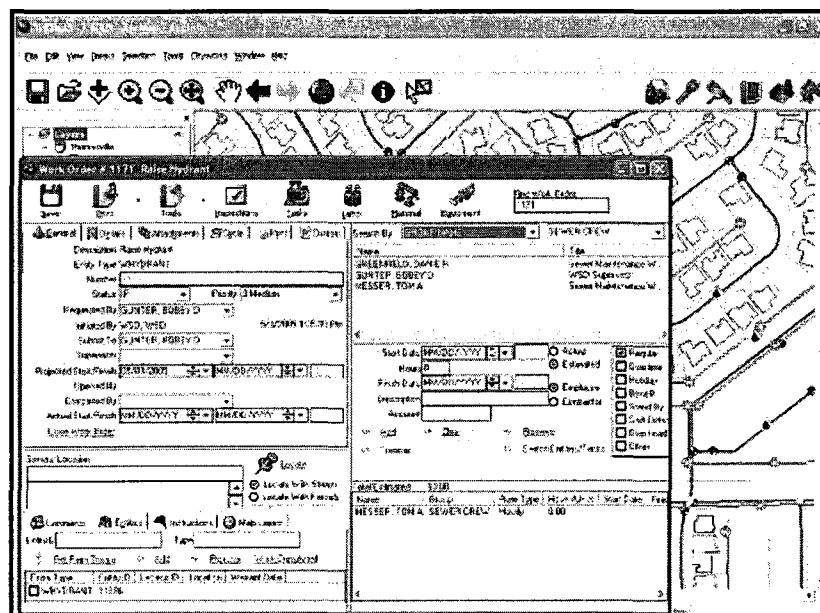


Figure 2.2 CityWorks add-on ArcGIS interface (CityWorks 2007)

### **2.2.2.3 Municipal Infrastructure Management System in Alberta**

The Municipal Infrastructure Management System (MIMS) is an initiative supported by Alberta Municipalities and Associations to enhance urban infrastructure management. The MIMS (2007) is primarily a data management system for water, wastewater, storm water, and road networks, including modules for managing gas pipelines and municipal buildings. The system is aimed to small and medium size municipalities. MIMS has extensive data import/export and reporting capabilities, and incorporates a wide range of pre-formatted reports. MIMS provides the users with a consistent set of forms and tools for managing different infrastructure assets. Each class of assets is broken down into its main components, which are in turn subdivided into asset types. Instead of interfacing with commercial GIS software, MIMS implements its own GIS functionality through the use of ESRI MapObjects COM components. For each asset type, MIMS supports four main functions: (1) *Locate* function activates the Map Viewer and zooms to the location of the current asset, (2) *Costing* function enables users to record cost data related to the asset, (3) *Functional Adequacy* function allows users to rate the overall performance and the level of service delivered by the asset (good, fair, or poor) for later use in the maintenance planning and prioritization process, and (4) *Condition* function allows users to rate the condition of the asset using a consistent standard for assessing the structural condition of the asset.

### **2.2.2.4 Real-time Asset Valuation Analysis (RIVA)**

RIVA (Real-time Asset Valuation Analysis 2007), developed by Loki Innovations ([www.loki.ca](http://www.loki.ca)), is a web based client-server application that can interface with most

common applications. The data can come from any ODBC-compliant source, so the user can leverage existing data and pull data from other databases, such as Hansen. RIVA has a modeling capability that can be used for asset valuation, determination of deferred maintenance, condition assessment, estimating remaining service life (RSL), and prioritization of maintenance and rehabilitation (M&R) processes. Deterministic and probabilistic models can be created using the Formula Builder tool.

The asset inventory and valuation capability within RIVA can retrieve information already in the user's existing databases and permit the user to set up new asset classes. The application can virtually support any number of user-defined asset classes. The hierarchical structure allows the user to specify the level of data detail required for each asset. RIVA also has a built-in valuation capability that uses integrated economic factors to calculate asset values. RIVA can also import condition assessment data from other sources, and directly link this data to assets at any level of detail. The application allows the user to set up models, based on user-defined parameters or on pre-existing systems to determine asset condition. RIVA supports both ESRI and Intergraph GIS applications. Figure 2.3 shows sample screens that demonstrate the GIS integration capabilities of RIVA. The limitation of RIVA is that it is not a work order system, it can import work order data and allow the user to see all of the work done on one asset. The main feature of RIVA is that RIVA has "thin client" architecture and does not require desktop installation. The data is typically stored on a server but can be hosted by RIVA for smaller municipalities.

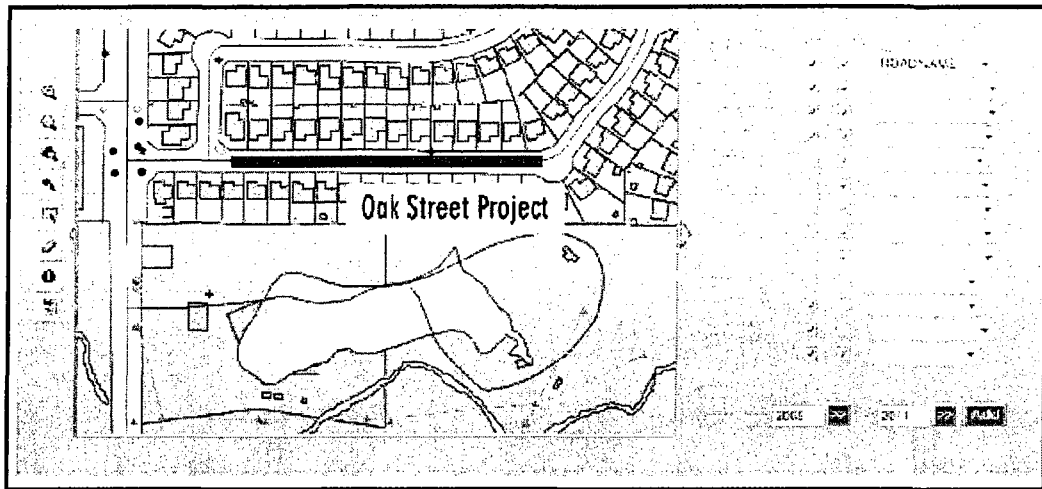


Figure 2.3 GIS integration capabilities of RIVA interface (Real-time Asset Valuation Analysis 2007)

#### **2.2.2.5 Hansen Information Technologies**

Hansen (2007) is a major asset management application developed by Hansen Information Technologies to provide capabilities for managing government operations including asset and property management, utility billing, permits, financial and human resources management. Hansen applications are typically used by medium to large municipalities or organizations. Hansen is a client-server application installed on individual desktops. Hansen's asset management tools are contained in two major modules: Public Works solutions and Transportation solutions. Hansen provides a detailed asset inventory and valuation capability using one or more of the following asset data models: hierarchical, directional, pressurized, segmented, point, area, linear and network models. Except for the pavement management module, Hansen does not have a built-in condition assessment capability. To address the strategic asset management gap, Hansen has recently teamed up with RIVA to extend Hansen's operational management



functionality with RIVA's long-term asset management capabilities. Figure 2.4 shows sample screens of elements of Hansen's public works and transportation modules.

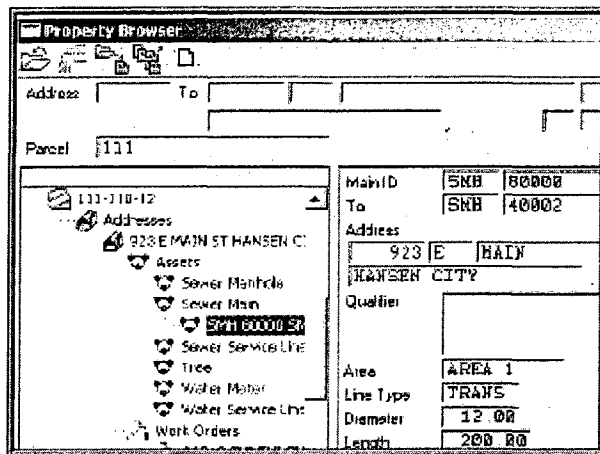


Figure 2.4 Hansen demo screen showing property browser from public works module

(Hansen 2007)

#### 2.2.2.6 Infrastructure 2000

Infrastructure 2000 (2007), developed by Vanasse Hangen Brustlin, Inc., provides capabilities for asset management planning, and is targeted at small to medium size organizations. Infrastructure 2000 consists of RoadManager 2000, with five asset management modules, and three work management tools. Infrastructure2000 is client-server application installed on individual desktops. It supports inventory data collection, asset valuation, and determination of deferred maintenance, condition assessment, estimating remaining service life, and prioritizing M&R options. The software can be integrated with popular GIS applications such as ESRI's ArcGIS. Currently, this system can only be used for pavement assets. For other types of assets, the user has to define the appropriate models and scenarios. Figure 2.5 shows sample screens from RoadManager2000, demonstrating the Pavement table options, and pavement condition assessment.

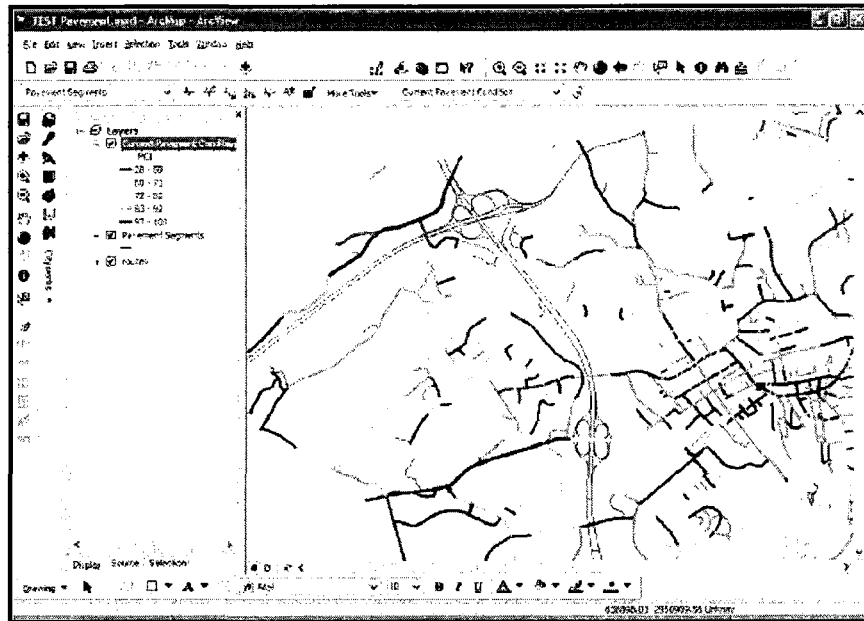


Figure 2.5 RoadManager 2000 pavement condition map (VHB 2008)

### 2.2.2.7 Harfan Infrastructure Management Approach (HIMA)

Harfan Infrastructure Management Approach (HIMA), developed by Harfan (2007) Inc., combines a five-step methodology with an Integrated Decision Support System that enhances decision making and provides for asset management through the systematic planning, control, expansion, and decommissioning of capital (infrastructure) assets. It integrates policy (strategic-level) management, system/network (mid-level) management, and operations management into one focused approach with high performance technology.

HIMA is geared to be a generic solution for long-term management of municipal assets. It attempts to be flexible in its design, so that it can be adapted to support: (1) extending the asset service life, and (2) optimizing the long-term investments. The software can be applied to water and sewer networks, roads, gas and telecommunications networks,

electricity networks, street lighting, buildings, marine assets, airports, and rail systems. Harfan uses an Oracle RDBMS, and allows one-way integration with the most popular GIS systems (e.g., Autodesk MapGuide and ESRI ArcGIS). Figure 2.6 shows a screen capture of integrated asset evaluation interface.

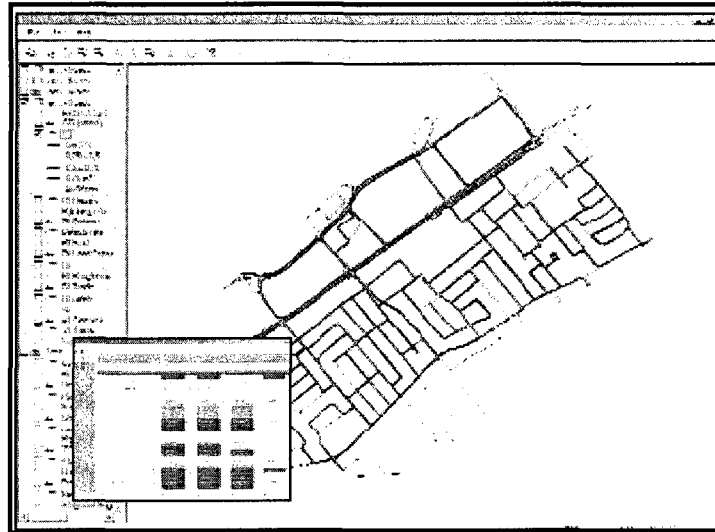


Figure 2.6 HIMA asset evaluation interface (HIMA 2008)

### 2.2.3 Summary of MIMS Software Applications

Based on the review of state-of-the-art MIMS applications, it is found that: (1) GIS is becoming an important information system for most modern MIMS and the vast majority of existing software tools has already implemented GIS functionality either internally or by supporting interfaces with GIS systems; (2) Web-based system architecture becomes a trend in MIMS because it allows users to access without time and location restriction; (3) System sharing and extensive capabilities with other assets system or applications is important for long-term system management; and (4) System modules based on functions, assets or decision making are important for system usage. On the other hand, the main limitation of these systems is that they were developed to function as individual systems

with limited or no capability of interoperability to share and exchange information with other systems. Insufficient data source is also an impediment to achieve an integrated MIMS system. The “islands of information” phenomenon exists in most of municipalities, which also can be attributed to the lack of consistent data model across the separate software tools. Moreover, the data quality issues have not been emphasized in most MIMS systems. Table 2.1 summarizes the review of municipal infrastructure management systems.

Table 2.1 Review of municipal infrastructure management systems

Resource	Synergen http://www.splwg.com	CityWorks http://www.azteca.com/	MIMS http://www.albertamims.org	RIVA www.loki.ca	HANSEN http://www.hansen.com/	Infrastructure2000 http://www.vhb.com	HIMA http://www.harfan.com
Type	General	General	Specific for water, wastewater, storm water and road networks	Specific for long-term asset planning	General (for managing government operation)	Specific for road management	General
Developer	Synergen, Inc.	Azteca systems Inc	Municipalities in Alberta Province	Loki Innovation	Hansen Information Technology	Vanasse Hhangen Brustlin Inc.	Harfan Technologies Inc.
Organization Size	large	large	small and medium	medium	medium and large	small and medium	medium
<b>Functionalities</b>							
GIS Integration	Available	Available	Available	Available	N/A	Available	Available
Architecture	web-based	GIS-based	GIS-based	Thin Client web-based	Client-server application	Client-server application	Oracle RDBMS
System Extension	Data import/export	GIS applications	ESRI Map Objects	Data import	GeoMedia	Other assets systems	GIS Applications
System Module	by function	Center Control	by assets	by function	by application	by assets	by decision steps

## **2.3 SPATIAL TECHNOLOGIES**

### **2.3.1 Geographic Information Systems (GISs)**

A GIS is usually designed to store spatially referenced data, which is on geographic scales of meters to thousands of kilometers. The term “Spatial Information System” is sometimes used to describe systems for storing the spatial ordering of protein chains, and other (typically very small scale) non-geographical data. In the 1980’s, GISs were defined as spatially referenced databases (Rogerson and Fotheringham 1994), which were initially developed as tools for the storage, retrieval and display of geographic information. Capabilities for the geographic analysis of spatial data were either poor or lacking in these early systems. However, with the significant development in computer power in the later 1980’s and 1990’s, GISs became more than spatially referenced databases. Worboys (1995) defined a GIS as “a computer-based information system that enables capture, modeling, manipulating, retrieval, analysis and presentation of geographically referenced data.” Meanwhile, from the business point of view, Environmental Systems Research Institute (ESRI 2008) defines a GIS as “An organized collection of computer hardware, software, geographic data, and personnel designed to efficiently capture, store, update, manipulate, analyze, and display all forms of geographically referenced information.” Therefore, GISs not only include the core spatial database, but also possess tools for the exploration, analysis and summarization of spatial data. A basic GIS provides: a core database system, spatial data analysis, and map display capabilities.

### **2.3.2 GIS Technology in MIMS**

A significant portion of municipal asset information is typically related to, or associated with, some spatial aspect of municipal assets. Spatial characteristics and inter-relationships of various assets are key factors in most municipal decision-making processes (NRC 2004). In order to support storing, organizing, querying, managing, analyzing, and reporting the asset data, Relational Database Management Systems (RDBMS) have been implemented in municipal asset management systems. A number of asset management add-on modules that interface with the database are used to perform such functions as maintenance planning, cost estimating, performance modeling, and lifecycle cost analysis. In different MIMSs, such as stormwater systems (Lior et al. 1991), sewer systems (Przybla and Kiesler 1991; Halfawy et al. 2000), cost estimating and risk assessment (Ashur et al. 1998), and water and sewer systems (Moutal and Bowen 1991), asset records are typically referenced by the asset's unique identification number. However, the disadvantage of these systems is the lack of graphical representation of the assets, which would require users to access the asset data through their unique IDs. Without graphical representation, the spatial characteristics and inter-relationships of the assets are communicated to the users in a textual form, which may impede the ease of data access, navigation, and analysis. Realizing the importance of representing and managing the spatial data of assets, many municipalities tried using Computer-Aided Design (CAD) platforms to graphically represent the asset spatial data, define and attach the non-graphic (attribute) data to the graphical objects, and use the DBMS to manage the asset data. An example of using CAD-based MIMS software is Archibus (QUT 2006), but the difficulty of updating and maintaining the linkages between the data stored in the

database and the drawing entities limits the development of CAD-based asset management systems' applications. Also, CAD models often lack the semantics needed to define the assets and their inter-relationships which restrict achieving various assets integration.

Unlike CAD, GIS technology has traditionally emphasized the definition and use of semantic rich object models that are linked to RDBMS to store and manage the attribute data associated with the spatial objects (or features), and topologies defined in maps and drawings (ESRI 2008). A GIS-based asset management system typically stores the graphical data (e.g. maps, themes, features) within the application itself, while using the RDBMS to store and manage the attribute (or non-graphical) data associated with the assets. The attribute data provide information about the assets, such as the type of entity, its properties, condition, and maintenance data (NRC 2004). For example, to retrieve information about a specific sewer pipe, an engineer could point to that pipe and in response, the attributes (e.g., length, diameter, material) of the pipe and other related documents (e.g. CAD drawing, Closed-Circuit Television (CCTV) videos, etc.) of that pipe will be displayed. Moreover, many GIS software programs have the capabilities to perform a wide range of spatial analysis to reveal spatial relationships between map features, such as network analysis, terrain modeling, and terrain analysis. These functionalities can significantly support a variety of purposes in MIMS, such as infrastructure planning, selection of suitable routing of utility lines, and so on. Although many CAD platforms (e.g. AutoCAD) have recently tried to support GIS-like functionalities such as introducing advanced capabilities for modeling semantic-rich data



and for data management, the ease of linking spatial and non-spatial data as well as the spatial analysis capabilities of GIS remain to be key features that are unparalleled by most CAD platforms. In summary, GIS platforms are more suitable for asset management applications than CAD platforms; a GIS platform can better support asset management applications by its combined data management, visualization, query, and analysis capabilities. An important survey of “*Current GIS Practices in Municipal Infrastructure Management*” (Vanier 2000) has indicated that many municipalities across Canada have already started using GIS to support a wide range of asset management processes, and that GIS is becoming an increasingly critical and integral component in almost all modern municipal asset management systems.

The most common commercial GIS software programs currently in use by municipalities across North America include ArcGIS (Environmental Systems Research Institute, ESRI), GeoMedia (Intergraph), MicroStation (Bentley), and MapGuide (AutoDesk) as shown in Appendix H. Also, there are some specific systems which have been introduced in Section 2.2.2 in municipal infrastructure management applications. In this thesis, the most widely used ESRI ArcGIS has been selected as the project platform based on its wide range of functions, including: (1) High-quality cartography, (2) Advanced editing tools, (3) On-the-fly projection, (4) Geocoding wizard-driven tools, (5) Support for metadata standards using XML, (6) Extensible architecture, (7) Direct read of more than 40 data formats, (8) Support for both 2D and 3D spatial data and network analysis, and (9) Support for WebGIS by GIS Server (ArcGIS 9 2008).

ESRI ArcGIS enables users to have direct access to GIS data by supporting most of framework data themes of Open Geospatial Consortium (OGC), including Web Feature Service (WFS), Web Coverage Service (WCS) and Web Map Service (WMS). Moreover, ArcGIS is an integrated collection of GIS software products which includes ArcView, ArcEditor, and ArcInfo. Each product includes progressively more functionalities.

### **2.3.3 Web-based GISs**

As mentioned above, GISs have become an important part of municipal infrastructure management. As a result, the interpretation of spatial data has become easy and increasingly simple to understand. Unfortunately, not all stakeholders have access to GIS, nor are they able to spend the time necessary to use it efficiently. Moreover, distributing maps without time and location restriction to users is another challenge. The Internet has made its way to many government organizations, and the ability to get information through the Internet encouraged spatial data providers to explore the Internet resources as a method for disseminating spatial information. Beyond desktop GIS, a GIS server provides connection with geographic information via Web applications and services. Organizations can use GIS servers to distribute maps and GIS capabilities over the Web to improve internal workflows, communicate vital issues, and engage stakeholders. Figure 2.7 shows an example of GIS server system architecture.

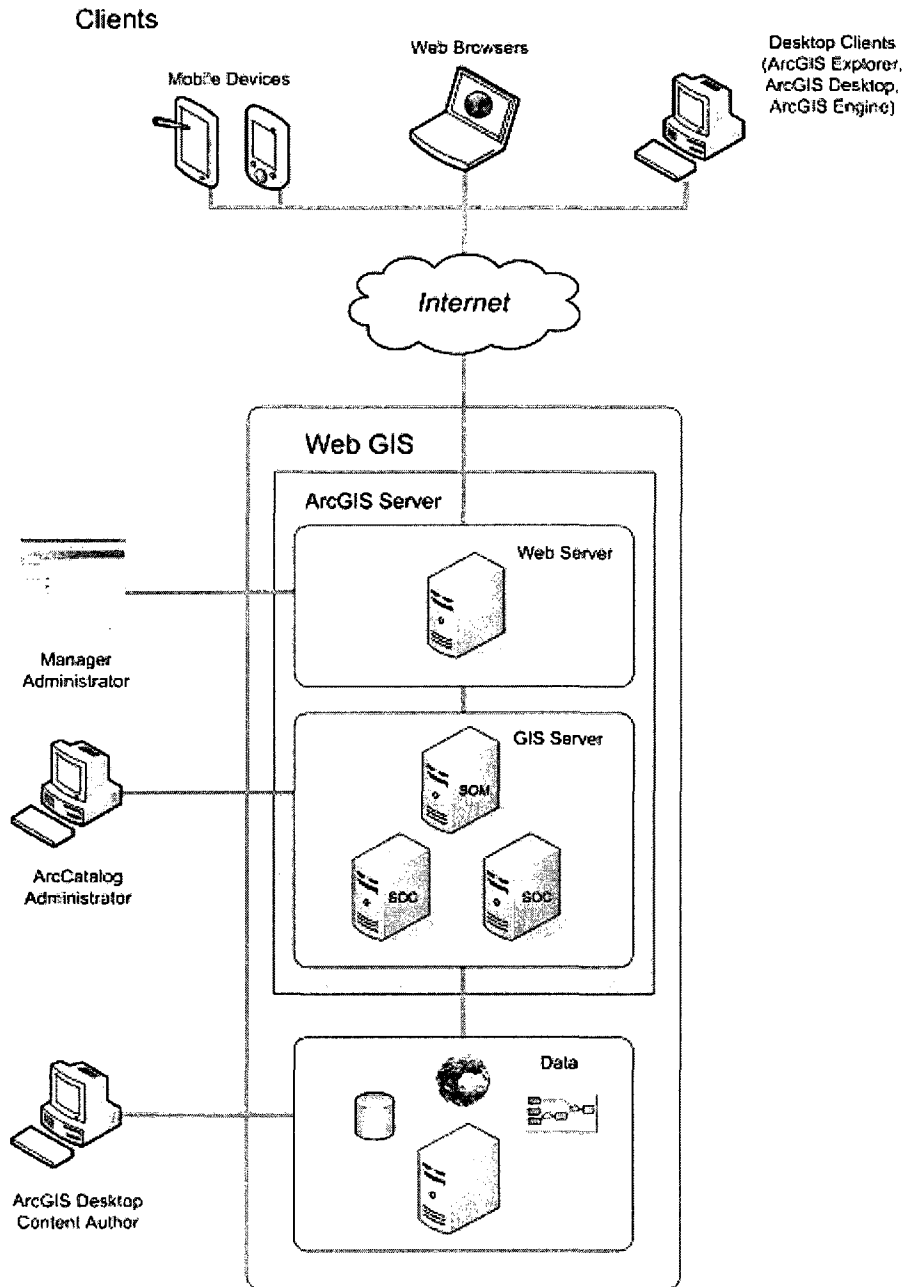


Figure 2.7 ArcGIS Server system architecture (ArcGIS Server 2008)

ArcGIS Server system architecture is a distributed system consisting of GIS server, Web server, and data. GIS server hosts and runs various services (i.e. map services, mobile data services, network analysis services, etc.) which can be operated by a server object manager (SOM) and one or more server object containers (SOCs). Web server hosts Web

applications and Web services that use the objects running in the GIS server. In the client side, Web browsers can be used to connect to Web applications running in the Web server. Desktop applications can connect either through HyperText Transfer Protocol (HTTP) to ArcGIS Web services running in the Web server, or connect directly to the GIS server. ArcGIS Server allows sharing GIS resources across an enterprise and across the Web. The GIS resources can be the maps, globes, address locators, and geodatabases. This information can be published on the GIS server, allowing client applications to use and interact with the resources. The main advantages of sharing GIS resources on a GIS server are: (1) provides synchronous data which will efficiently improve data accuracy and reduce data redundancy because the data is centrally managed, (2) supports multiple users, and (3) provides clients with the most up-to-date information. Thus, the GIS server not only allows sharing resources, but also allows accessing the embedded GIS functionality. For example, after a map has been shared through a server, the user can also interact with the maps, such as finding the closest facility and getting directions to it from his location. Furthermore, unlike GIS desktop which has add-on extension tools to apply various functionalities, GIS server just creates a GIS service to publish GIS resources.

On the other hand, a GIS server set different users authorities, such as Web application users and developers, to access the geographic information that efficiently improves data security. For example, a GIS analyst might author a map that contains a layer that references a map resource on a GIS server, and a developer may build a Web application

that uses a map and an address locator to schedule and route delivery trucks. For the general users, the details about GIS servers and GIS resources will be completely hidden.

### 2.3.4 Global Positioning System (GPS)

The GPS is a worldwide radio-navigation system formed from a constellation of 24 satellites and their ground stations (Trimble 2008). GPS is one positioning technology which is available anywhere within certain conditions and it measures the horizontal and vertical positions of the receiver from the GPS satellites. The satellites are arranged on 6 planes, each of them containing at least 4 slots (Figure 2.8) where satellites can be arranged equidistantly. Typically, more than 24 satellites orbit the earth, so that it can guarantee that the signals of at least four satellites can be received at any time all over the world and improve the availability of the system.

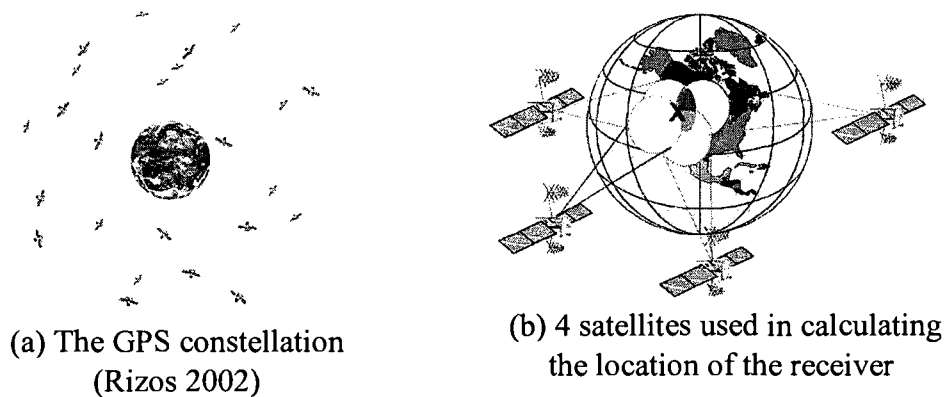


Figure 2.8 Satellites used for GPS

Each GPS satellite transmits data on two frequencies, L1 (1575.42 Mhz) and L2 (1227.60 Mhz). The atomic clocks aboard the satellite produce the fundamental L-band frequency, 10.23 Mhz. The L1 and L2 carrier frequencies are generated by multiplying the fundamental frequency by 154 and 120, respectively (Bean 2003). Two pseudorandom

noise (PRN) codes, along with satellite ephemerides (Broadcast Ephemerides), ionospheric modeling coefficients, status information, system time, and satellite clock corrections, are superimposed onto the carrier frequencies, L1 and L2. There are many factors affect the GPS accuracy. Firstly, the number and geometry of the received satellites influence the accuracy considerably, and in daily use, accuracy of about 20 m can be expected. More sophisticated GPS receivers for land survey achieve an accuracy of a few centimeters. The factors that affect GPS accuracy include multipath and atmospheric effects. The multipath effect is caused by reflection of satellite signals (radio waves) on objects.

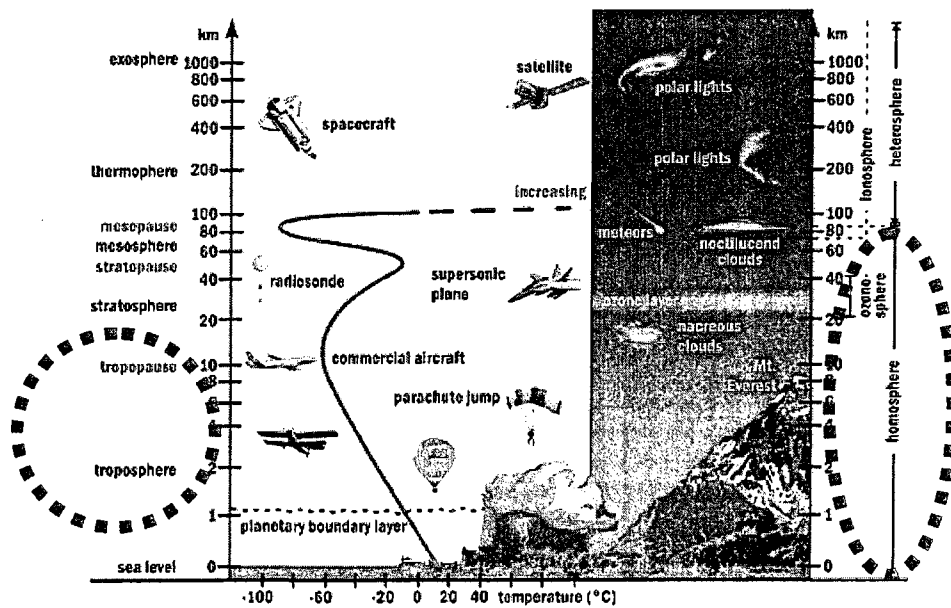


Figure 2.9 Ionosphere and troposphere in the vertical structure of the Earth's Atmosphere (adopted from GPS 2006)

For GPS signals this effect mainly appears in the neighborhood of large buildings or other elevations. The reflected signal takes more time to reach the receiver than the direct signal. The resulting error typically lies in the range of a few meters. Another source of inaccuracy is the reduced speed of propagation in the troposphere and ionosphere (Figure

2.9). In the ionosphere, at a height of 80 – 400 km, a large number of electrons and positive charged ions are formed by the ionizing force of the sun. It is known that electromagnetic waves are slowed down inversely proportional to the square of their frequency ( $1/f^2$ ) while passing the ionosphere. This means that electromagnetic waves with lower frequencies are slowed down more than electromagnetic waves with higher frequencies. These errors are mostly corrected by the receiver by calculations. The tropospheric effect is a further factor elongating the runtime of electromagnetic waves by refraction. The reasons for the refraction are different concentrations of water vapor in the troposphere, caused by different weather conditions. The error caused by the troposphere is smaller than the ionospheric error, but can not be eliminated by calculation. It can only be approximated by a general calculation model. In addition, GPS is also influenced by clock inaccuracies and rounding errors, receiver noise, and shifts in the satellite orbits. The errors of the GPS system are summarized in Table 2.2. The accuracy of a position is also a function of the geometry of the GPS constellation.

Table 2.2 Errors of the GPS system (Trimble 2008)

Type	Typical Error in Meters (per satellites)
Satellites' Clocks	1.5
Orbits Errors	2.5
Ionosphere	5.0
Troposphere	0.5
Receiver Noise	0.3
Multipath	0.6

### **2.3.5 Differential and Real-time Kinematics GPS**

One method to increase the accuracy of GPS is by using Differential GPS (DGPS). DGPS is based on correcting the effects of the pseudo-range errors caused by the ionosphere, troposphere, and satellite orbital and clock errors by placing a GPS receiver at a precisely known location, and then providing correction information to other receivers that are roving around. DGPS has a typical 3D accuracy of better than three meters and an update rate of 1-10 Hz. The DGPS corrections can be sent to the mobile GPS receivers in real time, or added later by post-processing of the collected data. Real-time kinematics GPS (RTK-GPS) receivers with carrier-phase ambiguity resolution can provide accuracy of about 1 cm, but are computationally intensive to run in real time (Kaplan 1996). The DGPS systems usually use long wave radio frequencies for broadcasting. In some situations, the streaming of DGPS or RTK correction data can be sent using a wireless communication method such as cell phones through the Internet when the radio infrastructure is not available in the local area (Lenz 2004). In Quebec, there are sixteen reference GPS stations which transmit data via the Internet (MRNF 2008).

### **2.3.6 Other Spatial Technologies**

Geographic data collection and inspection of underground utilities are more difficult to operate than other facilities. In this subsection, the latest spatial technologies for these utilities are introduced including Radio Frequency Identification (RFID), Ground Penetrating Radar (GPR), and Closed-Circuit Television (CCTV) as show in Appendix J.



## **2.4 SPATIAL DATA INTEROPERABILITY STANDARDS**

### **2.4.1 Interoperability of GISs**

Implementing municipal asset management systems involves defining a data model that represents the asset spatial and non-spatial data about many activities, such as inventory and inspection, data collection, condition assessment, performance evaluation, the prediction of future performance, planning and prioritizing maintenance and repair operations, and evaluating alternative technical and economic policies (NRC 2004). A standard data model could facilitate efficient data sharing and exchange, and improve the consistency and quality of asset information. To many municipalities, spatial data is not only accessed by its own employees, but also should be made available to several other organizations and agencies such as utility companies. A standard data model which provides the interoperability between different systems and organizations can efficiently eliminate the duplication of efforts in collecting and developing spatial data. In addition, it ensures the integrity and consistency of the data across various organizations or disciplines throughout the lifecycle of municipal assets.

During the past 20 years, the concepts, standards, and technologies for implementing GISs' interoperability have evolved through the following steps: (1) data conversion, (2) standard interchange formats, (3) open file formats, (4) application interfaces, (5) database management system, and (6) integration of standardized GIS Web services (ESRI 2003). The related approaches and technologies shown in Table 2.3 can be classified into two major areas: GIS data and technology interoperability and interoperability of GIS technology with other technologies and systems. On the

application level, GIS data and technology interoperability includes two parts: metadata support and spatial data standard. In order to build a strong spatial data infrastructure, metadata is crucial because it enables users to integrate data from multiple sources, organizations, and formats when applying data conversion, data sharing and data source definition. On the IT infrastructure level, the relationship between GISs and other applications systems, such as Enterprise Resource Planning (ERP), Customer Resource Management (CRM), Enterprise Application Integration (EAI), should also be considered to make the MIMS works in a compatible and interoperable way. Therefore, spatial data models and their related geo-relational databases, such as structure definition, maintaining performance, are included in data interoperability.

