Leak Detection in Pipelines using the Extended Kalman Filter and the Extended Boundary Approach

A Thesis Submitted to
the College of Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in the Department of Mechanical Engineering
University of Saskatchewan
Saskatoon

By

Kurtis L. Doney

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that the permission for copying this thesis in any manner, in whole or in part for scholarly purposes, may be granted by the professors who supervised my thesis work or, in their absence, by the Head of the Department or Dean of the College in which my thesis work was conducted. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis. Requests for permission to copy or to make other use of material in this thesis, in whole or part, should be addressed to:

Head of the Department Mechanical Engineering
University of Saskatchewan
Engineering Building
57 Campus Drive
Saskatoon, Saskatchewan S7N 5A9
Canada

Abstract

A model based algorithm of pipeline flow is developed and tested to determine if the model is capable of detecting a leak in a pipeline. The overall objective of this research is to determine the feasibility of applying the Extended Kalman Filter and a new technique defined as the Extended Boundary Approach to the detection of leakages in a physical water distribution system.

The demands on the water supply system increase as the human population grows and expands throughout the world. Water conservation is required to ensure an adequate supply of water remains for future generations. One way to conserve this water is by reducing the leakages in underground water distribution systems. Currently between 10 to 50 percent of the pumped water is lost due to unrecognized leakages. This results in a huge revenue loss of water, chemicals and energy that is required for transporting the water. The detection of underground leakages is a very complex problem because many leakages are small and go unnoticed by today's leak detection technology.

A model based leak detection technique is developed and tested in this thesis. The Method of Characteristics is used to develop a model of a single pipeline. This method is extensively used and provides the most accurate results of the two partial differential equations of continuity and momentum that describe pipe flow. The Extended Kalman Filter is used to estimate two "fictitious" leakages at known locations along the pipeline. In order to ensure the model is observable four pressure measurements are needed at equally spaced nodes along the pipeline. With the development of the Extended Boundary Approach only the upstream and downstream pressure measurements are required, however; the upstream and downstream flow measurements are also required. Using the information from the two "fictitious" leaks the actual leak location and magnitude can be determined. This method is only capable of detecting one leak in a single pipeline.

The results of the developed model show that the approach is capable of theoretically determining the leak location and magnitude in a pipeline. However, at this time, the feasibility of implementing the proposed leak detection method is limited by the required level of accuracy of the sensors which is beyond that found in today's technology. It was also found that the EKF used primarily steady state information to predict the leakage. It is recommended that further research explore alternate models which might better enhance the EKF approach using transient information from the pipeline. This may allow implementation on a real pipeline.

Acknowledgements

The author expresses his gratitude to his supervisors, Dr. Richard Burton and Dr. Saied Habibi for their guidance and support during the course of this research and the writing of this thesis. The technical assistance of Mr. D.V. Bitner is also gratefully acknowledged.

The author would also like to thank SaskWater for access to a real water distribution system. Also, thanks to the operators of the Melfort line for their time and patience in order to collect data for this research.

Also the author would like to thank the Canadian Water Network for financial support and providing a means to meet other professional working on water issues throughout Canada.

Lastly, the author wishes to thank his family and friends for support and encouragement throughout this research.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Tables

Chapter 5

Chapter 6

Chapter 7

Nomenclature

Pipe Modeling Nomenclature

A	Pipe Cross-sectional area	m^2
a	Wave speed	m/s
B	Pipe wave velocity constant	
C^+, C^-	Positive and negative characteristic equation sets	
c_1	Pipe loading condition	
D	Pipe diameter (inner)	m
D/Dt	Total derivative	
E	Young's Modulus	N/m^2
e	Wall roughness of pipe	mm
f	Darcy-Weisbach friction factor	
g_x	Axial body force or force of gravity	m/s^2
H	Hydraulic Head	m
P	Pressure	N/m^2
Q	Flow rate	m^3/s
R	Pipe friction constant	
r	Radial pipe position	m
u_r, u_θ, u_x	Radial, rotational and axial components of fluid velocity	m/s
V	Fluid velocity	m/s
x	Distance along length of pipe	m
z	Pipe Elevation	m

β	Bulk modulus of water	N/m^2
ε_T	Circumferential strain	
λ	Unknown multiplier used in Method of Characteristics derivation	m/s
μ	Fluid viscosity	Ns/m^2
ρ	Fluid density	kg/m^3
σ_θ	Circumferential pipe stress	N/m^2
σ_z	Axial pipe stress	N/m^2
τ	Shear Stress	N/m^2
υ	Poisson's ratio	

State and Estimation Nomenclature

$\mathrm{cov}[X,Y]$	Covariance of the random processes X and Y
$E[X]$	Expectation of a random process X
e_k^-	*A priori* error estimate
e_k	*A posteriori* error estimate
$f[x_k]$	Non-linear state function
G_k	Input matrix
H_k	Output matrix that linearly connects outputs and states
$h[x_k]$	Non-linear output matrix
I	Identity matrix
J_x	Jacobian of $f[x_k]$ evaluated at $\hat{x}_k$

J_h	Jacobian of $h[x_k]$ evaluated at $\hat{x}_{k+1}^-$
K_k	Kalman Gain at time t_k
P_k^-	Unrefined (*a priori*) estimate of covariance matrix at time t_k
P_k	Refined (*a posteriori*) estimate of covariance matrix at time t_k
Q_k	System noise covariance matrix
R_k	Measurement noise covariance matrix
u_k	Input vector
v_k	Vector of white measurement noise
w_k	Vector of white system noise
x_k	State vector at time t_k
$\hat{x}_k^-$	Predicted estimate of the state vector at time t_k
$\hat{x}_k$	Refined estimate of the state vector at time t_k
z_k	Vector of defined measurements
ϕ_k	System or transition matrix of constants for time t_k

Chapter 1: Introduction

1.1 The Importance of Water

The average family of four uses about 1320 litres of water per day (Water Quality and Health Council, 2006) which translates into using about 120,000 litres of water per person per year. It is not unreasonable to state that society's daily lives revolve around water. Indeed, society has been structured around the existence of a never-ending supply of clean drinking water (Gleick, 2005). Water is a key component of the modern Canadian economy: it is a fundamental resource for food production, plays an important role in virtually every modern industrial process and many recreational activities, and provides an essential element for urban development across the country (Environment Canada, 2006).

Having available water is one thing, but being able to adequately and safely distribute it, is another. If water is available but not potable, it is just as detrimental as not having any water at all (Water Quality and Health Council, 2006). Many parts of the world are suffering because of a lack of fresh water and as a result the world faces tremendous challenges in managing its water (Environment Canada, 2006). Water quality continues to deteriorate rapidly due to urbanization, agricultural practices, industrialization and over population (Environment Canada, 2006). As the world population grows, it is necessary to conserve the water resources and be aware of the water quality. One way to achieve this is to be aware of everything entering and leaving the water distribution systems at all times. The problem of depleting water resources and concern with safe drinking water indicates the need to be able to monitor water supplies and to minimize losses along distribution systems. This, in part, was the motivation for the study of leak detection in water distribution systems.

1.2 The Availability of the World's Water

Water resources are sources of water that are useful or potentially useful to humans (Wikipedia, 2006). There are many different uses for water; some of them agricultural, household, industrial, and recreational uses. Globally, the total amount of water available exceeds societies' requirements but only 2.5% of water on Earth is fresh water and over two thirds of that is frozen in glaciers and polar ice caps (Environment Canada, 2006). The small fraction of fresh water that is accessible, is extremely unevenly distributed in space and time; so unevenly that society spends billions of dollars every year to move water from wet areas to drier areas, to store it in wet seasons for coming dry periods, or to clean otherwise undrinkable sources (Gleick, 1995).

It has become common knowledge that many parts of the world are suffering because of a lack of clean water. Approximately 1.1 billion people do not have access to safe drinking water and 2.4 billion people do not have access to adequate sanitation (Environment Canada, 2006). The problem of accessing fresh water is expected to increase drastically in the future. Currently, agriculture is the largest user of fresh water resources. Therefore, as fresh water becomes in higher demand, agriculture may suffer due to a lack of fresh water, in turn reducing the availability of food and food products for human consumption (International Water Management Institute, 2000).

Fresh water is obtained from surface water, sub-surface water, desalination, and by melting frozen water. No practical approach has been proposed to melt the frozen water captured in the ice caps and glaciers and desalination is currently extremely expensive and will not solve the crisis for countries who cannot afford this technology. Therefore, this leaves only surface water and sub-surface water as an available resource. A number of different sources including sewage, fertilizer, and oil spills can pollute surface water. Since surface water only makes up for about 0.3% of the fresh water available, any polluting of this small source has a drastic effect on how much is left in a useable form (Environment Canada, 2006).

Ground water makes up about 33% of the fresh water supply and therefore is a large source of usable water (Environment Canada, 2006). However, groundwater supply is also being depleted and polluted. Ground water can become polluted when toxic substances are dissolved into the water and then leached, or carried down, into the aquifer. Ground water moves very slowly and therefore it can take years before a pollutant is detected in the water, usually far too late to rectify the situation. If ground water is extracted at a faster rate than the aquifer recharges, then the water supply can become depleted. This can have drastic effects on the users of the aquifer, as it will lower the water table therefore making it harder to draw water.

One very large aquifer in North America is the Ogallala Aquifer. This aquifer covers approximately 175,000 square miles in the United States. The aquifer runs north and south mainly through Texas, Oklahoma, Kansas, and Nebraska and provides water for agricultural, industrial and municipal requirements. The reliance on this aquifer has drastically increased since World War II and it is now being depleted much faster than it is being replenished. Increased irrigation for agricultural use is the main reason why this aquifer is being depleted (Guru, 2000). Since large-scale irrigation began in the 1940s, water levels have declined more than 30 meters (100 feet) in parts of the aquifer. The average water level decline has been about 3 feet per year since the 1980's. However, it is not known how long the aquifer will last. This causes a lot of uncertainty, and at times a lot of resentment, by the users.

According to the International Water Management Institute (IWMI) the world is approaching a world-wide water crisis in the near future. By 2025 IWMI forecasts that forty-five countries representing 33 percent of the population will be experiencing physical water scarcity. Also, countries containing 45 percent of the population will have substantially underdeveloped water resources, requiring 25 percent or more development of additional water supplies. Figure 1.1 show a graphical representation of the water scarcity predictions in 2025.

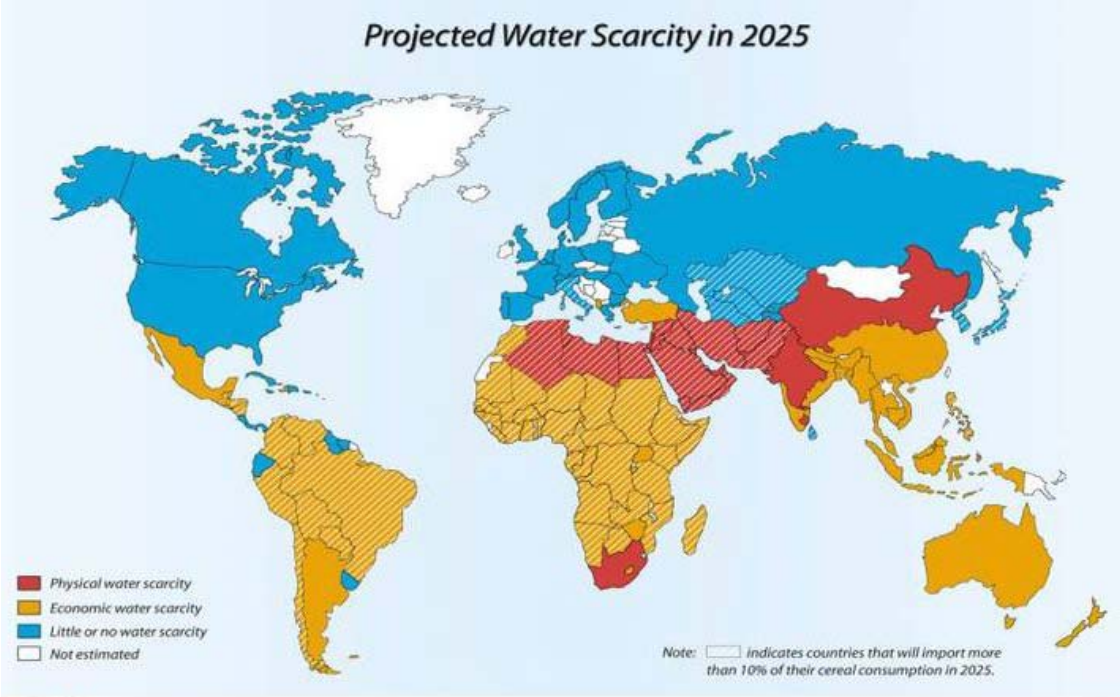

Figure 1.1: Projected Water Scarcity in 2025 (Permission granted by IWMI,2000)

In order to meet the projected world water needs in 2025 the world's primary water supply will need to increase by 22 percent, or 600 km^3 per year according to IWMI. To put this value into perspective, the annual average release of water from one of the largest dams in the world, the High Aswan Dam (HAD) of Egypt, is about 55 km^3 per year. Thus the additional storage to meet the 600 km^3 is equivalent to about twelve new HADs over the period or nearly one every 2 years (International Water Management Institute, 2000). In many of the most pump-intensive areas of India and Pakistan, water tables are falling at rates of 2 to 3 meters per year. The food security of India, Pakistan, China and many other countries in 2025 will largely depend on how they manage this groundwater problem (International Water Management Institute, 2000).

As water becomes scarcer nations will begin to look elsewhere for their water resources which can and probably will lead to conflict. As former UN Secretary-General, Kofi Annan stated "All too often, water is treated as an infinite free good. Yet even where supplies are sufficient or plentiful, they are increasingly at risk from pollution and rising

demand…Fierce national competition over water resources has prompted fears that water issues contain the seeds of violent conflict" (European Commission, 2002).

According to IWMI, by 2025, most of the world's population will live in urban and peri-urban areas. The people and industries in these areas will demand an increasingly large share of the total water available and much of this will be taken from irrigated agriculture. In India, the Philippines and many other countries, large irrigation areas are literally shut down, either permanently or in times of drought, by cities taking water from farmers with no compensation paid to them for loss of their livelihoods. A quote by Mark Twain "*Water flows uphill, toward power"* may fit the water crisis of the twenty first century unless necessary actions are taken sooner rather than later.

1.3 Underground Pipeline Infrastructure

As the world's population and economies grow, demand for water will also grow. Goals have been set by the United Nations to reduce the number of people without clean drinking water and clean sanitation by one-half, and to move on to full water security by 2025. It is estimated that global spending on new water infrastructure will need to more than double to approximately $180 billion per year over the next 20 to 25 years, from the current $80 billion/year (Environment Canada, 2006).

It is estimated that over the next 20 years the United States' 54,000 drinking water systems and 16,000 wastewater systems will require $1 trillion in infrastructure improvements (Kermit, 2001). According to U.S. Environmental Protection Agency (EPA), the replacement of existing water distribution pipes form the single largest category totalling about $77 billion. Leakage rates approaching 50% have been observed in aging and deteriorating pipe networks in the United States (Jowitt and Xu, 1990). In the United States alone, 24 percent of water-borne disease outbreaks reported in community water systems during the past decade were caused by contaminants that entered distribution systems (Tafuri, 2000). This is significant because water-borne diseases may enter the distribution system at leakage points during times when the

system experiences negative pressure such as valve closures or pipe flushing (Prevost, 2006).

The National Round Table on the Environment and Economy stated that the unmet water and wastewater infrastructure needs in Canada were $38-49 billion in 1996, and capital costs for the following 20 years would be in the order of $70-90 billion (Environment Canada, 2006). In many Canadian municipalities the infrastructure that helps to deliver quality water to households, and transports the waste water produced, is in need of major capital re-investment (Rollins, et al. 1997). Environment Canada estimates that $4.6 billion per year, over 10 years, will be required to maintain today's levels of water supply and quality (Environment Canada, 2006).

According to Pelletier et al. in 2003, water infrastructure systems are in poor condition and deteriorating rapidly. In a recent survey conducted by the Canadian National Research Council, cast iron pipes are rupturing at a rate of 35.9 breaks for every 100 kilometers of pipe in service per year. In addition, newer ductile iron pipes are averaging about 9.5 breaks per 100 kilometers per year. These numbers are significant because they translate into over 200,000 breaks every year in the U.S. and Canada (Water Quality and Health Council, 2006). It is not unusual for a water distribution system to have leakage rates of 20% or higher (Cheong, 1991). Europe's unaccounted for water is typically in the order of 9-30% of the total volume pumped (Cheong, 1991). In Pelletier's research it was found that the communities of Chicoutimi, Gatineau and Saint-Georges experienced 46, 36 and 19 pipe breaks per 100 km per year respectively. Based on the model that his team created, the communities of Chicoutimi, Gatineau, and Saint-George would see their annual pipe breaks increase by 63, 88, and 121% respectively over the next ten years (Pelletier, et al. 2003).

The pipes installed after 1960 have a much higher probability of break occurrence possibly because of the rapid urban growth resulting in poorer quality and installation techniques (Pelletier, et al. 2003). Leakage rates increase with pipe age; older networks tend to be affected most severely (Smith, Fields, Chen, Tafuri, 2000). At current-day

replacement rates, pipes will have to last 200 years - a period that far exceeds the design service life of the pipe, which is typically 100 years. (Tafuri, 2000). In terms of cost, taxpayers in Canada spend an average of $82 million every year to repair broken water mains -- an estimated $2,500 in repair costs alone for each pipe failure (Water Quality and Health Council, 2006).

It is quite apparent that significant amounts of money and resources are being spent on underground water distribution systems. However, recent events -- including waterborne disease outbreaks and extended "boil-water" notices in major cities -- have focused attention on the danger associated with contamination of public water supplies entering through unidentified leaks (Water Quality and Health Council, 2006). Therefore, more effort is still required to ensure the safety of drinking water and to conserve water by being able to identify the presence and location of leaks in an underground water distribution system.

1.4 Mathematical Modelling

One technique used to locate the magnitude and location of a leak in a water pipeline system is mathematical modeling. This approach involves transforming a model into a "calibrated" model which has been experimentally verified. A model in this sense is a program that accepts inputs and solves for a system's unknown variables. According to (Basmadjian, 1999) a model refers to the ensemble of equations which describe and interrelate the variables and parameters of a physical system or process. A calibrated model is one for which the output data accurately simulate actual field conditions for a given time frame (Cesario, 1995). Calibrating is the process of "fine-tuning a model until it simulates field conditions to an established degree of accuracy." (Cesario, 1995).

Before computers were readily available, mathematical analysis of water systems was limited because of the complex networks that make up water distribution systems. Most of the work in the area of fluid mechanics was performed in laboratories or involved field tests. Data were collected from this early work and tabulated into tables, which included flow characteristics at varying pressure, flow rates, and pipe materials (Cesario, 1995).

Hazen and Williams, in 1947, developed such tables and produced what became known as the "bible" for designers of pipelines (Williams and Hazen, 1947). The earliest network modeling began in the early 1960's. Network modeling involved developing a computerized mathematical model of a water distribution system and using the model to analyze and expand the existing system (Cesario, 1995). Initially, water utilities could not afford to produce/use such models. However, various forms of model water distribution systems are now frequently used for leak detection, cost optimization, etc.

Pipeline models allow engineers and planners to become more efficient because the model can perform tedious calculation, while the engineer or planner can anticipate what actions will be necessary in the future. A model allows engineers and planners to more accurately size pipelines and therefore reduces capital costs. A model also allows designers to predict when a new facility is needed and what will happen when disruptions are introduced during construction. Modeling is a tool to examine system performance and provide answers to "what if" questions (Cesario, 1995). Engineers in industry often use modeling as a starting point to the problem. Once a starting point is established then necessary changes can be made depending on the given situation. These are just a few applications of models for water distribution systems. Since the advancement of computers, models are now readily used in the water distribution industry.

1.5 Leak Detection Methods and Transient Modelling

Leakage control and demand management have become high priorities for water supply utilities and authorities (Covas, Ramos, Graham and Maksimovic, 2003). This is not only because there is a greater understanding of the economic and social costs associated with water losses, but there is also an imperative to make best possible use of the natural resource that is water (Covas, et al. 2003). Over the past century, there have been many different methods to detect underground pipeline leaks in water distribution systems. The ideal technology for assessing distribution system integrity should be non-intrusive, should not interrupt operations and should be adaptable to the complex array of materials and conditions present (Tafuri, 2000).

While small leaks may seem minor on an individual basis, taken collectively for long periods over large networks, they cause a significant loss of resource and revenue even for water utilities with well-managed distribution systems (American Water Works Association, 2003). Ideally, leak detection methods should be able to detect the smallest of leaks; however, practical concerns such as the fact that pipelines are buried at least six feet under the ground severely hamper the detection of small leaks.

For a single line with no distribution outlets or nodes, the easiest way to detect leaks in underground pipelines is to measure the flow entering and leaving. If there is a difference in these measurements (taking into account the known water distribution at the "nodes" if they do exist), then there is a leak in the system. However, this only indicates the presence and magnitude of a leak during steady state operation. Locating the leak is not possible with only this information.

Acoustic leak detection is the non-destructive evaluation method most commonly used by the water industry (Tafuri, 2000). The first phase involves using listening rods or aquaphones, to detect the sound of water escaping, by placing them in direct contact with the pipe or appurtenances (Tafuri, 2000). The second phase involves using ground microphones to listen to the leaks directly above the pipe. This method has shown some success; however it is susceptible to excess background noise, pipeline location and soil quality and is only somewhat reliable for larger leaks. A leak "correlator" can also be used to pinpoint the location of the leak. This involves placing magnetic sensors on either side of the suspected leak. The sensors send a signal to an amplifier and then onto a correlator which determines the relative location of the leak from each sensor. In long underground pipelines, there may be no access to pipelines for an extended distance and this method requires sensors to be placed at within at least 500 ft (152.4m) of each other (Tafuri, 2000). This method of leak detection is only suitable for metal pipelines and is a method that is being outdated as many new pipelines installed are made of plastic.

Many other leak detection methods include ground penetrating radar or infrared spectroscopy (Hunaidi, Giamou, 1998), transmission and reflection of pressure waves

(also known as time domain reflectometry) (Brunone, 1999) (Misiunas, Vitkovsky, Olsson, Simpson, Lambert, 2005) (Lennart, 1995), sequential statistical analysis (Buchberger, Nadimpalli, 2004), tracer compounds (Hargesheimer, 1985), nonlinear fluid model of finite dimension (Verde, 2004), frequency response methods (Mpesha, Gassman, Chaudhry, 2001) and transient damping methods (Wang, Lambert, Simpson, Liggett, Vitkovsky, 2002).

Buchberger and Nadimpalli, in 2004, have developed a new method for detecting the magnitude of leaks in pipe networks. The method is based on sequential statistical analyses of continuous high-resolution measurements taken at one location. If unsteady leakage rates are present, the method will detect maximum and minimum leakage rates. The basic idea of the algorithm is to repeatedly compute the mean and standard deviation of the measured flows and as soon as the flow values diverge from the statistical curves, the maximum network leakage rate is determined. This method does not locate the position of the leak; it only provides the magnitude of the leak, which could be used to determine which section of the pipe is in the worst condition. This method has not been field tested and therefore more work still needs to be completed in order to verify it.

Hargesheimer, in 1985, tried to identify water main leaks by using trihalomethanes (THMs) chloroform and dichlororbromomethane, which are normally present in chlorinated drinking water, as tracer compounds to identify the presence of treated water being leaked out of distribution systems. According to Heargesheimer, THM analysis provides a simple, specific, and sensitive means of identifying treated city water samples in seepage. However, this method only works on treated water and it does not give the magnitude of the leak, rather only a general location of the leak.

Verde, in 2004, presented a method for leak location in a pipeline, using flow and pressure sensors only at the ends of the pipeline. According to Verde, in 2004, the leak identification problem can be solved using a simple nonlinear model of the flow, assuming leak position with uncertainty, and combining static relationship between residual components and leak position error. The residual is the difference between the

model predicted location and the location input into the model. This means that the locations of the leaks are inputs into the model, with a predetermined error, and a static relationship between each residual and position error derived to determine the position of the actual leakage. This method was tested and it was found that the greatest error in this model occurred when the leak was positioned near the beginning or near the end of the pipeline. The model was tested using real noise data and was able to predict the location of a leak, provided that the location was not at the beginning or end of the pipeline. This model is useful in detecting the location of the leak, assuming the leak is a substantial distance from the boundaries, and does not attempt to determine the magnitude of the leak.

Misiunas, in 2005, presented a new continuous monitoring approach for detecting and locating breaks in pipelines. This new technique falls under a classification known as time domain reflectometry. Misiunas, in 2005, measured the pressure at one location in the pipeline to sense the negative pressure wave that was produced when a break occurred. The location of the break was determined by the timing of the initial and reflected transient waves produced by the break. The break size was estimated by the size of the magnitude of the transient wave. According to Misiunas, in 2005, the limitation to this approach includes the speed of the break opening. When the speed of the break is not fast (that is, when the break opening time is longer than the wave travel time from the break point to the closest boundary and back), the accuracy of locating the leaks drops substantially and the proposed methods becomes unsatisfactory.

Wang, in 2002, developed a method of leak detection using the damping of fluid transients. According to Wang et al. the overall leak-induced damping could be divided into two separate parts. The size of the leak was indicated by the magnitude of the damping, while the location of the leak was determined from different damping ratios. After experimentation it was found that leaks of 0.1% of a pipeline's cross sectional area, or smaller, could be detected and located using Wang's proposed model. This model is only applicable in single transmission lines because more complex distribution systems introduce complex waveforms which may be falsely recognized as leaks.