Table 2.3 Approaches of GIS interoperability (adopted from ESRI 2003)

**GIS data and technology interoperability**

<b>Methods</b>	<b>Functions</b>
Metadata support	Data conversion support
	Data sharing
	Data source definition
Spatial data standard	Content and accuracy of position
	Standard data exchange format

**Interoperability of GIS Technology with other technologies and systems**

<b>Methods</b>	<b>Functions</b>
Standard data models	Diverse requirements definition to support various disciplines
	Common semantics definition
The Geo-relational database	Geo-relational structures definition
	Maintaining performance for large data layers
	Elimination of “spatial data islands”

The Open Geospatial Consortium, Inc. (OGC) is an international industry consortium of 365 companies, government agencies and universities participating in a consensus process to develop publicly available interface specifications standards in spatial data

handling, modeling and display (OGC 2008). One of the key aims of the OGC is to enable interoperability which provides vendors (data suppliers, software suppliers, consultants) and users to recognize a common set of procedures at various levels, and use common representations for data. Therefore, OGC provides open interface specifications which enable content providers, application developers and integrators to focus on delivering more capable products and services to consumers in less time, at less cost, and with more flexibility. The ESRI supported OGC specifications are shown in Appendix I.

#### **2.4.2 XML Technology**

The Extensible Markup Language (XML) is a World Wide Web Consortium (W3C) initiative that allows information and services to be encoded with semantically meaningful structure that computers and humans can understand (w3.org 2008). XML is a simple, very flexible text format derived from Standard Generalized Markup Language (SGML) (ISO8879 2008). Originally designed to meet the challenges of large-scale electronic publishing (e.g., books listing, e-business pricing, etc.), XML technology plays a critical role in all software development projects in the exchange of a wide variety of data on the Web and elsewhere. XML is a language for describing other languages that we call *metalanguage*. Therefore, XML technology allows users to design their own markup languages for different types of documents (XML 2005). XML also improves the functionality of the web by letting users identify their information in a more accurate, flexible, and adaptable approach. Google Earth adopted the XML-based format and used KML (Keyhole Markup Language) to define locations and their description (Google Earth 2006).

### 2.4.3 Geography Markup Language (GML)

GML is an XML schema designed to represent geographical information (OGC 2008). This allows users and developers to describe generic geographic data sets that contain points, lines and polygons. GML version 1.0 was based on DTDs (Document Type Definitions) and is now obsolete. As of March 2001, version 2.0 was adopted which defined an XML schema for simple geographic features (point, polygon, and linestring). In January 2003, the much extended GML version 3.0 was released, which defines much more geometries (curve, surface, and coverage (grids)) in addition to the GML2. In April 2004, GML 3.0 was upgraded to a proposal for GML 3.1.

The composition of GML includes three elements: feature schema, geometry schema, and Xlink schema (OGC 2008) as shown in Figure 2.10.

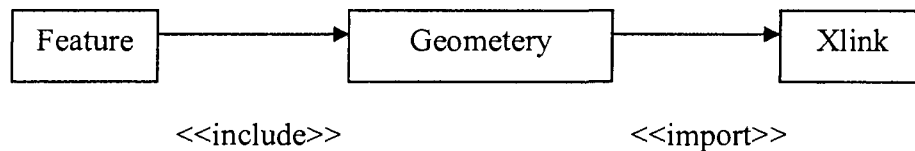


Figure 2.10 GML three elements (OGC 2008)

The features schema is used to describe geographical features and each feature may have one or more geometry objects; The geometry schema encodes the geometry of the objects which are to be presented in a particular spatial reference system such as points, polylines, and polygons; The Xlink schema is a W3C recommendation for linking elements in XML and allows cross referencing of data by Unified Resource Locator (URL). An example of the way that features and geometries are combined in practice is shown in Figure 2.11. This shows a new class pipe being defined (this is a specialisation of feature) which has

attributes name and materialtype, as well as a destination (which is itself a feature: a town) and a centerline which is a LineString. A LineString is a specialisation of the Geometry class.

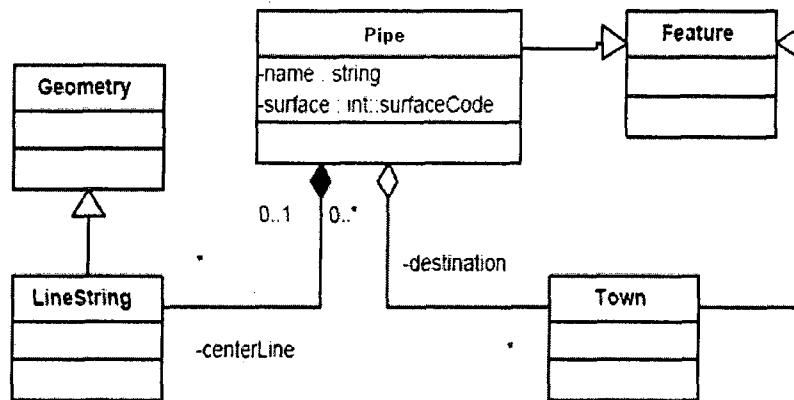


Figure 2.11 Partial example of GML (using a Unified Modeling Language) static structure diagram as applied to represent a pipe

#### 2.4.4 SVG (Scalable Vector Graphics)

Scalable Vector Graphics (SVG) is an XML specification and file format for describing two-dimensional vector graphics, both static and animated (W3C 2008). When generating maps, it is necessary to use suitable graphical representation. Traditionally, the Internet has relied on two raster (bitmap) formats: gif or jpeg. However, these are two compression formats for standard bit-mapped images which are not for vector data or large maps. Although there is no standard method for displaying 2D vector graphics over the Internet, Microsoft has tried to develop Vector Mark-up Language (VML) in Internet Explorer. Since it is limited to Internet Explorer, in January 2000 W3C has adopted SVG as a standard, and Adobe has developed a nice plug-in that can render SVG which can be embedded in an HTML (Hyper Text Markup Language). After that, SVG has rapidly become a worldwide standard to draw simple 2D features and link in very logical way

with the features for spatial data. In practice, GML requires several services to be used. This is illustrated in Figure 2.12. The GML schema defines how data is stored by some web feature server, and then passed to a rendering engine (i.e. Styling Engine) where a stylesheet is used to convert the map style data into engine. The target map will be portrayed in vector data on Web browser.

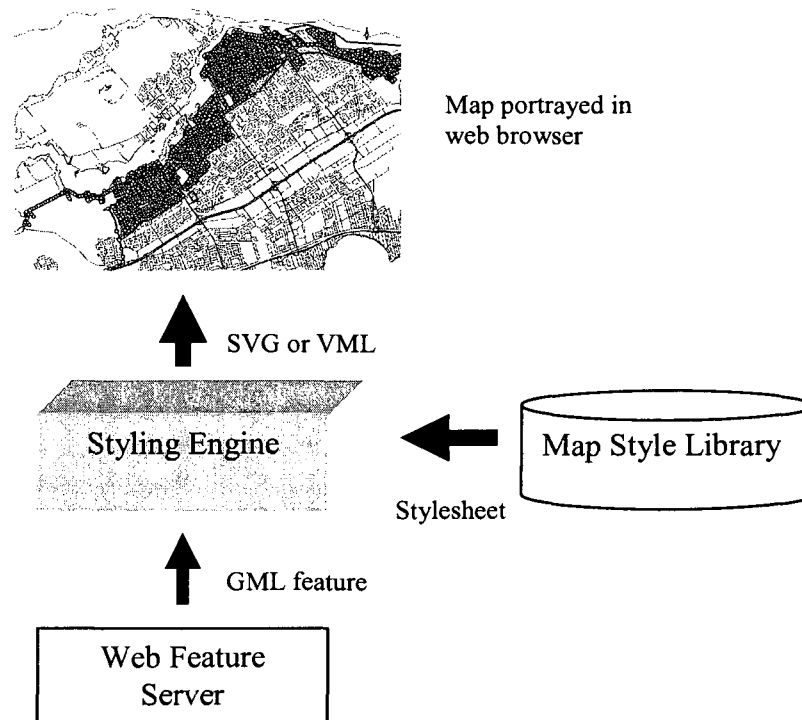


Figure 2.12 GML map making process

#### 2.4.5 Associated Technologies

GML forms the basis of a significant effort of a range of organizations and individuals to facilitate the sharing and improve the availability of geo-spatial data.

### 2.4.5.1 CityGML

CityGML is a common information model for the representation of 3D urban objects. It defines the classes and relations for the most relevant topographic objects in cities and regional models with respect to their geometrical, topological, semantical and appearance properties (CityGML 2008). CityGML is an OpenGIS Encoding Standard for the representation, storage and exchange of virtual 3D city and landscape models and it is implemented as an application schema of the GML version 3.1.1 (OGC 2008). The goal of CityGML are: (1) to establish high degree of semantic interoperability; (2) to represent 3D topography as observed; (3) to provide suitable spatial data infrastructures; and (4) to simplify using the same data by many applications. In the geospatial industry, remote sensing, photogrammetry, GIS, CAD, Assets Management/Facilities Management, navigation and spatial database management were considered different vertical markets served by different technology providers. However, there is an increasing desire and market pressure for geospatial technologies to converge. CityGML develops XML encodings to deal with various aspects of data exchange in the building lifecycle, and this convergence will have a major impact on professionals involved with infrastructure projects such as bridges and water systems. Infrastructure design, construction, and maintenance professionals need to share information as shown in Figure 2.13.

## Common user defined interoperable operating information

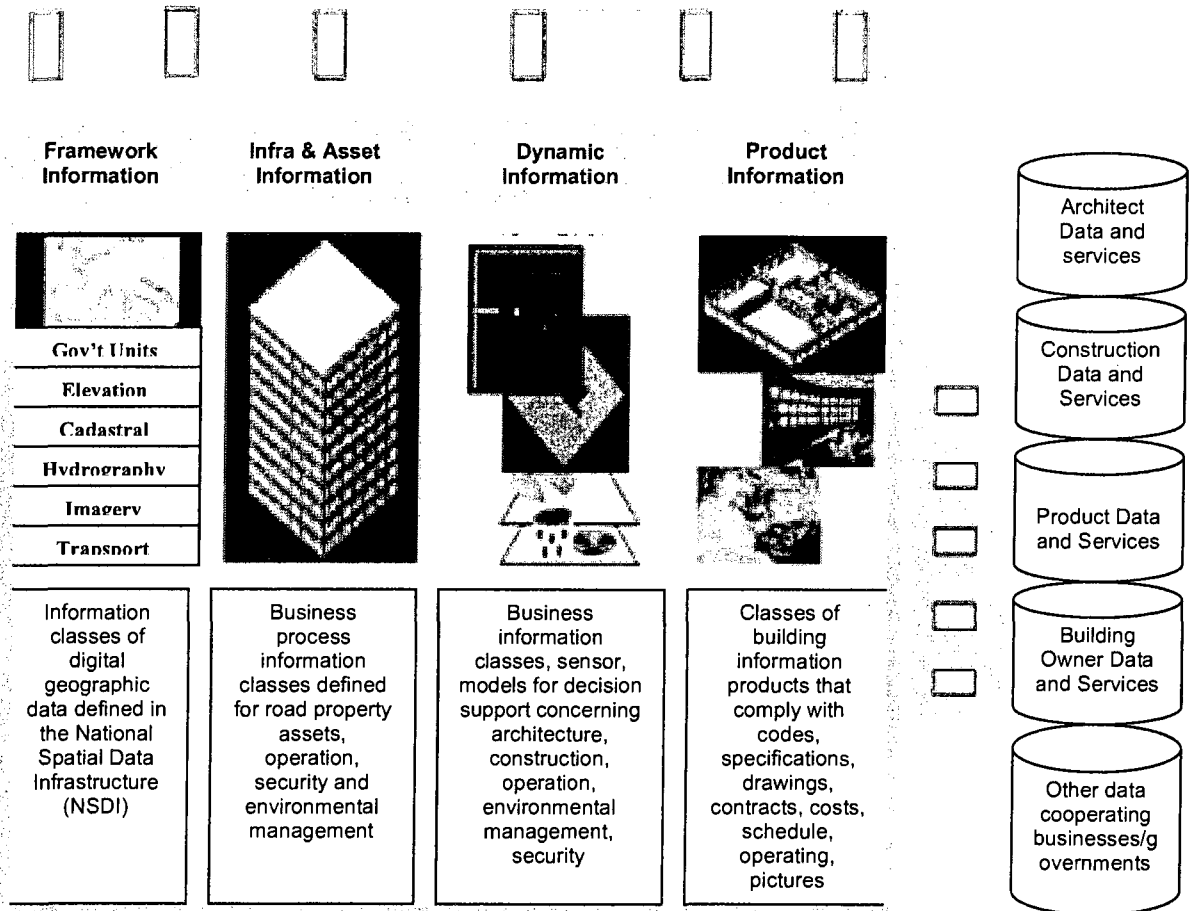


Figure 2.13 OGC, Building Smart Alliance (adopted from OGC 2008)



#### **2.4.5.2 Product Models**

Although water and sewer utilities are realizing the benefits of GIS technology for engineering, construction, and operations purposes, the semantics in different models have limited the sophistication of data analysis. Dependencies between different features were not explicitly defined. Therefore, inconsistency exists when different users modified the data. To avoid this limitation, some GIS data models (e.g. Halfawy et al. 2004b) represent spatial features using geometric primitives, define attribute data associated with these features, and create the inter-relationships between different features to ensure data correctness and consistency. One of the most important ISO data modeling standards that has been in active use within the AEC/FM (Architecture, Engineering and Construction/Facilities Management) in the last decade is the ISO 10303 Standard for the Exchange of Product Model Data (STEP) (Halfawy et al. 2000). A STEP-based standard data model provides a standard schema for representing the data and a neutral file format that enables different applications to efficiently interoperate by sharing and exchanging information. Many of the AEC/FM modeling effort in STEP is performed under ISO/TC184/SC4 ([www.tc184-sc4.org](http://www.tc184-sc4.org)). However, in the municipal asset management domain, the data modeling capabilities of UML and XML schema can achieve better and more cost-effective results. For example, the new XML 1.1 schema standards ([www.w3.org/XML/Schema](http://www.w3.org/XML/Schema)) can support the development of semantic rich object-oriented data models. Since UML models and XML schemas have been adopted by many commercial software tools, these models can be easily integrated with other forms of data, which potentially facilitate integrating data from different data sources. Moreover,

accessing and exchanging XML data over the Internet can be applied by using the web tools.

## **2.5 QUALITY ISSUES IN MUNICIPAL INFRASTRUCTURE**

### **2.5.1 Uncertainty of Spatial Data Quality**

For geotechnical applications, quality is a measure of the difference between the data and the reality that they present, if data are of poor quality, and use uncertain geographic information, they have little value. Twenty-four years ago, Nicholas Chrisman (1984) initiated a paper entitled “The role of quality information in the long-term functioning of a geographic information system (GIS)” in Ottawa. However, quality information was thought to be taken up as much space in data storage as geometric coordinates at that time and the quality concept had been ignored. After more than 20 years, efforts have been made by many GIS experts (i.e. Roger Tomlinson, Duane Marble, Michael F. Goodchild, etc.) to continuously characterize quality issues in spatial data management. In 1999, Tsai and Frost (1999) proposed the “Thoroughness as Measure of Investigation Quality” to address that the method for quantitatively evaluating for the quality of spatial data process as a whole issue for municipal infrastructure management still lacks. In 2005, Devillers and Jeansoulin (2006) pointed that the concept of data quality should be prominent in GIS and it is the responsibility for municipalities to develop a system that performs satisfactory while providing an appropriate level of safety and minimizing the use of financial and human resources. Moreover, assessment of variation in infrastructure projects is particularly important, such as water or sewer networks, because numerous excavation, surveying, tunneling, foundation, and environmental activities have

experienced cost overruns, unexpected delays, and occasional failure due to unforeseen variability in the subsurface.

### **2.5.2 Background of Six Sigma Principle**

In the pursuit of high spatial data quality, many quality aspects and methods have progressed considerably and we have witnessed: (1) the formal identification of spatial data quality elements (i.e. scale, visualization); (2) the inclusion of data quality information in data transfer standards (i.e. SDTS); (3) the introduction of international geographic data quality-related standards (i.e. FGDC); and (4) the widespread adoption of metadata entry tools for the production of metadata (Hunter et al. 2005). However, most of these methods focus on fell short of expectations of quality assurance aspect which aims for long-term spatial data management. In 2003, Lindermann et al. (2003) stated that Six Sigma principle is a statistics-based methodology, which relies on scientific methods to make significant reductions in customer-defined defect rates in an effort to eliminate defects from every product, process, and transaction. The Six Sigma principle can be represented on a normal distribution curve. The mean value is located at the center of this curve; the lower and upper limits are six times the standard deviation (sigma) from the mean. Since the normal distribution is symmetrical pattern, if the data set falls within  $\pm 6$  sigma from the mean as the upper and lower specification limits, it presents a 3.4 parts per million (ppm) as the quality level of  $\pm 6$  sigma. Motorola used this level (3.4 ppm) as the target level of the Six Sigma principle movement (Lindermann et al. 2003). It signifies only 3.4 defects/million parts or operations. Although the 3.4 ppm defect rate is an inappropriate goal for municipal infrastructure processes, the fundamental concept of taking the Six Sigma principle associated with minimizing the

defect rate along with different sigma levels and the continuous pursuit of productivity improvement can be considered an innovative approach for achieving a high level of process quality.

From a management point of view, Six Sigma principle is a problem-solving framework which seeks to identify and remove the causes of defects and errors in manufacturing and business processes (Snee 1999). Six Sigma has two key methodologies: DMAIC (define-measure-analyze-improve-control) and DMADV (define-measure-analyze-improve-design-verify) both inspired by Deming's Plan-Do-Check-Act cycle. DMAIC is used to improve an existing business process; DMADV is used to create new product or process designs (Defoe and Barnard 2005). In municipal infrastructure management, we suggest use DMAIC which emphasizes the identification and avoidance of variations. Moreover, Six Sigma principle underlines the explicit recognition of the root causes of defects and statistical process control to sustain continuous improvements (Abdelhamid 2003). Unlike manufacturing industry, construction work is characterized by its fragmented and project-oriented work processes. The use of the Six Sigma principle in the construction industry for quality control by assessing performance first appeared in 2000. Buggie (2000) introduced the Six Sigma principle as one of the approaches to augmenting productivity, which concentrated on reducing cycle time and eliminating any defects or errors engaged in the process. Cha and O'Connor (2005) described the Six Sigma principle as one of 44 state-of-the-practice tools for value management applicable to a construction project, based on the identification efforts conducted by the Construction Industry Institute (CII) project team 184. However, most municipalities that have

attempted to tackle data quality have tried tactical solutions such as defining regulations, giving the mandatory content according to the intended use, increasing the accuracy level of measurement, etc., to improve quality within a single application or within a single business process. Although this approach may mitigate part of the problem in the short-term, such limited initiatives generally fail to achieve long-term data quality improvement on a broad scale. To overcome this problem, the data quality issues for MIMS require an organization-wide approach that addresses organizational, process and technology infrastructure. A comprehensive municipal GIS should provide a common platform for data collection, storage, authorized and secure access to spatial and non-spatial data, harmonize the work flow of respective departments and disseminate information for the benefit of public at large (Subash and Pakak 2003).

## **2.6 SUMMARY AND CONCLUSIONS**

In this chapter, the literature about the current MIMS applications has been reviewed. There is significant development in using GIS to satisfy the MIMS requirements such as spatial data management, condition assessment, data sharing, etc. However, data collection in these applications is still done based on paper-based methods, which makes the collected data inaccurate and inefficient. In addition, the lack of quality control in spatial data and interoperability standards affects the integrated MIMS development. Therefore, the latest spatial technologies and spatial data interoperability standards should be addressed. This chapter introduced the latest spatial technologies, specifically review the data collection and inspection methods for underground infrastructure, and provides a comprehensive review about spatial data interoperability standards. This

review resulted in recommending the use of GIS to achieve MIMS integration and providing available spatial data collection approaches and data modeling standards to municipal infrastructure management.

Quality is another important issue in MIMS applications. In order to meet the requirement to ensure the overall reliability and interoperability of municipal infrastructure operations, the Six Sigma principle DMAIC methodology has been introduced. This is new initiative which applies Six Sigma principle in municipal infrastructure. Although there are some applications on construction material, safety or health, the operation in asset management is first proposed. The detailed spatial data quality process will be discussed in Chapter 3 and the related case studies will be presented in Chapter 4.

## CHAPTER 3: RESEARCH METHODOLOGY

### 3.1 INTRODUCTION

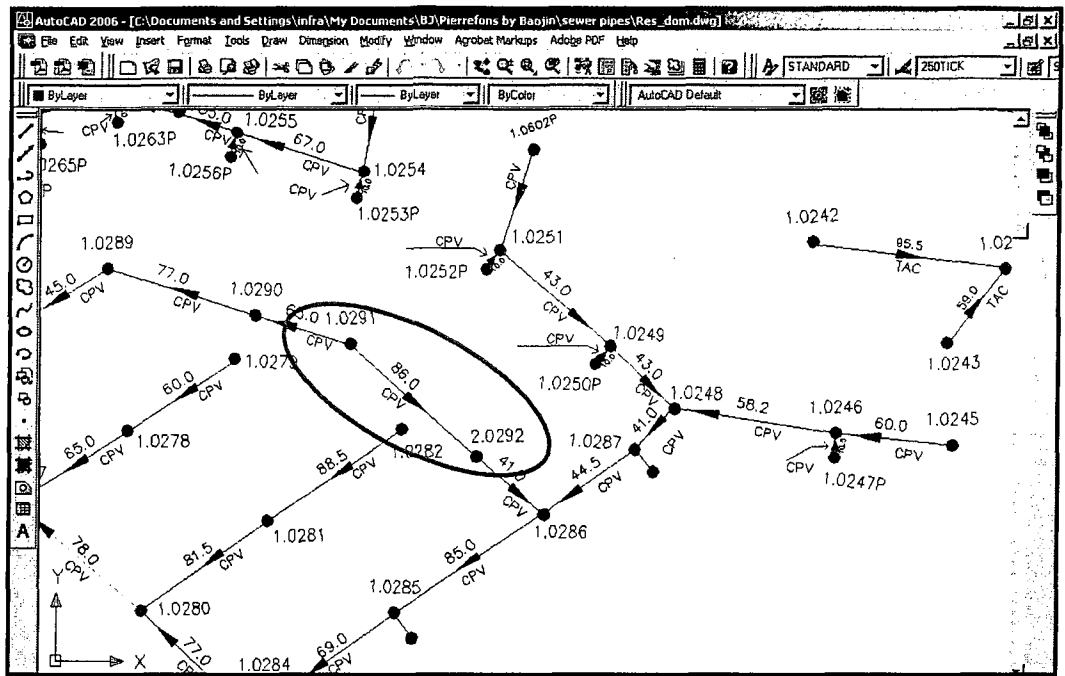
Applying quality in MIMS domain requires suitable spatial data representation and modeling methods. The user of such applications (e.g., facility inspector) needs to retrieve spatial information from drawing documents and interact with its related attribute (non-spatial) data. To achieve the full benefits of these data, it is important to improve the availability, completeness, accuracy, and integration of the data. This research proposes a comprehensive GIS data quality process. It can be used as a platform for driving an organization-wide data quality initiative for municipal infrastructure utilities, such as a sewer network. Furthermore, following a standard data representation is a key issue in such applications to enhance interoperability and achieve better results of minimizing errors and improving data quality.

As introduced in Chapter 2, GISs have been useful for many MIMS applications; however, the interoperability has not been well considered in those applications. A facility data comes from different sources of information, and each source has its own data format. For example, the facility spatial model can be in different formats, such as DWG, which is the defacto standard of Autodesk or the Shape file, which is the format of ESRI (1998). On the other hand, the process data can be represented in different formats such as EXL and MDB, which are the formats of Microsoft Excel and Access, respectively. Many municipalities are still using separate CAD files and databases to manage their infrastructures. In such a MIMS, the database records of each component of

an asset are referenced by the component's unique identification number (IDs), and use CAD software to graphically represent the asset's spatial data, define and attach the non-spatial (attribute) data to the graphical objects (Halfawy and Figueroa 2006). Working with different data sources that are not integrated may cause redundancy or data conflicts as in the following example (Figure 3.1). In Figure 3.1 (a), the map spatial data, which is recorded as a CAD file, presents that one sewer pipe starts from manhole 1.0291 and ends to manhole 2.0292; however, its attribute table shown in Figure 3.1 (b), which is saved in Excel, describes that this sewer pipe ends to manhole 1.0292.

In this chapter, a framework for data Quality Assurance (QA) process is introduced. Four main components: data collection, data accuracy analysis, data improvement and sharing, and data maintenance will be discussed in detail.





(a)

	RUE	No C.	DE REGA	A REGAR	Longueur la section	Diamèt (mm)	Matéria	Pent	Année de Construction	Radie AMONT	Radie AVAL	Dessu AMONT	Dessu AVAL	Année ins. Télévisé
68	BELVEDÈRE du	B085	RD.1.0488	RD.1.0489	32.0	300	TBA	N.D.	1959	N.D.	16.44	20.91	21.00	1994
69	BELVEDÈRE du	B085	RD.1.0489	RD.1.0490	69.5	300	TBA	N.D.	1959	16.44	N.D.	21.00	20.69	1994
70	BELVEDÈRE du	B085	RD.1.0490	RD.1.0491	94.0	300	TBA	N.D.	1959	N.D.	N.D.	20.69	20.66	1994
71	BERGERON	B106	RD.1.0338	RD.1.0371	63.3	300	TBA		1961					
72	BERGERON	B105	RD.1.0371	RD.1.0372	53.3	300	TBA	-0.60%	1961	20.82	21.14	23.66	23.80	
73	BERGERON	B105	RD.1.0338	RD.1.0373	51.8	300	TBA	-1.00%	1961	20.06	20.58	23.30	23.51	
74	BERGERON	B105	RD.1.0373	RD.1.0374	51.8	300	TBA	-1.00%	1961	20.58	21.10	23.51	23.67	
75	BOISJOLI	B185	RD.1.0276	RD.1.0288	77.0	250	CPV	-0.58%	1987	20.37	20.82	24.10	23.80	1988
76	BOISJOLI	B185	RD.1.0288	RD.1.0289	45.0	250	CPV	-0.58%	1987	20.82	21.08	23.80	24.10	1988
77	BOISJOLI	B185	RD.1.0289	RD.1.0290	77.0	250	CPV	-0.58%	1987	21.12	21.57	24.10	24.55	1988
78	BOISJOLI	B185	RD.1.0290	RD.1.0291	53.0	250	CPV	-0.58%	1987	21.57	21.88	24.55	24.84	1988
79	BOISJOLI	B185	RD.1.0291	RD.1.0292	86.0	250	CPV	0.52%	1987	21.88	21.43	24.84	24.40	1988
80	BOISJOLI	B185	RD.1.0292	RD.1.0286	41.0	250	CPV	0.51%	1987	21.43	21.22	24.40	24.20	1988
81	BONNY	B190	RD.1.0163	RD.1.0164	79.3	300	TBA	N.D.	1959	N.D.	N.D.	N.D.	N.D.	
82	BONNY	B190	RD.1.0164	RD.1.0165	79.3	300	TBA	N.D.	1959	N.D.	23.86	N.D.	26.60	
83	BONNY	B190	RD.1.0165	RD.1.0166	33.6	300	TBA	0.27%	1959	23.86	23.77	26.60	26.60	
84	BONNY	B190	RD.1.0166	RD.1.0167	26.9	300	TAC	0.23%	1959	23.77	23.71	26.60	26.50	
85	BONNY	B190	RD.1.0167	RD.1.0168	51.8	300	TAC	0.10%	1959	23.71	23.66	26.50	26.70	
86	BONNY	B190	RD.1.0168	RD.1.0169	35.1	300	TAC	0.31%	1959	23.66	23.55	26.70	26.90	
87	BONNY	B190	RD.1.0169	RD.1.0170	35.1	300	TAC	0.26%	1959	23.55	23.46	26.90		
88	BONNY	B190	RD.1.0170	RD.1.0177	39.6	300	TAC	0.25%	1959	23.46	23.36	26.90	26.50	
89	CERES	C080	RD.1.0245	RD.1.0246	60.0	200	CPV	1.00%	1993	23.20	22.60	25.10	25.10	

(b)

Figure 3.1 Example of data conflict caused by different data sources

### **3.2 FRAMEWORK FOR MIMS DATA QUALITY PROCESS**

To date, most municipalities that have attempted to tackle data quality have tried tactical solutions, such as defining regulations, giving the mandatory content according to the intended use, increasing the accuracy level of measurement, etc., to improve quality within a single application or within a single business process. Although this approach may mitigate part of the problem in the short-term, such limited initiatives generally fail to achieve long-term data quality improvement on a broad scale. To overcome this problem, the data quality issues for MIMS require an organization-wide approach that addresses organizational, process and technology infrastructure. The proposed spatial data QA process (Figure 3.2) is an all-encompassing management approach combining technical aspects of quality, qualitative methods, and human resources, in a system designed to meet all requirements of spatial data workflows, procedures, and standards. This process is inspired by Six Sigma DMAIC concept (DMAIC 2008), focusing on spatial data related information about three aspects of a geographic feature: topology (the type of geographic feature) (Toth and deLima 2005), location, and spatial dependency. A GIS topology is a set of rules and behaviors that model how points, lines, and polygons share geometry (GIS Topology 2005). For example, adjacent soil polygons share edges. The issue of topological integrity is highly crucial in internal error-checking procedures. Location is also an important aspect in this process because location uncertainty may seriously affects the positional accuracy, such as the difference in map projections, map scale, data acquisition, and data processing (Kim 2006). Spatial dependency is "the propensity for nearby locations to influence each other and to possess similar attributes" (Goodchild 1992). Different from ordinary data, spatial data are characterized by their

value and by their location. The value can be a quantitative value, such as groundwater depth, degree of pollution, or a qualitative label, such as land use, land cover, or the type of building, separately. Therefore, spatial dependency is a measure for the degree of associative dependency between independently measured values in situ. ordered set. Underground network features, such as material types and slope, for instance, are more likely to be similar at points two meters apart than at points two kilometers apart. The process is made up of four parts: data collection, data accuracy analysis, data improvement and sharing, and data maintenance. It can be used as a platform for driving an organization-wide data quality initiative for municipal infrastructure utilities.

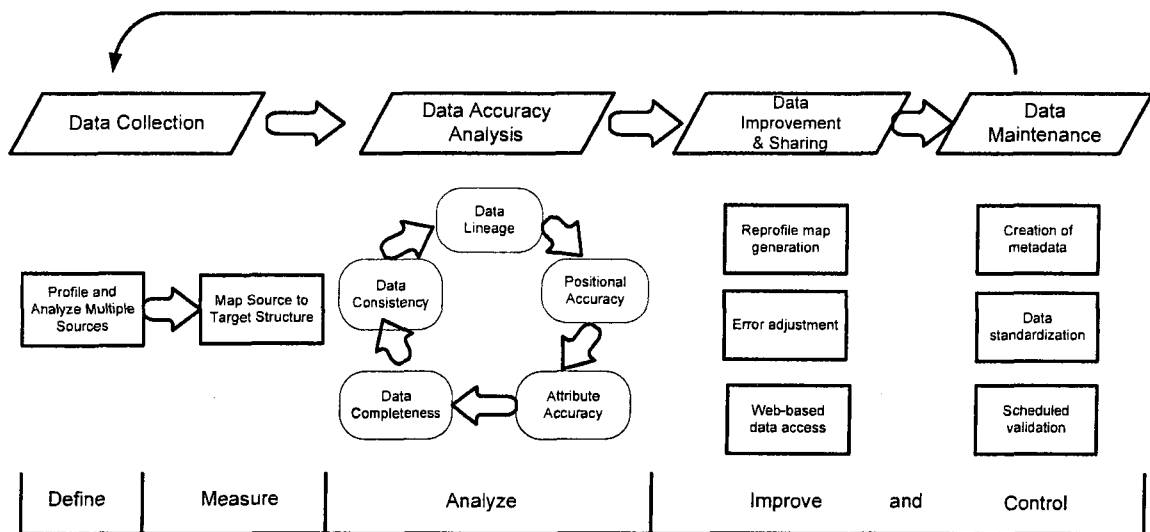


Figure 3.2 Proposed spatial data quality assurance process

In the QA process, the data collection is firstly considered because there are various types of data sources such as data collected from surveying (e.g., GPS), data purchased from special organization or data downloaded online. In fact, data will be migrated, converted, inherited, and derived during lifecycle of facility operation. Using the proposed QA approach which allows GIS data expectation to match all requirements of data workflows

is extremely important. Therefore, in data collection part, we separated actions into two steps: profile and analyze multiple data sources, and map data sources to target structure. Once the target data structure is established, assessing or analyzing accuracy in GIS data is necessary. One way to assess quality is to find and record errors. However, data accuracy measurements are difficult to quantify because it depends on how the data will be used. For example, the attribute information represented in a CAD file, such as the length of a pipe, is added as a character string. That means the value could not be used in calculations.

Geographical observations describe phenomena with spatial, temporal, and thematic components (Berry 1964; Sinton 1978). Unlike manufactured products which have physical characteristics that allow quality to be easily assessed, geospatial data have intangible properties such as *completeness* and *consistency*. Moreover, no map is 100% correct and the highest accuracy possible will not be any better than the least accurate input overlay. Therefore, the objective of dealing with inaccuracy should be not to eliminate it but to manage it. Numerous standards for spatial data quality have been proposed, one of the earliest and most widely used standards is the US Spatial Data Transfer Standard (SDTS) (2008). It provides standard definitions of data elements, standardization for data transfer and descriptive metadata about database contents. There are five components of data quality which have been defined (Veregin 1999): (1) Lineage (provenance of data); (2) Positional accuracy (accuracy of positional information); (3) Attribute accuracy (accuracy of attribute information); (4) Data completeness refers to the degree to which all features are included in a data set. This completeness can be

defined over space, time, or theme. For example, considering a database which depicts the locations of sewers in Montreal City, including the sewers of Pierrefonds Borough which were installed before 2002, this database would be incomplete if it included only sewers in Pierrefonds Borough (incompleteness in space, since Pierrefonds Borough covers only a portion of Montreal), or only sewers placed before 2002 (incompleteness in time, since sewers may have been added after 2002), and (5) Logical consistency (absence of logical contradictions in data); for example, only one point may exist at a given location, lines must intersect at nodes, polygons are bounded by lines.

After data accuracy analysis, data can be improved and shared by other users. In this process, the Web-based data access method will be introduced. Finally, the data quality criteria will be set in data maintenance part which includes creation of metadata, data standardization, and scheduled validation.

In this process, the DMAIC methodology of Six Sigma quality has been applied as the main principle in this framework management not only because this methodology can be utilized to examine quality issues and make recommendations for improving a target organization's quality but also because DMAIC is a systematic, closed-loop process for continued improvement that is scientific and fact based. In the Define phrase, the problem/ opportunity, process and requirements have to be defined. Therefore, investigation multiple data sources becomes priority in this step. In the Measure phrase, well compiling and displaying data is necessary for further data analysis. In the Analyze phrase, it is a scrutinize step to find improvement opportunities and advanced statistic

analysis could be applied to provide fact based result of data accuracy. As we discussed before, spatial data will be examined from five components of data quality. Finally, regarding the proposed methodology for general application, we combine Improve and Control phrases together for generating solutions and tracking the stability of the improvements. When applying to different activities, these two phrases could be separated in detail based on different circumstances.