Kiuchi, in 1993, developed a leak localization method in pipelines by using a fluid transient model. The model uses the continuity and momentum equations and the following assumptions are made: isothermal flow, steady-state friction, negligible pipe wall expansion and a constant slope over a pipe segment. The model computes the inlet and outlet flows, which are calculated from pressure measurements at the inlet and outlet of the pipeline. When there is no leak in the system the inlet and outlet flows from the model match the flows measured from the system. When a leak is present in the system the model computed pressure and flows do not change because the model assumes the pressure at the inlet and outlet are constant. The location of the leak is determined by integrating the pressure and flow profiles over the pipeline in two segments: from the beginning of the pipeline to the unknown leak location and from the unknown leak location to the end of the pipeline. The first segment uses the model flow outputs, as that section has no leak. The second segment uses the measured flows, as that section has the leak. Theoretically, Kiuchi, in 2003, showed that this method is capable of locating a leak in a pipeline; however, field applications showed that more work in the area of model tuning is necessary.

Mpesha, in 2002, developed a leak detection frequency response method. The transient flow was analyzed using the two continuity and momentum equations and the method of characteristics was applied. The results were transformed into the frequency domain by using the Fast Fourier transform. For detecting a single leak, a frequency response graph was created for the system without a leak and with a leak. A secondary pressure amplitude peak in between the primary pressure amplitude peaks indicated the leak. The technique required measuring pressure and discharge at one location in the pipeline. Mpesha stated that this technique could be applied to real-life pipe systems; however currently there has not been an attempt to do so.

Carpentier and Cohen, in 1991, completed one of the most extensive studies on water distribution lines. This study was entitled “State estimation and leak detection in water distribution networks” and involved studies in Paris, France. According to Carpentier and

Cohen one of the main problems in water distribution systems is determining if the available set of measurements is sufficient information to determine the full system state. The aim of their study was to determine which variables are obtainable from the available measurements. In their work, a graph-theory approach was used to determine the obtainable variables. Carpentier and Cohen concluded that without knowing the state of the valves and other devices and due to some errors in the technical data, it was difficult to acquire an accurate network model. However, from their study they produced better knowledge regarding the network and in turn increased the efficiency of the network. Since they needed highly accurate instrumentation, within a reasonable cost range, it was difficult to select instrumentation. The detection of leaks was successfully performed, however, it was noted that a better method would incorporate "estimating consumption rates every five minutes and analyzing the results using statistical filters to produce a more accurate diagnosis" (Carpentier and Cohen , 1991).

Benkherouf and Allidina, in 1988, describe a method for detecting and locating leaks in long gas pipelines using the Extended Kalman Filter. The proposed model had artificial leakage states at predefined positions along the pipeline and the filter was implemented to detect the magnitude of these leakages. The location of the leakages was determined by a linear interpolation from the magnitude and location of the artificial leakages. The model simulated a 90km pipeline and the filter was successful at locating a leak at 50km. No attempt was made to detect a leak closer to the boundaries of the pipeline or to experimentally verify the results. Results have shown that a large discretisation step can be used in the filter design without reducing the accuracy in the estimate of the leak position. Using a larger discretisation step reduces computational effort and allows the model to be implemented on a microcomputer.

One specific model was developed by Lesyshen, in 2005. Lesyshen developed a theoretical model-based algorithm to determine the magnitude and location of a leak in a single water distribution pipeline. The algorithm applied the Extended Kalman Filter (see Chapter 3) for the purpose of leak detection. The algorithm introduced two "artificial or

fictitious" leakages within the system. The leakages were placed at pre-defined locations; however the magnitudes of these artificial leaks were the states estimated by the model. The estimates of these fictitious leakage states were then used to locate the actual position and magnitude of leakage within the distribution line.

The first part of the model involved estimating the magnitude of the two fictitious leaks. The Extended Kalman Filter (EKF) was able to combine the system measurements (upstream and downstream head measurements) with the prediction from the model using an optimal gain factor, known as the Kalman gain.

The Kalman gain is determined by minimizing the mean squared error terms. The mean squared error terms are the squared difference between the actual states and the states predicted by the model. The minimizing process is accomplished by taking the derivatives of the mean squared error terms and setting them equal to zero. For the simple pipeline considered by Lesyshen, in 2005, the model predicted 12 states including: head measurements at (four) nodes, flow rates, and the flow rate of two fictitious leaks at the interior nodes.

The Kalman filter requires a model of the pipeline to estimate states or parameters. The model Lesyshen used was based on the transient theory presented in (Wylie and Streeter, 1984), (Chaudry, 1987), and (Watters, 1984). The equations that the model used to determine the head, flow rate and leakage rates were derived from the continuity and momentum equations. However, the model must be represented in a state space form, with the present state being only a function of prior states and inputs. This was required for the implementation of the Extended Kalman Filter.

In order to numerically solve these two equations (continuity and momentum equations) an approximation had to be made. This approach is known as the Method of Characteristics and involves neglecting the spatial terms, in both equations, when compared to the other terms. This is only valid when both the spatial and time variation appears, for pressure (P) and flow (Q), in the same equation. This approximation was

valid because the spatial term is much smaller than the time term (Watters, 1984). The equations developed for head and flow rate in the model, at the two interior nodes, included leakage at each node. Therefore, the model was "told" there was a leak at the two interior nodes and through the Kalman Filter the magnitudes were determined.

The second part of the leak detection process involved applying the concept of equivalent systems to determine the actual position and magnitude of the leakage within the distribution pipeline. A relationship, based on the fundamentals of momentum and continuity, was derived for determining the estimate of the leak position and magnitude from the two artificially modeled leak states. The magnitude was determined as the sum of the two leakage estimates and the position was given by a first order location equation via a linear interpolation. This method was capable of locating one leak within the line. The combined pipeline model, the Kalman Filter Estimation of the leakage magnitudes and the interpolation approach to determine the location, formed Lesyshen's model-based algorithm.

As stated in Lesyshen, in 2005, it was shown through simulated results that "the estimation process theoretically produces accurate estimates of the position and magnitude of leaks when the leakage was as small as 5% of the total flow. The results displayed a linear increase in the standard deviation of the position estimate with a decrease in percent leakage flow, from 6.3 m deviation for 20% leak flow to 23.9 m deviation for 5% leak flow. The standard deviation in the position estimate then grew exponentially for leak flows between 5% and 1%. The magnitude of the leakage was estimated with an error of approximately 1% for leakage flows ranging from 5% to 20%". The standard deviation was calculated by taking the square root of the variance in the data. The variance was calculated by taking the average of the square of the difference between each estimate and the actual mean.

Lesyshen, in 2005, also stated that "the standard deviation or randomness in the flow magnitude estimates were all identical. However, the standard deviations in the position measurements was similar but not identical. This was attributed to the fact that the

position estimate equation was an interpolation between two estimated points. Therefore, when the distance between the actual leakage and the two fixed fictitious estimates increased, which was the case when the leakage was located near the boundaries of the pipeline, the error in the position estimate increased." The model developed by Lesyshen, in 2005, was evaluated using a theoretical model of a pipeline only. Subsequently, it was the goal of this research study to determine the feasibility of using Lesyshen's approach on a real water distribution system provided by SaskWater[1].

1.6 Research Scope and Objectives

The prime objective of this research was to determine if Lesyshen's model-based algorithm could be applied to a "real" water distribution system (provided by Sask Water) with the intent of determining if this algorithm could accurately detect the location and magnitude of an underground leak.

The water distribution line selected was located north of Melfort, Saskatchewan. This water distribution line was selected it is a simple distribution system (one water inlet and one outlet under normal conditions). In addition, it was possible to physically "simulate" a leak along the distribution line simply by opening and closing a valve located between the inlet and outlets. The philosophy adopted in this study was based on the belief that the use of the Extended Kalman Filter to detect the leak must be successful on a simple line before it can be applied to more complex water distribution lines.

The scope of this research project involved (1) examining the literature on leak detection, (2) using Lesyshen's model to theoretically detect a leak in an underground pipeline (3) modifying Lesyshen's model to satisfy observability conditions, (4) further refinement of the algorithm to compensate for problems encountered in simulation studies, (5) performing field work on the Melfort pipeline to collect the necessary data, and (6) determining if the expanded Lesyshen's algorithm could detect the leak in a "real" underground pipeline. Based on these studies, it was a secondary objective to modify the approach to accommodate practical considerations.

1. SaskWater is the utility company in Saskatchewan that provides quality water and wastewater services to Saskatchewan municipalities, industry, First Nations and rural water groups.

It should be noted that although the initial intent of the study focused on the practical implementation of the algorithm (to the Melfort line), the reality of the study was that only simulated pipelines were considered in detail. As will be shown, this was a consequence of a failure of Lesyshen's original algorithm to accurately predict leakage locations in simulated pipelines under a wide variety of conditions.

1.7 Thesis Outline

This thesis will follow a similar outline as the research that was conducted to complete this project. The first Chapter was an introduction that defined the need of the project and laid out the objectives of this project. Chapter 2 presents the theoretical model developed and the equations that were used and developed to describe the transient behaviour in pipelines. Chapter 3 discusses the theory of the Kalman and the Extended Kalman Filter. The following chapter, Chapter 4, describes the Melfort site, which was the location of the testing and the implementation of the Extended Kalman Filter for leak detection. Chapter 5 presents the results of the Extended Kalman Filter applied to the pipeline model. Chapter 6 presents the results of a new approach known as the Extended Boundary Approach. Chapter 7 presents the result of the model applied to a small hydraulic line. Finally, Chapter 8 discusses the project summary, conclusion and recommendations for future work.

Chapter 2: Transient Pipe Flow Equations

2.1 Introduction

The approach adopted by (Lesyshen, 2005) and followed in this study made use of the Extended Kalman Filter (EKF) to estimate leakage at specified locations along the line. The EKF requires a model of the pipeline for implementation and as such, the governing equations for a fluid transmission line are required. This section will present the governing equation that will be used by the EKF and that will act as the "plant" equations upon which feasibility studies could be conducted.

Flow in pipelines is unsteady and therefore the typical steady flow pipeline equations do not fully explain the conditions in the pipeline. The transient-state flow in pipelines can be explained by the equations for conservation of mass and momentum. These equations are referred to as the continuity and momentum equations. These equations are partial differential equations because the flow and pressure in transient conditions are a function of both time and distance. The pressure and flow are dependent variables and the time and distance are independent variables. The goal is to determine the dependent variables (pressure and flow) as a function of time and space. The following chapter derives the momentum and continuity equations.

Once the equations are derived a numerical method known as the Method of Characteristics is introduced. This method, along with boundary conditions, may be used to solve the momentum and continuity equations. The objective of the method of characteristics is to convert the two partial differential equations into four ordinary differential equations that can be easily solved. These four equations are derived and presented in Chapter 2. The equations derived in this chapter are based on the work completed by Lesyshen, in 2005, and Chaundry, in 1987. The derivations are repeated in order to clarify the remaining work in this thesis.

2.2 Continuity Equation

The conservation of mass is used to derive the continuity equation. The conservation of mass states that the mass flowing into a control volume is equal to the mass flowing out of a control volume. The control volume is shown in Figure 2.1.

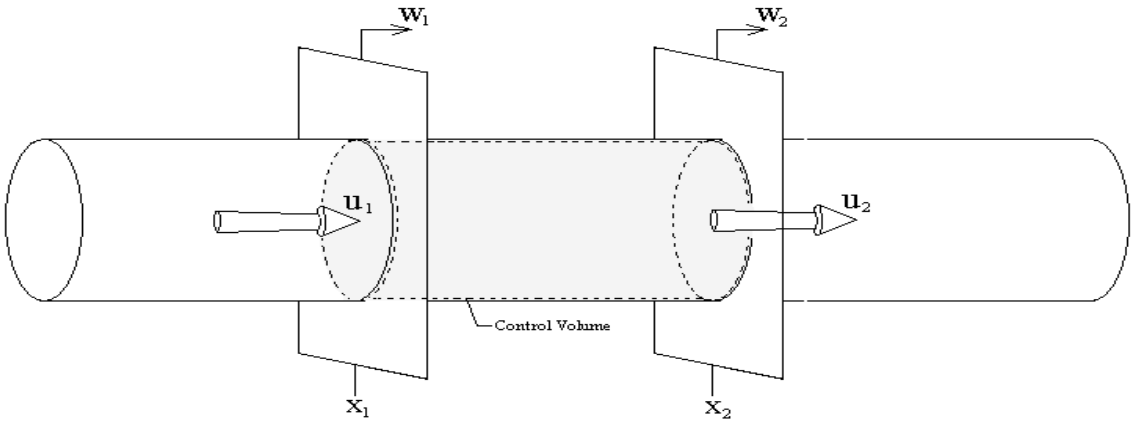

Figure 2.1: Control Volume for Continuity Equation (Lesyshen, 2005)

The control volume has three different control surfaces. The positive direction is defined as the direction of the flow or in the downstream direction. Two control surfaces are represented by x_1 and x_2. These two control surfaces are the circular ends of the control volume. The third control surface is represented by the dashed line in Figure 2.1. This control surface is the inner wall of the control volume. The incoming and outgoing fluid flows are represented by u_1 and u_2, respectively. The control volume may shorten or elongate due to pressure fluctuations. The velocities of the contraction or elongation of the control volume are represented by w_1 and w_2. If the radial velocity is small, the radial contraction and expansion can be assumed negligible. The flow is one dimensional and the pressure is uniform at the end sections of the control volume.

Applying the Reynolds transport theorem for the conservation of mass gives:

$$\frac{d}{dt}\int_{x1}^{x2}\rho A dx + \rho_2 A_2(u_2 - w_2) - \rho_1 A_1(u_1 - w_1) = 0\,, \qquad [2.1]$$

Chaudhry, in 1987, defines the Leibnitz's rule as:

$$\frac{d}{dt}\int_{f_1(t)}^{f_2(t)} F(x,t)dx = \int_{f_1(t)}^{f_2(t)} \frac{\partial}{\partial t} F(x,t)dx + F(f_2(t),t)\frac{df_2}{dt} - F(f_1(t),t)\frac{df_1}{dt},$$

where f_1 and f_2 are differential functions of *t* and *F(x,t)* and $\partial f/\partial t$ are continuous in space and time. Applying the Leibnitz's rule to Equation [2.1] results in:

$$\int_{x1}^{x2} \frac{\partial}{\partial t}(\rho A)dx + \rho_2 A_2 \frac{dx_2}{dt} - \rho_1 A_1 \frac{dx_1}{dt} + \rho_2 A_2 (u_2 - w_2) - \rho_1 A_1 (u_1 - w_1) = 0 \qquad [2.2]$$

Since x_1 and x_2 are fixed in time, their derivatives with respect to time are equal to the wall velocities.

$$\frac{dx_1}{dt} = w_1 \text{ and } \frac{dx_2}{dt} = w_2 \qquad [2.3]$$

When Equation [2.3] is substituted into Equation [2.2] the wall velocities cancel out and Equation [2.2] becomes:

$$\int_{x1}^{x2} \frac{\partial}{\partial t}(\rho A)dx + \rho_2 A_2 u_2 - \rho_1 A_1 u_1 = 0 \qquad [2.4]$$

The mean value theorem states that:

$$\int_{x_1}^{x_2} F(x)dx = (x_2 - x_1)F(\xi), where: x_1 \le \xi \le x_2$$

Applying the mean value theorem to the first term in Equation [2.4] results in:

$$\frac{\partial}{\partial t}(\rho A)\Delta x + (\rho A u)_2 - (\rho A u)_1 = 0 \qquad [2.5]$$

where $$\Delta x = (x_2 - x_1)$$

Dividing Equation [2.5] by Δx results in:

$$\frac{\partial}{\partial t}(\rho A) + \frac{(\rho_2 A_2 u_2 - \rho_1 A_1 u_1)}{\Delta x} = 0 \qquad [2.6]$$

If the limit of the second term on the left side of Equation [2.6] is taken:

$$\lim_{\Delta x \to 0} \frac{(\rho_2 A_2 u_2 - \rho_1 A_1 u_1)}{\Delta x} = \frac{\partial}{\partial x}(\rho A u) \qquad [2.7]$$

The result is the partial derivative $(\rho A u)$ with respect to *x*. Therefore Equation [2.6] can be simplified to:

$$\frac{\partial}{\partial t}(\rho A) + \frac{\partial}{\partial x}(\rho A u) = 0 \qquad [2.8]$$

Expanding the second term in Equation [2.8] by partial fractions gives:

$$\frac{\partial}{\partial t}(\rho A) + u\frac{\partial}{\partial x}(\rho A) + \rho A\frac{\partial u}{\partial x} = 0 \qquad [2.9]$$

Since the above equations deal with spatial and temporal terms the total derivative is introduced to simplify the equations. The total derivative of a function F that varies spatially and temporally, F = f(x,t), is represented as:

$$\frac{D}{Dt}F = \frac{\partial F}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial F}{\partial t}, \qquad [2.10]$$

where $\frac{D}{Dt}F$ represents the total derivative of function *F*. Allowing $\partial x/\partial t = u$, Equation [2.10] becomes:

$$\frac{D}{Dt}F = u\frac{\partial F}{\partial x} + \frac{\partial F}{\partial t}. \qquad [2.11]$$

Referring back to Equation [2.9] the first two terms represent the total derivative of (ρA) with respect to the fluid velocity, u. Therefore Equation [2.9] is rewritten as:

$$\frac{D}{Dt}(\rho A) + \rho A\frac{\partial u}{\partial x} = 0 \qquad [2.12]$$

Dividing Equation [2.12] by ρA and expanding by partial fractions gives:

$$\frac{1}{\rho A}\left(\left(\frac{D}{Dt}\rho\right)A + \rho\left(\frac{D}{Dt}A\right)\right) + \frac{\partial u}{\partial x} = 0 \qquad [2.13]$$

In order for Equation [2.13] to be used to relate pressure and flow, the elastic theory related to pipelines is used. The fluid's bulk modulus of elasticity, β, is defined as:

$$\beta = \rho_o\left(\frac{\partial P}{\partial \rho}\right), \qquad [2.14]$$

The partial derivative may be represented by a dot over the dependent variable which may also be rewritten as:

$$\frac{\dot{\rho}}{\rho} = \frac{\dot{P}}{\beta}. \qquad [2.15]$$

The circumferential strain, $\dot{\varepsilon}_T$, is related to the pipe wall expansion per unit area, per unit time ($\dot{A}/A$) by the following expression:

$$\frac{\dot{A}}{A} = 2\dot{\varepsilon}_T \,. \qquad [2.16]$$

Hooke's law relates the stress and strain and in a pipe it is expressed as:

$$\dot{\varepsilon}_T = \frac{1}{E}\left(\dot{\sigma}_\theta - \upsilon \dot{\sigma}_z \right), \qquad [2.17]$$

where σ_θ is the circumferential stress, σ_z is the axial stress, υ is Poisson's ratio and E denotes Young's modulus of the pipe. Substituting Equations [2.15], [2.16] and [2.17] into Equation [2.13] results in:

$$\frac{\dot{P}}{\beta} + \frac{2}{E}\left(\dot{\sigma}_\theta - \upsilon \dot{\sigma}_z\right) + \frac{\partial u}{\partial x} = 0 \,. \qquad [2.18]$$

The circumferential stress, in Equation [2.18], is related to pressure by:

$$\sigma_\theta = \frac{PD}{2e} \,,$$

therefore,

$$\dot{\sigma}_\theta = \frac{\dot{P}D}{2e} \,. \qquad [2.19]$$

where D is the pipe diameter and e is the thickness of the pipe walls. The diameter is assumed constant during transient flow because it changes very little when compared to pressure. The axial stress, σ_z, in Equation [2.18] is shown below for three different cases.

Case A: Pipe anchored at upstream end only: $\dot{\sigma}_z = \frac{\dot{P}D}{4e}$,

Case B: Pipe anchored throughout: $\dot{\sigma}_z = \upsilon \dot{\sigma}_\theta$, [2.20]

Case C: Pipe anchored with expansion joints throughout: $\dot{\sigma}_z = 0$.

Consider Case C and substituting equations [2.19] and [2.20] into [2.18] yields:

$$\frac{\dot{P}}{\beta} + \frac{2}{E}\left(\frac{\dot{P}D}{2e} - 0\right) + \frac{\partial u}{\partial x} = 0$$

$$\text{or, } \dot{P}\left(1 + \frac{\beta D}{Ee}\right) + \beta\frac{\partial u}{\partial x} = 0$$

$$\dot{P} + \left(\frac{\rho\frac{\beta}{\rho}}{1 + \frac{\beta}{E}\frac{D}{e}}\right)\frac{\partial u}{\partial x} = 0$$

$$\text{from which } \dot{P} + \rho a^2\frac{\partial u}{\partial x} = 0 \qquad [2.21]$$

where:
$$a^2 = \frac{\beta/\rho}{1 + \left[\left(\beta/E\right)\left(D/e\right)\right]c_1} \qquad [2.22]$$

a is a constant that is calculated from properties of the fluid, pipe properties and represents the wave speed in a pipe and in this case c_1 =1. If Case C is not used then the constant, c_1, has a non-unity value and is dependent on how the pipe is supported; the three cases are shown below.

$$\begin{aligned} &\text{Case A:} && c_1 = 1 - \upsilon/2, \\ &\text{Case B:} && c_1 = 1 - \upsilon^2, \\ &\text{Case C:} && c_1 = 1. \end{aligned} \qquad [2.23]$$

Expressing Equation [2.21] in terms of flow and expanding the total derivative the continuity equation may be expressed as:

$$\frac{Q}{A}\frac{\partial P}{\partial x}+\frac{\partial P}{\partial t}+\frac{\rho a^2}{A}\frac{\partial Q}{\partial x}=0. \qquad [2.23]$$

where $Q = VA$.

2.3 Momentum Equation

The Navier-Stokes equation is used to derive the momentum equation in this section. Navier-Stokes equations are used extensively to describe many different fluid flows such as internal flows (pipe flow) and external flows (flow over an airplane wing). Figure 2.2 shows the pipeline geometry.

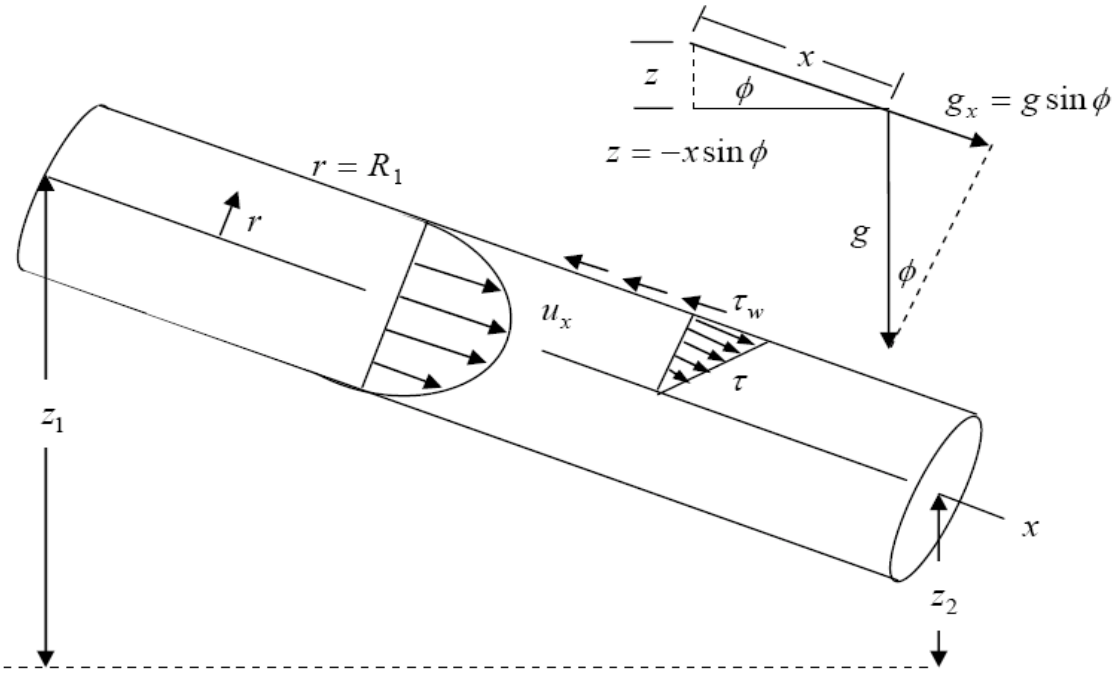

Figure 2.2: Pipeline Geometry (Lesyshen, 2005)

Pipeline flow is assumed one-dimensional and therefore only the x component of the Navier-Stokes equations is considered. The axial direction equation in cylindrical coordinates is:

$$\rho\left(\frac{\partial u_x}{\partial t}+u_r\frac{\partial u_x}{\partial r}+\frac{u_\theta}{r}\frac{\partial u_x}{\partial \theta}+u_x\frac{\partial u_x}{\partial x}\right)=\rho g_x-\frac{\partial P}{\partial x}+\mu\left(\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u_x}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2 u_x}{\partial \theta^2}+\frac{\partial^2 u_x}{\partial x^2}\right) \quad [2.24]$$

where ρ is the fluid density, u is the fluid velocity which consists of radial u_r, rotational u_θ and axial u_x components, g_x is axial component of the gravity acceleration, P is pressure and μ is fluid viscosity.