### **3.3 DATA COLLECTION**

#### **3.3.1 Sources of Data Errors**

There are many sources of error that may affect the quality of a dataset. Some are obvious, such as data redundancy which can be automatically identified by the software itself, but others can be difficult to discern. Therefore, particular care should be devoted to error sources. Error recognition and what level of error is tolerable and affordable must be acknowledged and accounted for by GIS users. Burrough (1986) divides sources of error into three main categories: (1) Obvious sources of error: age of data, area cover, map scale, density of observations, relevance, format, and accessibility; (2) Errors resulting from natural variation or from original measurements: positional accuracy and accuracy of content; and (3) Processing errors: these are the most difficult to detect by users and must be specifically looked for. They require knowledge of the information and the systems used to process it, because they can occur in multiple sets of data being manipulated in a project. These errors can be subcategorized into numerical errors, errors in topological analysis, classification and generalization problems, and digitizing and geocoding errors. Applying topological analysis helps to find drawing mistakes, e.g. a polygon should be a closed area; classification can be utilized to distinguish different

types of sewer pipes; generalization can help to extract the common character between water pipes and sewer pipes (Bedell et al. 1997).

### **3.3.2 Data Collection Methods**

Collecting accurate asset data includes spatial and non-spatial information collection. Generally, spatial data represent a geographical location of the features and non-spatial data provide the descriptive information of the facility, such as the name, number, function, construction year, diameter, material, inspection date(s), and length. These data were generally gathered over the years and mainly were collected after response to citizen's claims (Qasem and Hammad 2006). Although many municipalities have developed their own proactive tools for managing these inventories based on their experience and their administrative regulations, some missing information still exists and will lead to data inconsistency in the future. Therefore, it is important to investigate the possible data sources. Based on the data classification from NRC (2003), *physical data*, *operational data* and *land use and environment data* have been defined. The potential data sources could be: (1) network inventory plans, (2) drawing maps, (3) existing reports, (4) operation and maintenance reports, and (5) inspection reports. Moreover, using efficient and measurable data collection technologies should be considered to acquire the necessary data with acceptable quality.

Regarding the quality of spatial data, various new technologies for data collection have been introduced to enable more efficient and timely data gathering, such as the GPS, Automated Vehicle Location (AVL), Road Weather Information Systems (RWIS),

CCTV, and so on. However, each of these technologies has certain limitations. For example, GPS receivers are unable to collect accurate coordinates effectively while they are moving; even the best GPS receivers can only provide centimeter level results while moving at high speeds. In this case, AVL technology is more suitable because AVL devices use specialized data collectors/ transmitters mounted on moveable assets, such as emergency vehicles, snowplows, and dump trucks (NRC 2004). Depending on the variability of the attribute data, such as utility age, diameter, and material, it is important to select a data collection system and define what can be collected with this system. For example, RWIS technology can be applied to collect data about road surface temperature, visibility, solar radiation, road surface status (dry, wet, icy), and type and amount of precipitation. In the case of underground facilities, such as sewer and water networks, CCTV technology efficiently provides sufficient records which supplements GIS attribute data.

### **3.3.3 Data Availability**

The availability of infrastructure database containing all the information required to assess the current situation of infrastructure facilities is considered a major challenge. It would be a tedious process to start collecting and creating this information from scratch. On the other hand, one promising way to start building this huge database is through investigating the available data about infrastructure facilities from different perspectives, such as the data available from *Info-excavation*, *Common Ground Alliance* (CGA 2005), *APISQ* (Québec Alliance for the Protection of Underground Facilities 2008), and the Ontario Regional Common Ground Alliance (ORCGA 2007). These alliances and similar



ones in North America and Europe aim to protect infrastructure facilities from any damage caused by any type of excavation (Info-excavation 2006). They provide their services through a client call center which takes some initial information about the location and type of the expected excavation activity. Then, they refer to their infrastructure facility database and inform the caller about the possible presence of infrastructure facilities under the excavation area. The center will then inform the infrastructure facility owner about the client request. The facility owner will send a locator to indicate exact location of his company's facility.

### **3.3.4 Data Conversion**

Recently, realizing the inefficiencies in the use of paper-based records, many municipalities have converted map drawings from paper maps to CAD maps (estimated at 50,000 plans/year) (Halfawy and Figueroa 2006). Instead of scanning the drawings, most maps have been redrawn by a software to enable better QA and QC (Quality Control). As the demand for CAD data grew higher, CAD records exhibited some serious limitations. First, data access was limited to CAD users, while others had to rely on mass printouts kept at the drafting counter. Second, the CAD models represented features as geometric primitives that lack the semantics needed to verify and validate the data. Third, most CAD systems exhibit limited data management functionality, which impeded the ability to efficiently maintain the linkages between the spatial objects and their attribute data (Halfawy and Figueroa 2006). As was discussed in Section 2.3.1, unlike CAD drawings, GIS emphasize the definition and use of semantic maps that are linked to a database to store and manage the attribute data. By integrating spatial and non-spatial data, GIS enable efficient and partially automated data verification and validation; both

are critical for ensuring the quality and reliability of the asset data. The flowchart of implementing GIS-based spatial data quality management is illustrated in Figure 3.3. From the flowchart, we can see that spatial and attribute data are integrated into a GIS system. Then, the extracted layers can be analyzed by different QA methods, such as visual and automated inspections as will be explained in the following sections (McCain and Master 1998). Based on the analysis, maps' errors can be reported by classifying different sources of errors. Data improvement is realized by adjusting fixable errors into the appropriate use. Maps are regenerated to allow users to integrate with other systems, such as ERP and CRM. Meanwhile, data quality could be maintained by creating metadata, data standardization and scheduled validation.

### **3.4 DATA ACCURACY ANALYSIS**

#### **3.4.1 Position Measurement**

Position measurement is an important technological challenge when collecting accurate spatial information for surveying objects because a meaningful accuracy assessment requires an accurate reference dataset (Foody 2001). How spatial dataset representation will crucially affect the result of further data accuracy analysis. Components of data quality typically include the positional accurate measurement for well-defined points, polygons, quality of transfer functions and the linking of data quality of inputs to data quality of the results. Regarding to position measurement, we focus on the accuracy of point's coordinates which is the necessary information describing an object in this QA process. Two dimensions (x and y) point coordinates have been measured using GPS as will be discussed in Chapter 4.

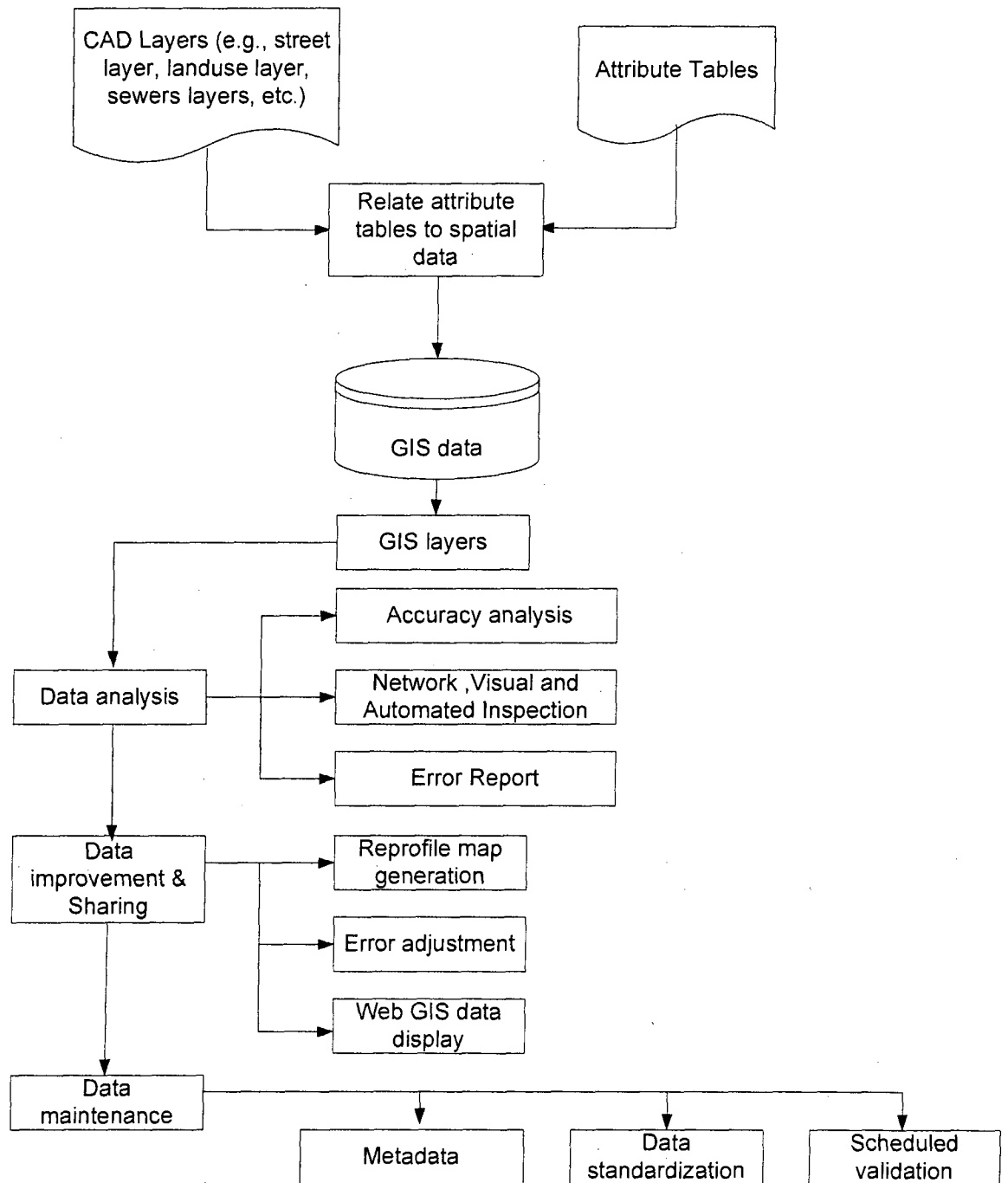


Figure 3.3 GIS-based spatial data quality management flowchart for MIMS

### 3.4.2 Network Analysis

Network analysis allows creating and managing sophisticated network datasets and generating routing solutions among other functions (Olaleye and Sangodine 2002). It is a powerful method for routing, providing a framework for network-based spatial analysis (i.e., location analysis, drive time analysis, and spatial interaction modeling). In this QA process, network analysis is mainly used to find the topology errors when analyzing polylines (sewer pipe). By applying network analysis to find the shortest route between a manhole and a sink (sewer treatment plant). If there is gap or disconnection in polylines, the network analysis will automatically detect this gap as shown in Figure 3.4. Therefore, network analysis is very useful when investigating the topology of spatial data.

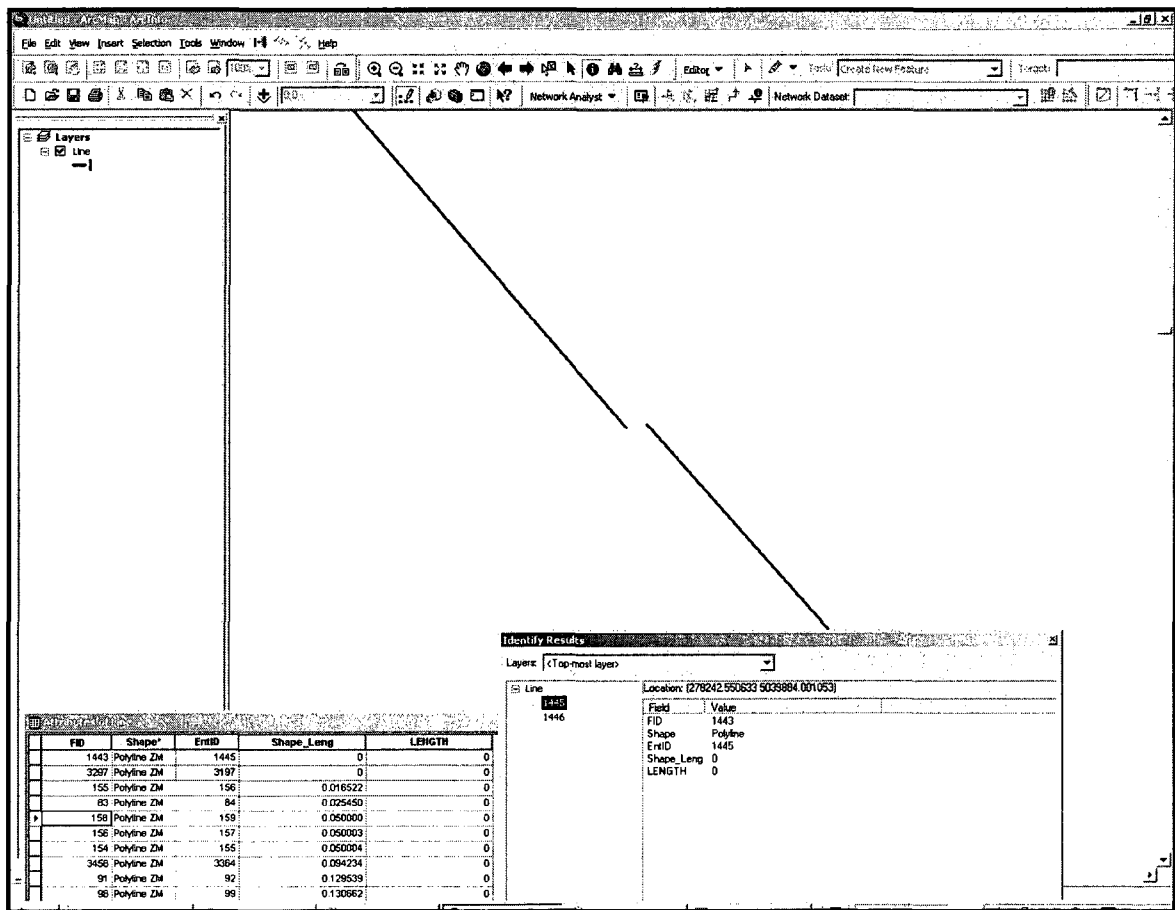


Figure 3.4 Polyline's disconnection when applying network analysis

### 3.4.3 Accuracy Measurement

Spatial data accuracy is an umbrella term that encompasses concepts such as uncertainty, standardization, positional tolerance, confidence intervals and others (Burkholder 2004). How to correctly measure the spatial data position accuracy depends on dimensionality. Metrics are well defined for point entities, but widely accepted metrics for lines and areas have yet to be developed. For points, an error is usually defined as the discrepancy (normally Euclidean distance) between the encoded location and the location as defined in the specification. Errors can be measured in any one of, or in combinations of, the three dimensions of space.

The most common measures are horizontal error (distance measured in x and y simultaneously) and vertical error (distance measured in z) (Figure 3.5) (Veregin 1999). For lines and areas, the situation is more complex since there is no simple statistical measure of error that can be adopted. Errors in lines arise from the errors in the points that define those lines. However, as these points are not randomly selected, the errors present at points cannot be regarded as typical of errors present in the line (Goodchild 1991).

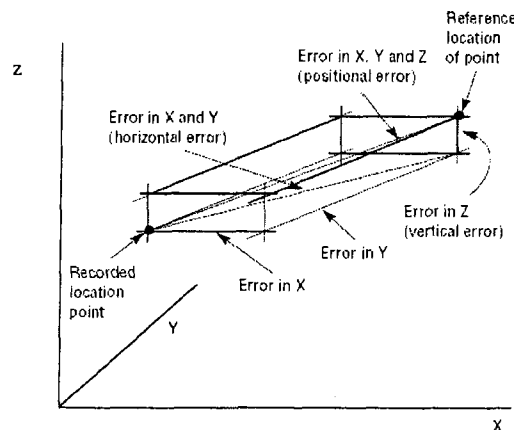


Figure 3.5 Positional error measurements (Veregin 1999)

There are various metrics that can measure the spatial error for sets of points. One common metric is root mean squared error (RMSE), which is computed as the square root of the mean of the squared errors (Beard 1989). From a statistics point of view, RMSE is useful in describing differences between two datasets (Rasford et al. 2003) and is widely used to describe positional accuracy, including horizontal and vertical accuracy (FGDC 1998). RMSE is a measure of the magnitude of error but it does not incorporate bias since the squaring eliminates the direction of the error (Veregin 1999). Moreover, the National Standard for Spatial Data Accuracy (NSSDA) defines positional accuracy of spatial data by implementing a well-defined statistics (i.e., 95% confidence level), in both digital and graphic form, as derived from sources such as aerial photographs, satellite imagery, or other maps (Infotech 2007)

### **3.5 DATA IMPROVEMENT AND SHARING**

#### **3.5.1 Data Sharing and Access**

The client-server architecture of most GIS software enables browser-based thin clients to efficiently access the data over Internet/Intranet networks. The browser-based architecture can support mass deployment of spatial data across various departments and stakeholders within municipalities in a cost-effective manner. Therefore, in order to achieve the integration among departments in a municipality, these departments must have well organized and standardized asset management procedures. In order to share GIS data more productively and efficiently, web-based GIS technology enables application managers to distribute maps, models, and tools to all application users in an organization in a simple way that is similar to browsing the Internet. Then, any

authorized users can query accurate, up-to-date data without the need to install desktop GIS software. Server GIS complements desktop GIS by allowing GIS analysts to author cost-effective maps and geoprocessing tasks on their desktops and publish them via a server using integrated tools. GIS functions can then be delivered as services throughout the enterprise to a large number of users over the Internet/Intranet. Enterprise GIS users connect to a central GIS server using traditional desktop GIS as well as Web browsers and mobile computing devices. For example, ESRI ArcGIS server (ESRI 2008) is a software package compliant and interoperable with other enterprise software such as CRM or ERP systems. It offers support for interoperability standards in the GIS domain (Open Geospatial Consortium, 2008).

### **3.5.2 Coordinate Systems and Map Projections**

When collecting point coordinates, the projection and coordinate system should be addressed. All spatial data must be properly registered to the earth's surface. A projection system must be used in conjunction with a coordinate system in order to compile 3D data onto a 2D flat surface map. The most general coordinate system to use would be spherical polar coordinates that are two angles and a distance from the centre of the Earth (WolframMathworld 2008). These coordinates are more familiar by latitude and longitude giving the angle from the centre of the Earth, in terms of the south north angle across the equator (latitude), and west east angle from Greenwich (longitude). However the Earth is not a perfect sphere, and there are several factors which affect the shape of the Earth, such as the locations of the continents and the pull of the moon. Therefore, there are different versions of reference ellipsoids to define the distortion caused by the

continents. The most commonly used coordinate is the WGS-84 (World Geodetic System) datum, which is an earth-centered datum used to support the GPS and it gives the equatorial radius as 6378.137 km and the polar radius as 6356.752 km (Madhwal 2002). Therefore, when acquiring data from two different sources, the two data sets should have a common geoid model to define the projections of both data sets. If we do not account for the uses of different geoids, serious errors can result.

Meanwhile, various projections are available, such as the Universal Transverse Mercator (UTM) and Modified Transverse Mercator (MTM). Each projection method may cause distortion of the area, distance or shape of spatial objects. Particular projections are more suitable for specific locations and help to control distortion, but no map projection completely rules out distortion. The decision on which projection to use will depend on the region which is to be represented.

### **3.5.3 WebGIS Application**

As discussed in Chapter 2, development of the Web and expansion of the Internet provide two key capabilities that can greatly help municipalities or geoscientists. First, the Web allows visual interaction with data. By setting up a Web Server, clients can produce maps. Since the maps and charts are published on the Internet, other clients can view these updates, helping to speed up the evaluation process. Second, the geospatial data can be widely accessible. Clients can work on it from almost any location where there is Internet. However, to provide a successful WebGIS application is more difficult than simply buying the appropriate hardware and software. Several strategies have been proposed to



provide successful implementation. The implementation strategies should be scientifically assessed and modified so that the requirement of the project can be met with minimum cost and time (Alesheikh et al. 2002). Figure 3.6 shows the WebGIS development cycle, which is described in terms of 7 major activities starting with the requirement analysis and ending with on-going use and maintenance of the WebGIS system.

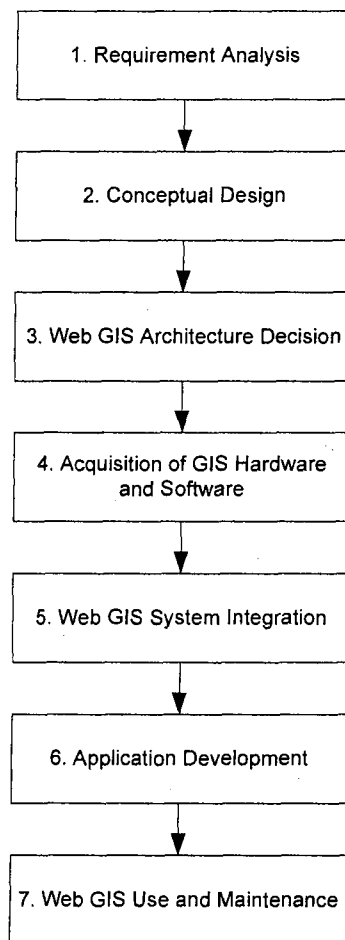


Figure 3.6 WebGIS Development Cycle (adopted from Alesheikh et al. 2002)

### 3.5.4 WebGIS Architectures

Based on the ArcGIS Server 9.2 (ESRI 2008) framework shown as Figure 3.7, WebGIS is similar to the client/server typical three-tier architecture. The geo-processing is

breaking down into server-side and client-side tasks. A client typically is a Web browser. The server-side consists of a Web Server, WebGIS software and a database. In this framework, the sever simply has the proprietary GIS running, and adds a client interface at the client side and a middleware of the server side to communicate between the client and the proprietary GIS software. The development of Object-Oriented Programming (OOP) makes it possible to produce software components, and sends them to the client before running them in the client machine, such as Java classes (Helali 2001). Therefore, both thin and thick-client system could be considered to design WebGIS architecture.

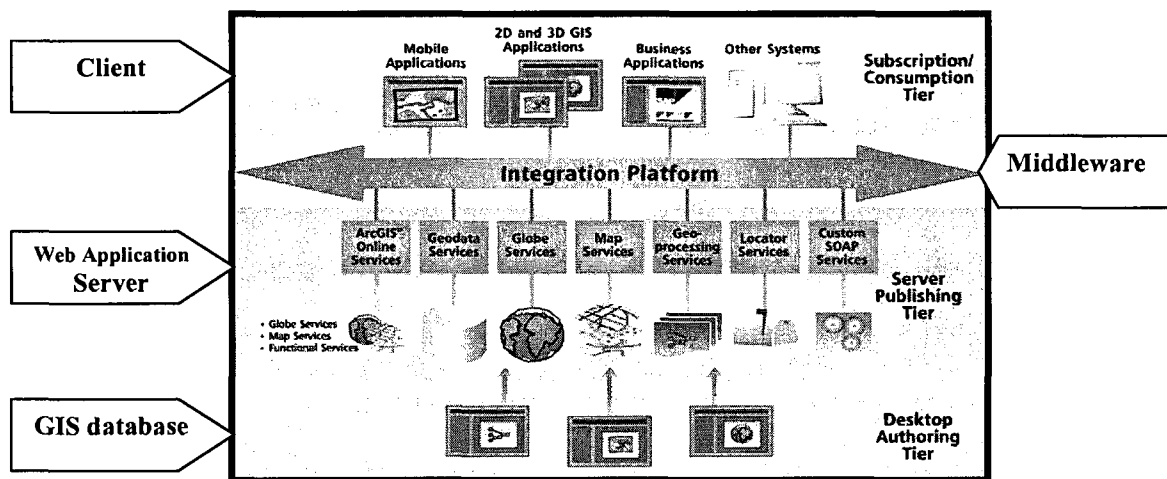


Figure 3.7 Web services framework (adopted from ESRI 2008)

In typical thin client architecture, the clients only have user interfaces to communicate with the server and display the results, whereas the server manages the centralized resources. Figure 3.8 shows schematic communication between Web Browser, Web Server and GIS server. On the Web Server side, Common Gateway Interface (CGI), Web Server Application Programming Interface (API), Active Server Pages (ASP), Java Server Pages (JSP), and Java-Serverlets have been connected to the GIS system. Using this architecture, the users on the client side do not need to know about the linkage of the

IMS (Internet Map Server) at the server side. Major advantages of this architecture include ease for data updating, keeping the last version, database centralization, data updating flexible, easy construction and integration, and less cost. The disadvantages are less functions, larger data volume, longer response time, less interactive on client side, and limited data formats.

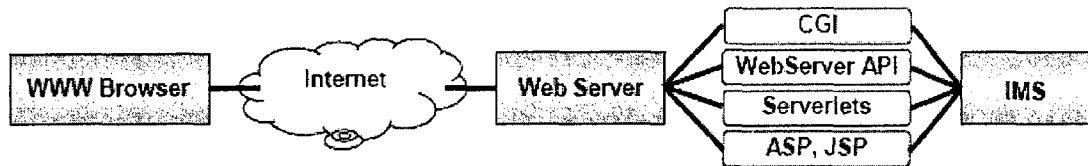


Figure 3.8 Thin client architecture applications (Alesheikh 2002)

On the other hand, Thick Client Architecture allows the browser's functionality to be extended because it focuses on the user interface functionality on the clients side. Most browser applications offer a mechanism that allows third tier programs to work together with the browser as a plug-in (Alesheikh 2002). In Thick Client Architecture shown as Figure 3.9, the user interface functionality has developed from single document to more interactive applications which could be Java script or Java applets to provide client-side functionality (Byong et al. 1998).

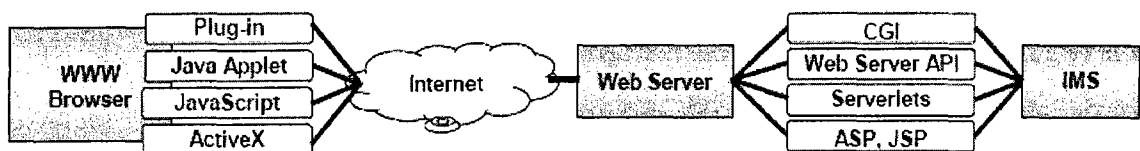


Figure 3.9 Thick client architecture applications (Alesheikh 2002)

Major advantages of this architecture include interactivity to various data formats, modern interface is possible and document standards are not required operations. The disadvantages are: additional software is required on the client side, platform/browser may be incompatible, and non-conformance limits.

### 3.6 DATA MAINTENANCE

#### 3.6.1 Surveying Standards

An accurate surveying standard will be able to maintain assets with new and historical details in datasets, and broaden the value of the cadastre, such as addresses and streets, which is crucial to city operations (ESRI 2008). Moreover, surveyors need the ability to electronically access survey data, upload it to their field equipment, record new data, and then manage it. Although there are some electronic field equipment such as Tripod Data System (TDS 2008) and Merrick (2008) offering this capability as does most surveying software, they lack the ability to properly manage the data when the surveyor completes the fieldwork and computations. Therefore, creating a survey data model is prior to set a fieldwork to find problems or errors in plans. The important role of data surveying standards and data standardization across whole departments is shown as in Figure 3.10.

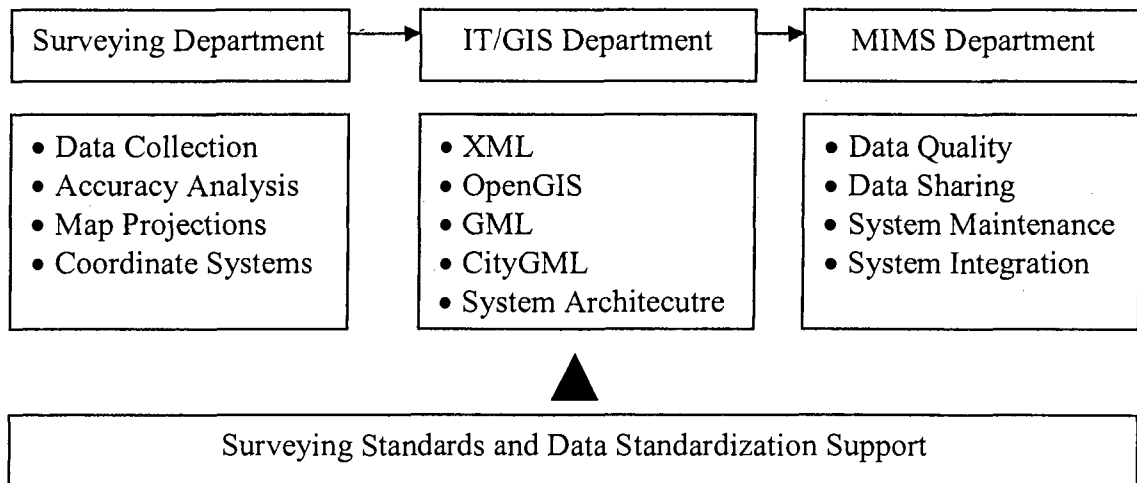


Figure 3.10 Surveying standards and data standardization across departments

Once the survey is completed, a surveyor builds up a survey database. This information is suitable for reuse in future surveys in the same area. This saves on searching and data entry time and cost of new surveys. An example of survey report is shown in Appendix K.

### 3.6.2 Data Standardization

The capability to index and reference various forms of asset information (Halfawy et al. 2004a) needs standardized methods to model, represent, and exchange municipal asset data. A standard infrastructure data model provides a schema for representing and exchanging the multi-disciplinary data of the infrastructure assets. It includes assets spatial and non-spatial objects, their attributes and their inter-relationships across different disciplines of asset planning, design, construction, operation, and maintenance. The standardization should be related to the acquisition, management, exchange and data applications (Takashi and Koji 2004). The Municipal Infrastructure Data Standard (MIDS) has been developed with a primary focus on public works by the *Tri-Committee for the Utilization of Information Technology*. The latest version MIDS 3.0 was published in 2005. The MIDS is a set of detailed specifications and rules for storing data about physical municipal infrastructure. The types of infrastructure that can be represented within MIDS include pavements, water valves, sanitary and storm water maintenance holes, rail crossings, playground structures and recreational facilities. Additional types can be added as easily as adding additional data records to the MIDS database, and data structure changes are not typically required (MIDS 2005). MIDS 3.0 provides a “blueprint” for municipalities and other infrastructure owners to build an asset management database. The unique advantage of MIDS is that it provides the foundation to design a database that can be shared by many application systems without data conversion or the duplication of information, thus reducing the risks, costs and time to obtain vital information for decision-making. Furthermore, MIDS promotes strategic management of information, enabling the rapid and accurate exchange of infrastructure

information with service delivery partners, regulatory agencies and other stakeholders (MIDS 2005).

Another way for standardizing municipal data, procedure and guidelines is to provide a standard guideline which presents a classifying system for the quality of existing subsurface utility data. This standardization process will allow design engineers, subsurface utility engineers, constructors and utility owners to develop strategies to optimize utility rehabilitation processes, reduce cost and manage risks caused by existing subsurface utilities. The ASCE (2003) suggests guidelines to standardize a unified system for collecting and classifying utility data. These guidelines contain four data quality levels used to categorize utility information ranging from level D to level A as follows: (1) Quality level D: information derived from existing record or oral recollections; (2) Quality level C: information obtained by surveying and plotting visible above ground utility features and by using professional judgment in correlating this information to quality level D; (3) Quality level B: information is obtained through the application of appropriate surface geophysics methods to determine the existing and approximate horizontal position of subsurface utility. This information is surveyed to applicable tolerance defined by the project and reduced on plan documents; (4) Quality level A: specific horizontal and vertical location of a utility obtained by the actual coverage (or verification of previously exposed surveyed utilities), and preceding measurements of surface utility, usually at a specific point. Minimally intrusive excavation equipment is typically used to minimize the potential for utility damage. Precise horizontal and vertical locations, as well as other utilities attributes are shown on plan documents. Accuracy is

set to 15 mm vertical and applicable horizontal survey and mapping accuracy as defined or expected by project owner. Based on the associated quality level, the risk will be evaluated for each level using risk management techniques. By applying this procedure, the risk associated with unknown subsurface utility that may exist at the time of construction will be mitigated and minimized.

### **3.6.3 Metadata**

Metadata literally means data about data (FGDC 2006). A useful feature of a data translation program is the ability to generate metadata. The resulting output file(s) will contain not only the attribute values that were entered in the field by the user, but also a variety of additional attribute data that are generated by the translation program (Gilbert 1996). Some of these generated attributes are common to the entire dataset, such as the GPS receiver type used to collect the data or the name of the original data file. Then, GIS users, such as municipalities, can search and retrieve spatial information by searching meta-databases that contain information about the content, quality, and characteristics of spatial data sets.

The Federal Geographic Data Committee Metadata Content Standard (FGDC 2006) defines a standard schema to document the content and structure of the metadata of spatial data. A metadata record is a file of information, usually presented as an XML document, which captures the basic characteristics of a data or information resource. It answers the questions of *who*, *what*, *when*, *where*, *why* and *how* related to the resources. Geospatial metadata is a type of metadata that is applicable to objects that have an

explicit or implicit geographic extent. Such objects may be stored in a GIS or may simply be documents, datasets, images or other objects, such as a data directory, data inventory, etc. Metadata could also provide information about the precision, consistency, and integrity of the data. The metadata schema defines a large list of metadata elements. However, only a subset of these elements, known as the core elements, must be defined by data creators. Other elements are left optional (Lange and Gilbert 1999).

#### **3.6.4 Scheduled validation**

Automated inspection and visual inspection are reinforced by QC program which checks for adherence to database design, attribute accuracy, logical consistency, and referential integrity. Automated inspection allows quick inspection of large amounts of data. It will report inconsistencies in the database that may not appear during the visual inspection process. During this process, both random and systematic errors are detected. After reviewing, studying and evaluating the quality of this information, the corrected data can be correlated with other GIS available data, such as road and water networks. Moreover, Scheduled database validation is also a must for large multi-user databases to identify some very important and potentially costly errors. These errors may point to a lack of control during the database update process. Errors or last minute changes in business rules, bugs in the maintenance application or inconsistent editing methods can all be detected during scheduled validation. It is always cheaper to fix a bad process than to correct hundreds or even thousands of errors that may have been introduced into a database (McCain and Masters, 1998). Several data inspection techniques will be introduced in Chapter 4.



### **3.7 SUMMARY AND CONCLUSIONS**

MIMS are facing unprecedented challenges due to the increasing number of aging infrastructure assets combined with declining maintenance budgets. Improving the quality of assets information is considered a crucial strategy to address these challenges. In this chapter, a framework of spatial data quality assurance process has been proposed. This process is summarized by the following four parts: data collection, data accuracy analysis, data improvement and sharing and data maintenance. The sources of data errors and data collection methods have been discussed. Data quality management flowchart and data accuracy analysis have been presented. Moreover, in the data sharing part, web-based GIS for an MIMS has been introduced.

## **CHAPTER 4: CASE STUDY**

### **4.1 INTRODUCTION**

Based on the framework of spatial data quality assurance as discussed in Chapter 3, a case study to apply data availability and completeness, data accuracy, and interoperability using data standards for data exchange to support MIMS is developed. In order to proof the efficiency of the proposed framework, four examples related to each step within the framework will be used to demonstrate this process through a case study on Pierrefonds Borough sewer pipe network in Montreal.

### **4.2 SELECITION OF DEVELOPMENT TOOLS**

Desktop GIS is used as an applied and developed environment for creating, editing, and analyzing geographic knowledge because it allows revealing data patterns, analyzing data relationships, and improving decision making. ArcGIS Desktop developed by ESRI (ArcGIS 2008) with related extensions (e.g. Network Analyst) is chosen because its features help create, edit, import, map, query, analyze, and publish geographical information.

Meanwhile, ArcGIS Server which offers centralized management of geodata, imagery, process models, and GIS-based Web services and applications has been chosen as the WebGIS application to share data. ArcGIS Server is a scalable, interoperable, and securable GIS Server.

In the data collection implementation, Trimble 5700 GPS receiver has been selected to collect the data coordinates because it can receive both L1 and L2 frequencies, which were introduced in Section 2.3.5, from GPS satellites to provide precise position data. It has centimeter-accuracy real-time positioning and up to 10 Hz position updates. The operating temperature is from -40° C to 65° C. More details about equipment setting, software configuration and troubleshooting guide are included in Appendix A.

In the spatial data analysis implementation, Visual Basic for Applications (VBA) programming for data integration and data inspection has been used. This approach is used to link data between AutoCAD and Excel.

For the creation of metadata, DataLogr (version 2.11), a data documentation software, has been used to describe data and allows metadata to be entered into established forms.