The flow is assumed to be Newtonian, irrotational, one-dimensional, and incompressible with constant properties. Therefore the following simplifications are made:

$$\begin{aligned} u_r &= 0, \\ u_\theta &= 0, \\ u_x &= u(r,x). \end{aligned}$$

The above expressions state that the flow is only a function of the radial and axial position in the pipe. Since the flow is one-dimensional, the axial fluid velocity is now simply stated as u. Equation [2.24] can therefore be reduced to:

$$\rho\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}\right)=\rho g_x-\frac{\partial P}{\partial x}+\frac{\mu}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) \quad [2.25]$$

where the terms on the right hand side represent the gravity forces, pressure forces and shear stresses, respectively. The first term on the left hand side of Equation [2.25] represents the local acceleration. This is the acceleration due to opening or closing of a valve and therefore is the inertial or transient effects of the flow. The second term on the left hand side of Equation [2.25] is the convective acceleration. This is the acceleration due to changes in pipe geometry such as elbows. The shear stress can be stated as:

$$\tau=\mu\frac{\partial u}{\partial r}, \quad [2.26]$$

Substituting Equation [2.26] into Equation [2.25] results in:

$$\rho\left(\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}\right)=\rho g_x-\frac{\partial P}{\partial x}+\frac{1}{r}\frac{\partial(r\tau)}{\partial r}. \qquad [2.27]$$

Generally the factors that affect shear stress in a pipe are wall roughness, viscosity, fluid velocity and radial position. Assuming that wall roughness, and viscosity are constant the shear stress is then only a function of fluid velocity and radial position, $\tau = f(r,u)$. According to Chaudhry, in 1987, it is an accepted practice to assume that for the same flow rate the head loss (friction), during transient conditions, is the same as steady state conditions. This allows the shear stress in Equation [2.27] to be evaluated as an ordinary differential equation and it allows the friction forces to be determined by the Darcy-Weisbach equation. The Darcy-Weisbach equation that relates the shear stress at the wall to the friction factor f is given as:

$$\tau_w=\frac{\rho f u|u|}{8} \qquad [2.28]$$

The flow is written as $u|u|$ to take care of reverse flows. Equation [2.27] can be rearranged and integrated from the centre of the pipe, r = 0, to the wall, r = R. The results are:

$$\int_0^R \frac{d(r\tau)}{dr}dr=\int_0^R r\left(\rho\frac{\partial u}{\partial t}+\rho u\frac{\partial u}{\partial x}-\rho g_x+\frac{\partial P}{\partial x}\right)dr,$$

$$R\tau=\frac{R^2}{2}\left(\rho\frac{\partial u}{\partial t}+\rho u\frac{\partial u}{\partial x}-\rho g_x+\frac{\partial P}{\partial x}\right),$$

$$\tau=\frac{R}{2}\left(\rho\frac{\partial u}{\partial t}+\rho u\frac{\partial u}{\partial x}-\rho g_x+\frac{\partial P}{\partial x}\right). \qquad [2.29]$$

According to Figure 2.3 the shear stress is acting in the negative x-direction and therefore after substituting Equation [2.28] into Equation [2.29] and rearranging the resulting equation is:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} - g_x + \frac{1}{\rho}\frac{\partial P}{\partial x} + \frac{fu|u|}{4R} = 0\,. \qquad [2.30]$$

Elevation, z, may be introduced into the equation by relating it to the body forces, g_x. Referring to Figure 2.2, $g_x = g\sin\phi$, and substituting the diameter for the radius, $R = \frac{D}{2}$, Equation [2.30] becomes:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} - g\sin\phi + \frac{1}{\rho}\frac{\partial P}{\partial x} + \frac{fu|u|}{2D} = 0 \qquad [2.31]$$

Equation [2.31] can be rearranged to group the pressure and body force term as:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{1}{\rho}\frac{\partial}{\partial x}\left(P - \rho g x\sin\phi\right) + \frac{fu|u|}{2D} = 0\,. \qquad [2.32]$$

From Figure 2.2, $x\sin\phi = -z$ and Equation [2.32] becomes:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{1}{\rho}\frac{\partial}{\partial x}\left(P + \rho g z\right) + \frac{fu|u|}{2D} = 0\,. \qquad [2.33]$$

Making the substitution for velocity, $u = \frac{Q}{A}$, the momentum equation that represents transient pipe flow can be written as:

$$\frac{\partial Q}{\partial t} + \frac{Q}{A}\frac{\partial Q}{\partial x} + \frac{A}{\rho}\frac{\partial}{\partial x}\left(P + \rho g z\right) + \frac{fQ|Q|}{2DA} = 0\,, \qquad [2.34]$$

2.4 General Remarks on the Continuity and Momentum Equations

The transient-state flow in pipelines is described by the continuity and momentum equation (Equation [2.23] and [2.34]). These two equations are a set of first-order, partial differential equations. The equation types must be determined in order to choose a suitable numerical method which can be accomplished by examining the roots of the characteristic equation (Chaundhry,1987). Rewriting Equation [2.23] and [2.34] in matrix form results in:

$$\frac{\partial}{\partial t}\begin{Bmatrix} Q \\ \mathrm{P} \end{Bmatrix} + \frac{1}{A}\begin{bmatrix} Q & A^2/\rho \\ \rho a^2 & Q \end{bmatrix}\frac{\partial}{\partial x}\begin{Bmatrix} Q \\ \mathrm{P} \end{Bmatrix} = \begin{Bmatrix} -Ag\dfrac{\partial z}{\partial x}_0 - \dfrac{fQ|Q|}{2DA} \end{Bmatrix}. \qquad [2.35]$$

According to Chaundhry, in 1987, the eigenvalues, λ, of the square matrix on the left hand side of Equation [2.35] give the roots of the equation set and therefore determine the type of the equation set. The eigenvalues are determined as:

$$B = \begin{bmatrix} Q & A^2/\rho \\ \rho a^2 & Q \end{bmatrix},$$

$$D(\lambda) = \det(B - \lambda I) = \begin{bmatrix} Q-\lambda & A^2/\rho \\ \rho a^2 & Q-\lambda \end{bmatrix} = 0,$$

$$(Q-\lambda)^2 = A^2 a^2,$$

$$\lambda = Q \pm Aa.$$

The eigenvalues and therefore the roots of the characteristic equation are real, distinct and therefore Equation [2.23] and [2.34] are a set of non-linear hyperbolic partial differential equations which represents the phenomenon of wave propagation (Chaundhry, 1987). An exact solution of these equations does not exist; however through approximations, solutions can be obtained. The most popular numerical method for these equations is known as the Method of Characteristics and is discussed below.

2.5 The Method of Characteristics

The Method of Characteristics is used, in this thesis, to numerically solve the momentum and continuity equations. The momentum and continuity equations are repeated below to facilitate discussion.

Momentum Equation $$\frac{\partial Q}{\partial t}+\frac{Q}{A}\frac{\partial Q}{\partial x}+\frac{A}{\rho}\frac{\partial}{\partial x}\left(P+\rho g z\right)+\frac{fQ|Q|}{2DA}=0\,, \qquad [2.34]$$

Continuity Equation $$\frac{Q}{A}\frac{\partial P}{\partial x}+\frac{\partial P}{\partial t}+\frac{\rho a^2}{A}\frac{\partial Q}{\partial x}=0\,. \qquad [2.23]$$

An approximation must be made in order to be able to numerically solve the above equations. The spatial terms of Q and P may be neglected when both the spatial and transient terms appear in the same equation because the spatial variation is much smaller than the time-varying equation (Watters, 1984). Implementing this assumption Equation [2.34] and [2.23] can be rewritten as:

$$L1=\quad \frac{\partial Q}{\partial t}+\frac{A}{\rho}\frac{\partial P}{\partial x}+Ag\frac{\partial z}{\partial x}+\frac{fQ|Q|}{2DA}=0\,, \qquad [2.36]$$

$$L2=\quad \frac{A}{\rho}\frac{\partial P}{\partial t}+a^2\frac{\partial Q}{\partial x}=0\,. \qquad [2.37]$$

Since $\partial z/\partial x$ represents the slope of the pipeline and does not vary in time, it can be written as an ordinary differential equation dz/dx. Equations [2.36] and [2.37] can be combined in a linear fashion and Equation [2.38] is one possible combination:

$$L=\lambda L1+\ L2=\lambda\left(\frac{\partial Q}{\partial t}+\frac{A}{\rho}\frac{\partial P}{\partial x}+Ag\frac{\partial z}{\partial x}+\frac{fQ|Q|}{2DA}\right)+\frac{A}{\rho}\frac{\partial P}{\partial t}+a^2\frac{\partial Q}{\partial x}=0 \qquad [2.38]$$

Two equations equivalent to the momentum and continuity equations can be determined with any two real, distinct values of λ.

Combining like terms in Equation [2.38] results in:

$$\left(\lambda\frac{\partial Q}{\partial t}+a^2\frac{\partial Q}{\partial x}\right)+\frac{A}{\rho}\left(\frac{\partial P}{\partial t}+\lambda\frac{\partial P}{\partial x}\right)+\lambda Ag\frac{\partial z}{\partial x}+\frac{\lambda fQ|Q|}{2DA}=0. \qquad [2.39]$$

The two bracketed terms are the total derivatives with respect to some velocity, dx/dt. Replacing the first bracket with the total derivative times an unknown λ results in:

$$\left(\lambda\frac{\partial Q}{\partial t}+a^2\frac{\partial Q}{\partial x}\right)=\lambda\frac{dQ}{dt}=\lambda\left(\frac{\partial Q}{\partial t}+\frac{dx}{dt}\frac{\partial Q}{\partial x}\right)=\lambda\left(\frac{\partial Q}{\partial t}+\frac{a^2}{\lambda}\frac{\partial Q}{\partial x}\right)$$

and,

$$\frac{dx}{dt}=\frac{a^2}{\lambda}. \qquad [2.40]$$

Next the second bracketed terms in Equation [2.39] gives:

$$\frac{dP}{dt}=\left(\frac{\partial P}{\partial t}+\frac{dx}{dt}\frac{\partial P}{\partial x}\right)=\left(\frac{\partial P}{\partial t}+\lambda\frac{\partial P}{\partial x}\right),$$

and,

$$\frac{dx}{dt}=\lambda. \qquad [2.41]$$

Combining Equation [2.40] and [2.41] results in:

$$\lambda^2=a^2 \qquad or \qquad \lambda=\pm a, \qquad [2.42]$$

Substituting Equation [2.42] into Equation [2.39] and using the total derivatives notation results in:

$$a\frac{dQ}{dt}+\frac{A}{\rho}\frac{dP}{dt}+Aag\frac{dz}{dx}+\frac{afQ|Q|}{2DA}=0. \qquad [2.43]$$

Dividing Equation [2.43] by the wave speed, a, results in:

$$\frac{dQ}{dt}+\frac{A}{a\rho}\frac{dP}{dt}+Ag\frac{dz}{dx}+\frac{fQ|Q|}{2DA}=0. \qquad [2.44]$$

From the derivation above, Equation [2.44] is valid for

$$\frac{dx}{dt}=\pm a. \qquad [2.45]$$

Equation [2.44] can be written as two separate equations because it is valid for $dx/dt=\pm a$. These two equations are referred to as the compatibility equations (Equation [2.46] and [2.48]) and are shown below. This results in two sets of equations known as the C^+ and C^- equations.

$$\left.\begin{aligned}\frac{dQ}{dt}+\frac{A}{a\rho}\frac{dP}{dt}+Ag\frac{dz}{dx}+\frac{fQ|Q|}{2DA}=0,\\ dx/dt=+a,\end{aligned}\right\}\ C^+ \qquad [2.46],\ [2.47]$$

$$\left.\begin{aligned}\frac{dQ}{dt}-\frac{A}{a\rho}\frac{dP}{dt}+Ag\frac{dz}{dx}+\frac{fQ|Q|}{2DA}=0,\\ dx/dt=-a.\end{aligned}\right\}\ C^- \qquad [2.48],\ [2.49]$$

The above equations can be simplified by introducing the piezometric head term given below as:

$$H=\frac{P}{\rho g}+z \qquad [2.50]$$

Therefore Equation [2.46] and [2.48] become:

$$\frac{dQ}{dt} + \frac{Ag}{a}\frac{dH}{dt} + \frac{fQ|Q|}{2DA} = 0 \qquad [2.51]$$

$$\frac{dQ}{dt} - \frac{Ag}{a}\frac{dH}{dt} + \frac{fQ|Q|}{2DA} = 0 \qquad [2.52]$$

These equations are referred to as the C^+ and C^- equations because they are valid along the C^+ and C^- characteristic lines. Referring to Figure 2.3 the idea of characteristic lines is graphically shown on a characteristic grid. Equation [2.47] plots one of the characteristic lines and similarly Equation [2.49] plots the other characteristic line. The compatibility equations are only valid along the corresponding characteristic lines. The pipeline shown in Figure 2.3 is divided into a number of "reaches" of length Δx. The time step, Δt, is calculated based on Equation [2.47]. Therefore data for every time step (Δt interval) can be calculated for each position (Δx interval).