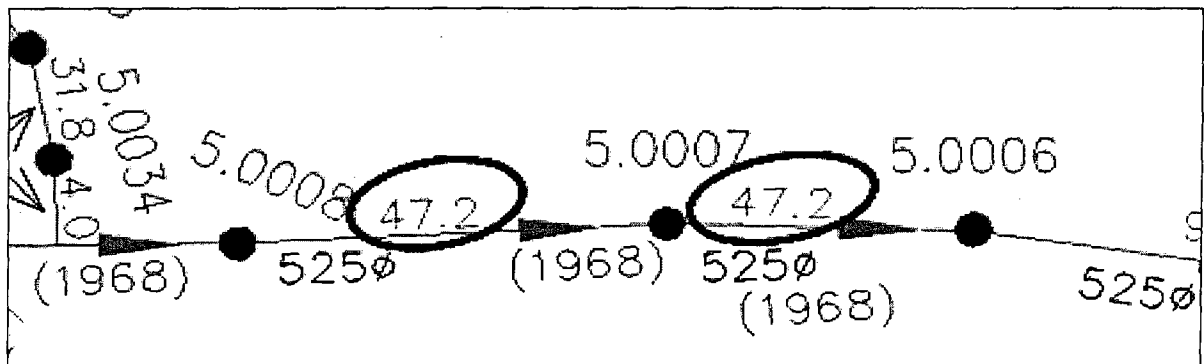
#### **4.3 BACKGOURND OF CASE STUDY**

Pierrefonds Borough sewer pipe system is like most sewer systems in many municipalities which mostly rely on CAD files to document city maps. However, the current situation has been challenged as the demand for CAD data grew higher. First, data redundancy increases when the number of map layers increases. Second, CAD-based maps do not emphasize the semantic object-oriented model, therefore, they limit the integration of spatial data and attribute data; the later can be managed with an Access database or an Excel sheet. Therefore, inconsistencies between the CAD maps and the attributes are inevitable as shown in the example of Figure 4.1 where the sewer pipe

length between manholes RD.5.0007 and 5.0008 in the Excel file is recorded as 67.1 meter. However, this length is shown on the CAD map as 47.2 meter. Third, CAD-based system cannot support MIMS data exchange and this weakness can lead to the generation of more inconsistent data and affect decision making across different asset management applications at the municipal level.

Road	Road ID	Manhole From	Manhole to	Length of pipe (meter)	Diameter (mm)	Material	Slope	Year of construction
GOUIN	G065	RD.5.0005	RD.5.0006	90.8	525	TBA	0.25%	1968
GOUIN	G065	RD.5.0006	RD.5.0007	47.2	525	TBA	0.23%	1968
GOUIN	G065	RD.5.0007	RD.5.0008	67.1	525	TBA	0.25%	1968
GOUIN	G065	RD.5.0008	RD.5.0009	74.7	525	TBA	0.24%	1968
GOUIN	G065	RD.5.0009	RD.5.0050	63.0	525	TBA	0.29%	1968

(a) Sewer attribute table



(b) CAD drawing of sewer map

Figure 4.1 Example of errors: sewer lengths in CAD map and Excel file do not match

#### 4.4 CASE STUDY IMPLEMENTATIONS

##### 4.4.1 Data collection and data conversion

As discussed in Chapter 3, improving data quality for Pierrefonds sewer network requires gathering complete data from different sources and performing quality analysis.

Therefore, it is essential to inspect data sources and convert these data to a GIS system. In this study, Pierrefonds sewer network data can be acquired from the following sources: (1) scanned design drawings of sewers network (i.e. Figure 4.2), (2) hard copies of inspection or maintenance reports (i.e. intervention plan for sewer and water network for Pierrerronds), and (3) softcopies of CAD drawings of the network (i.e. Figure 4.3 and 4.4).

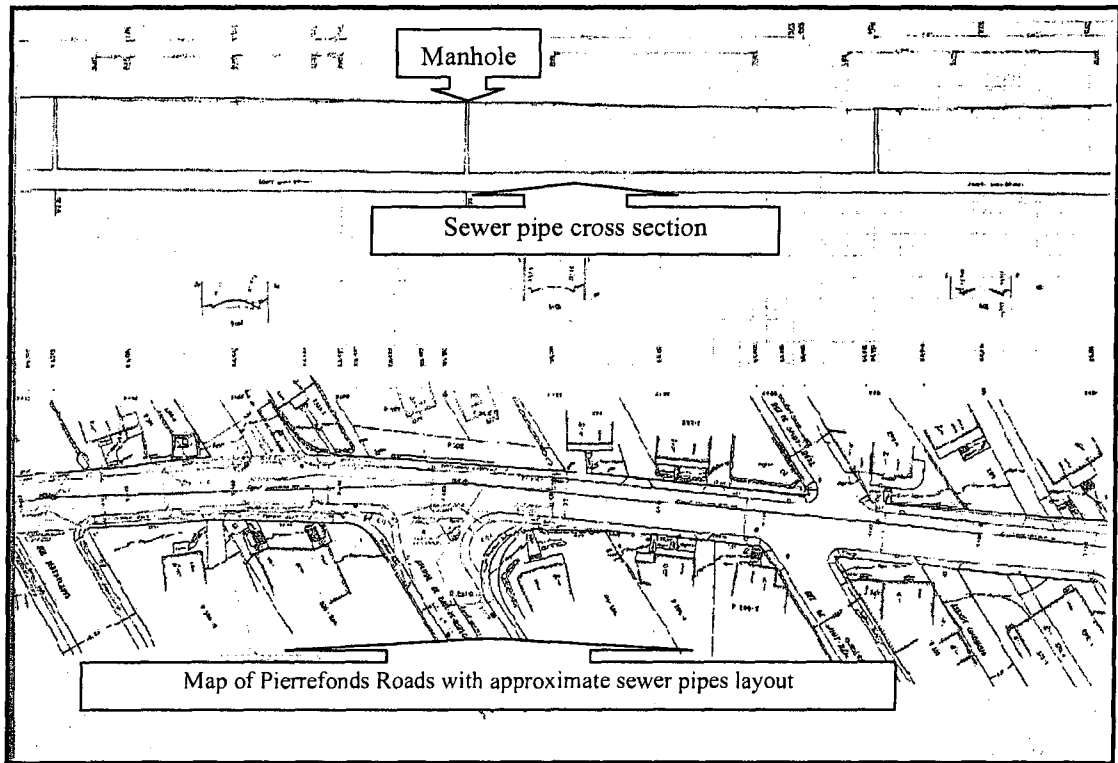


Figure 4.2 Scanned design drawing of sewers network for Pierrefonds Borough

After reviewing the available data, these sources were classified under categories as inventory data which include pipe length, diameter, material, etc., and operation data such as intervention plans or maintenance reports. Based on the discussion in Chapter 3, in addition to inventory and operational data, technical specifications such as Road Geobase (Montreal 2003) can be used as metadata reference.

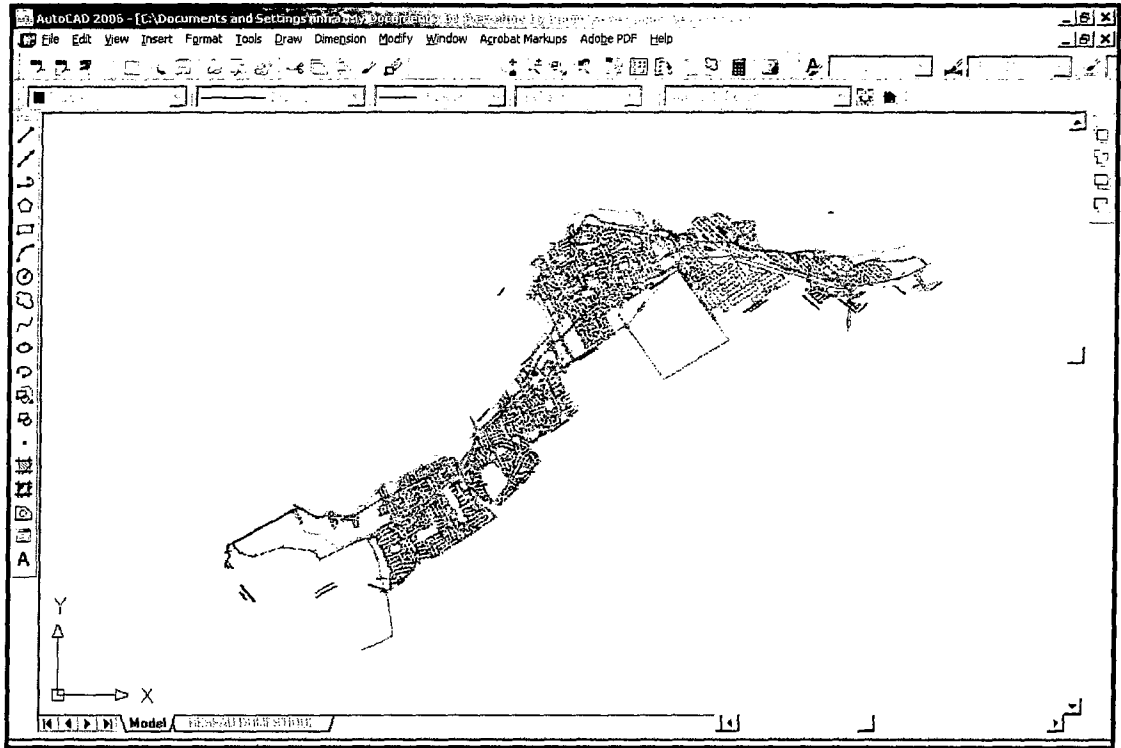


Figure 4.3 CAD drawing of sewer network of Pierrefonds

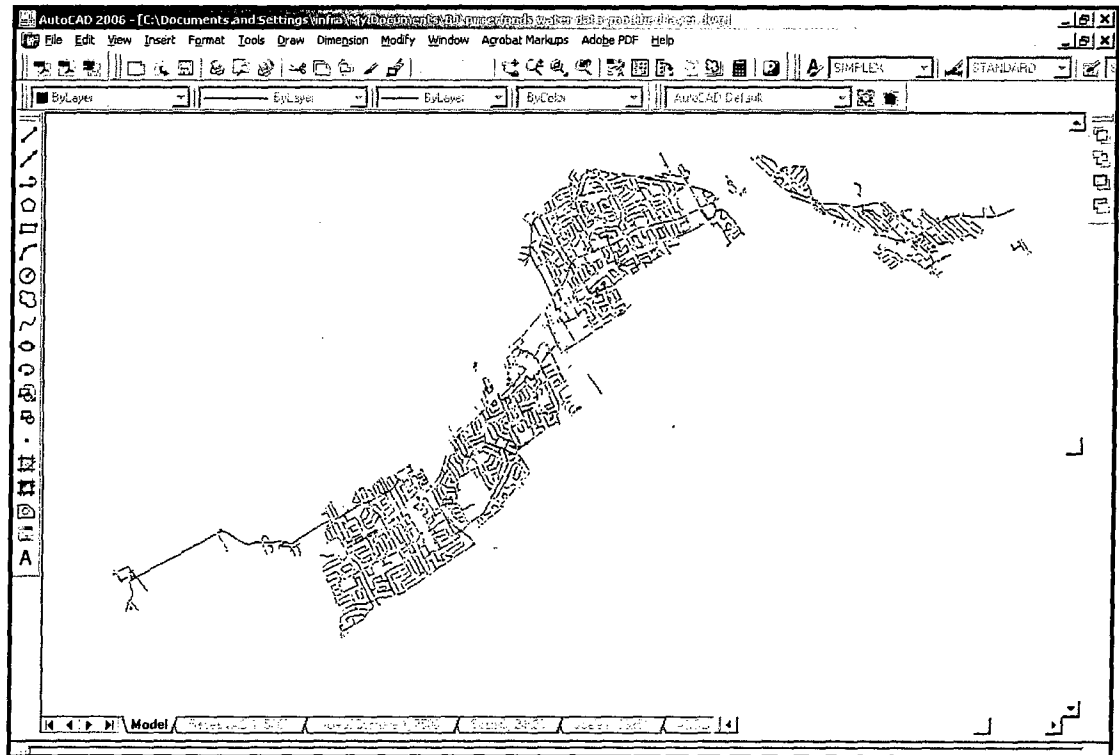


Figure 4.4 CAD drawing of water network of Pierrefonds

During the conversion, different map layers have been examined when studying the drawing map in CAD file and referring to related specifications. In Figure 4.5(a), there are approximately 36 layers which have been shown in different colors in the original Pierrefonds sewer network CAD file. However, only few layers are related with completed sewer information shown in Figures 4.5(b) and (c), while other layers are related to future planning.

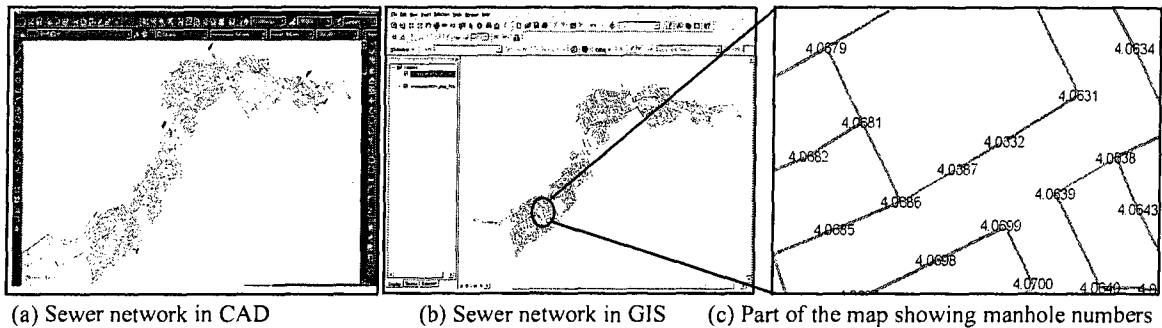


Figure 4.5 Pierrefonds sewer pipes and manholes converted from CAD to GIS

The purpose of converting CAD data to GIS data is to integrate spatial and non-spatial data for further data analysis and data interoperability. Before converting CAD data to GIS data, it is necessary to understand both spatial and attribute dataset. In this case, 36 layers of sewer network have been identified. Appendix G shows the detail explanation of each layer. In this table, layer “DOM\_Agglomeration” and layer “DOM\_NO\_REG” represent the main sewer pipes and manholes, respectively. Meanwhile, the related attribute data which include each pipe starting manhole, ending manhole, length, material, construction year, and slope information are recorded in an EXCEL file as shown in Figure 4.6. Therefore, the main target in data collection step is to connect non-spatial data (EXCEL file) with spatial data (CAD drawing) and convert both dataset into GIS system. We can verify the relationship between the pipe slope and the depth of manholes by the calculation shown in Figure 4.7.





RUE	No. De	DE REGARD	À REGARD	Longueur de la section	Pente	Radier AMONT	Radier AVAL
Road	ID	From manhole ID	To manhole ID	Length (m)	Slope	Upstream (m)	Downstream (m)
GOUIN	G065	RD.5.0007	RD.5.0008	PL 67.1	$\alpha$ -0.25%	$h_1$ 21.16	$h_2$ 21.33

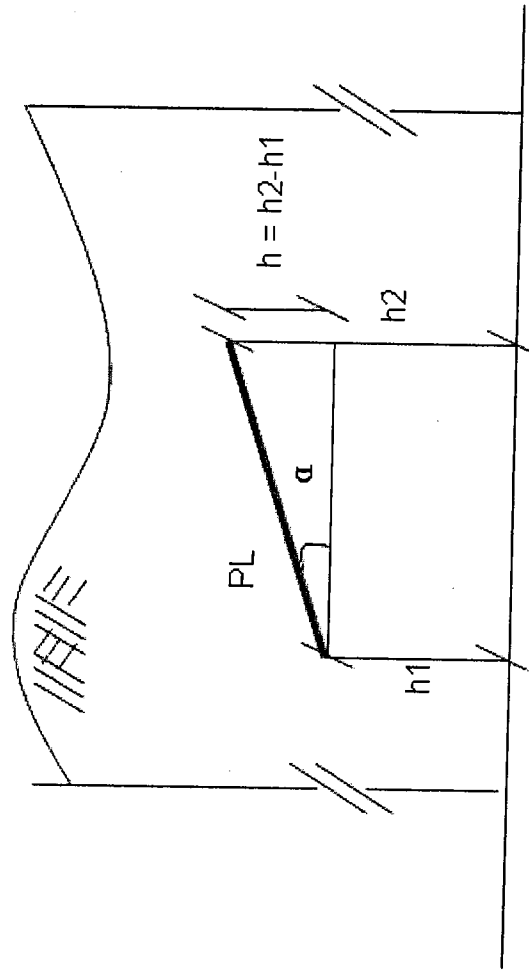


Figure 4.7 Sewer pipe slope calculation

### (1) Converting spatial data from AutoCAD to ArcMap

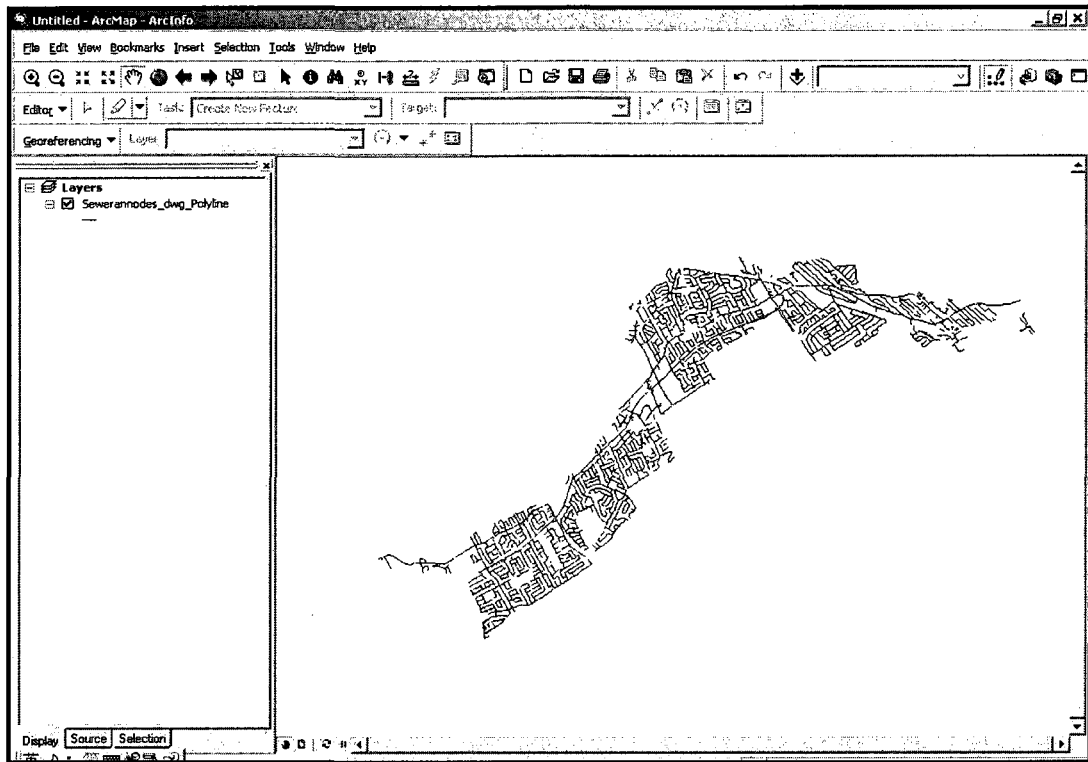
The Pierrefonds sewer network spatial data file “Res\_dom.dwg” is provided by Pierrefonds Municipality as an AutoCAD file. After the layers have been identified, the main layers (manhole and sewer pipe layers) are converted into GIS layers for further analysis by using “Conversion tools to shape file” in the ArcMap tool box. Figure 4.8 shows the Pierrefonds sewer pipe and manhole layers presented in ArcMAP as polyline and point layers, respectively.

### (2) Adjusting sewer attribute data type

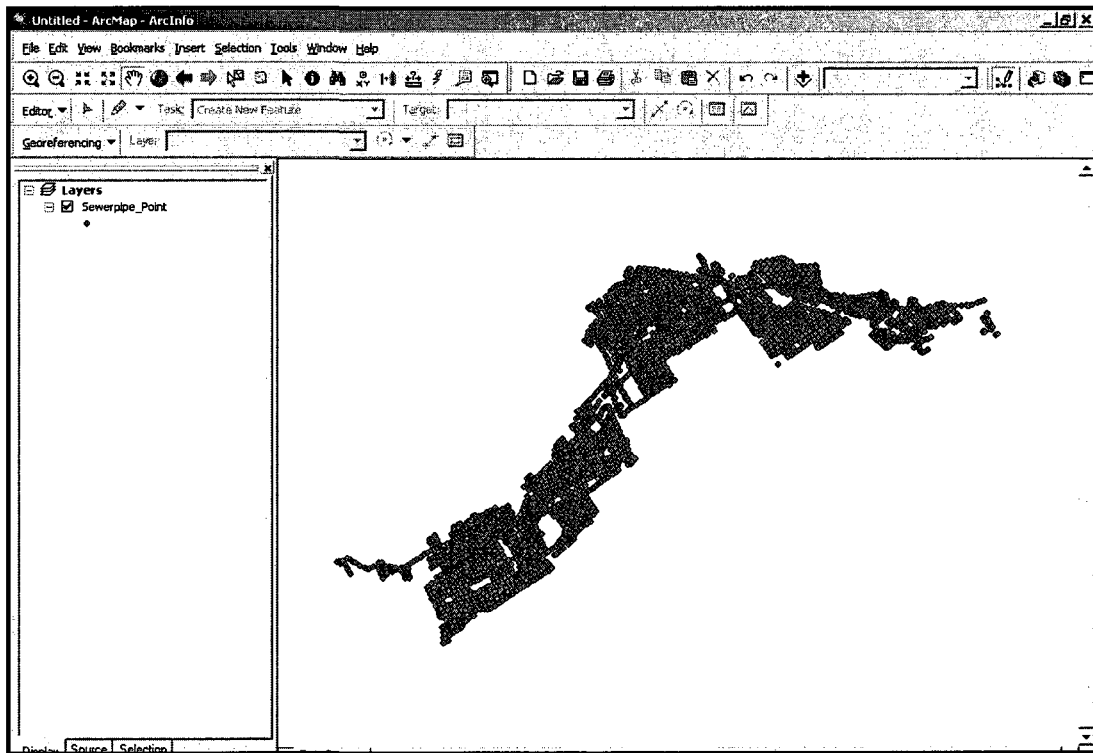
Pierrefonds sewer network non-spatial (attribute) data “pirrefondssewerdata.exl” is provided as an Excel file. Changing the field type of manhole ID from *string* to *number* is necessary in order to link the attribute table to the spatial data file. The Excel file is then saved as a database file in “pirrefondssewerdata.dbf”.

### (3) Linking spatial data and attribute data in GIS

In ArcMap application, the “join” function allows linking a spatial data set and an attribute data set. However, since attribute data (e.g. length and material of pipes) are recorded based on the starting and ending manhole IDs of pipes and these IDs are not given in the pipe layer, but in the manhole layer. Each pipe’s starting and ending manhole IDs have to be identified before linking the spatial data and attribute data. The process to extract the starting and ending manhole IDs is shown in Figure 4.9.



(a) Pipe polyline layer



(b) Manhole point layer

Figure 4.8 Pierrefonds sewer pipe and manhole layers in ArcMAP

For each pipe in the pipe layer, a buffer is created with certain width (e.g. 2m) around the pipe. Then, the two manholes at the ends of the pipe are extracted from the manhole layer, and their IDs are retrieved. This process is repeated for all pipes. Figure 4.10 shows the IDs of starting and ending manholes of pipes. The ID of a pipe is created by concatenating (joining) these two IDs. After each pipe's two manhole IDs have been identified, the field of manhole IDs "NOREG" is chosen as key index to link spatial and attribute dataset as shown in Figure 4.11.

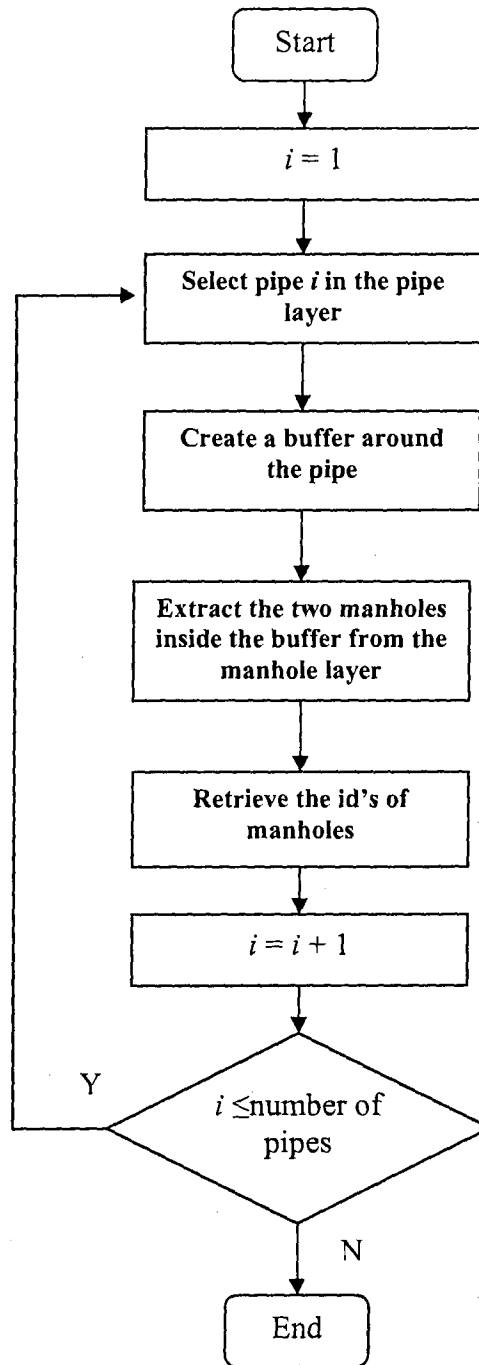


Figure 4.9 Process to retrieve pipe's starting and ending manhole by using "buffer"

function

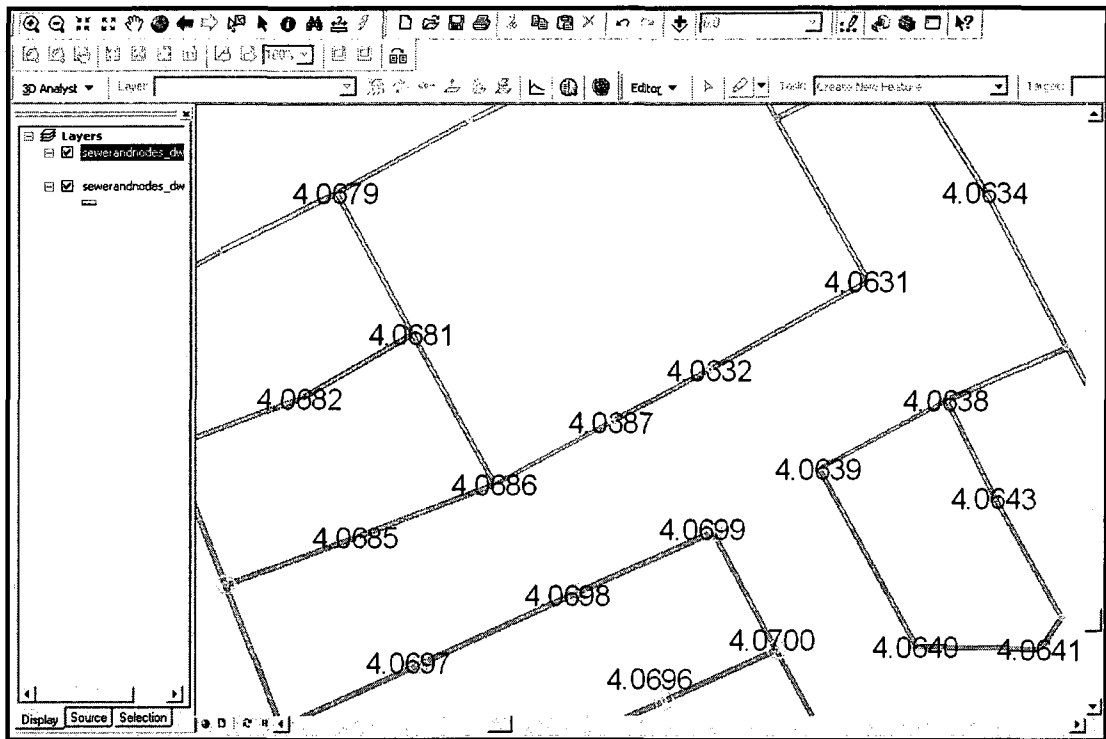


Figure 4.10 IDs of starting and ending manholes of pipes

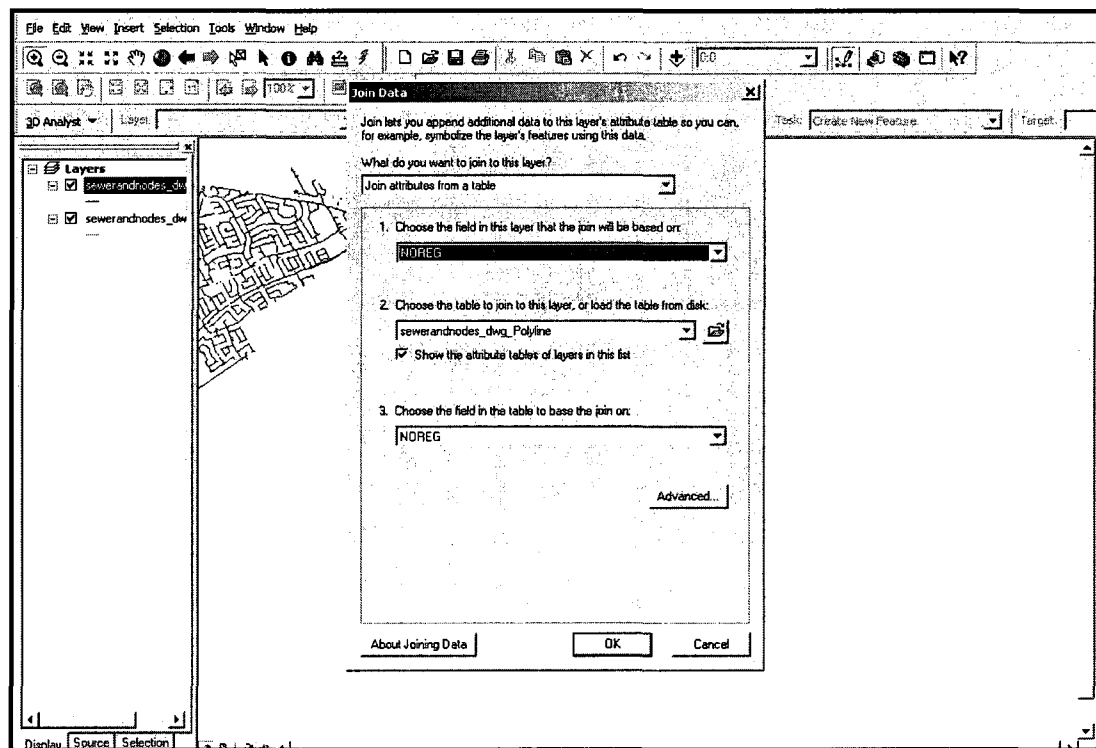


Figure 4.11 Append manhole attribute data with graphic file by using “join” function in ArcMap

#### 4.4.2 Data Inspection and Spatial Data Error Analysis

The main objective of data inspection is to identify errors and improve the accuracy of the map. In accomplishing this objective, it was necessary to evaluate the current available data sources, which are the sewer spatial data set and the related attribute data set independently. Figure 4.12 shows Pierrefonds sewer network in AutoCAD and Figure 4.13 shows detailed CAD drawing of Pierrefonds sewer network in sector 4.

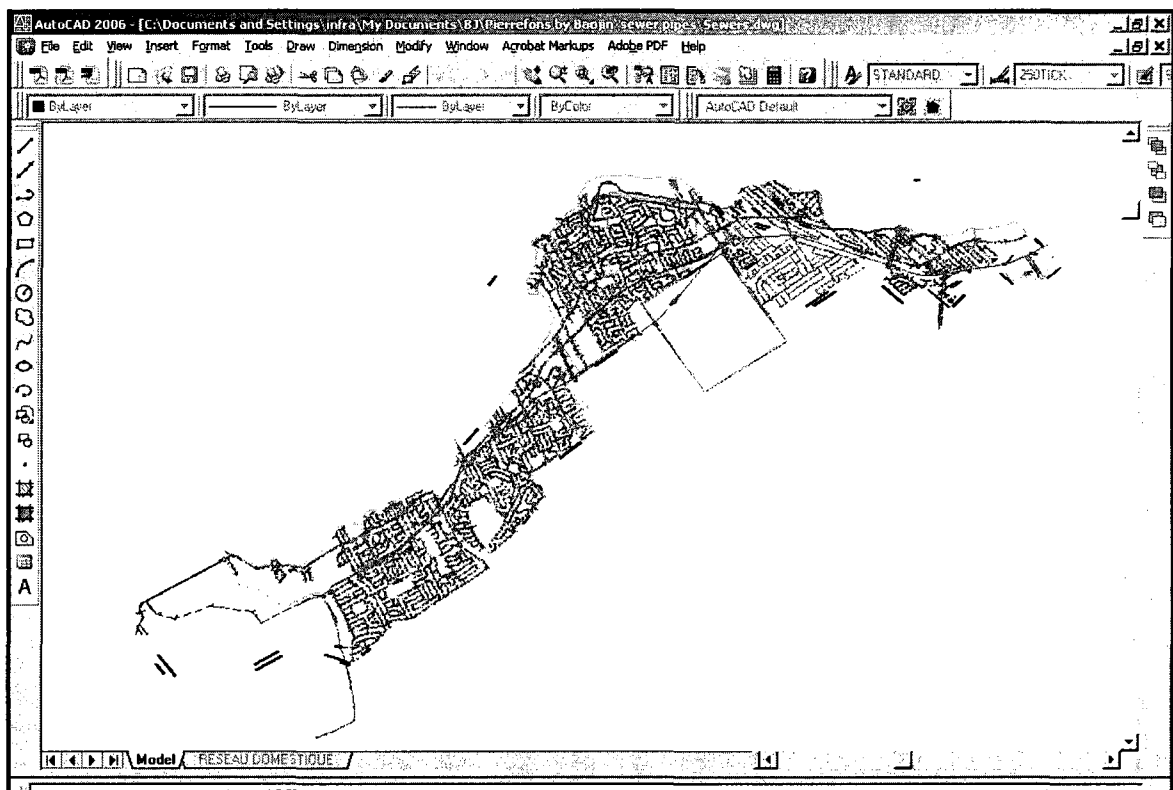


Figure 4.12 Pierrefonds sewer network in AutoCAD

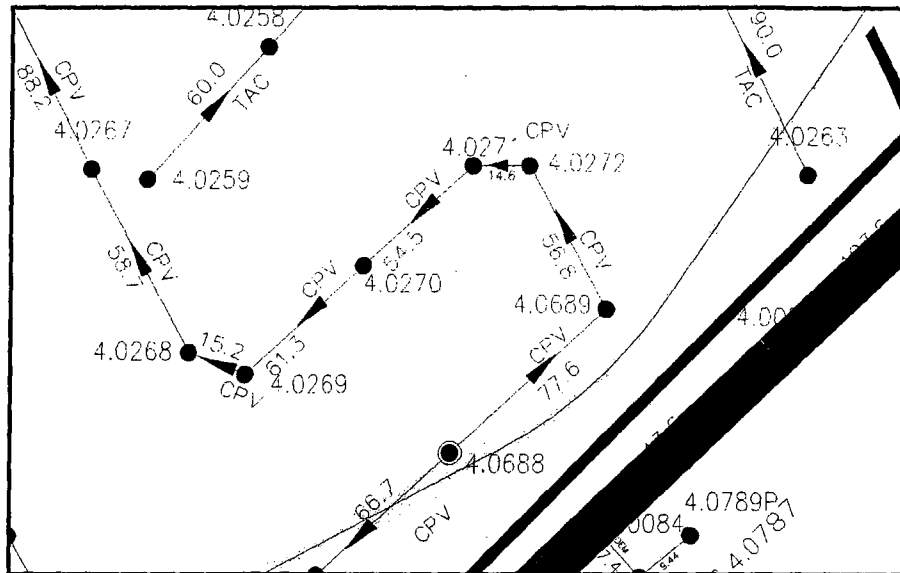


Figure 4.13 Detailed CAD drawing of partial sewer network

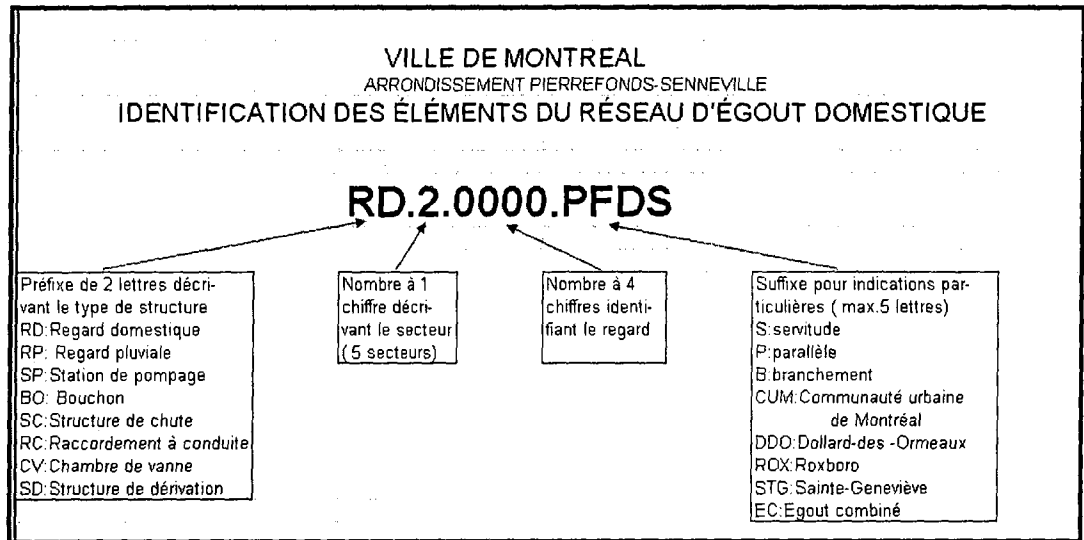
As shown in the figure 4.13 the spatial information about the sewer network system includes the IDs of the manholes represented by codes (e.g., 4.0272 where 4 is the sector number and 0272 is the manhole ID), the directions represented by the purple arrows, and pipes represented by lines between any two manholes. The file contains many layers (Appendix G) for different part of the map. Those layers can be hidden or displaced as desired by the user.