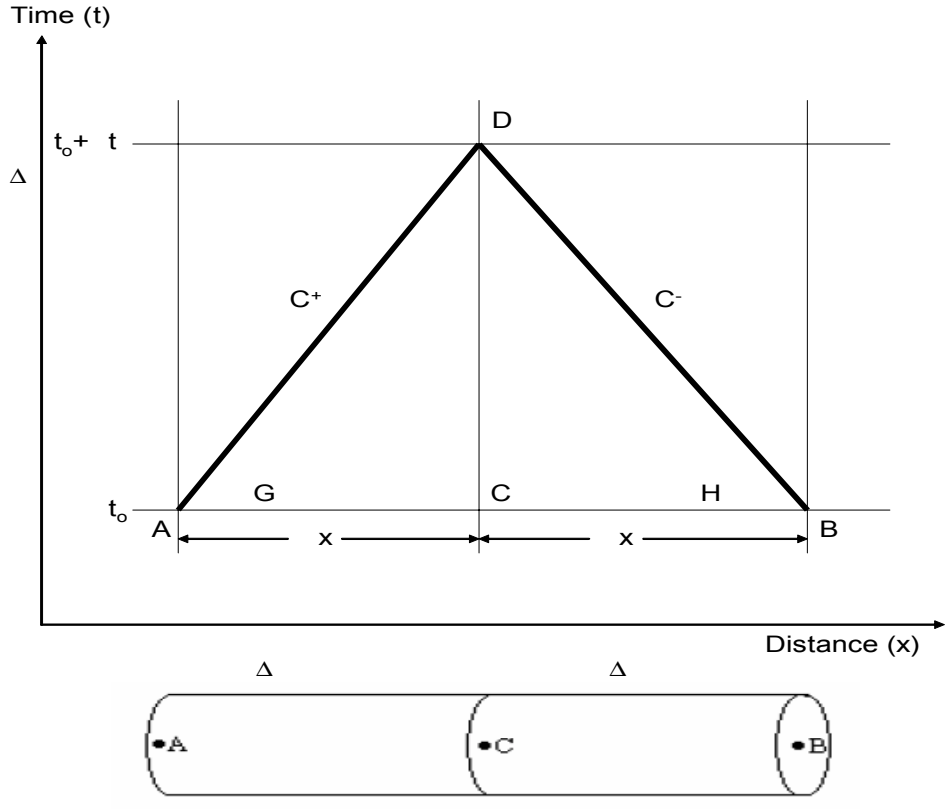

Figure 2.3: Method of Characteristics Grid

As mentioned above the C^+ line that joins points A and D, is plotted by Equation [2.47]. Therefore, Equation [2.46] (or the equivalent Equation [2.51]) can be integrated from point A to D if the pressure and flow at point A is known. The result of the integration will produce an equation in terms of the unknown pressure and flow at point D. Similarly the C^- line, that joins point D and B, is plotted by Equation [2.49]. Therefore, Equation [2.48] (or the equivalent Equation [2.52]) can be integrated from point B to D if the pressure and flow at point B is known. This will result in a second equation at point D in terms of the unknown pressure and flow. This process results in two equations and two unknowns (P_D and Q_D) at point D. Therefore performing this process allows pressure and flow to be calculated at each Δx or "reach" and Δt. This results in building a dynamic picture of the pressure and flow throughout the pipeline. In order to start the integration process all the pressures and flows at each Δx must be known. The next section shows how the equations are integrated along the characteristic lines.

2.5.1 Discretization

Equations [2.51] and [2.52] can be integrated if the initial conditions are known at point A and B. Equation [2.51] may be integrated along the C^+ characteristic line after it is multiplied by $adt = dx$.

$$a\int_{Q_A}^{Q_D}\frac{dQ}{dt}dt + Ag\int_{H_A}^{H_D}\frac{dH}{dt}dt + \frac{f}{2DA}\int_{X_A}^{X_D}Q|Q|dx = 0 \qquad [2.53]$$

The last term in Equation [2.53] is the spatial variation of the flow, Q. This term is unknown and therefore an approximation must be made. A first order approximation is satisfactory where the friction term does not dominate. If this assumption is questionable, a numerical test can be performed (see Section 2.4.2). The last term under the integral is approximated as $Q_A|Q_A|\Delta x$. The remaining terms in Equation [2.53] are easily integrated and result in:

$$a(Q_D - Q_A) + Ag(H_D - H_A) + \frac{f\Delta x}{2DA}Q_A|Q_A| = 0\,, \qquad [2.54]$$

Rearranging Equation [2.54] and solving for H_D gives:

$$C^+: \quad H_D = H_A - \frac{a}{Ag}(Q_D - Q_A) - \frac{f\Delta x}{2gDA^2} Q_A |Q_A| \qquad [2.55]$$

Similarly Equation [2.52] can be integrated along the C^- characteristic line (line BD) resulting in:

$$C^-: \quad H_D = H_B + \frac{a}{Ag}(Q_D - Q_B) + \frac{f\Delta x}{2gDA^2} Q_B |Q_B| . \qquad [2.56]$$

In order to clearly define the variables for a general case, a subscript notation is introduced. The spatial and temporal locations of measurement D are shown by the subscripts i and j shown as Q_{ij} or H_{ij}. Introducing the new notation, Equation [2.55] and [2.56] are simplified to:

$$C^+: \quad H_{ij} = C_P - BQ_{ij} , \qquad [2.57]$$

$$C^-: \quad H_{ij} = C_M + BQ_{ij} . \qquad [2.58]$$

where the constants are:

$$C_P = H_{i-1,j-1} + Q_{i-1,j-1}\left(B - R\left|Q_{i-1,j-1}\right|\right), \qquad [2.59]$$

$$C_M = H_{i+1,j-1} - Q_{i+1,j-1}\left(B - R\left|Q_{i+1,j-1}\right|\right), \qquad [2.60]$$

$$B = \frac{a}{Ag} , \qquad [2.61]$$

$$R = \frac{f\Delta x}{2gDA^2} . \qquad [2.62]$$

Once the initial conditions are known, Equations [2.57] to [2.62] can be easily programmed into computer code and the flow and head can be determined throughout the pipeline. A steady state solution is used to determine the initial conditions in order to start the iterative process. However, boundary conditions are also required to solve these equations. Since the C^+ equation requires information from the previous "reach, Δx" it cannot be used at the pipe inlet and a boundary equation is required. Also the C^- equation requires information at the next "reach, Δx" therefore it cannot be used at the pipe outlet and another boundary condition is required. Usually, boundary conditions include reservoirs, pumps, and valves. A complete solution to the transient conditions in the pipeline exists once the appropriate boundary conditions are determined. Chapter 4 discusses the boundary conditions chosen for this thesis.

2.5.2 Stability

In order to assure the discretization technique is stable it must satisfy the Courant-Friedrich-Lewy (CFL) stability condition (Chaundry, 1987 and O'Brien, 1951). According to the CFL, the discretization technique will be stable if: $\Delta x \geq a\Delta t$. From this the Courant number (Cn) is defined as "the ratio of the actual wave speed to the numerical wave speed, $Cn = \frac{a}{\Delta x / \Delta t} = \frac{a\Delta t}{\Delta x}$" (Chaundry, 1987). Therefore for the discetization to be stable Cn must be $\leq$ 1. Referring to Figure 2.4, this condition states that the characteristic lines through Point D must intersect the line AB between AC and CB. If Cn is larger than 1 than the slope of the characteristic lines would decrease and characteristic lines would not intersect line AB between AC and CB causing the discretization technique to become unstable.

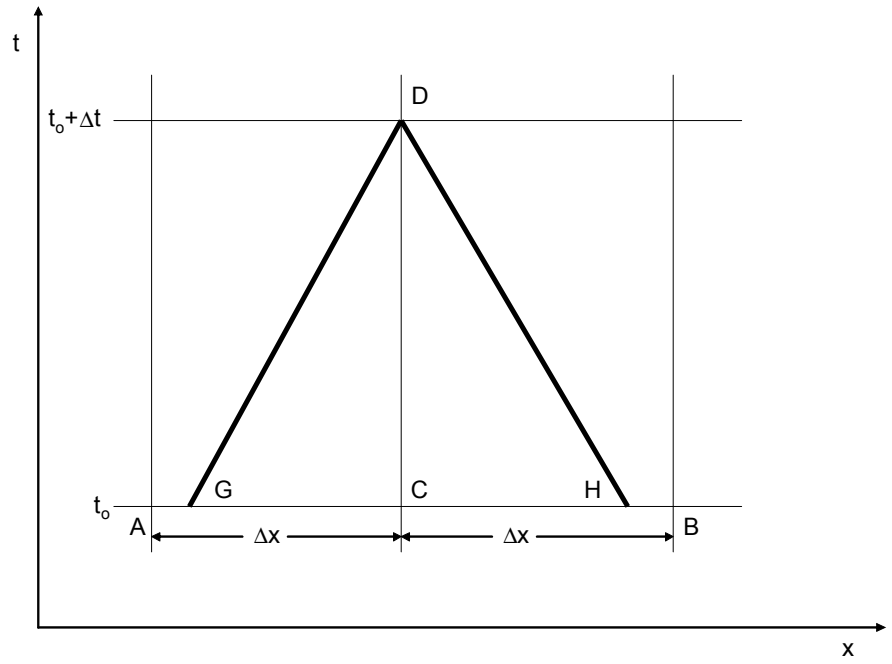

Figure 2.4: Courant Stability Condition

The first-order approximation in Equation [2.54] will produce accurate results if the above stability condition is satisfied and the friction term in Equation [2.54] is small. If the friction term is not considered small then a first order approximation may not result in accurate results. In cases where there is doubt in the first-order approximation of the friction the approximation may be checked by reducing the time-step. If the same results occur for a reduced time step then the first order response may be accepted with confidence (Wylie and Streeter, 1983). Also a stability criteria for the first order model was developed by O'Brien et.al, in1951, which states:

$$\frac{f\Delta tQ}{4DA} \leq 1 \qquad [2.63]$$

If the left hand side of Equation [2.63] is less or equal to 1 then the accuracy of the first-order integration is not in doubt (Wylie and Streeter, 1983).

Chapter 3: The Extended Kalman Filter

3.1 Introduction

This chapter discusses the theory behind the Kalman Filter (KF) and the Extended Kalman Filter (EKF). The EKF is used to estimate pipeline leakage, given a pipeline model and set of inputs. Since the equations that represent the conditions in the pipeline are non-linear, the EKF is required (as opposed to the Kalman Filter). However a complete explanation of the KF is required to understand the EKF.

The Kalman Filter was first introduced by Rudolph Kalman in 1960 when he published his paper describing a *"recursive solution to the discrete-data linear filtering problem"* (Kalman, 1960). *"One of the most well-known and often-used tools for stochastic estimation from noisy sensor measurements is known as the Kalman Filter"* (Welsch and Bishop, 2001). Systems that are corrupted due to noise or have an associated random portion are referred to as stochastic systems. The Kalman Filter has been used for navigation, tracking and guidance, fault detection, and many other applications where parameter estimation is required. The Kalman Filter is a set of mathematical equations that are transformed into a recursive state estimation tool that is able to provide robust, unbiased solution of the least square method. The Kalman Filter has become more useful for very complicated real time applications with the development of high speed computers. It is possible to process measurements online because of the recursive property of the Kalman Filter. The Kalman Filter is able to take information from measurements and models to provide an optimal solution. In the Kalman filter, each new estimate is formed as a "*blend of the old estimate and the current measurement*" (Brown, 1997). A review of probability and statistics is included in Appendix B in order for unfamiliar readers to better understand the principles behind Kalman Filtering.

3.2 The Kalman Filter

The Kalman Filter is a technique to estimate the *"instantaneous state of a linear dynamic system perturbed by white noise by using measurements linearly related to the state, but corrupted with white noise"* (Grewal and Andrews, 1993). The Kalman filter minimizes

the estimated error covariance in a linear stochastic system. There are two types of noise associated with stochastic estimation, process noise and measurement noise. Process noise can be explained as the difference between the real system and the model. The measurement noise is the noise associated with the sensors and instrumentation. The Kalman filter is capable of handling situations with a lot of noise, or high uncertainty in the data; therefore, it is believed to be suitable for pipeline leak detection.

3.2.1 Discrete State Space Model

State-space models are basically a notational convenience for estimation and control problems, developed to make *"tractable what would otherwise be a notation-intractable analysis"* (Welsch and Bishop, 2001). Equation [3.1] shows how the new state is modeled by a combination of the prior state, inputs and random noise. The state and measurement equations are given as:

$$x_{k+1} = \phi_k x_k + G_k u_k + w_k, \qquad [3.1]$$

$$z_{k+1} = H_k x_{k+1} + v_{k+1}, \qquad [3.2]$$

where, $x_k = (n \times 1)$ state vector at time t_k,

$\phi_k = (n \times n)$ system or transition matrix for time t_k,

$G_k = (n \times r)$ input matrix,

$u_k = (r \times 1)$ input vector

$w_k = (n \times 1)$ vector of random system disturbances characterized by zero mean white noise,

$z_{k+1} = (p \times 1)$ vector of defined measurements,

$H_k = (p \times n)$ output matrix that linearly connects outputs and states, and

$v_{k+1} = (p \times 1)$ vector of white measurement noise.