The attribute data set of Pierrefonds sewer network in the Excel data sheets includes: streets names, streets numbers, starting manhole code, ending manhole code, length, diameter, material, and slope of the pipe, year of construction, television inspection date, etc.

However, the codes referral of the manholes in the Excel sheets are different than those in the CAD file in that they include a prefix and suffix letters (e.g. RD 4.0272PFDS). Figure



4.14 shows the French codes and the English translation (e.g. type of the structure and type of connection).



the one digit number indicates sectors  
the four digit number indicates the manhole ID.

<b>PREFIX</b>	(only 2 letters describing the type of structure)	<b>SUFFIX</b>	(max 5 letters)
RD	Domestic Manhole	S	Service
RP	Rain Manhole	P	Parallel
SP	Pumping Station	B	Branch
BO	Pipe Close	CUM	Urban Community of Montreal
SC	Water Pressure Drop	DDO	Dollard-des-Ormeaux Region
RC	Driving Join	ROX	Roxboro Region
CV	Gate Room	STG	Sainte-Genevieve Region
SD	Distributary Structure	EC	Combined Sewer
		PFDS	Pierrefonds Region

Figure 4.14 Codes of manholes in Excel data sheet

The approach of data automated inspection is to merge the above two files in one integrated system to give results in one form of municipality codes, the key parameter between the two files is the manhole's id. The following steps show the principle of this approach application and Figure 4.15 summarizes the process of error detection by using automated inspection.

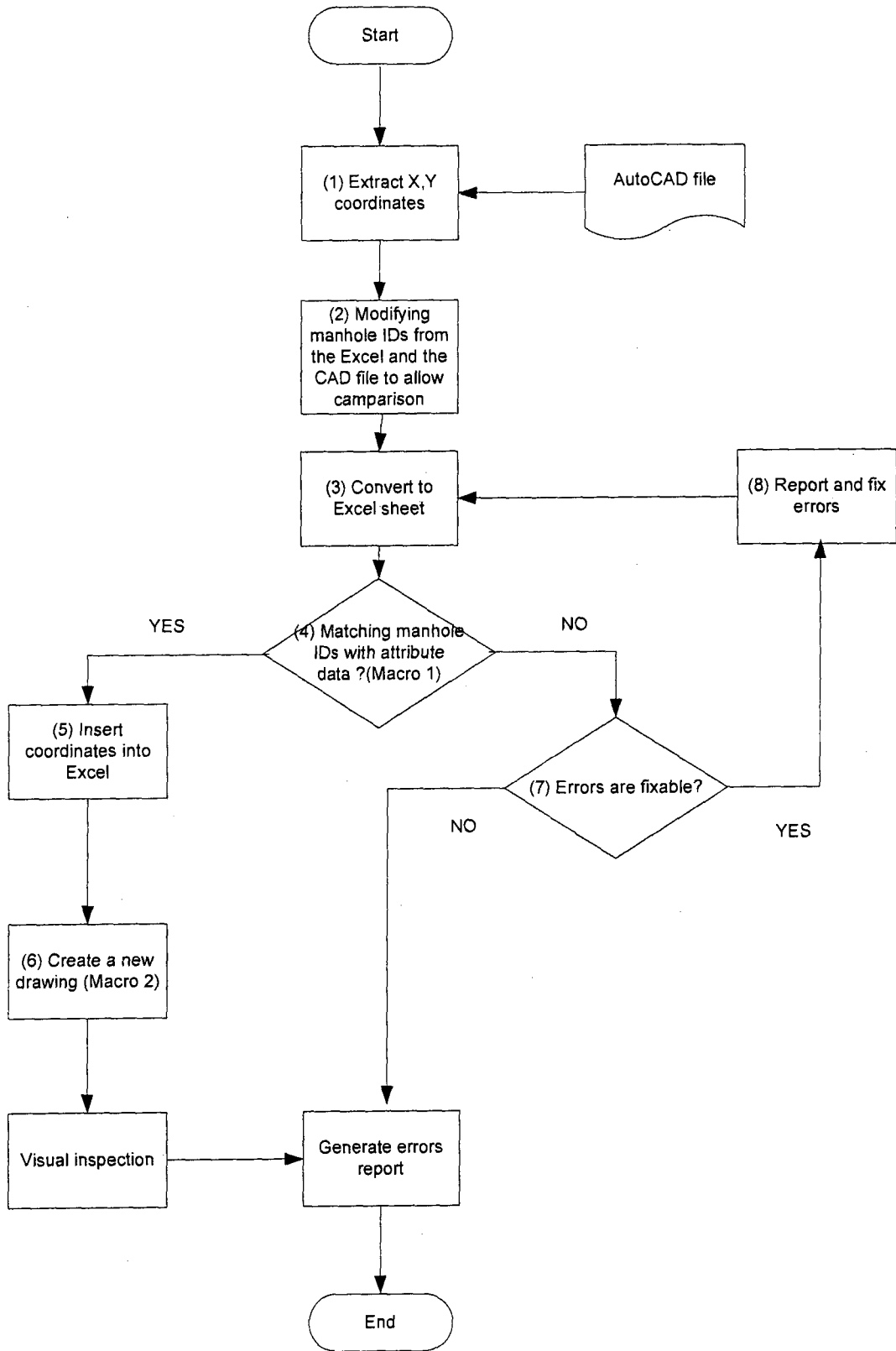


Figure 4.15 Error automated inspection process

(1) Extracting manholes' X, Y coordinates in the CAD file

X, Y coordinate extraction is implemented in AutoCAD tools menu by using attribute extraction setting wizard to get the manholes' IDs and their X, Y coordinates shown in Figure 4.16.

(2) Modifying manhole IDs from the Excel and the CAD file to allow comparison

- Modifying manholes' codes to remove the prefix and suffix using the REPLACE function in the Excel menu (e.g. change manhole ID codes from column C, D into P and Q in Figure 4.17).
- Converting manholes' codes obtained from the CAD file using the FIXED formula in order to allow data comparison (e.g. data format conversion from column X into column Y in Figure 4.17).

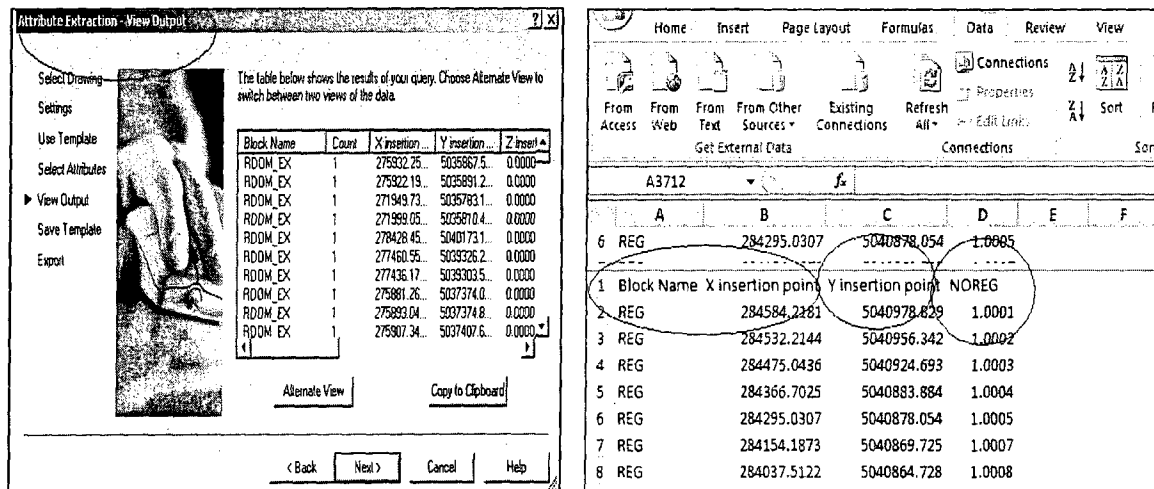


Figure 4.16 Attribute data extraction from CAD to EXCEL

(3) Converting manholes' X, Y coordinates into an Excel sheet

The original Excel sheet contains all the information of pipes which includes the name of pipe, the IDs of the starting and ending manholes. However, the information extracted

from the AutoCAD file represents only the manholes' IDs and their X, Y coordinates. Therefore, the principle for this step is to generate the same format between the extracted dataset and original attribute dataset using the following steps.

(4) Matching manholes' all IDs between X,Y coordinates in Excel file and attribute dataset,

Since the original attribute dataset is based on pipes which include starting and ending manholes and the extracted X, Y coordinates are based on manholes, it is necessary to add the X, Y coordinates to the manholes of each pipe.

(5) Insert coordinates into Excel

By creating a VBA program (Macro 1) included in APPENDIX E, each pipe starting and ending manholes have been identified then the related X, Y coordinates were added to each pipes (i.e., adding starting manhole X, Y coordinates in columns R and S, and ending manhole X, Y coordinates in columns T and U in Figure 4.17).

	C	D	O	P	G	R	S	T	U	X	Y	Z	Position X
1	IDENTIFICATION DES ÉLÉ												
2	DE REGARD	À REGARD	Année de	REGARD	EGARD	X	Y	X	Y	NOREG			
3	No	No	Colmatage	No									
4	RD.1.0034	RD.1.0185	-	1.0034	1.0185	283013.96	5040535.324	283200.851	5040351.734	1.0001	1.0001	284584.218	504097
5	RD.1.0177	RD.1.0178	-	1.0177	1.0178	283172.38	5040378.273	283200.851	5040351.734	1.0002	1.0002	284532.214	504098
6	RD.1.0177	RD.1.0179	-	1.0177	1.0179	283172.38	5040378.273	283150.003	5040398.239	1.0003	1.0003	284475.044	504099
7	RD.1.0179	RD.1.0182	-	1.0179	1.0182	283150	5040398.239	283077.745	5040463.774	1.0004	1.0004	284366.703	504088
8	RD.1.0184	RD.1.0184	-	1.0182	1.0184	283077.75	5040463.774	283077.745	5040463.774	1.0005	1.0005	284295.031	504087
9	RD.1.0184	RD.1.0032	-	1.0184	1.0032	283017.97	5040518.538	283017.97	5040518.538	1.0006	1.0006	284154.187	504088
10	RD.1.0210	RD.1.0209	-	1.0210	1.0209	283285.6	5040278.538	283285.6	5040278.538	1.0008	1.0008	284037.512	504088
11	RD.1.0204	RD.1.0205	1993	1.0204	1.0205	283303.14	5040249.217	283303.14	5040249.217	1.0009	1.0009	283943.339	504087
12	RD.1.0205	RD.1.0206	-	1.0205	1.0206	283308.55	5040224.911	283317.454	5040196.776	1.0010	1.0010	283841.665	504088
13	RD.1.0206	RD.1.0207	-	1.0206	1.0207	283308.55	5040224.911	283355.386	5040147.407	1.0011	1.0011	283814.996	504088
14	RD.1.0207	RD.1.0208	-	1.0207	1.0208	283317.45	5040196.776	283142.583	5040389.121	1.0012	1.0012	283770.469	504088
15	RD.1.0179	RD.1.0197	-	1.0179	1.0197	283150	5040398.239	283142.583	5040389.121	1.0013	1.0013	283759.992	504098
16	RD.1.0197	RD.1.0198	-	1.0197	1.0198	283142.58	5040398.239	283102.563	5040344.315	1.0014	1.0014	283708.155	504088
17	RD.1.0198	RD.1.0199	-	1.0198	1.0199	283102.56	5040398.239	283068.267	5040344.315	1.0015	1.0015	283653.318	504087
18	RD.1.0199	RD.1.0200	-	1.0199	1.0200	283068.27	5040398.239	283068.267	5040344.315	1.0016	1.0016	283640.81	504088
19	RD.1.0200	RD.1.0201	1993	1.0200	1.0201	283355.39	5040398.239	283355.39	5040398.239	1.0017	1.0017	283593.823	504088
20	RD.1.0201	RD.1.0202	-	1.0201	1.0202	283392.64	5040398.239	283392.64	5040398.239	1.0018	1.0018	283540.81	504088
21	RD.1.02	RD.1.0214	-	1.0214	1.0214	283470.62	5040398.239	283470.62	5040177.854	1.0019	1.0019	283500.823	504088
22	RD.1.02	RD.1.0215	-	1.0214	1.0215	283470.62	5040398.239	283470.62	5040177.854	1.0020	1.0020	283470.62	504077
23	RD.1.02	RD.1.0216	-	1.0214	1.0216	283502.5	5040398.239	283502.5	5040206.251	1.0021	1.0021	283470.62	504077
24	RD.1.02	RD.1.0221	-	1.0208	1.0221	283555.38	5040398.239	283555.38	5040240.91	1.0022	1.0022	283470.62	504077
25	RD.1.02	RD.1.0222	-	1.0222	1.0222	283355.38	5040398.239	283355.38	5040090.015	1.0023	1.0023	283470.62	504077
26	RD.1.02	RD.1.0223	-	1.0222	1.0223	283290.5	5040398.239	283290.5	5040059.861	1.0024	1.0024	283470.62	504077
27	RD.1.03	RD.1.0416	-	1.0222	1.0358	283257.88	5040398.239	283257.88	5040031.28	1.0025	1.0025	283470.62	504077
28	RD.1.04	RD.1.0417	-	1.0416	1.0417	282044.4	5040398.239	282044.4	5041003.97	1.0026	1.0026	283470.62	504077
29	RD.1.04	RD.1.0418	-	1.0416	1.0418	282044.4	5040398.239	282044.4	5041003.97	1.0027	1.0027	283470.62	504077
30	RD.1.04	RD.1.0419	-	1.0416	1.0419	282112.48	5040398.239	282112.48	5040933.29	1.0028	1.0028	283470.62	504077
31	RD.1.0503	RD.1.0510	1995	1.0418	1.0503	282134.55	5040398.239	282134.55	5040933.159	1.0029	1.0029	283470.62	504077
32	RD.1.0503	RD.1.0510	1995	1.0503	1.0503	280822.2	5040398.239	280822.2	5041566.127	1.0030	1.0030	283470.62	504077
33	RD.1.0503	RD.1.0510	1995	1.0503	1.0503	280822.2	5040398.239	280822.2	5041566.127	1.0031	1.0031	283470.62	504077

Figure 4.17 Adding coordinates of manholes from CAD data

(6) Creating a new drawings by using the data in the updated Excel sheet

In this step, a new map is redrawn by using another VBA program. In this program, manholes are recreated using X, Y coordinates and the pipes have been redrawn based on starting and ending manholes. The program is included in Appendix F. Then visual inspection has been applied to detect errors as explained in Section 3.6.4.

(7) Categorizing errors

In Chapter 3, three types of data errors have been introduced. By using automated inspection and visual inspection, the obvious errors and processing errors can be defined and categorized. For example, in Step (5), non-matching manholes (i.e., manhole for which the X, Y values were not added) have obvious errors that can be detected automatically, and some of these errors can be fixed by manually checking the records in the Excel sheet as will be explained in the next step. Another example of errors that can be discovered only by using visual inspection is the error in the direction of a pipe, as will be explained below.

(8) Report and fixing errors

Errors are classified as following:

***Obvious errors***

These errors are found when the macro program (Macro 1) running in Excel sheet cannot match IDs in the original Excel sheet with those from the CAD file due to type mismatch. For example, Figure 4.18 shows an ID error caused by wrong punctuation. This error can

be fixed automatically by replacing the wrong punctuation. Figure 4.19 shows an ID error by typing wrong sector code in the CAD file. This error can be fixed automatically.

RUE	No. De	DE REGARD	À REGARD	Longueur de	Diamètre
	RUE	No	No	la section	(mm)
ÉRABLES,des	D085	RD.4.0393	RD.4.0394	11.7	300
ÉRABLES,des	D085	RD.4.0394	RD.4.0395	32.4	300
ÉRABLES,des	D085	RD.4.0395	RD.4.0396	91.5	300

Figure 4.18 Example of ID errors by using wrong punctuation

DE REGARD No	À REGARD No	From X	From Y	To X	To Y
1.0288	1.0287	283392.64	5040177.854	283470.617	5040196.21
1.0287	1.0288	283470.62	5040196.21	283502.503	5040206.251
1.0288	1.0289	283502.5	5040206.251	283564.452	5040240.91
1.0289	1.0290	283355.39	5040147.407	283290.503	5040090.015
1.0290	1.0291	283290.5	5040090.015	283257.886	5040059.861
1.0291	1.0292	283257.89	5040059.861	283223.448	5040031.29
1.0292	1.0286	282044.4	5041029.911	282044.868	5041003.97
1.0163	1.0164	282044.87	5041003.97	282112.483	5040933.29
1.0164	1.0165	282112.48	5040933.29	282134.549	5040933.159

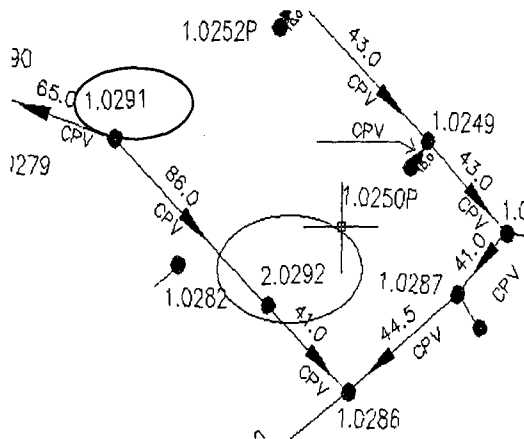


Figure 4.19 Example of ID error by typing wrong sector code in the CAD file

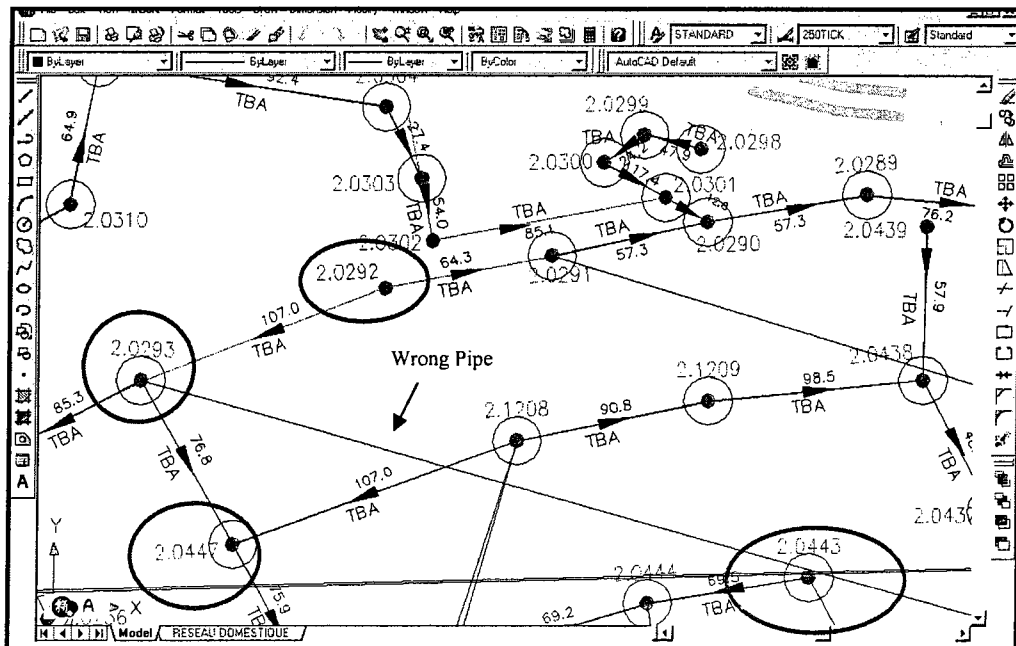
### Processing errors

After running Macro 2 program, new CAD file is drawn. Comparing this file with the original one using visual inspection, some process errors such as logical inconsistency

(e.g. Figure 4.20), wrong direction (e.g. Figure 4.21), and duplication of manhole IDs (e.g. Figure 4.22) can be found out. In Figure 4.20, the attribute table shows that there are four pipes connecting with manhole RD.2.0293, (pipe RD.2.0292 – RD.2.0293, pipe RD.2.0293 – RD.2.0294, pipe RD.2.0293 – RD.2.0477, and pipe RD.2.0293 – RD.2.0443). However, when applying macro 2 program to redraw the map, we found the pipe the pipe RD.2.0293 – RD.2.0443 has logical inconsistency because the sewer pipes cannot go across with other pipes. Therefore, it is categorized as processing error.

RUE	No. De RUE	DE REGARD No	À REGARD No	Longueur de la section	Diamètre (mm)	Material
PAVILLON	P050	RD.2.0291	RD.2.0292	64.3	300	TBA III
PAVILLON	P050	RD.2.0292	RD.2.0293	106.2	300	TBA IV
PAVILLON	P050	RD.2.0293	RD.2.0294	85.3	300	TBA IV
PAVILLON	P050	RD.2.0293	RD.2.0477	51.2	300	TBA
PAVILLON	P050	RD.2.0304	RD.2.0303	17.4	300	TBA
PAVILLON	P050	RD.2.0303	RD.2.0302	12.8	300	TBA
PAVILLON	P050	RD.2.0293	RD.2.0443	24.2	300	TBA

(a) Attribute table

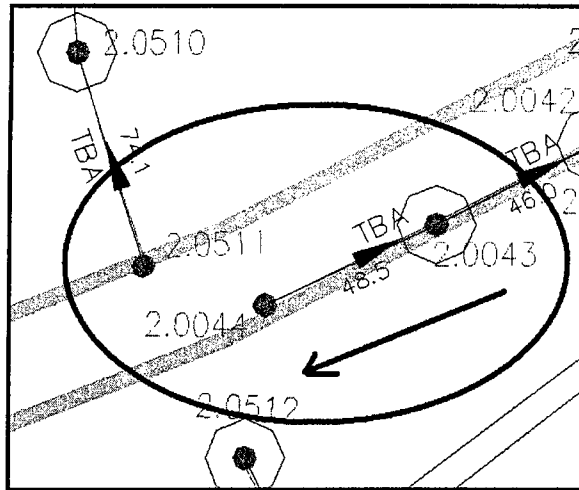


(b) Drawing map

Figure 4.20 Example of processing error in logical inconsistency



In Figure 4.21, we can see that the direction in the CAD file is incorrect not only because the sewer flow direction can be figured out by checking flow topology which can be automatically captured from end-users to mains. In Figure 4.22, there is duplication of manhole ID at two locations in sector 4.



(a) Pipe direction from RD.2.0044 to RD.2.0043 in CAD file

RUE	DE REGARD No	À REGARD No	Longueur de la section	Diamètre (mm)	Material	Slop
GOUIN	RD.2.0042	RD.2.0043	46.0	300	TBA	-
GOUIN	RD.2.0043	RD.2.0044	48.5	300	TBA	-
GOUIN	RD.2.0045	RD.2.0046	35.3	300	CPV	0.42%
GOUIN	RD.2.0046	RD.2.0047	52.2	300	TBA	0.40%
GOUIN	RD.2.0047	RD.2.0048	74.2	300	TBA	0.39%

(b) Pipe direction from RD.2.0043 to RD.2.0044 in Excel file

Figure 4.21 Example of processing error in wrong direction

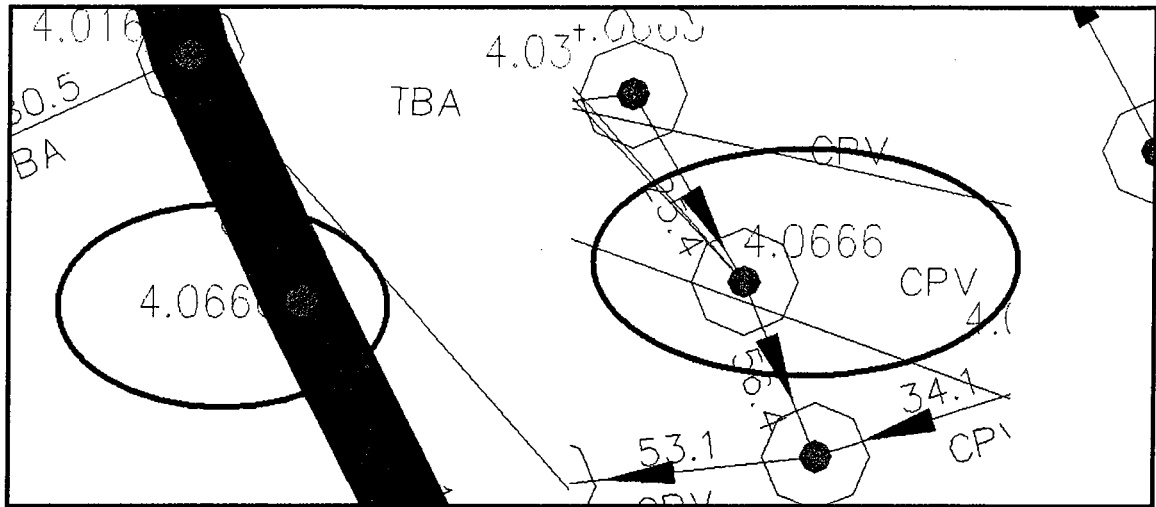


Figure 4.22 Examples of duplication of manhole ID

Based on both automated and visual inspections, there are many similar errors in Pierrefonds sewer pipe network and most errors can be efficiently detected and corrected. In this implementation, 4333 sewer pipes have been investigated. We found that 99 pipes have not used appropriated coordinates and 40% of these errors result from typing mistakes, 22% from wrong direction, and 38% from logical inconsistency. However, not every error can be found or fixed automatically because of the variety of types of errors. Therefore, finding error patterns and scheduled validation is necessary to identify very important and potentially costly errors.

#### 4.4.3 Positional Accuracy Analysis

The objective of this example is to analyze and determine the positional accuracy of Pierrefonds Borough manholes because manholes are at the end of each pipe and they are also an independent layer in original CAD file. Based on the first implementation, the Pierrefonds manhole layer has already been transferred to GIS shape file and connected to sewer pipe network. Therefore, to investigate the accuracy of manhole position is

equal to analyze sewer pipe network accuracy. The investigation will be achieved by the following steps as accuracy measurement discussed on Section 3.4.3: (1) Determining which test to use, (2) Selecting test data set, (3) Selecting an independent data set, (4) Calculating the position accuracy statistics, and (5) Preparing an accuracy statement and including the accuracy report in metadata. These steps are explained in the following:

**(1) Determining which test to use**

The first step in applying data accuracy analysis is to identify the spatial characteristics of the data set being tested. Based on the National Standard for Spatial Data Accuracy (NSSDA 1999) handbook, there are two types of positional accuracy: horizontal and vertical accuracy measurement. Since this study focuses on planimetric (the x, y coordinate accuracy) dataset evaluation, the horizontal accuracy analysis is applied. However, in future, if the 3D position accuracy (i.e. depth of pipe) is needed, the vertical accuracy statistic could be considered.

**(2) Selecting test data set (from GIS)**

A dataset's accuracy is tested by comparing the coordinates of a set of points within the dataset to the coordinates of the same points from an independent dataset of greater accuracy (ANSI 1998). The dataset of manhole features in Pierrefonds Borough is obtained from original CAD file which is created by Pierrefonds Borough engineer in 2006, using a scale of 1:4000. After the manhole dataset is transferred into GIS, it is found that there are 6788 points covering the whole Pierrefonds area, although many of these points are not manholes. Therefore, we selected points whose attributes tables contain manhole information and found that there are 3496 points related to manhole as

shown in Figure 4.8(b). In order to analyze manhole location accuracy, desired certainty, sample size and required hypothesis testing parameter should be decided. As discussed in Section 3.4.3, 95% confidence level is recommended by testing methodology. However, since there are more than 4000 pipes in Pierrefonds, it is so time consuming that 95% certainty requires a sample size of 384 according to the following statistic sampling formula (Schwalbe 2007):  $\text{Sample size} = 0.25 * (\text{Certainty Factor} / \text{Acceptable Error})^2$ . Therefore, we choose 90% certainty instead which results in a sample size of 68 for the test. Besides, we assume that the acceptable mean error between map data and observations is 10 meter and exceeding 10 meter is not acceptable.

### **(3) Selecting an independent data set (from GPS)**

The independent dataset must be acquired separately from the dataset being tested. It should be of the highest accuracy available (FGDC 1998). In the implementation, the collecting method for manhole x and y coordinates uses a Trimble GPS 7500 receiver (Trimble 2008). It is a 24-channel dual-frequency RTK GPS receiver which has the Trimble Maxwell technology for superior tracking of GPS satellites, increased measuring speed, longer battery life through less power use, and optimal precision in outdoor environments. It can receive available differential correction signals with accuracy of approximately 1 centimeter. Based on the previous test hypothesis, sixty-eight well-defined manhole points which are located on Pierrefonds from manhole RD.2.0100 to RD.2.0168 have been chosen to collect their coordinates as shown in Figure 4.23. The detailed manhole coordinates collecting procedure is included in Appendix B.

### **(4) Calculating the position accuracy statistics**

Based on the step 3, the independent data of points have been collected using the software ArcMap. Meanwhile, using VBA programming, the x, y coordinates' values can be extracted and transferred into a DBF file as explained in Appendix C.



Figure 4.23 Independent dataset collections

The result of position accuracy statistics is shown in Table 4.1. The appropriate hypotheses are:

$$\begin{aligned} H_0 : \mu &= 10 \\ H_1 : \mu &> 10 \end{aligned} \quad Z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{11.45 - 10}{5.8 / \sqrt{68}} = 2.0626 \quad (4-1)$$

Since we define 90% certainty, the type I error of  $\alpha=0.1$ , then from Cumulative Standard Normal Distribution Table (Appendix L), we find  $Z_{\alpha}=Z_{0.1}=1.29$ . Therefore, we reject  $H_0 : \mu = 10$  and conclude that the mean error exceeds 10 meter.

Table 4.1 Positional accuracy statistic from manhole RD.2.0100 to RD.2.0168

Point number	x		y		diff in x	(diff in x) <sup>2</sup>	y		diff in y	(diff in y) <sup>2</sup>	(diff in x) <sup>2</sup> + (diff in y) <sup>2</sup>		Diff in X & Y
	(from GIS)	(from GPS)	(from GIS)	(from GPS)			(from GIS)	(from GPS)			(diff in x) <sup>2</sup>	(diff in y) <sup>2</sup>	
RD.2.0100	278504.701	278505.107	5040556.266	5040569.066	-0.405271566	0.164245	5040556.266	5040569.066	-12.8004	163.8509	164.0151196	12.80683878	
RD.2.0101	275932.256	275933.540	5035867.523	5035872.760	-1.283186332	1.646567	5035867.523	5035872.760	-5.23733	27.4296	29.07616636	5.392232039	
RD.2.0102	275922.191	275922.890	5035891.222	5035891.768	-0.699171825	0.488841	5035891.222	5035891.768	-10.5456	111.21	111.6988309	10.56876676	
RD.2.0103	271949.732	271954.528	5035783.145	5035797.115	-4.795775484	22.99946	5035783.145	5035797.115	-13.9705	195.1736	218.1731062	14.77068401	
RD.2.0104	271999.058	272003.486	5035810.445	5035824.749	-4.428016679	19.60733	5035810.445	5035824.749	-14.3047	204.6258	224.2330925	14.9744146	
RD.2.0105	278428.457	278433.244	5040173.188	5040176.588	-4.786900702	22.91442	5040173.188	5040176.588	-3.40007	11.56045	34.4748733	5.871530746	
RD.2.0106	277460.552	277463.831	5039326.276	5039342.224	-3.279559942	10.75551	5039326.276	5039342.224	-15.9472	254.3126	265.0681461	16.28091355	
RD.2.0107	277426.179	277428.556	5040000.000	5040000.000	2.374000000	5.604444	5040000.000	5040000.000	0.00000	0.00000	5.604444	2.374000000	
RD.2.0108	279931.781	279940.302	5040592.866	5040604.002	-8.520853219	72.60494	5040592.866	5040604.002	-11.1364	124.0183	196.6232588	14.02224158	
RD.2.0109	279892.896	279900.439	5040648.331	5040663.540	-7.542806219	56.89393	5040648.331	5040663.540	-15.2091	231.3182	288.2121595	16.97681241	
RD.2.0110	279863.635	279869.983	5040690.547	5040700.539	-6.34841092	40.30232	5040690.547	5040700.539	-9.99179	99.83595	140.1382672	11.83800098	
RD.2.0111	279837.023	279838.279	5040731.968	5040736.659	-1.255869634	1.577209	5040731.968	5040736.659	-4.69133	22.00854	23.58575288	4.856516538	
RD.2.0112	279802.968	279803.523	5040783.358	5040793.385	-0.55457469	0.307553	5040783.358	5040793.385	-10.0273	100.5458	100.8533912	10.04257891	
RD.2.0113	279634.782	279643.512	5041085.996	5041103.787	-8.729996651	76.21284	5041085.996	5041103.787	-17.7914	316.5338	392.746645	19.81783654	
RD.2.0114	279594.454	279594.942	5041159.841	5041168.782	-0.487550522	0.237706	5041159.841	5041168.782	-8.94113	79.94384	80.18155046	8.954415138	
RD.2.0115	279555.593	279560.576	5041234.788	5041239.297	-4.982814616	24.82844	5041234.788	5041239.297	-4.50885	20.32973	45.15817305	6.719983114	
RD.2.0116	279514.995	279523.749	5041310.142	5041319.788	-8.754102164	76.6343	5041310.142	5041319.788	-9.6464	93.05302	169.6873219	13.02640863	
RD.2.0117	279497.713	279502.843	5041339.944	5041346.374	-5.129764815	26.31449	5041339.944	5041346.374	-6.42958	41.33948	67.65396344	8.225202942	
RD.2.0118	279467.332	279476.832	5041380.411	5041383.639	-9.499719751	90.24468	5041380.411	5041383.639	-3.22764	10.41764	100.6623163	10.03306116	
RD.2.0119	279437.794	279445.738	5041426.216	5041434.293	-7.944433822	63.11403	5041426.216	5041434.293	-8.07729	65.24267	128.3567011	11.32946164	
RD.2.0120	279433.584	279433.767	5041461.959	5041475.476	-0.183134723	0.033538	5041461.959	5041475.476	-13.5171	182.7128	182.7463105	13.51836937	
RD.2.0121	279401.778	279403.858	5041521.908	5041527.735	-2.079777587	4.325475	5041521.908	5041527.735	-5.82729	33.95727	38.28274521	6.187305165	
RD.2.0122	279646.994	279648.892	5041412.002	5041426.408	-1.9081279	3.640952	5041412.002	5041426.408	-14.406	207.5319	211.1728261	14.53178675	
<b>SUM</b>												<b>778.76957976</b>	
<b>MEAN</b>												<b>11.45249382</b>	
<b>STD</b>												<b>5.805744443</b>	

### (5) Preparing an accuracy statement and including the accuracy report in metadata

Since the positional accuracy of a test dataset has been determined, it is important to report that value in a consistent and meaningful way. Reporting the positional accuracy in a complete description of the data set will be useful for future data maintenance. Metadata, described as “data about data” (NISO 2004), has been used to record the content, quality, condition, history and other characteristics of a data set. Based on the data sharing discussion in Chapter 3, Pierrefonds sewer pipes network metadata set is created by DataLogr software (version 2.11) shown as Figure 4.24. The result of Pierrefonds sewer pipe network metadata is included in Appendix D.

SECTION: CONTACT INFORMATION	
DataLevel: 2	
TITLE/ BRIEF DESCRIPTION	Pierrefonds Borough
GEOGRAPHIC AREA NAME(S)	Sever Pipe Network
ARCHIVE/TRANSFER FILE NAME	Res_dom.dwg
ARCHIVE/TRANSFER FILE TYPE	Self-Extracting Compress
DATA FILE/COVER NAME(S)	sewerpoit_attt.dbf, sewer
DATA FILE(S) FORMAT	ArcView ShapeFile
LAST FILE EDIT DATE	2008 OCT 31

REPEATED GROUP: CONTACT	
PERSON and TITLE	BaoJin Hu
ORGANIZATION	Concordia Unive
MAILING STREET/P.O. BOX	ba_hu@encs.conc
CITY	Montreal
STATE	Quebec
ZIP CODE	
PHONE	514-8482424 * 7

Figure 4.24 Create metadata for Pierrefonds sewer pipe network by using DataLogr

#### 4.4.4 WebGIS application

Sharing the collection of geographic information synchronously in MIMS with others is an essential part in the GIS data quality assurance process. The users of MIMS could



include architects, cost estimators, engineers, manufacturers, constructors, surveyors, and facility managers. Therefore, efficiently communicating geographic information and reducing errors within the system become significantly important. Based on the discussion in Chapter 3, a Pierrefonds sewer network WebGIS application has been implemented in the following steps:

### **(1) Requirement Analysis**

The objective of this application is to disseminate error-free Pierrefonds sewer pipe network system and integrate with other systems (e.g. Montreal road network) through the Internet, to ensure that any user, who can access the Internet, can easily browse the system. The requirement analysis step has been performed through interviewing potential users. This step produced two critical pieces of information:

- A list of functions that are needed: The required functions are the basic visualization functions (e.g. Pan and Zoom), and more advanced functions such as object identification, search, and shortest path. Clients can use these functions to view road information and search the target object by using keywords.
- A list of available/needed geographic data. In this case study, Pierrefonds sewer pipes, manholes and Montreal road layers have been included.
- Acquisition of hardware and software: selecting suitable software is an important step in a successful implementation. Software was evaluated based on functionality and performance, and independency of the hardware and operating system. Based on the Alesheikh (2002), ArcGIS Server 9.2 has been chosen for running the WebGIS because of its powerful functionalities.

## (2) Conceptual Design and WebGIS Architecture

Once the required data have been investigated, the data model that identifies the entities and their relationships were designed. Since, the data will be delivered through a central server, and clients will have access to raster formats, thin client architecture as shown in Figure 4.25 was chosen providing the users with access to interactive maps.

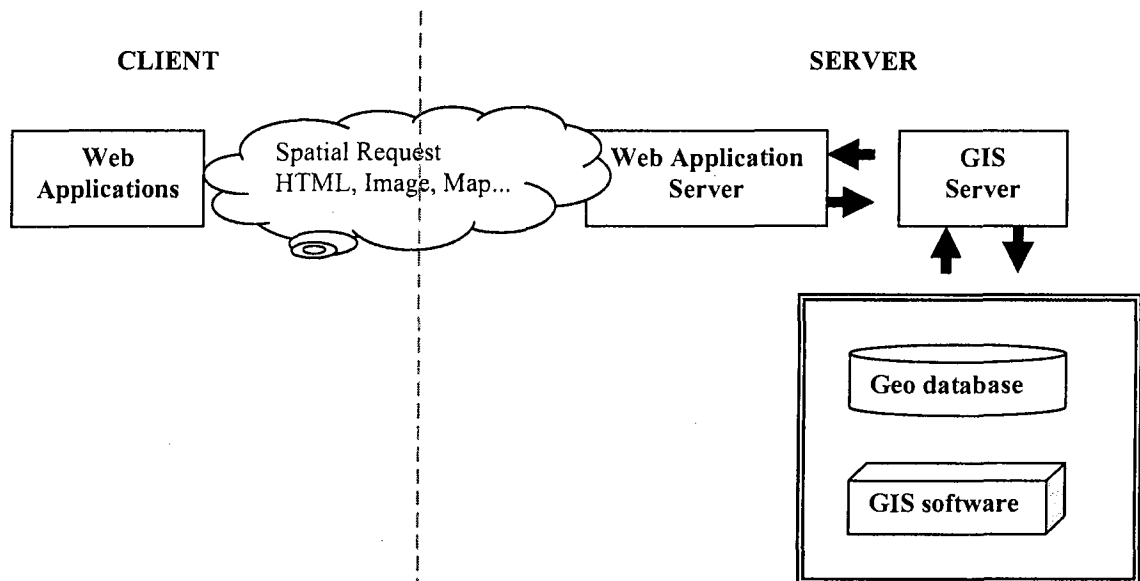


Figure 4.25 Pierrefonds sewer network WebGIS architecture

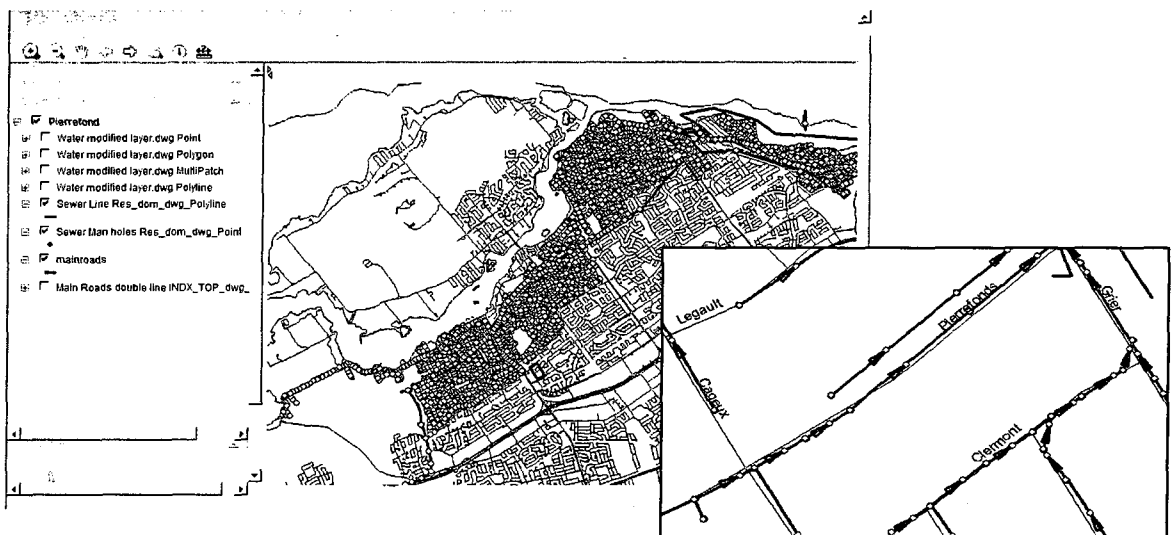
## (3) WebGIS System Integration

The objective of this step is to integrate different systems such as ERP, CRM to make sure they work as expected and to initiate all procedures necessary to use the GIS.

## (4) Application Development

The initial requirement analysis contained some applications of a complex nature. However, the majority of initial applications was straightforward, and can be implemented using basic functionalities (e.g. display) that are part of the WebGIS software. The more complex applications were not supported by the basic functions of

WebGIS but have to be programmed. Ease of use, user-friendliness, and reducing the volume of data transfer are the critical issues considered in the development. Figure 4.26 shows how Pierrefonds sewer pipes network displays in WebGIS and integrate with Montreal road system. The detailed steps to publish map on WebGIS have been explained in Appendix M. The benefits of integration will help to correctly detect the underground pipes buried under specific roads and efficiently find the routing during assets operation and maintenance activities.



(a) Accessing Pierrefonds sewer network using Internet browser

(b) Pierrefonds sewer and road networks

Figure 4.26 Pierrefonds sewer system network with Montreal roads presented in WebGIS

### (5) WebGIS Use and Maintenance

The final step in WebGIS implementation is to put the system to use. With system integration and testing completed and all applications available for use, the system should be released to users. Two activities should be done:

- User support and service, in which new applications will be determined

- System maintenance (database, hardware, software) where the WebGIS must run smoothly

#### **4.5 SUMMARY AND CONCLUSION**

Based on the innovative framework and methods discussed in Chapter 3, this chapter implemented data quality assurance process by using GIS technology to enhance municipal asset management. Quality issues grouped into four domains have been applied: data collection, data accuracy, data sharing and data maintenance. The sources of data errors and data collection methods have been discussed. Data conversion flowchart and data accuracy analysis have been presented. Moreover, in the data sharing part, WebGIS for an Integrated Municipal Infrastructure Management System has been introduced.

## **CHAPTER 5 SUMMARY, CONCLUSIONS, AND FUTURE WORK**

### **5.1 SUMMARY**

The core part of MIMS is to build up a quality-oriented system for whole infrastructure management process including planning, construction, operation, and maintenance of various assets. Through applying various quality control methods such as automated and visual inspection, using GIS, etc., in data collection, data accuracy analysis, data sharing, and data maintenance, spatial and attribute information can be efficiently managed and integrated with other systems.

As introduced in the literature review, there is significant development in using GIS to manage MIMS. Many spatial technologies (i.e. data collection and inspection methods) and data standards have been applied in latest MIMS applications. However, data quality issues have not been well emphasized and the requirement to ensure the overall reliability and interoperability of municipal infrastructure operations is important for facilities long-term using. Therefore, the new initiative framework of spatial data quality assurance process which applies Six Sigma principle in municipal infrastructure has been proposed. Data collection, data accuracy, data improvement and sharing and data maintenance have been discussed in detail. The source of data errors has been investigated and data collection methods have been compared and discussed. Furthermore, WebGIS has been introduced for data sharing. Based on the proposed methodology, a case study of Pierrefonds sewer network system is applied using the framework, and four implementations have been demonstrated. The first implementation not only discusses

data availability but also deals with the data inconsistency between attribute tables and CAD drawing maps. Moreover, the spatial and attribute data have been integrated into a GIS system. The second implementation focuses on data completeness by applying both automated and visual inspections to detect various errors for spatial data like data redundancy. In the third implementation, the positional data accuracy has been examined. The horizontal test has been applied and sample spatial data were collected by using GPS. Furthermore, metadata were created and developed by using DataLogr. In addition, a web-based GIS system for Pierrefonds underground infrastructure system, which includes both sewer and water network systems, has been designed and developed.

## **5.2 CONCLUSIONS AND CONTRIBUTIONS**

The conclusions and contributions of this research are grouped into the following items:

- (1) An innovative municipal infrastructure management approach has been developed which first apply the Six Sigma DAMIC methodology to municipal infrastructure spatial data focusing on three aspects of geographic features: typology, location, and spatial dependency. This new approach addresses technical aspects of quality and quantitative methods. Furthermore, it presents solutions for spatial data quality issues and requirements in municipal infrastructure management including data availability and completeness, data accuracy, and data attribute by integrating GIS maps, CAD models, and databases of the facilities' information.
- (2) A technical method has been proposed for linking different facility layers without the necessary information like key index. This method is based on using ArcMap to

retriever the attribute table related information and to save this information in a spatial database.

- (3) Visual and automated inspection methods, implemented in VB program, have been developed to detect both spatial and attribute data errors when applying data inspection.
- (4) The process of spatial data collection for implementing positional accuracy analysis has been developed and the horizontal positional accuracy analysis is applied in practice.
- (5) A new integrated WebGIS municipal infrastructure system has been designed and developed which demonstrates the general application for facilitating the exchange of data between individuals and organizations.

### **5.3 FUTURE WORK**

While pursuing this research, several limitations have been identified related to the requirements and the performance of the developed methods and techniques. Future work will consider the following issues:

- (1) The usability of the proposed framework for MIMS needs more testing. Further development and testing in practical situations are necessary to improve the functionalities and usability of the system.
- (2) The interoperability of the system can be enhanced by extending the MIMS representation in the system databases to cover more aspects of municipal infrastructure management (e.g., cost estimation and scheduling of maintenance activities, etc.).

(3) Creating vertical positional analysis can be applied for underground facilities including the slope measurement to show the accuracy of facilities in three dimensions.



## REFERENCES

- 5700/5800 GPS Receiver Guide, 2003, Engineering and Construction Division, Trimble Navigation Limited, Dayton, Ohio, USA.
- Abdelhamid, T. S., 2003, "Six-Sigma in Lean Construction Systems: Opportunities and Challenges." Proc., 11th Annual Conf. of the International Group for Lean Construction (IGLC-11), Blacksburg, Va.
- Alesheikh, A., Helali, H., and Behroz, H., 2002, "Web GIS: Technologies and Its Applications", Symposium on Geospatial Theory, Processing and Applications, Ottawa, CA.
- ANSI (American National Standards Institute), 1998, Information Technology - Spatial Data Transfer Standard (SDTS), ANSI-NCITS 320, New York.
- APISQ, 2008, Québec Alliance for the Protection of Underground Facilities<[www.apisq-qcga.ca](http://www.apisq-qcga.ca)>, [cited on Feb., 2008].
- ArcGIS 9, 2008, the Complete Geographic Information System 9, an ESRI white paper, California, USA.
- ASCE (American Society of Civil Engineers), 2003, Standard Guidelines for the Collection and Depiction of Existing Subsurface Utility Data, ASCE Standard N. 38-2 CI.
- AutoDesk, 2008, Autodesk MapGuide Services & Support, AutoDesk Inc., <<http://usa.autodesk.com/adsk/servlet/home?siteID=123112&id=129446>>, [cited on December, 2008].
- Bean, J. E., 2003, Factors Affect GPS Position Accuracy, Effective Use of the ConnDOT GPS Base Station, the Connecticut Department of Transportation, University of Connecticut, Feb. 2003.
- Beard, M. K., 1989, User Error: The Neglected Error Component. Proceedings, AutoCarto, 9, pp. 808-17.
- Bedell, R. L., Lei, Y., and Heston, G. S., 1997, Spatial Data Accuracy and Its Importance to Effective Mineral Exploration, Integrated Exploration Information Management, Fourth Decennial International Conference on Mineral Exploration, Nevada, U.S.A..
- Bentley, 2008, Mapping with All the Power and Accuracy of MicroStation, Bentley Inc., <<http://www.bentley.com/en-US/Products/Mapping>>, [cited on December, 2008].

- Berry, B., 1964, Approaches to Regional Analysis: a Synthesis, *Annals of the Association of American Geographers*, pp. 54:2-11.
- Buggie, F. D., 2000, "Beyond six sigma", *Manage. Eng.*, 16(4), pp. 28–31.
- Burkholder, E. F., 2004, Fundamentals of Spatial Data Accuracy and the Global Spatial Data Model (GSDM), Global COGO, Inc., Las Cruces, NM, April, 2004.
- Burrough, P. A., 1986, Principles of Geographical Information Systems for Land Resource Assessment. Clarendon Press. Oxford.
- Byong, L., Young, K., and Jin, Y., 1998, Web Interface for GIS in Agriculture, Agricultural Information Technology in Asia and Oceania, Kyounggi Province, Rep. of Korea.
- Cha, H. S., and O'Connor, J. T., 2005, "Optimizing Implementation of Value Management Processes for Capital Projects." *J. Constr. Eng. Manage.*, 131(2), pp. 239–251.
- Chrisman, N. R., 1984, the Role of Quality Information in the long-term functioning of a geographic information system, *Cartographica*, Vol. 21, pp.78-87.
- CityGML, 2008, CityGML, <[www.citygml.org/1523/](http://www.citygml.org/1523/)>, [cited on August, 2008].
- CityWorks, 2007, Cityworks Overview white paper, Azteca Systems, inc., 11075, South State Street, Sandy, UT.
- CGA (Common Ground Alliance), 2005, Common Ground Alliance, <[www.commongroundalliance.com](http://www.commongroundalliance.com)>, [cited on June 12, 2007].
- Cumberbatch, S., 2005, What lies beneath? Surveyor, SVY0505 pp.14-15.
- CWWA (Canadian Water and Wastewater Association), 2004, Water Infrastructure: Research for Policy and Program Development, < [http://www.infc.gc.ca/research-recherche/results-resultats/rs-rr/rs-rr-2004-01\\_01-eng.html](http://www.infc.gc.ca/research-recherche/results-resultats/rs-rr/rs-rr-2004-01_01-eng.html)>, [cited on December 2006].
- Danylo, N. and Lemer, A., 1998, Asset Management for the Public Works Manager: Challenges and Strategies, Findings of the APWA Task Force on Asset Management.
- Defoe, J. and Barnard, W., 2005, JURAN Institute's Six Sigma Breakthrough and Beyond - Quality Performance Breakthrough Methods, Tata McGraw-Hill Publishing Company Limited.
- Devillers, R. and Jeansoulin, R., 2006, Fundamentals of Spatial Data Quality, ISTE, Great Britain and the United States, pp. 17.

- Dziadok, K., Sommerville, J., and Kumar, B., 2008, RFID Based 3D Buried Assets Location System, Glasgow Calcdonain University, Scotland, ITCon, Vol. 13, pp.55.
- ESRI, 1998, ESRI Shapefile Technical Description, an ESRI White Paper, July 1998, <<http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>>, [cited on July, 2007].
- ESRI, 2003, Spatial Data Standards and GIS interoperability, an ESRI White Paper, January 2003, <<http://www.esri.com/library/whitepapers/pdfs/spatial-data-standards.pdf>>, [cited on June, 2007].
- ESRI, 2008a, Environmental Systems Research Institute Data Models, <<http://support.esri.com/index.cfm?fa=downloads.dataModels.gateway>>, [cited on March, 2008].
- ESRI, 2008b, ESRI-Supported Open Geospatial Consortium and ISO/TC 211 Standards, an ESRI White Paper, June, 2008.
- FGDC (Federal Geographic Data Committee), 1998, Geospatial Positioning Accuracy Standards; Part 3: National Standard for Spatial Data Accuracy, FGDC-STD-007.3. Federal Geographic Data Committee. Washington, D.C.
- FGDC (Federal Geographic Data Committee), 2006, Federal Geographic Data Committee, <[www.fgdc.gov/metadata](http://www.fgdc.gov/metadata)>, [cited on March, 2008].
- Foody, G., 2001, GIS: the Accuracy of Spatial Data Revisited, Progress in Physical Geography, 25(3), pp.389-398.
- GeoModel, 2008, Ground Penetrating Radar, Basic Operating Principles, <[www.geomodel.com/gprtext.htm](http://www.geomodel.com/gprtext.htm)>, [cited on August, 2008].
- Gilbert, C., 1996, the Software with Your GPS Part 1 – Data Export, Mapping Awareness, 10(8), pp. 32–4.
- GIS Topology, 2005, an ESRI White Paper, ESRI 380 New York St., Redlands, CA, USA, July 2005, <[http://www.esri.com/library/whitepapers/pdfs/gis\\_topology.pdf](http://www.esri.com/library/whitepapers/pdfs/gis_topology.pdf)>, [cited on March, 2007].
- Goodchild, M. F., 1991, Issues of Quality and Uncertainty. In Müller J-C (ed.) Advances in cartography, Oxford, Elsevier Science, pp. 111–39.
- Goodchild, M. F., 1992, Geographical Information Science, International Journal of Geographic Information Systems, pp.31-45.
- Google Earth, 2006, Google Earth KML Documentation, <<http://earth.google.com/kml>>, [cited on February, 2006].

- GPS, 2006, Global Positioning System, Kowoma, German, <<http://www.kowoma.de/en/gps/index.htm>>, [cited on March 4, 2007].
- Halfawy, M. R., Pyzoha, D., Young, R., Abdel-Latif, M., Miller, R., Windham, L., and Wiegand, R., 2000, GIS-based sanitary sewer evaluation survey, 20th Annual ESRI International User Conference, Jun., San Diego, CA.
- Halfawy, M. R., Vanier, D. J. and Hubble, D., 2004a, Integration of Municipal Information Systems for Sustainable Management of Infrastructure Assets, Environmental Informatics Archives, Vol. 2, ISEIS.
- Halfawy, M., 2004b, The Interoperability of Geographic Information Systems for Municipal Asset Management Applications, Municipal Infrastructure Investment Planning (MIIP) Project Report, Part II, National Research Council Canada.
- Halfawy, M., Newton, L., and Vanier, D., 2005, Municipal Infrastructure Asset Management Systems: State-of-the-Art Review, CIB W78 Conference on Information Technology in Construction, Dresden, Germany, pp. 1-8.
- Halfawy, M. and Figueroa, R. T., 2006, Developing Enterprise GIS-Based Data Repositories for Municipal Infrastructure Asset Management, Joint International Conference on Computing and Decision Making in Civil and Building Engineering.
- Hansen, 2007, Hansen Information Technologies, <[www.hansen.com](http://www.hansen.com)>, [cited on July 10, 2008].
- Harfan, 2007, Harfan Infrastructure Management Approach, <[www.harfan.com/eng/Home/index2.asp](http://www.harfan.com/eng/Home/index2.asp)>, [cited on June 28, 2007].
- Helali, H., 2001, Design and Implementation of a Web GIS for the City of Tehran. M. Sc thesis, Department of Geodesy and Geomatics Engineering K. N. Toosi University of Technology, Tehran, Iran.
- HIMA, 2008, L'approche Harfan Engestion des Infrastructures, Plan d'Intervention, Pont-Rouge, Quebec.
- Hunter, G. J., Hope, S., Sadiq, Z., Boin, A., and Marinelli, M., N, 2005, Next-Generation Research Issues in Spatial Data Quality, CRC fro Spatial Information Department of Geomatics, University of Melbourne, Parkville.
- Infrastructure 2000, 2007, Vanasse Hangen Brustlin Inc., < [www.vhb.com](http://www.vhb.com)>, [cited on July 10, 2007].
- Infrastructure Canada, 2004, Assessing Canada's Infrastructure Needs: "A Review of Key Studies", Research and Analysis Infrastructure Canada.

- Info-Excavation, 2006, 1600 Henri-Bourassa Blvd. West, Office 340, Montreal, QC, <[www.infox.com/index.php?option=com\\_content&task=view&id=74&Itemid=164](http://www.infox.com/index.php?option=com_content&task=view&id=74&Itemid=164)>, [cited on May, 2007].
- Infotech, 2007, Accuracy Assessment, Infotech Enterprises America, Geospatial Division, Loudoun County Phase IV, <[www.loudoun.gov](http://www.loudoun.gov)>, [cited on December, 2008].
- Intergraph, 2008, the Power to See, Products, <<http://www.intergraph.com/default.aspx>>, [cited on December, 2008].
- ISO8879, 2008, International Organization for Standardization, <[www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=2855](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=2855)>, [cited on March, 2008].
- Kaplan, E., 1996, Understanding GPS Principles and Applications, Artech House Publishers, New York, USA.
- Kim, L., 2006, Fundamentals of Spatial Data Quality. ISTE Ltd., UK and USA, 2006, pp. 91.
- Lange, A. F. and Gilbert, C., 1999, Using GPS for GIS Data Capture, Geographical Information Systems: Principles and Technical Issues (2nd. ed.), John Wiley & Sons, Vol.1:475.
- Lenz, E., 2004, Networked Transport of RTCM via Internet Protocol (NTRIP): Application and Benefit in Modern Surveying Systems, FIG Working Week 2004, Athens, Greece.
- Lindermann, K., Schroeder, R. G., Zaheer, S., and Choo, A. S., 2003, "Six sigma: A Goal-Theoretic Perspective." J. Operations Manage, 21, pp. 193–203.
- Lior, S.K., Chambers, C.L., and Spollen, J.F., 1991, Norfolk uses a GIS to satisfy its stormwatermanagement information needs, Proceedings: Conference on Civil Engineering Applications of Remote Sensing and Geographic Information Systems, May, Washington,DC, pp. 97-104.
- Madhwal, H. B., 2002, Precise Positions with GPS in WGS 84 Datum- Accuracy Evaluation of Positions, Survey of India, Dehradun-248001, India.
- McCain, M. J. and Masters, W. C., 1998, Integrating Quality Assurance into the GIS Project Lifecycle, ArcUser, July-September, pp. 42-5.
- Merrick, 2008, Merrick Advanced Remote Sensing (MARS), <[www.merrick.com/servicelines/GIS/surveying.aspx](http://www.merrick.com/servicelines/GIS/surveying.aspx)>, [cited on March, 2008].
- Microlise, 2002, Consideration for Successful RFID Implementations, ZDNet, UK.

- MIDS (Municipal Infrastructure Data Standard) 3.0, 2005, Tri-Committee for the Utilization of Information Technology in Public Works, Road Geobase acquire from Montreal City, February, 2005.
- MIMS, 2007, Municipal Infrastructure Management Systems, <[www.albertamims.org](http://www.albertamims.org)>, [cited on July 10, 2007].
- Montgomery, D. C., 2005, Introduction to Statistical Quality Control, Fifth Edition, John Wiley & Sons, pp.716
- Moutal, H. P. and Bowen, D. R., 1991, Updating New York City's sewer and water main distribution systems: practical application of GIS, Proceedings of the Conference on Civil Engineering Applications of Remote Sensing and Geographic Information Systems, May, Washington, DC, pp. 155-164.
- MRNF (Ressources Naturelles et Faune Québec), 2008, GPS en temps réel par téléphonie cellulaire, Ressources Naturelles et Faune Québec, <[www.mrnf.gouv.qc.ca/territoire/outils/outils-cellulaire.jsp](http://www.mrnf.gouv.qc.ca/territoire/outils/outils-cellulaire.jsp)>, [cited on July, 2008].
- NISO, 2004, Understanding Metadata, National Information Standards Organization, NISO Press, Bethesda, MD 20814, USA.
- NMEA, 2007, NMEA 0183 Standard, <[www.nmea.org/pub/0183](http://www.nmea.org/pub/0183)>, [cited on May, 2007].
- NRC, 2003, Selection of Technologies for Sewer Rehabilitation and Replacement, The National Guide to Sustainable Municipal Infrastructure, NRCC-48198.
- NRC, 2004, Municipal Infrastructure Investment Planning (MIIP) report, Geographic Information Systems (GIS) and Interoperability of Software for Municipal Infrastructure Applications, B-5123.3.
- NSSDA, 1999, Positional Accuracy Handbook, Using the National Standard Spatial Data Accuracy to Measure and Report Geographic Data Quality, Minnesota Planning Land Management Information Center, pp. 3-8.
- OGC, 2008, Open Geospatial Consortium Inc., <[www.opengeospatial.org](http://www.opengeospatial.org)>, [cited on May, 2008].
- Olaleye, J. B. and Sangodine, J. O., Network Analysis in GIS: A Tool for Transportation and Utility Management – A Case Study of Nigerian Transportation and Utility Management, FIG XXII International Congress, Washington, D.C. USA, April, 2002.
- ORCGA, 2007, Ontario Regional Common Ground Alliance , <[www.orcga.com/home.asp](http://www.orcga.com/home.asp)>, [cited on Jan., 2008].
- Przybla, J. and Kiesler, C. L., 1991, Extending GIS capability for enhanced sewer systems modelling, Proceedings of the Conference on Civil Engineering Applications

- of RemoteSensing and Geographic Information Systems, May, Washington, DC, pp. 105-114.
- Qasem, A., and Hammad, A., 2006, Towards Integrated Municipal Infrastructure Management Systems, INFRA 2006: 12th Annual Week on Urban Infrastructure, Quebec City, Canada.
- QUT, 2006, Design Standards and Guidelines Edition 4, A3, Appendix 2 of the Facilities Management Procedures Manual, Queensland University of Technology, Facilities Management, September 2006.
- Rasford, W., Cai, H., Tilley, C., Brun, S., Karimi, H. and Robson, F., 2003, Transportation Distance Measurement Data Quality, Computing in Civil Engineering, 17(2), pp.75-87.
- Real-time Asset Valuation Analysis, 2007, Real-time Infrastructure Valuation Analysis, <[www.rivaonline.com](http://www.rivaonline.com)>, [cited on July, 2007].
- Rogerson, P. and Fotheringham, S., 1994, GIS and Spatial Analysis: Introduction and Overview, in Spatial Analysis and GIS (S. Fotheringham and P. Rogerson, eds.), Taylor and Francis, London, pp. 281.
- Scanseal, 2008, Survey and Drain Repairs, <[www.scanseal.com/cctv\\_surveys.shtml](http://www.scanseal.com/cctv_surveys.shtml)>, [cited on July, 2008].
- Schwalbe, K., 2007, Information Technology Project Management, Fifth Edition, Chapter 8: Project Quality Management, pp.320-1.
- SDTS, 2008, "What is SDTS", US Spatial Data Transfer Standard, SDTS Task Force U.S. Geological Survey, 1400 Independence Road Rolla, MO 65401.
- Shawn, M. S., and Kevin, R. P., 2008, the GIS Nucleus: A SynerGISTIC Mechanism for Wastewater Operations Management, ESRI, 2008.
- Sinton, D. F., 1978, The Inherent Structure of Information as a Constraint in Analysis, In Dutton, G.(ed.) Harvard papers on geographic information systems, Reading (USA), Addison-Wesley.
- Snee, R. D., 1999, "Why Should Statisticians Pay Attention to Six Sigma?" Quality Progress, September, pp. 100-3.
- Statistics Canada, 2006, the Age of Public Infrastructure in Canada, Analytical Paper, National Research Council Canada, Ottawa.
- Subash, S. and Padak, A., 2003, Enterprise GIS for Municipalities – An Integrated Approach, Map India Conference, 2003.

- Synergen, 2007, Synergen Infrastructure Management Systems, Synergen Inc., <[www.splwg.com/solutions/SYNERGEN/synergen.asp](http://www.splwg.com/solutions/SYNERGEN/synergen.asp)>, [cited on July 12, 2007].
- Takashi, T. and Koji, M., 2004, Tendency of Standardization of Spatial Data and Present State of Its Application to Production Data in Japan, Kokusai Kogyo, Co., Ltd., <[www.kkc.co.jp/english](http://www.kkc.co.jp/english)>, [cited on December , 2008].
- TDS, 2008, Tripod Data System, <[www.tdsway.com/products/land\\_surveying](http://www.tdsway.com/products/land_surveying)>, [cited on March, 2008].
- Thomas, H. R., Horman, M. J., Minchin, E., and Chen, D., 2003, “Improving Labor Flow Reliability for Better Productivity as Lean Construction Principle.” *J. Constr. Eng. Manage.*, 129(3), pp. 251–261.
- Timble, 2008, Timble GPS Tutorial, <[www.trimble.com/gps/whatgps.shtml](http://www.trimble.com/gps/whatgps.shtml)>, [cited on March, 2008].
- Toth, K. and de Lima, V. N., 2005, Data Quality and Scale in Context of European Spatial Data Harmonization, ESDI: Setting the Framework, 11<sup>th</sup> EC-GIS Workshop, Alghero, Sardinia, July, 2005.
- Tsai, Y. C., and Frost, J. D., 1999, “Using Geographic Information System And Knowledge Base System Technology For Real-Time Planning Of Site Characterization Activities.” *Can. Geotech. J.*, Vol. 36, pp. 300-312.
- Vanier, D., 2000, Municipal Infrastructure Investment Planning (MIIP) Project: Statement of Work, Inst. for Research in Construction, National Research Council Canada.
- Vanier, D. J., and Rahman, S., 2004, A Primer on Municipal Infrastructure Asset Management, MIIP, Client Report B5123.3, National Research Council Canada, Ottawa.
- Veregin, H., 1999, Data Quality Parameters, *Geographical Information Systems: Principles and Technical Issues* (2nd. ed.), John Wiley & Sons, Vol.1:184.
- VHB, 2008, RoadManager GPMS, Geographic Pavement Management System Infrastructure Asset Management System, 54 Middletown, Connecticut.
- W3.org, 2008, Extensible Markup Language, <[www.w3.org/XML](http://www.w3.org/XML)>, [cited on Aug., 2008].



WolframMathworld, 2008, Spherical Coordinates from Wolfram MathWorld, <<http://mathworld.wolfram.com/SphericalCoordinates.html>>, [cited on March, 2008].

Worboys, M.F., 1995, GIS: A Computing Perspective, Taylor & Francis, London.

XML, 2005, Focus Area News, XML org., <[www.xml.org](http://www.xml.org)>, [cited on Nov., 2007].

## APPENDIX A: GPS Configuration

In order to collect spatial data and upload spatial data to ArcGIS, it is very important to configure GPS and connect GPS receiver to GIS properly. The detail setting steps about Trimble GPS 7500 receiver are describe as following:

### A.1 GPS physical connection

(1) The GPS should be off before connecting. A serial cable to the GPS receiver, a computer to run ArcGIS, and a male-to-male gender changer is required. Since the COM connection cable of GPS is female, link COM1 in computer, as shown in Figure A.1. If computer does not have a COM port, link USB port with the USB-COM convert cable.

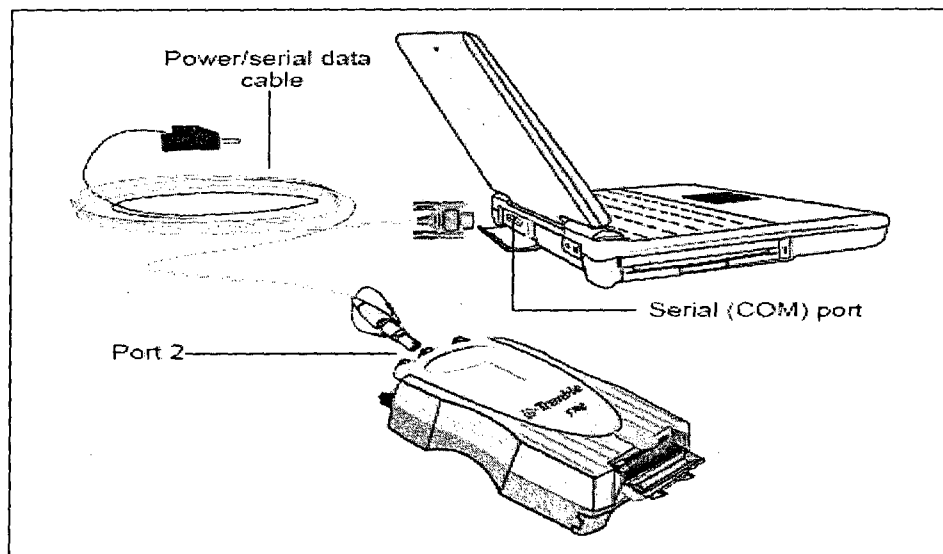


Figure A.1 Trimble GPS 5700 receiver physical connection (5700/5800 GPS Receiver User Guide 2003)

(2) As shown in Figure A.2, connecting the black wire (with a serial connection at the end) into socket 2. The serial connection will be connected to the computer;

meanwhile, connecting the yellow wire to the socket marked GPS. The other end of this wire will be screwed into the connection on the antennae.