The Kalman Filter requires a discrete state space model. The transient equations, from Chapter 2, are discretized using a finite difference method known as the Method of

Characteristics. The process and measurement noise, w_k and v_k are assumed white. The term "white" noise is referred to as a signal that has equal power at all frequencies and is completely uncorrelated temporally, random and zero mean sequence. Mathematically, these correlations are denoted by Q_k (process noise) and R_k (measurement noise).

$$E[w_k] = 0 \qquad E[v_k] = 0\,, \tag{3.3}$$

$$\mathrm{cov}[w_k, w_j] = E[w_k, w_j^T] = \begin{cases} Q_k, & j = k \\ 0, & j \neq k \end{cases}, \tag{3.4}$$

$$\mathrm{cov}[v_k, v_j] = E[v_k, v_j^T] = \begin{cases} R_k, & j = k \\ 0, & j \neq k \end{cases}, \tag{3.5}$$

where, $E[\ ]$ denotes the expectation, and $\mathrm{cov}[\ ]$ denotes the covariance.

3.2.2 The Filtering Process

The Kalman filter is used to extract (estimate) states that cannot be measured directly from a physical "plant". The Kalman Filter works by blending the information from system measurements and model predictions using the Kalman Gain, K_k. Initial "guesses" of the plant states are made (defined as "a priori" or "unrefined" estimates) and projected to the output space. The predicted and measured outputs from the system are then compared. The associated error is then multiplied by the Kalman gain and used to refine the a priori state estimate. The Kalman gain is considered to be optimal and yields a better state estimate (defined as "refined" or "a posterior" estimates).

The measurements from the system provide valuable information that is not known by the assumed model. By incorporating system measurements into the prediction of the states, it allows the model to better predict real world situations. The Kalman Filter is therefore a two stage filter known as the predictor-corrector algorithm which is explained in this section. It should be noted that in the following discussion, the general form of the Kalman Filter equations will be introduced to illustrate the steps. Derivations of the equations are deferred until Section 3.2.3.

The Kalman filter estimates a process state by using a form of feedback control: "the filter estimates the process state at some time and then obtains feedback in the form of (noisy) measurements" (Welsch and Bishop, 2001). The Kalman Filter is therefore comprised of two sets of equations; the measurement update (corrector) equations and the time update (predictor) equations. The estimation process begins by acquiring an unrefined guess of the system states, $\hat{x}_k^-$ and the error covariance matrix, P_k^- (formally defined in the next section). The unrefined error covariance matrix is used to calculate the Kalman Gain, K_k. (see Equation[3.6]) The difference between the measurements and the predicted output of the system obtained from the unrefined states is multiplied by the Kalman Gain and added to the unrefined estimates to improve the estimated states (see Equation [3.7]). The refined error covariance estimate is also determined by using the Kalman Gain and the unrefined error covariance (see Equation [3.8]).

In summary, the corrective phase equations are given as follows:

$$K_k = P_k^- H_k^T \left(H_k P_k^- H_k^T + R_k \right)^{-1}, \qquad [3.6]$$

$$\hat{x}_k = \hat{x}_k^- + K_k \left(z_k - H_k \hat{x}_k^- \right), \qquad [3.7]$$

$$P_k = \left(I - K_k H_k \right) P_k^- . \qquad [3.8]$$

Because a model of the plant is known, the estimate of the states and error covariance matrices for the next step $\hat{x}_{k+1}^-$, P_{k+1}^- can be predicted using Equations [3.9] and [3.10], respectively (Grewal and Andrews, 1993).

$$\hat{x}_{k+1}^- = \phi_k \hat{x}_k + G_k u_k , \qquad [3.9]$$

$$P_{k+1}^- = \phi_k P_k \phi_k^T + Q_k . \qquad [3.10]$$

These projected states then become the "unrefined" states and error covariance for the next step at time k+1. This procedure is repeated for every time step.

The error covariance matrix, P_k^-, along with all of the other Kalman Filter equations will be further explained and derived in Section 3.2.3. The Kalman Filter requires initial estimates of the states, inputs and error covariance matrix. Because of the robust nature of the Kalman Filter, the initial conditions may be assigned to zero and the filter will still converge if the assumption of Gaussian noise is correct, observability conditions satisfied, and the error is not too large to cause numerical problems due to matrix inversion (Grewal and Andrews, 1993). Equations [3.6] to [3.10] are a recursive process and are graphically shown in Figure 3.1.

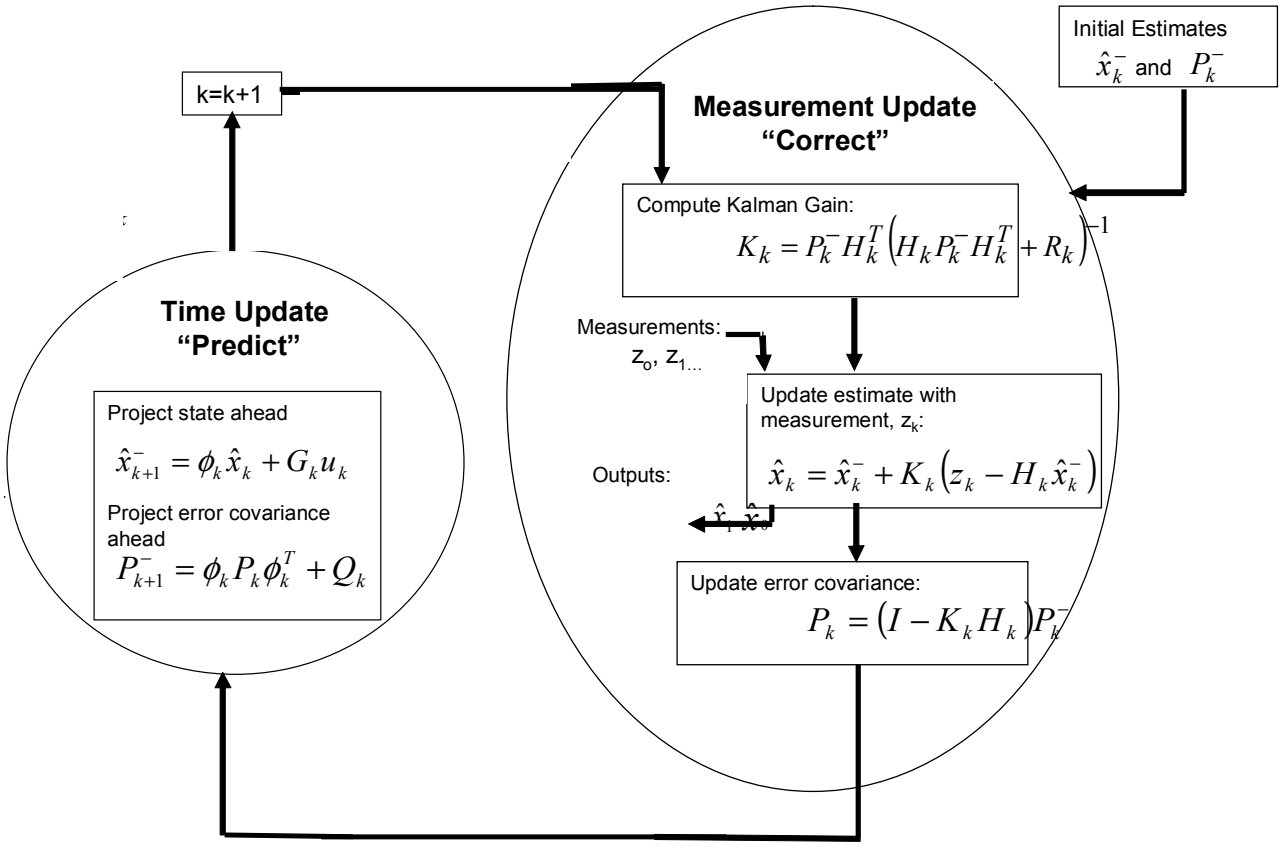

Figure 3.1: Kalman Filter Recursive Loop

3.2.3 Derivation of the Kalman Filtering Equations

The filter works by minimizing the error covariance matrices. Appendix B contains some statistical definitions including covariance. *"There are two estimate errors, modelling error and measurement error, that are defined as a priori and a posteriori"* (Welsch and Bishop, 2001). The difference between the actual states and the unrefined states predicted by the model is known as the *a priori* error. The *a posteriori* error is the difference between the actual states and the refined states (refined because they have knowledge of the system by way of measurements). Mathematically the *a priori* and *a posteriori* errors are defined as:

$$e_k^- \equiv x_k - \hat{x}_k^- \text{, } (a\ priori) \qquad [3.11]$$

$$e_k \equiv x_k - \hat{x}_k \text{ . } (a\ posteriori) \qquad [3.12]$$

The mean squared error terms are represented by the covariance matrices P_k^- and P_k, and are associated with the errors defined above. These matrices are therefore referred to as the error covariance matrices. The *a priori* and *a posteriori* error covariance matrices are given as:

$$P_k^- = E\left[e_k^- e_k^{-T}\right] = E\begin{bmatrix} (e_{1k}^-)^2 & e_{1k}^- e_{2k}^- & \cdots & \\ e_{2k}^- e_{1k}^- & (e_{2k}^-)^2 & & \\ \vdots & & \ddots & \\ & & & (e_{nk}^-)^2 \end{bmatrix} = E\left[\left(x_k - \hat{x}_k^-\right)\left(x_k - \hat{x}_k^-\right)^T\right], \qquad [3.13]$$

$$P_k = E\left[e_k e_k^T\right] = E\begin{bmatrix} (e_{1k})^2 & e_{1k} e_{2k} & \cdots & \\ e_{2k} e_{1k} & (e_{2k})^2 & & \\ \vdots & & \ddots & \\ & & & (e_{nk})^2 \end{bmatrix} = E\left[\left(x_k - \hat{x}_k\right)\left(x_k - \hat{x}_k\right)^T\right]. \qquad [3.14]$$

The diagonal of Equation [3.13] and [3.14] represent the mean squared error terms. This is the squared error of one state. The non-diagonal elements in Equations [3.13] and [3.14] are the combined errors of two different states, either prior or posterior. The

overall goal of the Kalman Filter is to determine the Kalman Gain, K_k, which will minimize the mean squared error terms in an optimal manner subject to the assumption of Gaussian noise.

In order to minimize the mean squared error terms in the error covariance matrices an expression must be derived to relate the Kalman Gain to the error covariance. By combining Equations [3.13] and [3.7] an expression relating the Kalman Gain and the error covariance may be determined as:

$$P_k = E\left[\left(x_k - \hat{x}_k^- - K_k\left(z_k - H_k\hat{x}_k^-\right)\right)\left(x_k - \hat{x}_k^- - K_k\left(z_k - H_k\hat{x}_k^-\right)\right)^T\right] \qquad [3.15]$$

Substituting in Equations [3.2] and [3.11] gives:

$$P_k = E\left[\left(e_k^- - K_k\left(H_k x_k + v_k - H_k\hat{x}_k^-\right)\right)\left(e_k^- - K_k\left(H_k x_k + v_k - H_k\hat{x}_k^-\right)\right)^T\right]. \qquad [3.16]$$

Simplifying the $e_k^- - K_k\left(H_k x_k + v_k - H_k\hat{x}_k^-\right)$ expression of Equation [3.16]:

$$\begin{aligned} e_k^- - K_k\left(H_k x_k + v_k - H_k\hat{x}_k^-\right) &= e_k^- - K_k H_k x_k - K_k v_k + K_k H_k\hat{x}_k^- \\ &= e_k^- - K_k H_k\left(x_k - \hat{x}_k^-\right) - K_k v_k \\ &= e_k^- - K_k H_k e_k^- - K_k v_k \\ &= \left[I - K_k H_k\right]e_k^- - K_k v_k . \end{aligned} \qquad [3.17]$$

The second expression $\left[e_k^- - K_k\left(H_k x_k + v_k - H_k\hat{x}_k^-\right)\right]^T$ of Equation [3.16] may be simplified to:

$$\begin{aligned} \left[e_k^- - K_k\left(H_k x_k + v_k - H_k\hat{x}_k^-\right)\right]^T &= e_k^{-T} - \left(H_k x_k + v_k - H_k\hat{x}_k^-\right)^T K_k{}^T \\ &= e_k^{-T} - \left(H_k e_k^- + v_k\right)^T K_k{}^T \end{aligned}$$

$$= e_k^{-T} - \left(e_k^{-T} H_k^{\ T} + v_k^{\ T} \right) K_k^{\ T}$$

$$= e_k^{-T} \left(I - H_k^{\ T} K_k^{\ T} \right) - v_k^{\ T} K_k^{\ T} . \qquad [3.18]$$

Substituting Equations [3.17] and [3.18] back into Equation [3.16] gives the error covariance matrix as:

$$P_k = E\left[\left[[I - K_k H_k] e_k^- - K_k v_k \right] \left[e_k^{-T} \left(I - H_k^{\ T} K_k^{\ T} \right) - v_k^{\ T} K_k^{\ T} \right] \right]$$

$$= E\left[[I - K_k H_k] e_k^- e_k^{-T} \left[I - H_k^{\ T} K_k^{\ T} \right] + K_k v_k v_k^{\ T} K_k^{\ T} \right]$$

$$P_k = [I - K_k H_k] E\left[e_k^- e_k^{-T} \right] \left[I - H_k^{\ T} K_k^{\ T} \right] + K_k E\left[v_k v_k^{\ T} \right] K_k^{\ T} . \qquad [3.19]$$

The a priori error covariance was stated previously as Equation [3.13], $P_k^- = E\left[e_k^- e_k^{-T} \right]$. The measurement noise covariance matrix was also stated as Equation [3.5], $R_k = E\left[v_k v_k^T \right]$. Rewriting P_k gives:

$$P_k = [I - K_k H_k] P_k^- \left[I - H_k^{\ T} K_k^{\ T} \right] + K_k R_k K_k^{\ T} . \qquad [3.20]$$

Equation [3.20] is the general expression for the error covariance and is valid for any gain, K_k. However, the Kalman gain is that particular gain, K_k, which minimizes the individual terms along the major diagonal of P_k. Since the diagonal or trace of P_k represents the mean square error, setting the derivative of the trace of P_k equal to zero will result in the optimal gain K_k.