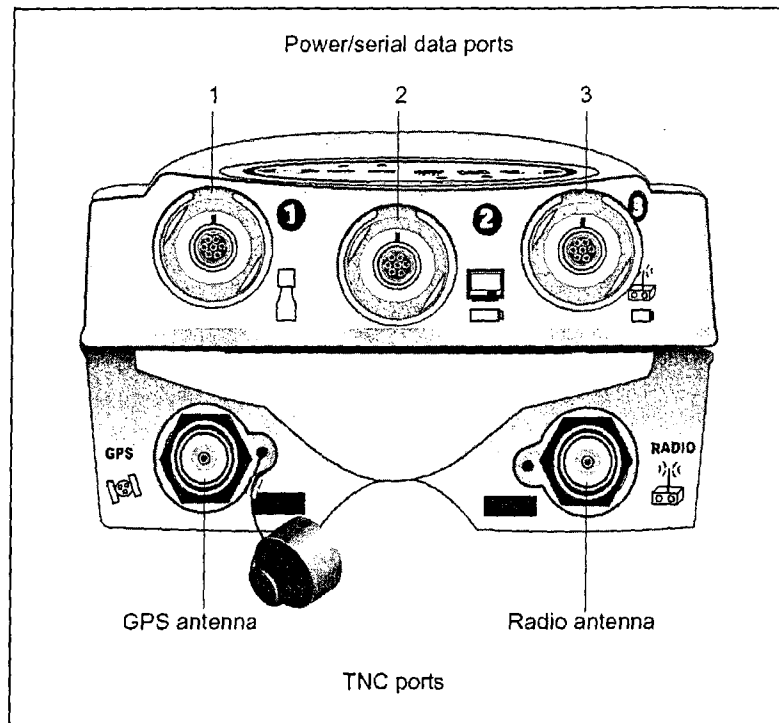


Figure A.2 Trimble GPS 5700 receiver top panel (5700/5800 GPS Receiver User Guide 2003)

## A.2 GPS receiver configuration setting

By default, most handheld GPS receivers are configured to not output any GPS data; therefore, we need to configure GPS receiver to output data by setting the NMEA protocol (NMEA 2007). In our case, we use configuration software such as Trimble GPS Configurator, Trimble Configuration Toolbox to set Trimble 5700. We set the port 2 as the output port and output frequency is 44.2500 MHz Hz/sec. Since NMEA 0183 GGA which is essential fix data providing 3D location and accuracy data, we add the output protocol as the NMEA-GGA for port 2 and also set the baud rate of port 2 as 9600 as shown in Figure A.3.

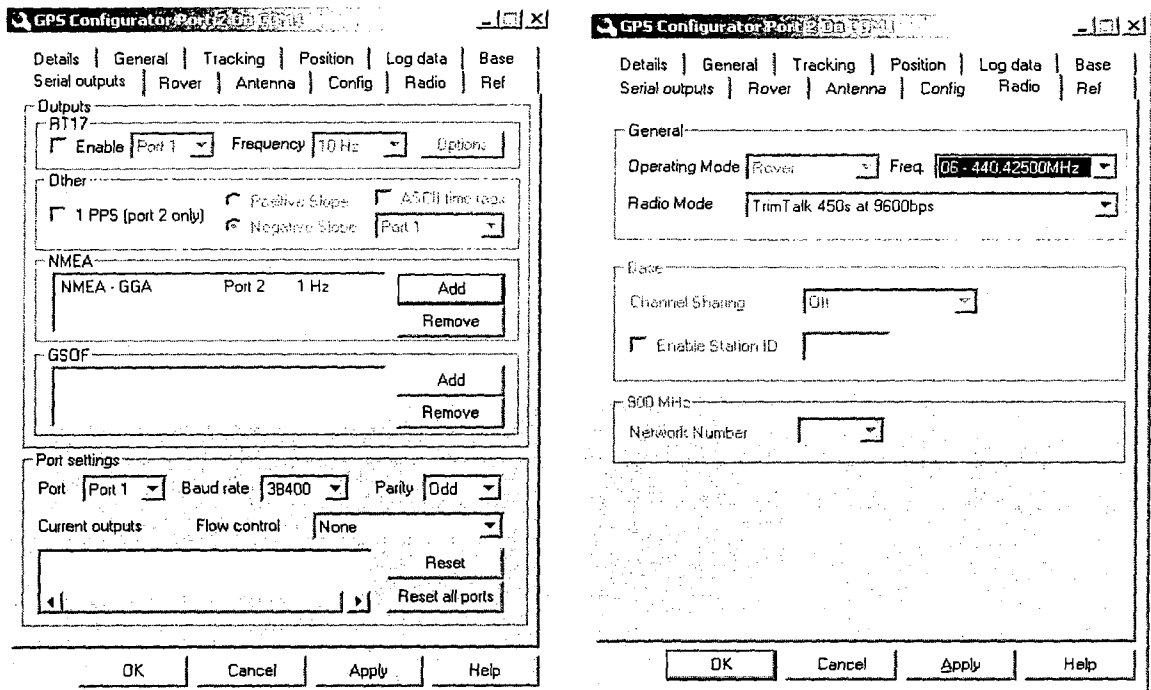


Figure A.3 Trimble GPS 5700 receiver configuration setting

### A.3 Connection GPS to ArcPad

Before activating GPS, we have to set the GPS communication parameters in ArcPad to match the parameters set on GPS receiver. In ArcPad, from the icons on the toolbar, select from the GPS position option>GPS Tracklog. This will start to log data to the GPS.

If the GPS doesn't connect, select from the tools menu>find GPS. There are two pages of communication parameters in the ArcPad Options dialog box: The protocol page and GPS page. They have to be set according to GPS receiver configuration as shown in Figure A.4.

### A.4 Testing data by using Accessories "Hyper Terminal"

In order to verify if GPS equipment is properly connect, the Hyper Terminal testing has been applied before collecting data onsite. We firstly open the Window accessories

“Hyper Terminal”, configure the port setting as shown in Figure A.5, and then connect to GPS as shown in Figure A.6.

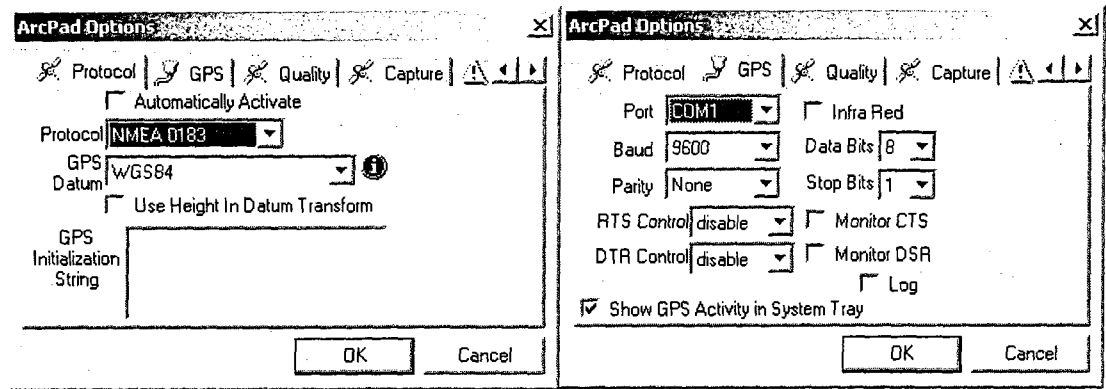


Figure A.4 ArcPad communication parameters setting

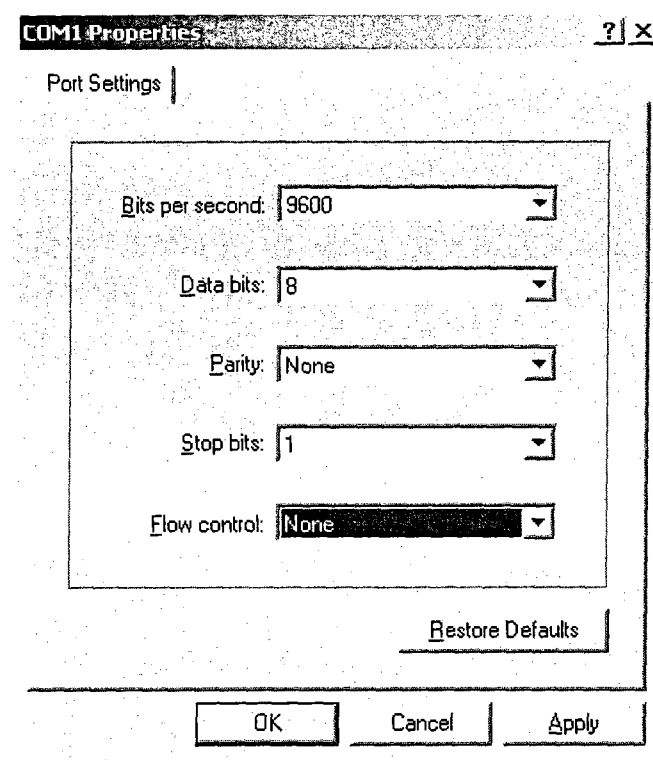


Figure A.5 Port settings

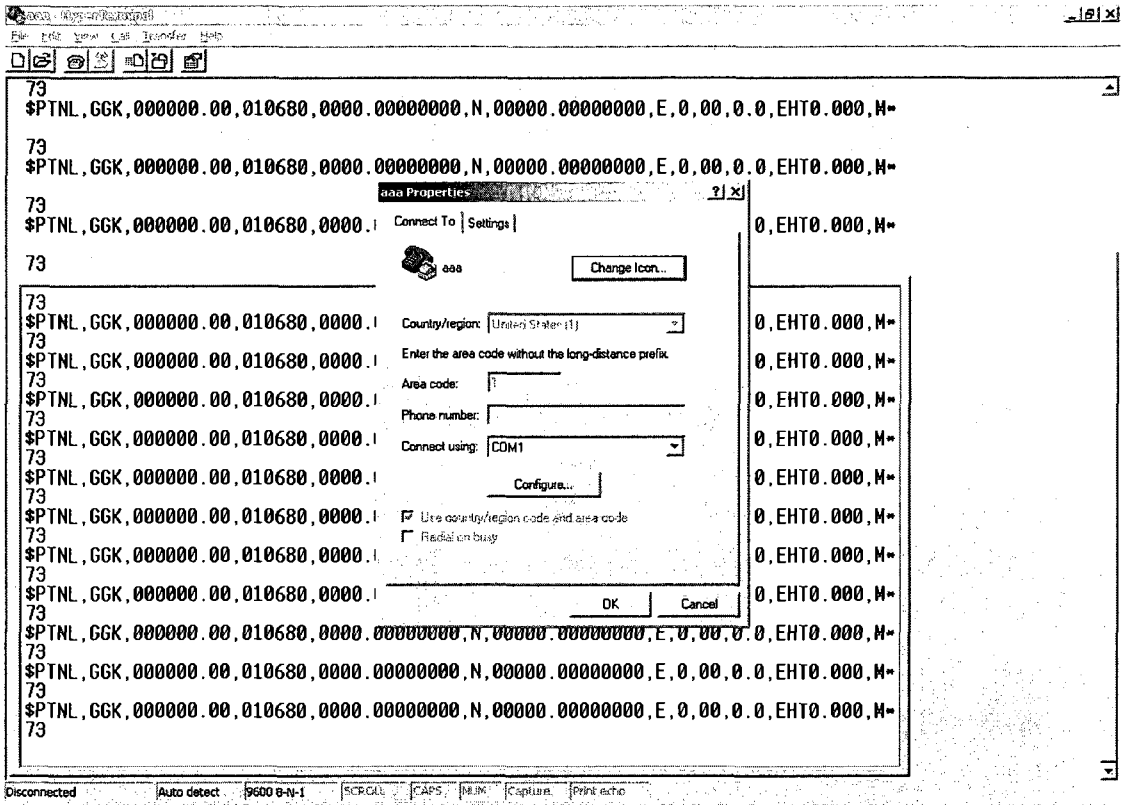


Figure A.6 Connecting to GPS

## **A.5 Troubleshooting**

### ***Is the GPS able to connect at all with the computer?***

Sometimes there is a problem with the physical connection, or the computer may be trying to connect with a different port.

To see if the GPS will connect to the computer use the Trimble GPS configuration software. To open this program select from the start menu > programs > Trimble>GPS Configurator>GPS Configurator> choose 5700 from the list of Trimble devices and select connect. If the GPS can connect to the computer, you will be able to connect using any of the programs. If you cannot connect via this method, there is a problem with the cables.

### ***When you are trying to connect with ArcPad, did somebody change the default setting?***

In order to restore the settings, we have to select from the start menu > prgrams > Trimble > Configuration Toolbox > Configuration Toolbox. From the File drop down menu, we select Open. From the AppFiles Directory (this is the default directory) select the file named CURRENT to open. A dialog box will open. Select Transmit. Select Save and ten select Close. Once these settings are restored, The GPS will be able to be used in ArcPad.

### ***How to connect the GPS in the GPS program?***

1. The GPS must be turned off before it is connected to the computer.
2. When the GPS is off, connect the serial or USB port to the computer.
3. Load the GPS program.
4. When the access style window opens, choose local and click OK.
5. On the Login Viewer window, choose building option and click login button.
6. Click "Show 2D Map" to display the target map.
7. Click "GPS Track" to engage the GPS.
8. To record data click "Record Locations".

### ***When you are trying to connect with GPS Program, do you use a USB or Serial Port?***

The program may not be looking in the right place for a connection with the GPS. You must change the code manually.

If you are using a serial port, the program must look for the GPS connection in COM 1; If you are using a USB cable, the Program may find the connection in either COM 3, 4 or 5. You must find the correct COM number to change.

In computer directory, select properties, hardware, device manager, and ports, you just need to make note of the number of the com port so that you can adjust it accordingly in the GPS program.

Now that you know the right COM port, you will change it in the program

***Does the battery of the GPS have power?***

Check the battery level of the GPS in the GPS Configurator under the receiver tab.



## APPENDIX B: Manhole's Coordinates Collecting Procedure by Using GPS

### B.1 Collecting data in ArcPad by registering a proper projection

Once the data collecting on site, it is very important to select proper projection to guarantee the spatial data is registered to the same surface. In this case study, the projection “WGS84” has been chosen on ArcPAD when collecting data on site as shown in Figure B.1.

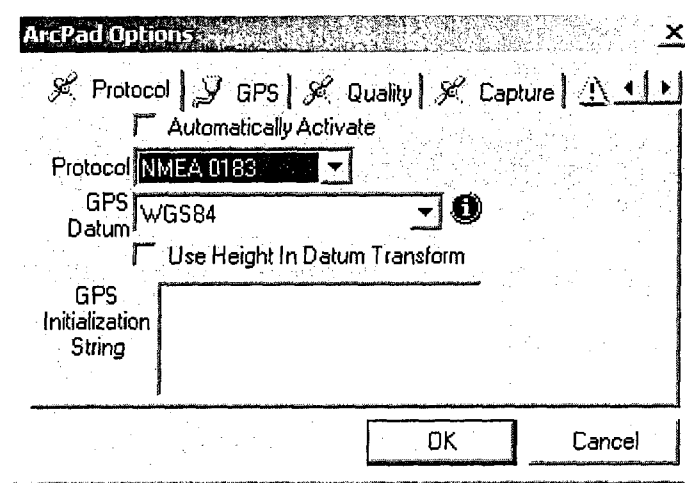


Figure B.1 Selecting “WGS84” as track data projection

### B.2 Upload tracking data to target map in ArcMap

Meanwhile, setting coordinate and projection should be considered when adding tracking layer to existing road map. In this case, a series of setting steps have to be conducted as following:

- (1) Overlaying Montreal road map to GPS data map by putting them in the same x, y coordinate system. Open the ArcMap and add the Montreal mainroad shape and TrackLog data in the same layer. However, since not selecting a projection, the two spatial data are apart far away as shown in Figure B.2.

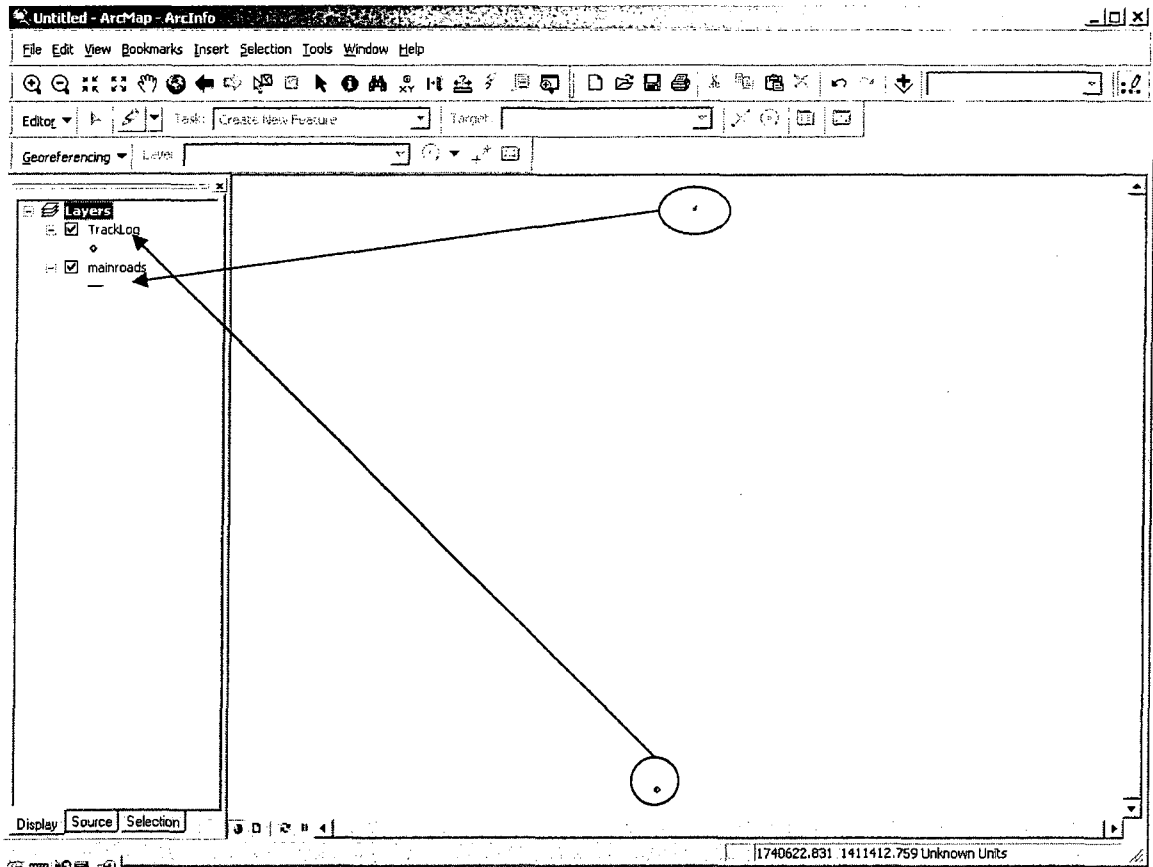


Figure B.2 Montreal mainroads and TrackLog layers are apart before projection

- (2) Open the properties as shown in Figure B.3 and choose the coordinate system and insert the MTM in the name space and enter the value.

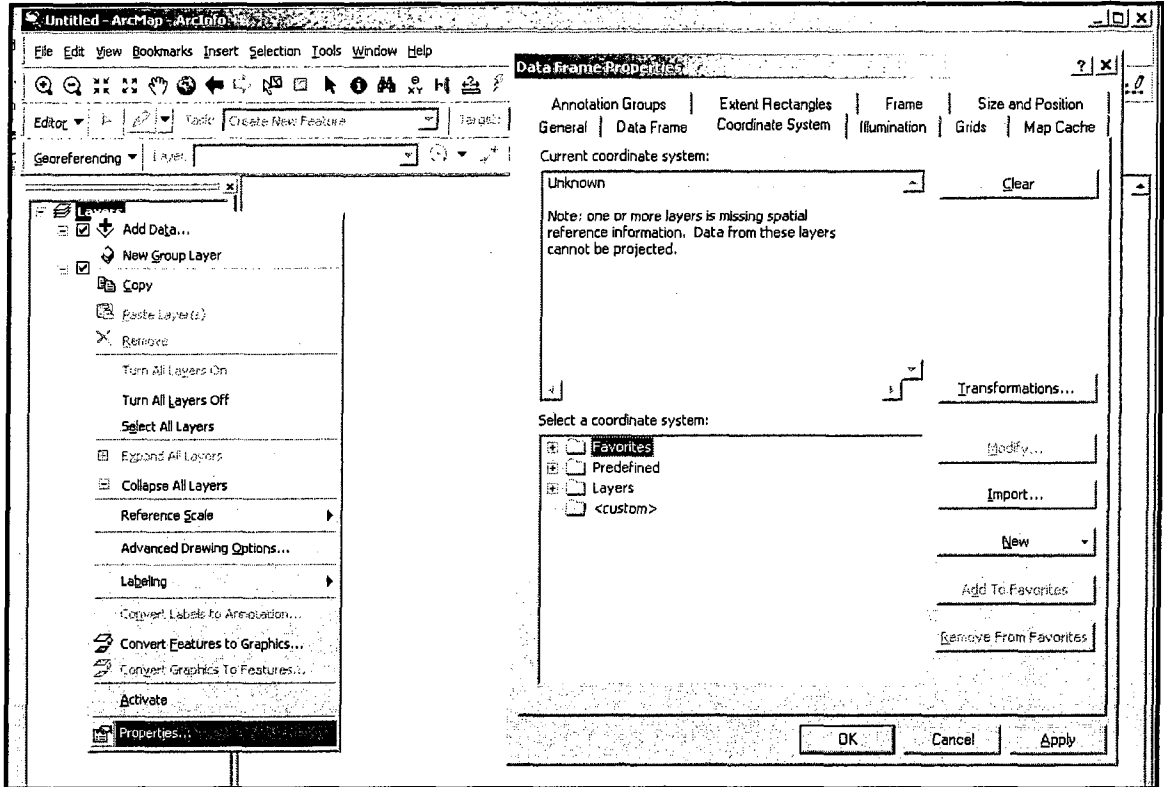


Figure B.3 Open layers properties

(3) Creating a new MTM value as shown in Figure B.4

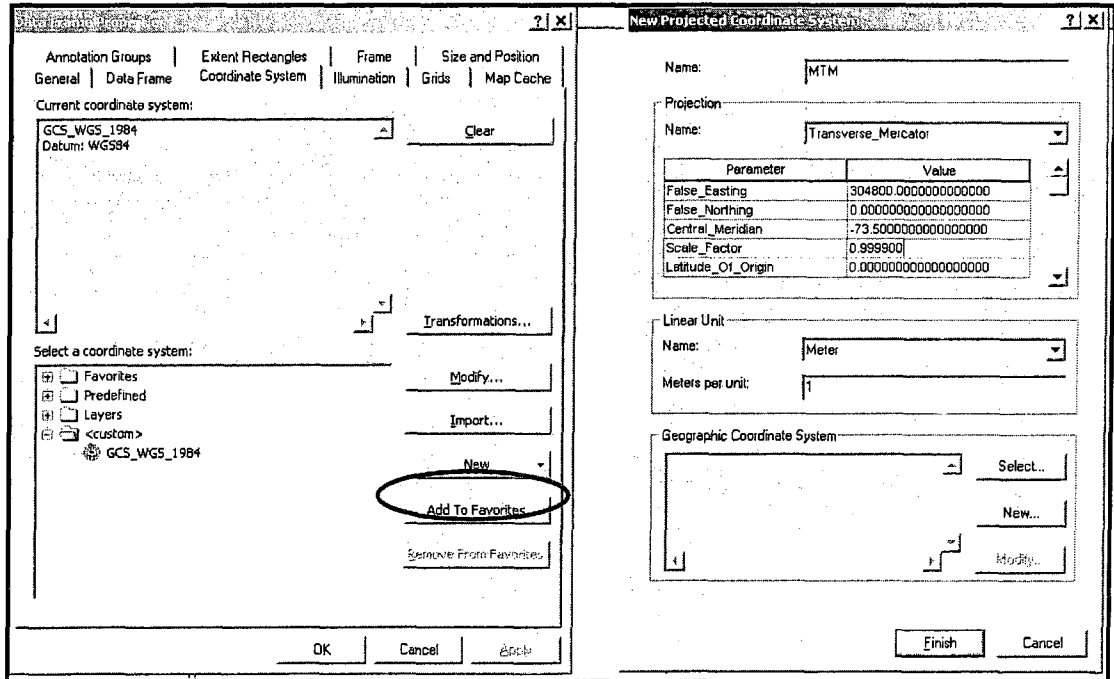


Figure B.4 Create new Transverse Mercator

(4) Choosing North American 1983 HARN as target “Coordinate System” as shown in

Figure B.5

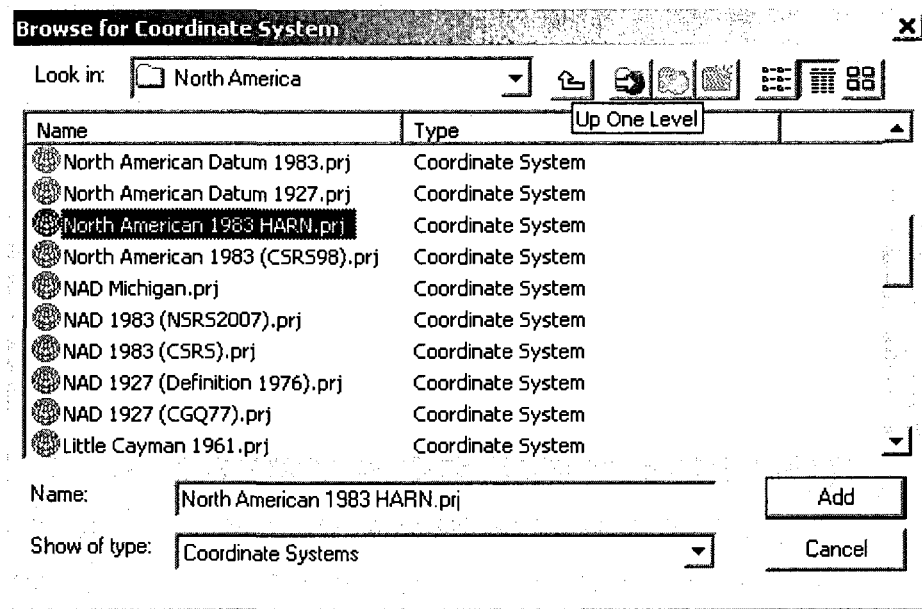


Figure B.5 North American 1983 HARN as “Coordinate System”

(5) After projection, the Montreal mainroad and TrackLog data are in the same layer as shown in Figure B.6.

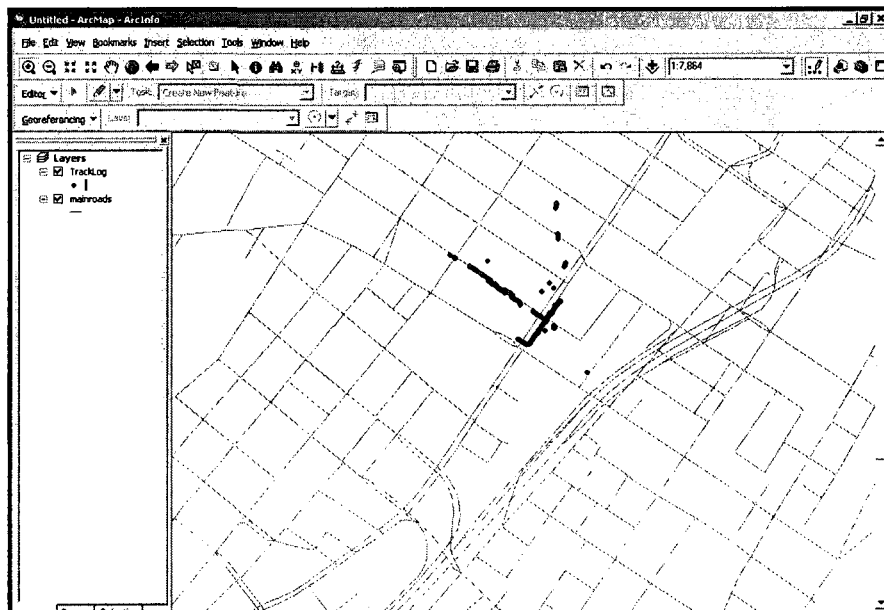


Figure B.6 Montreal mainroad and TrackLog data are in the same layer

## APPENDIX C: Calculating the X, Y Coordinates in Attribute Table

1. Open target layer attribute table and selection “Add Field” in options as shown in Figure C.1

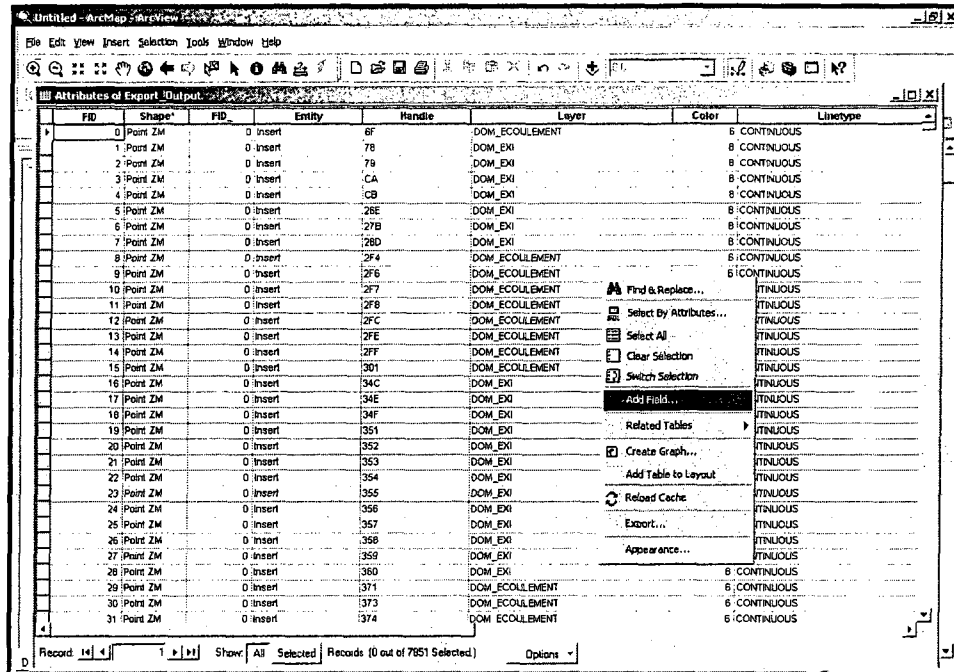


Figure C.1 Add field in attribute table

2. Make sure the attribute table is in “Editing Start” status, insert two fields into attribute table for X, Y coordinates’ value as shown in Figure C.2

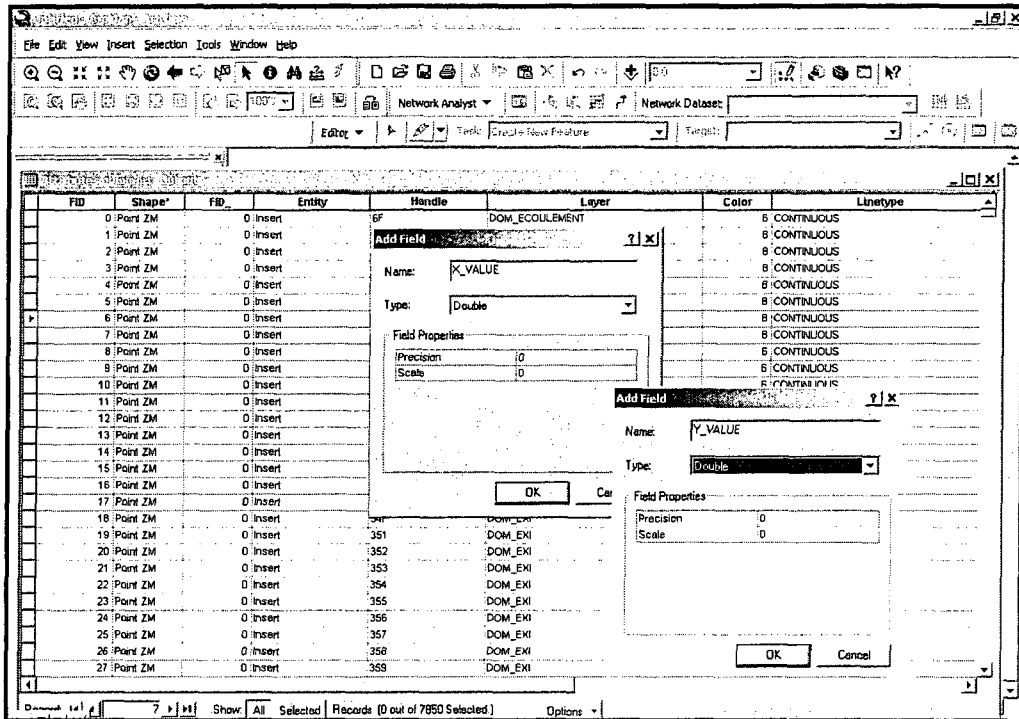


Figure C.2 Insert two fields in attribute table for X and Y coordinates' value

3. Calculate coordinates' value by writing VBA code in "Field Calculator" as shown in

Figure C.3 and the VBA code is written as below:

```
//For calculate x coordinate value
Dim dblX As Double
Dim pPoint As IPoint
Set pPoint = [Shape]
dblX = pPoint.X

//For calculate y coordinate value
Dim dblY As Double
Dim pPoint As IPoint
Set pPoint = [Shape]
dblY = pPoint.Y
```

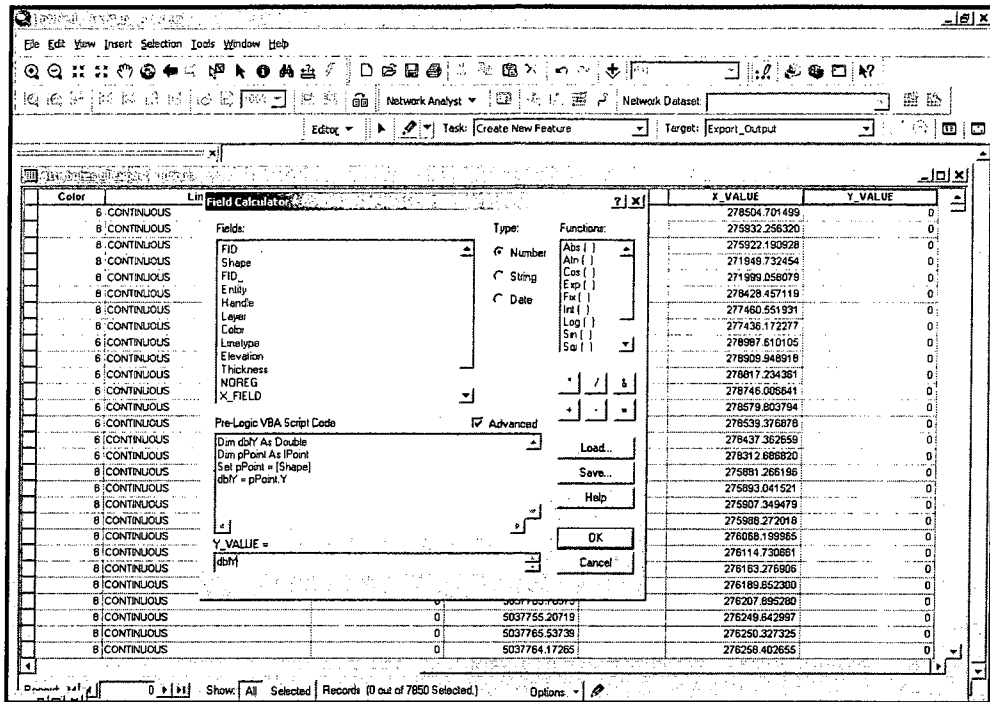


Figure C.3 Calculate x, y coordinates' value

- As x, y coordinate' value has been calculated in attribute table as shown in Figure C.4, we have to transferred it into dbf file for further analysis by using "Export Data".

X_VALUE	Y_VALUE
278504.701499	5040566.26575
275932.25632	5035867.52261
275922.190928	5035891.22224
271949.732454	5035783.14475
271999.058079	5035810.44459
278426.457119	5040173.18809
277460.551931	5039326.27845
277436.172277	5039303.54765
278987.810105	5040423.58355
278909.948918	5040392.53834
278817.234351	5040354.83813
278746.006641	5040326.02217
278579.803794	5040262.11085
278539.376878	5040235.21978
278437.362659	5040178.52981
278312.68682	5040107.15669
275881.268195	5037374.03278
275883.041521	5037374.83597
275907.349479	5037407.6068
275988.272018	5037499.34707
276068.199865	5037574.42808
276114.730661	5037617.55147
276163.278906	5037664.83887
276189.6523	5037636.28806
276207.89538	5037703.78375
276249.642897	5037755.20719
276250.327325	5037765.53739
276268.402655	5037764.17265
276007.259099	5037479.58571
275771.02856	5037445.55957
275963.429559	5037524.43082
275974.777159	5037484.02475
275983.922361	5037483.2193
276092.911118	5037587.27838
276140.120231	5037582.22981
276207.186316	5037718.54849
282206.16516	5041047.30145
275091.454269	5035980.44649

Figure C.4 the result of x, y coordinates' value



## APPENDIX D: Pierrefonds Sewer Pipe Network Metadata

```
0, "DataLogrV2", "SAMPLE.FTF"
1, -1, "GENERAL INFORMATION", ""
2, 1, "TITLE/ BRIEF DESCRIPTION", "Pierrefonds Borough"
2, 1, "GEOGRAPHIC AREA NAME(S)", "Sewer Pipe Network"
2, -2, "ARCHIVE/TRANSFER FILE INFORMATION", ""
3, 1, "ARCHIVE/TRANSFER FILE NAME", "Res_dom.dwg"
3, 1, "ARCHIVE/TRANSFER FILE TYPE", "Self-Extracting Compressed File."
2, -2, "DATA FILE/COVERAGE INFORMATION", ""
3, 1, "DATA FILE/COVER NAME(S)", "sewerpoit_attt.dbf,
sewerspolyline_attt.dbf"
3, 1, "DATA FILE(S) FORMAT", "ArcView ShapeFile"
3, 1, "LAST FILE EDIT DATE", "2008 OCT 31"
2, 1, "SPATIAL DATA TYPE", "Vector GIS"
2, 1, "REGISTERED TO BASE", "Not Applicable"
2, 1, "INTENDED USE/LIMITATIONS", "Township/City planning, evaluation, or
management. Not intended for site specific use."
1, -1, "TECHNICAL INFORMATION", ""
2, 1, "COORDINATE SYSTEM", "MTM"
2, 1, "PLANAR DISTANCE/MAP UNITS", "Meters"
2, 3, "SOURCE(S) OVERVIEW", "(1) scanned design drawings of sewers network
"
"(2) hard copies of inspection or maintenance reports (i.e.
intervention plan for sewer and water network for Pierreronds/
Roxboro)"
"(3) softcopies of CAD drawings of the network"
1, -1, "CONTACT INFORMATION", ""
2, -9, "CONTACT", "REPEATING GROUP, Occurrence #1"
3, 1, "PERSON and TITLE", "BaoJin Hu"
3, 1, "ORGANIZATION", "Concordia University"
3, 1, "MAILING STREET/P.O. BOX", "ba_hu@encs.concordia.ca"
3, 1, "CITY", "Montreal"
3, 1, "STATE", "Quebec"
3, 1, "PHONE", "514-8482424 * 7074"
2, -99, "END REPEATING GROUP", ""
```

### Data Structure Explanation

A DataLogr file is a comma-delimited ASCII file with each line formatted as follows:

**<Nesting Level>, <## Lines/non-field CODE>, "<Field Label>", "<Data>"**

**<Nesting Level>** The hierarchy of the data. The second DataLogr line (i.e., first data line) should be set to 1, sections should be set to 1, and all fields within a section should be set to 2 or greater. Numbers should never be skipped when nesting into deeper levels, but it is fine to skip numbers when 'de-nesting' (e.g., level sequence of 1-2-2-3-4-2-3-1)

<## **Lines/non-field CODE**> The number of lines this field occupies in the DataLogr file, or a code indicating the type of non-field line that this is (e.g., section title, repeating groups header). Values greater than zero are valid for field data. Values of zero or less are non-field codes. Valid Values:

**1, 2, 3 ... n:** How many lines in this DataLogr file does this field occupy (the current line counts as line 1)

**0:** The first line of the DataLogr file (DataLogr File Code, Setup Table used)

**-1:** A Section title (all data below belong to this section until the next '-1' code is encountered. On printouts with PrintStyle '0' (the old Version 1 style), this title will be spaced apart from other items accordingly. (60 characters maximum)

**-2:** A Sub-section title (all data following this sub-section and NESTED FURTHER than this sub-section belongs to this sub-section (until the next sub or data at the Sub-section's nesting level is encountered. On printouts, this sub-section title will be printed as it is encountered IF there is any data filled out within this section. (60 characters maximum)

**-3:** Other labels the user wants mingled in with the data. These will be printed as they are encountered with no spacing added before or after them.

**-9:** The start of a repeating group instance (if not in a group currently, this signals the first of possibly several groups - if in a group, this is the beginning of the next group instance).

**-99:** This signals the end of a repeating group section. Along with the '-99' code, the field label will be set to "END REPEATING GROUP" and the data line will be blank.

<**Field Label**> The label for this field, as defined in the fields template file (FTF). This information will be surrounded by double quote marks.

**<Data>** The first line of data associated with this field. This information will be surrounded by double quote marks. If the **<## Lines>** code is greater than one, then more lines of Data will follow, each line by itself on a line in the file, surrounded by double quotes. For example:

2, 3, "MAILING ADDRESS", "IMAGIN" (1<sup>st</sup> of 3 data lines)

"414 Main Street, Suite 211B" (2<sup>nd</sup> of 3 data lines)

"Rochester, MI 48307" (3<sup>rd</sup> of 3 data lines)

2, 1, "PHONE", "(248) 608-8781" (with 1 data line)

## APPENDIX E: Macro 1 to Insert X, Y Coordinates to Each Manhole

```

Sub dandan()
Dim n As Long
Dim m As Long
n = 4
m = 2
Do Until n=4340
    regid = Cells(n, 16).Value
    Do Until m=3498
        regch = Cells(m, 25).Value
        If regid = regch Then
            ssss = Cells(m,26).Value
            bbbb = Cells(m,27).Value
            Cells(n,18).Value = ssss
            Cells(n,19).Value = bbbb
        Else
            End IF
        m = m+1
    Loop
    m = 2
    n = n+1
Loop

n = 4
m = 2
Do Until n = 4340
    regid1 = Cells(n,17). Value
    Do Until m = 3498
        regch1 = Cells(m,25).Value
        If regid1 = regch1 Then
            ssss1 = Cells(m,26).Value
            bbbb1 = Cells(m,27).Value
            Cells(n,20).Value = ssss1
            Cells(n,21).Value = bbbb1
        Else
            End If
        m = m+1
    Loop
    m = 2
    n = n+1
Loop
End Sub

```

n: pipe record number (4 to 4340)  
regid retrieves starting manhole ID value from column P  
m: manhole record number  
regch retrieves manhole ID value from column Y  
If regid value is equal to regch, insert manhole related X, Y coordinates from columns Z and AA to columns R and S, respectively.  
Otherwise, add row number and check through all starting manhole IDs in column Y  
Add row number and insert all starting manholes' X,Y coordinates.

n: pipe record number (4 to 4340)  
regid1 retrieves ending manhole ID value from column Q  
m: manhole record number  
regch1 retrieves manhole ID value from column Y  
If regid1 value is equal to regch1, insert manhole related X, Y coordinates from columns Z and AA to columns T and U, respectively.  
Otherwise, add row number and check through all ending manhole IDs in column Y  
Add row number and insert all ending manholes' X,Y coordinates.

## APPENDIX F: Macro 2 to Create a Drawing File by Using X, Y

### Coordinates

```
Private Sub cmf1_Click()  
`THIS COMMAND WILL CREATE A MAP FROM DATA EXISTING IN EXCEL FILE  
Dim EXELAPP As Excel.Application  
Dim WBKOBJ As Workbook  
Dim SHIOBJ As Worksheet  
On Error Resume Next  
UserForm1.Hide  
Err.Clear  
Set EXCELAPP = GetObject(,"EXCEL.APPLICATION")  
EXCELAPP.Visible = True  
  
Set WBKOBJ = Workbooks.Open(FileName:="O:\infrastructure/dandan.xls")  
Set SHIOBJ = EXCELAPP.Worksheets(1)  
EXCELAPP.VISIBLE = True  
n = 4  
    Do Until n = 4340  
        xs = SHIOBJ.Cells(n,18)  
        ys = SHIOBJ.Cells(n,19)  
        xe = SHIOBJ.Cells(n,20)  
        ye = SHIOBJ.Cells(n,21)  
  
        Dim dblCenter(0 To 2) As Double  
        Dim startp(0 To 2) As Double  
        Dim endp(0 To 2) As Double  
  
        Dim radius As Double  
        radius = 10  
        dblCenter(0) = xs : dblCenter(1) = ys : dblCenter(2) = 0  
        startp(0) = xs : startp(1) = ys : startp(2) = 0  
        endp(0) = xe : endp(1) = ye : endp(2) = 0  
        Set CircleObject =  
        ThisDrawing.ModelSpace.AddCircle(dblCenter, radius)  
        Set lineObject = ThisDrawing.ModelSpace.AddLine(startp, endp)  
        n = n+1  
    Loop  
End Sub
```

Define the Excel file that contain the information needed

Get starting and ending manhole X, Y coordinates values

Create circles according to related coordinates to representing the manholes

Create lines representing the pipes

## APPENDIX G: Pierrefonds Borough Sewer Network Different Layers

### Explanation

	Name (full name)	Explanation	Feature (Geographic /text/symbol)
1	ARP RELEVE	n/a	
2	ARPENTAGE	n/a	
3	DEFPOINTS	n/a	
4	DOM_ABAND ( <i>abandoner</i> )	Abandoned pipes	G+T
5	DOM_Agglomeration	Main pipes	G+T
6	DOM_ANNEE	Pipe construction year	T
7	DOM_BASSIN	Basin feature in the map	G
8	DOM_CUM ( <i>Communauté Urbaine Montreal</i> )	cumulative pipes	G
9	DOM_DDO	Pipe in DDO region	G
10	DOM_DIA ( <i>Diametre</i> )	Pipes diameter	T
11	DOM_ECOULEMENT	Flow direction	S
12	DOM_EXI ( <i>Existant</i> )	All Pipes	G+T
13	DOM_EXI_TXT ( <i>Texte Existant</i> )	n/a	
14	DOM_LONG ( <i>Longeur</i> )	Length of pipes	T
15	DOM_MAT ( <i>Material</i> )	Pipes materials	T
16	DOM_NO_REG ( <i>Regard</i> )	Manhole codes	T
17	DOM_PRO ( <i>Propose</i> )	Not Sure	G+T
18	DOM_PROP. GOUIN	Not sure	G+T
19	DOM_PROP_UIN TEXTE	Diameter of DOM_PRO	T
20	DOM-RADIER	n/a	
21	DOM_REFOUL ( <i>Refoulement</i> )	Back up pipes	G+T
22	DOM_SERVITUDE	Partly match pipes	G
23	DOM_SOUS_BASSIN	The underground water network	G
24	DOM_STRUC ( <i>Structure</i> )	Pipes in structure	G

25	DOM STRUCTURE	n/a	
26	EAU EXI	Water pipes	G
27	EGO SANIT SURFACE	n/a	
28	EGO SAINIT E TUYAUX	n/a	
29	LIM MUN TXT	Name of district	T
30	Limite ville	n/a	
31	Noms rue	Name of the road	T
32	PLUV EXI	Rain water manhole	S
33	PP TEXTE	Geographic feature	T
34	STATION POMPAGE EXI	Pumping station	S+T
35	STATION P PAGE INEXL	Old pumping station	S+T
36	TRC TRONCON S	n/a	

## APPENDIX H: Commercial GIS Software Vendors

ESRI (ESRI 2008)	Intergraph (Intergraph 2008)	Bentley (Bentley 2008)	AutoDesk (AutoDesk 2008)
ArcView	Geomedia	Microstation PowerDraft	MapGuide
ArcEditor	Geomedia Professional	PowerSurvey	MapGuide Migration
ArcInfo			
ArcSDE/ ArcIMS	Geomedia WebMap	PowerOnline	MapGuide Server Mid-Tier



## APPENDIX I: ESRI Supported OGC Specifications (ESRI 2008b)

Specification	Provider	Consumer
Web Mapping Service (WMS) 1.1.1	<ul style="list-style-type: none"> <li>• ArcIMS® 9.2</li> <li>• ArcGIS® Server 9.2</li> <li>• ArcIMS 9.3</li> <li>• ArcGIS Server 9.3</li> </ul>	<ul style="list-style-type: none"> <li>• GIS Portal Toolkit 3.1</li> <li>• ArcGIS Desktop 9.2</li> <li>• ArcGIS Desktop 9.3</li> <li>• ArcGIS Explorer</li> <li>• ArcGlobe™ 9.2</li> <li>• ArcGIS 9.2 .NET and Java™ Application Developer Framework (ADF™)</li> <li>• ArcGlobe 9.3</li> <li>• ArcGIS 9.3 .NET and Java ADF</li> </ul>
Web Mapping Service (WMS) 1.3	<ul style="list-style-type: none"> <li>• ArcIMS 9.2</li> <li>• ArcGIS Server 9.2</li> <li>• ArcIMS 9.3</li> <li>• ArcGIS Server 9.3</li> </ul>	<ul style="list-style-type: none"> <li>• GIS Portal Toolkit 3.1</li> <li>• ArcGIS Desktop 9.2</li> <li>• ArcGIS 9.2 .NET and Java ADF</li> <li>• ArcGIS Desktop 9.3</li> <li>• ArcGIS 9.3 .NET and Java ADF</li> </ul>
Styled Layer Descriptor (SLD) 1.0	<ul style="list-style-type: none"> <li>• ArcIMS 9.2</li> <li>• ArcIMS 9.3</li> <li>• ArcGIS Server 9.3</li> </ul>	<ul style="list-style-type: none"> <li>• GIS Portal Toolkit 3.1</li> <li>• ArcGIS Desktop 9.3</li> </ul>
Web Feature Service (WFS) 1.0	<ul style="list-style-type: none"> <li>• ArcIMS 9.2</li> <li>• ArcIMS 9.3</li> </ul>	<ul style="list-style-type: none"> <li>• ArcGIS Desktop 9.2</li> <li>• ArcGIS Data Interoperability 9.2</li> <li>• ArcGIS Desktop 9.3</li> <li>• ArcGIS Data Interoperability 9.3</li> </ul>
Web Feature Service 1.1 Transaction	<ul style="list-style-type: none"> <li>• ArcIMS 9.3</li> <li>• ArcGIS Server 9.3</li> </ul>	<ul style="list-style-type: none"> <li>• ArcGIS Desktop 9.3</li> <li>• ArcGIS Data Interoperability 9.3</li> <li>• GIS Portal Toolkit 3.1</li> </ul>
Web Coverage Service (WCS) 1.0	<ul style="list-style-type: none"> <li>• ArcGIS Server 9.3</li> </ul> <p>WCS capability available with/without Image Server extension</p>	<ul style="list-style-type: none"> <li>• GIS Portal Toolkit 3.1</li> <li>• ArcGIS Desktop 9.3</li> </ul>
Catalog Services 1.0—Z39.50	<ul style="list-style-type: none"> <li>• ArcIMS 9.2</li> <li>• ArcIMS 9.3</li> <li>• GIS Portal Toolkit 3.1</li> </ul>	<ul style="list-style-type: none"> <li>• GIS Portal Toolkit 3.1</li> <li>• GIS Portal Toolbar for ArcGIS 9.3</li> </ul>

Catalog Services—Web (CS—W) 2.0	<ul style="list-style-type: none"> <li>• ArcIMS Metadata Server— 9.1</li> <li>• GIS Portal Toolkit 2.0</li> </ul>	<ul style="list-style-type: none"> <li>• GIS Portal Toolkit 2.0 Metadata</li> <li>• Explorer—9.3</li> </ul>
Simple Features 1.1	<ul style="list-style-type: none"> <li>• SQL—ArcSDE®</li> <li>• OLE/COM—ArcGIS</li> </ul>	<ul style="list-style-type: none"> <li>• ArcGIS</li> </ul>
Simple Features GML	<ul style="list-style-type: none"> <li>• ArcGIS Desktop 9.2</li> <li>• ArcGIS Data Interoperability 9.2</li> <li>• ArcGIS Desktop 9.3</li> </ul>	<ul style="list-style-type: none"> <li>• ArcGIS Desktop 9.2</li> <li>• ArcGIS Data Interoperability 9.2</li> <li>• ArcGIS Desktop 9.3</li> </ul>

## **APPENDIX J: Other Spatial Technologies**

### **(1) Radio Frequency Identification (RFID)**

RFID is a wireless automatic identification (Auto-ID) technology that allows for contactless reading, and is effective in manufacturing and other hostile environments (Dziadak et al. 2008). RFID includes devices: the transponders (tags), antennae and the transceivers (readers). An electromagnetic field which supplies passive tags with the necessary power exchange data between the tag and the reader, whereas active tags use a battery to exchange the data with the reader. The RFID technology is based on a range of radio frequencies (RF) such as, 125 kHz - low frequency (LF), 13.56 MHz - high frequency (HF), 868-915 MHz - ultra high frequency (UHF) and 2.45 GHz – microwaves (Cumberbatch 2005). The accurate location of buried assets is very important before beginning the excavation process. Less accurate positioning of buried assets frequently forces utility contractors to dig trial holes with great caution, which results in delays in road works, affecting the public. Based on various applications and situations, different frequencies are used. For example, a lower frequency means less absorption by moisture, better omni-directional capability, less impact from the presence of metal, but a shorter signal range and a slower reading. A higher frequency means a longer reading range, higher speed, but more interference from metal. The efficiency of the energy transferred from the reader to the tag and the data rate are also affected by the frequency (Microlise 2002). Dziadat et al. (2008) investigated the feasibility of using RFID for locating buried assets in 3D.

### **(2) Ground Penetrating Radar (GPR)**

Ground penetrating radar (GPR) produces a continuous cross-sectional profile or record of subsurface features, without drilling, probing, or digging (GeoModel 2008). GPR is a nondestructive geophysical method which is used for evaluating the location and depth of buried objects and to investigate the presence and continuity of natural subsurface conditions and features. GPR transmits pulses of UHF radio waves into the ground through a transducer (also called an antenna). The transmitted energy is reflected from various buried objects or distinct contacts between different earth materials (GeoModel 2008). The antenna then receives the reflected waves and stores them in the digital control unit. For each reflected wave, the radar signal changes polarity twice. Therefore, three bands on the radar profile for each interface contacted by the radar wave are produced by these polarity charges. Usually GPR waves can reach depths up to 100 feet (30 meters) in low conductivity materials such as dry sand or granite. However in moist clays, shale, and other high conductivity materials, the waves may attenuate because the material absorbs GPR signals, greatly decreasing the depth of penetration to 3 feet (1 meter) or less. Meanwhile, the GPR antenna also can affect the depth of penetration. The low frequencies of Antennas (which range from 25 to 200 MHz) obtain subsurface reflections from deeper depths (about 30 to 100 feet or more). These low frequency antennas are used for locating sinkholes or fractures, and large, deep buried objects. In underground asset applications, the 3D image of underground pipes can be drawn by using GPR data as shown in Figure J.1.

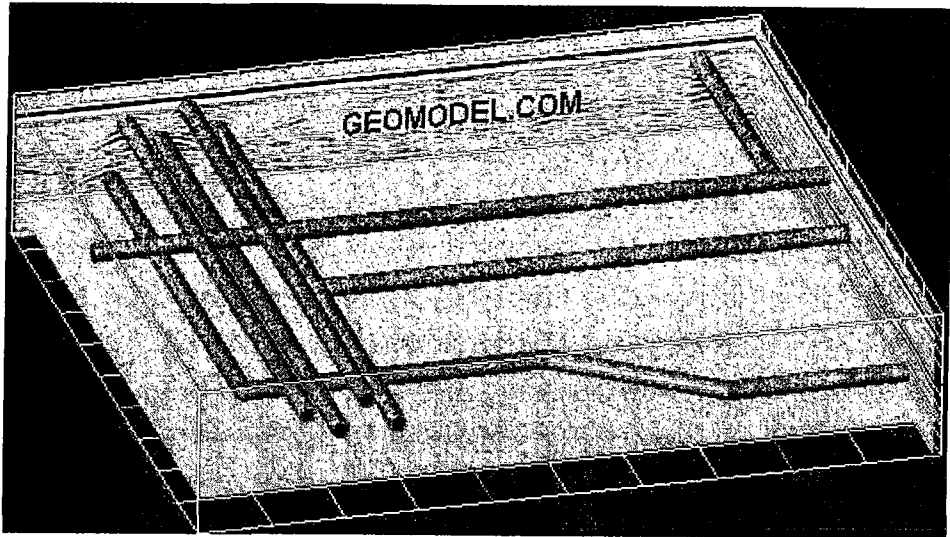


Figure J.1 3D image of underground pipes can be drawn by GPR data (GeoModel 2008)

There are many GPR instruments for utility location by using a GPR digital control unit and various antennas as shown in figure J.2.



(a) Using a magnetometer to detect ferrous metal utilities



(b) Using an utilocator measures to reduce current through metal pipes, cables, or PVC pipes with metal wires installed



(c) Other electromagnetic conductivity instrument detects metal utilities

Figure J.2 Examples of GPR applications (GeoModel 2008)

### **(3) Closed-Circuit Television (CCTV)**

One of the important maintenance and inspection tools for underground utilities is CCTV inspection. The data collected through CCTV inspection are very valuable information

and are important for the authority's fulfillment of its regulatory requirements. The information collected from the CCTV can be downloaded and copied on a CD and loaded into the CCTV program as shown in Figure J.3. An example of linking CCTV videos inspection with a GIS is CCTV Truck Link program where a customized ArcMap project and corresponding Access database utility stores and prints the text and video data from the inspections (Shawn and Kevin 2008). The ArcMap CCTV project resides on one workstation but the data created are stored on the GIS server.

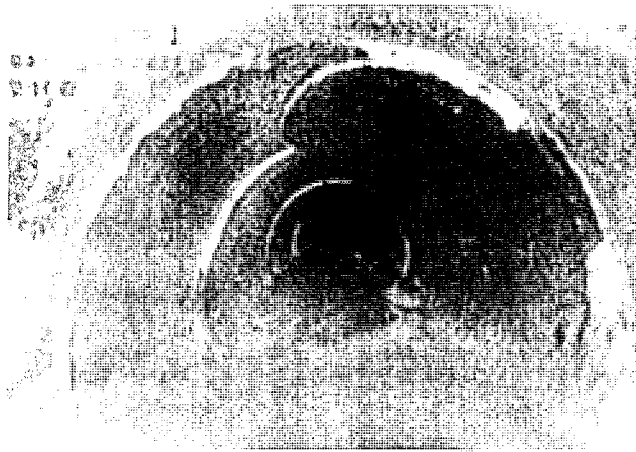
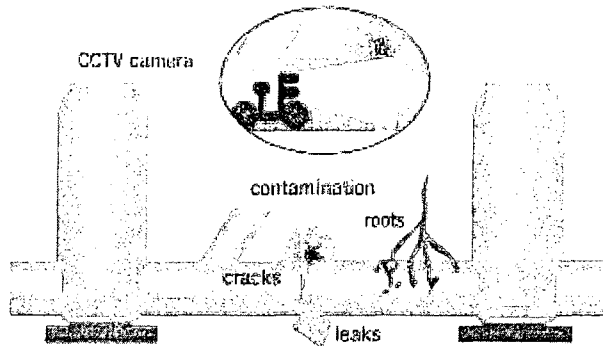
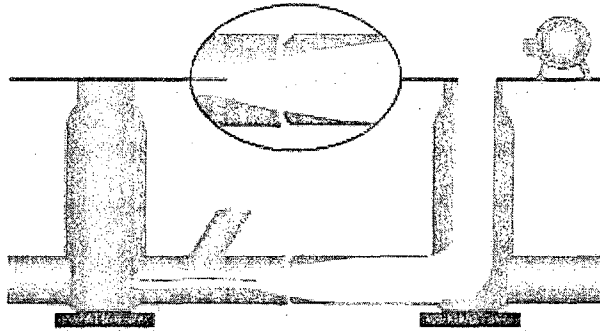


Figure J.3 Pipe inspection's image in CCTV video (Scanseal 2008)

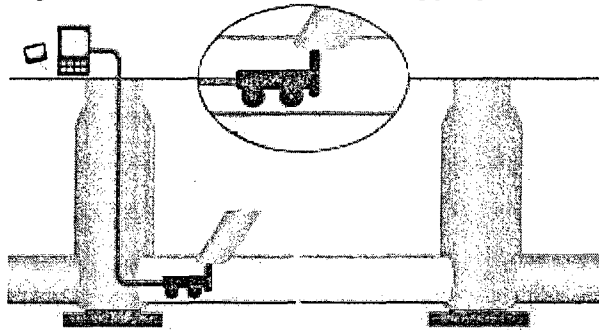
CCTV survey of pipelines is used to inspect pipes' structural and service conditions. This inspection includes: (1) Locating, accessing and the inspection of pipelines; (2) Reporting on pipeline condition and serviceability; (3) Providing city councils with DVD showing video footage of CCTV survey, Condition Assessment Report and photographs of significant defects. The process of CCTV inspection before and after repair is shown in Figure II.4.



(a) CCTV inspection can locate a number of reasons that causes pipe failure including leaking and open joints, root intrusion and cracks



(b) Pipe can then be fixed with an appropriate method



(c) Once fixed the pipe can be checked with another CCTV inspection

Figure J.4 Examples of CCTV application (Scanseal 2008)



# APPENDIX K: Survey Report Example

## Survey Recovery Form

City of Everett  
Department of Public Works

State: Washington  
County: Snohomish

Station Name: \_\_\_\_\_

Horizontal Datum: NAD 83/91  
Vertical Datum: NAVD 88  
Control Point ID: Book: Page: Pt#:

Established by:  
Date:

Section: \_\_\_\_\_ Township: \_\_\_\_\_ Range: \_\_\_\_\_

LATITUDE: (5 decimal places) \_\_\_\_\_

LONGITUDE: (5 decimal places) \_\_\_\_\_

ORTHOMETRIC HEIGHT: \_\_\_\_\_ Feet  
\_\_\_\_\_ Meters

### Washington State Plane Coordinates – North Zone

	Northing (4 decimal places)	Easting (4 decimal places)	Point Scale Factor (to 2 decimal places)	Convergence Deg Min Sec
Feet				
Meters				

#### HORIZONTAL DATA:

Horizontal position established by:  
 \_\_\_\_\_ resection survey, comments: \_\_\_\_\_  
 \_\_\_\_\_ GPS survey, GPS adjustment software used: \_\_\_\_\_  
 \_\_\_\_\_ other, explain: \_\_\_\_\_  
 surveyor's estimated relative positional accuracy: \_\_\_\_\_ ft @ \_\_\_\_\_ % confidence  
 surveyor's estimated absolute positional accuracy: \_\_\_\_\_ ft @ \_\_\_\_\_ % confidence

#### VERTICAL DATA:

Vertical position established by:  
 \_\_\_\_\_ differential leveling to \_\_\_\_\_ order accuracy  
 \_\_\_\_\_ GPS survey, accuracy within \_\_\_\_\_ ft \_\_\_\_\_ meters  
 surveyor's estimated relative positional accuracy: \_\_\_\_\_ ft @ \_\_\_\_\_ % confidence  
 surveyor's estimated absolute positional accuracy: \_\_\_\_\_ ft @ \_\_\_\_\_ % confidence

#### MONUMENT DESCRIPTION

#### TO REACH

#### REFERENCES:

Az from North	Distance	Reference Description
---------------	----------	-----------------------



## APPENDIX M: WebGIS Application (ArcGIS Server Help 2009)

Distributing geospatial information on the Internet is an important part of data sharing and integration because Internet allows all level of clients to access geospatial information without time and location limitation. The detail setting steps about how to publish map on WebGIS are describe as following:

### M.1 Publishing GIS resources

Publishing a GIS resource is the quickest way to create a service. This option prompts for the minimum amount of information needed to publish a GIS resource as a service. When you publish a service with this option, the service's properties such as pooling, recycling, and timeouts are all set to the default values as shown in Figure M.1

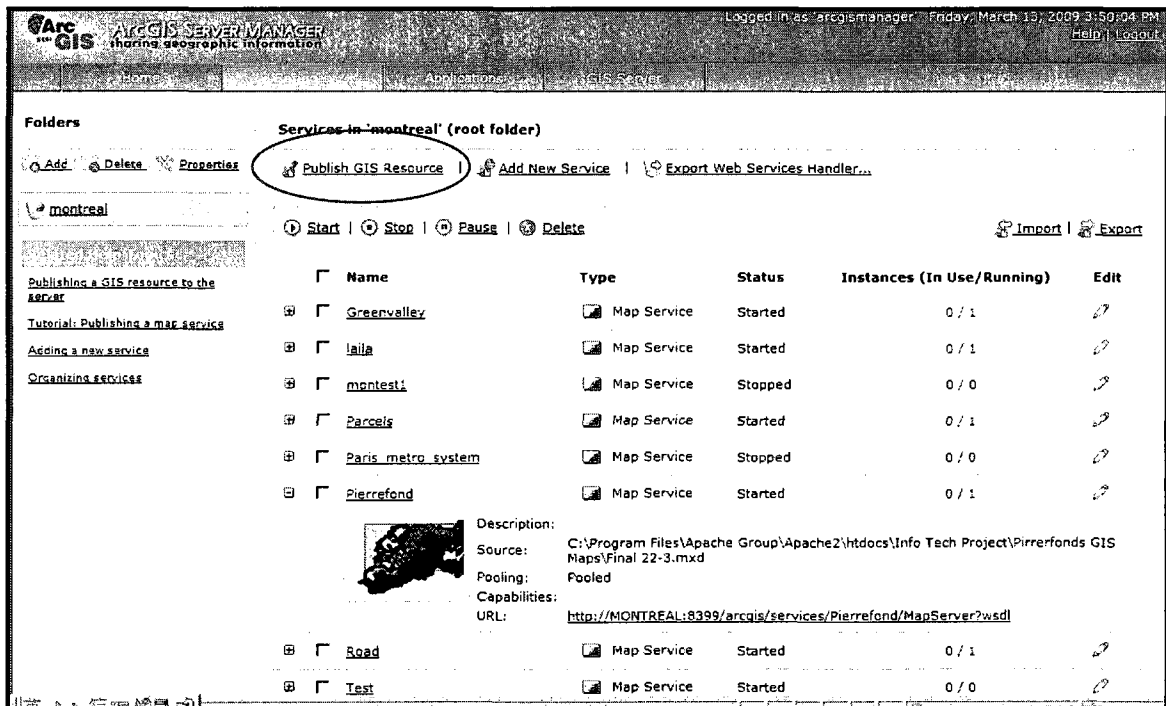
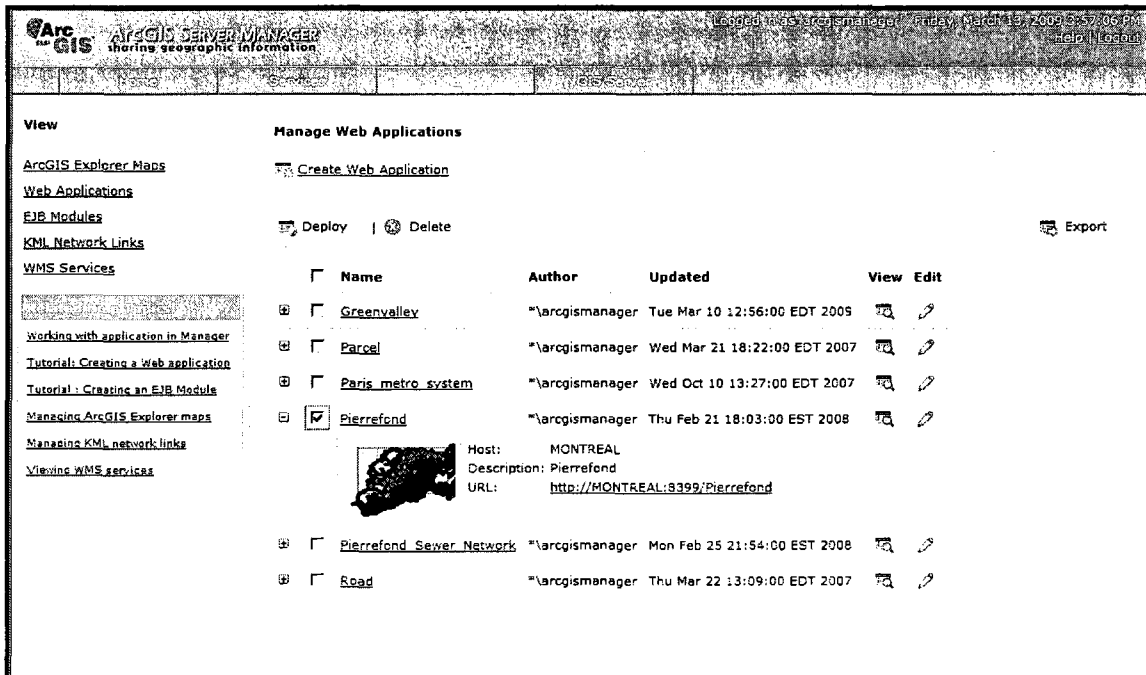


Figure M.1 Publishing GIS resources

## M.2 Creating Web Application

Once publishing GIS resources, it is very important to create its related web application. ArcGIS Server Manager enables create web applications to publish ArcGIS Explorer maps on the GIS Server and create network links as shown in Figure M.2.



The screenshot displays the ArcGIS Server Manager web interface. At the top, the ArcGIS logo and 'ArcGIS SERVER MANAGER' text are visible, along with the tagline 'sharing geographic information'. The user is logged in as 'arcgismanager' on Friday, March 13, 2009, at 3:57:06 PM. The main content area is titled 'Manage Web Applications' and includes a 'Create Web Application' button, 'Deploy' and 'Delete' icons, and an 'Export' button. A table lists several web applications:

<input type="checkbox"/>	Name	Author	Updated	View	Edit
<input type="checkbox"/>	Greenvalley	"\arcgismanager	Tue Mar 10 12:56:00 EDT 2009		
<input type="checkbox"/>	Parcel	"\arcgismanager	Wed Mar 21 18:22:00 EDT 2007		
<input type="checkbox"/>	Paris_metro_system	"\arcgismanager	Wed Oct 10 13:27:00 EDT 2007		
<input checked="" type="checkbox"/>	Pierrefond	"\arcgismanager	Thu Feb 21 18:03:00 EST 2008		
<input type="checkbox"/>	Pierrefond Sewer Network	"\arcgismanager	Mon Feb 25 21:54:00 EST 2008		
<input type="checkbox"/>	Road	"\arcgismanager	Thu Mar 22 13:09:00 EDT 2007		

Below the selected 'Pierrefond' application, a small thumbnail image is shown. To its right, the following details are listed:

- Host: MONTREAL
- Description: Pierrefond
- URL: <http://MONTREAL:8399/Pierrefond>

Figure M.2 Creating Web Application

## **APPENDIX N: List of Publications**

Hu, B. and Hammad, A. (2008). **“Data Quality Issues in Municipal Infrastructure Management”**, 37<sup>th</sup> CSCE (Canadian Society for Civil Engineering) Annual Conference, June, 2008, Québec City, Québec, Canada.