To determine the derivative of Equation [3.20] two matrix differentiation formulas are used which as taken from Brown and Hwang,1997.

$$\frac{d[trace(AB)]}{dA} = B^T \quad (AB \text{ must be square}), \qquad [3.21]$$

$$\frac{d[trace(ACA^T)]}{dA} = 2AC \quad (\text{if } C \text{ is symmetric}). \qquad [3.22]$$

Expanding Equation [3.20] results in:

$$P_k = P_k^- - P_k^- H_k{}^T K_k{}^T - K_k H_k P_k + K_k H_k P_k^- H_k{}^T K_k{}^T + K_k R_k K_k{}^T \qquad [3.23]$$

Since P_k^- is symmetric and equal to P_k^{-T}, and referring back to Equation [3.13], the trace of $P_k^- H_k{}^T K_k{}^T$ is equal to the trace of the transpose $K_k H_k P_k^-$. Simplifying Equation [3.23], and implementing previous comment, results in:

$$P_k = P_k^- - 2K_k H_k P_k^- + K_k \left[H_k P_k^- H_k{}^T + R_k\right] K_k{}^T. \qquad [3.24]$$

Differentiating the trace of Equation [3.24], considering [3.21] and [3.22] gives:

$$\frac{d}{dK_k}(trace(P_k)) = -2P_k^- H_k{}^T + 2\left[H_k P_k^- H_k{}^T + R_k\right] K_k. \qquad [3.25]$$

Setting this derivative equal to zero and solving for the gain K_k gives:

$$\begin{aligned} K_k &= \frac{P_k^- H_k{}^T}{H_k P_k^- H_k{}^T + R_k} \\ K_k &= P_k^- H_k{}^T \left(H_k P_k^- H_k{}^T + R_k\right)^{-1} \end{aligned} \quad . \qquad [3.26]$$

The optimal gain is given by Equation [3.26]. The covariance matrix that is expressed in terms of the optimal gain may now be determined by substituting Equation [3.26] into Equation [3.24] as:

$$P_k = P_k^- - 2\frac{P_k^- H_k{}^T}{H_k P_k^- H_k{}^T + R_k} H_k P_k^- + P_k^- H_k{}^T \frac{H_k P_k^-}{H_k P_k^- H_k{}^T + R_k},$$

$$= P_k^- - P_k^- H_k{}^T \left(H_k P_k^- H_k{}^T + R_k\right)^{-1} H_k P_k^-,$$

$$= P_k^- - K_k H_k P_k^-,$$

$$P_k = \left(I - K_k H_k\right) P_k^-. \qquad [3.27]$$

Equation [3.27] is the same as [3.8]. Equation [3.27] is only valid for the optimal gain condition.

The next stage in the Kalman Filter is the prediction stage. Now the *a priori* matrix must be determined by using the *a priori* definition as stated in Equation [3.11] which is repeated here for convenience.

$$e_k^- \equiv x_k - \hat{x}_k^-, \text{ (}a\ priori\text{)} \qquad [3.11]$$

Since this is the prediction stage the time step has increased by one and k is replaced by k+1. Substituting Equation [3.1] and [3.9] into Equation [3.11] results in:

$$e_{k+1}^- \equiv \left(\phi_k x_k + G_k u_k + w_k\right) - \left(\phi_k \hat{x}_k + G_k u_k\right),$$

$$\equiv \left(\phi_k x_k + w_k\right) - \phi_k \hat{x}_k,$$

$$\equiv \phi_k \left(x_k - \hat{x}_k\right) + w_k$$

$$e_{k+1}^- \equiv \phi_k e_k + w_k. \qquad [3.28]$$

From Equation [3.13] the *a priori* covariance matrix is $E\left[e_k^- e_k^{-T}\right]$ and substituting Equation [3.28] into [3.13] gives:

$$P_{k+1}^- = E\left[e_{k+1}^- e_{k+1}^{-T}\right] = E\left[\left(\phi_k e_k + w_k\right)\left(\phi_k e_k + w_k\right)^T\right],$$

$$= E\left[\left(\phi_k e_k + w_k\right)\left(e_k{}^T \phi_k{}^T + w_k{}^T\right)\right],$$

$$= E\left[\phi_k e_k e_k{}^T \phi_k{}^T + \phi_k e_k w_k{}^T + w_k e_k{}^T \phi_k{}^T + w_k w_k{}^T\right]. \quad [3.29]$$

Since $Q_k = E\left[w_k w_k{}^T\right]$, and $P_k = E\left[e_k e_k^T\right]$ the *a priori* covariance matrix can be simplified to:

$$P_{k+1}^- = E\left[\phi_k P_k \phi_k{}^T + \phi_k e_k w_k{}^T + w_k e_k{}^T \phi_k{}^T + Q_k\right] \quad [3.30]$$

However, Equation [3.3] states that the process noise is a zero mean process, or $E[w_k] = 0$. Therefore the *a priori* estimate of the covariance matrix may be simplified to:

$$P_{k+1}^- = \phi_k P_k \phi_k{}^T + Q_k. \quad [3.31]$$

All five Kalman filter equations have now been derived above. These five equations form a recursive loop of correction (measurement update) and prediction (time-update). The equations may be summarized as follows:

Prediction Equations:

$$\hat{x}_{k+1}^- = \phi_k \hat{x}_k + G_k u_k, \quad [3.9]$$

$$P_{k+1}^- = \phi_k P_k \phi_k^T + Q_k. \quad [3.31]$$

Corrector Equations:

$$K_k = P_k^- H_k^T \left(H_k P_k^- H_k^T + R_k\right)^{-1}, \quad [3.26]$$

$$\hat{x}_k = \hat{x}_k^- + K_k\left(z_k - H_k \hat{x}_k^-\right), \quad [3.7]$$

$$P_k = (I - K_k H_k) P_k^-. \quad [3.27]$$

One of the main reasons the Kalman Filter is so successful is because it blends information from the measurements and the model. This is done by using the noise covariance matrices, R_k and Q_k, as weighting factors in determining the Kalman Gain. The value of the Kalman Gain depends on R_k and P_k^-, which can be seen from Equation

[3.26]. However, P_k^- depends on Q_k and therefore the calculation of the Kalman Gain depends on both R_k and Q_k. If $Q_k << R_k$ the state predictions will only be based on the model predictions and measurements will be neglected. However, the opposite is also true, if $R_k << Q_k$ then the state predictions will only be based on the measurements and the model predictions will be neglected. Selecting the appropriate values for R_k and Q_k is often done by trial and error. As indicated in the literature, *"there does not exist one single cure for all numerical problems and the choice of* R_k *and* Q_k*, in some instances, may be very arbitrary"* (Brown and Hwang, 1997). Setting the measurement noise may be done by determining the statistical variance of the measurement device. Determining the system noise is much more difficult as it is the uncertainty in the process model. Unfortunately many times the system noise is determined through trial and error as stated above.

3.3 The Extended Kalman Filter (EKF)

When the dynamic equations of a system are non-linear the "Extended Kalman Filter" must be implemented (Grewal and Andrews, 1993). The Extended Kalman Filter works the same way the Kalman Filter works except the Extended Kalman Filter must linearize the state equations around the most recent state estimate for each time step. The time-update equations are linearized around the *a posterior* ($\hat{x}_k$) state estimates and the measurement equations are linearized around the *a priori* ($\hat{x}_k^-$) state estimates. The linearization process is done by using a Taylor series approximation.

The non-linear model may be described by the following state equations:

$$x_{k+1} = f[x_k] + u_k + w_k, \quad \text{(State equation)} \qquad [3.32]$$

$$z_k = h[x_k] + v_k, \quad \text{(Measurement equation)} \qquad [3.33]$$

where w_k and v_k once again represent the process and measurement noise. Notice that the time varying input matrix (G_k) is dropped in Equation [3.32] because the inputs for this model are constant. $f[x_k]$ and $h[x_k]$ may both be nonlinear functions. Since the functions are nonlinear the Taylor series approximation is used (Grewal and Andrews, 1993).

The first step is to calculate the correction (measurement update) equations. The first step is to linearize the measurement Equation [3.33]. The most recent state is *a priori* estimate, $\hat{x}_{k+1}^-$, and so linearization is performed around this state. In general a nonlinear function, $h[x_k]$, may be linearized about $\hat{x}_{k+1}^-$ by a Taylor series approximation as:

$$h[x_k] \approx h\left[\hat{x}_{k+1}^-\right] + J_h\left[x_k - \hat{x}_{k+1}^-\right]. \qquad [3.34]$$

in which J_h is the Jacobian of $h[x_k]$ evaluated at $\hat{x}_{k+1}^-$. The Jacobian operation is given as:

$$J_h = \left.\frac{\partial h[x_k]}{\partial x}\right|_{\hat{x}_{k+1}^-} = \begin{bmatrix} \frac{\partial h[x_{1k}]}{\partial x_1} & \frac{\partial h[x_{1k}]}{\partial x_2} & \cdots \\ \frac{\partial h[x_{2k}]}{\partial x_1} & \frac{\partial h[x_{2k}]}{\partial x_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}_{\hat{x}_{k+1}^-}. \qquad [3.35]$$

The Kalman gain may now be computed using the linearized measurement Jacobian.

$$K_k = P_k^- J_h^T \left(J_h P_k^- J_h^T + R_k\right)^{-1}, \qquad [3.36]$$

This equation is developed similar to the development of Equation [3.26]. The *a posterior* state matrix can be computed by:

$$\hat{x}_k = \hat{x}_k^- + K_k\left(z_k - H_x\hat{x}_k^-\right). \qquad [3.37]$$

The next step in the Extended Kalman Filter is to compute the *a posteriori* covariance estimate as:

$$P_k = (I - K_k J_h)P_k^- . \qquad [3.38]$$

In general, a nonlinear function, $f[x_k]$, may be linearized about $\hat{x}_k$ by a Taylor series approximation as:

$$f[x_k] \approx f[\hat{x}_k] + J_x[x_k - \hat{x}_k], \qquad [3.39]$$

in which J_x is the Jacobian of $f[x_k]$ evaluated at $\hat{x}_k$. The Jacobian operation is given as:

$$J_x = \left.\frac{\partial f[x_k]}{\partial x}\right|_{\hat{x}_k} = \begin{bmatrix} \frac{\partial f[x_{1k}]}{\partial x_1} & \frac{\partial f[x_{1k}]}{\partial x_2} & \cdots \\ \frac{\partial f[x_{2k}]}{\partial x_1} & \frac{\partial f[x_{2k}]}{\partial x_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}_{\hat{x}_k} . \qquad [3.40]$$

The final step for the filter is to estimate predictor equations. The *a priori* estimate of the state vector is given as:

$$\hat{x}_{k+1}^- = f[\hat{x}_k] + u_k . \qquad [3.41]$$

in which $f[\hat{x}_k]$ is the solution of the nonlinear function about the most current state estimate $\hat{x}_k$. The next step is to calculate the covariance matrix of this preliminary state estimate. Once Equation [3.41] is linearized about its current state $\hat{x}_k$, the a priori

covariance matrix can be computed similar to the a prior covariance matrix in the Kalman Filter section. The *a priori* covariance matrix is therefore given as:

$$P_{k+1}^{-} = J_x P_k J_x^T + Q_k . \qquad [3.42]$$

The Extended Kalman Filter uses the Jacobian matrix instead of the linear transition matrix.

Figure 3.2 shows a graphical representation of the Extended Kalman Filter.

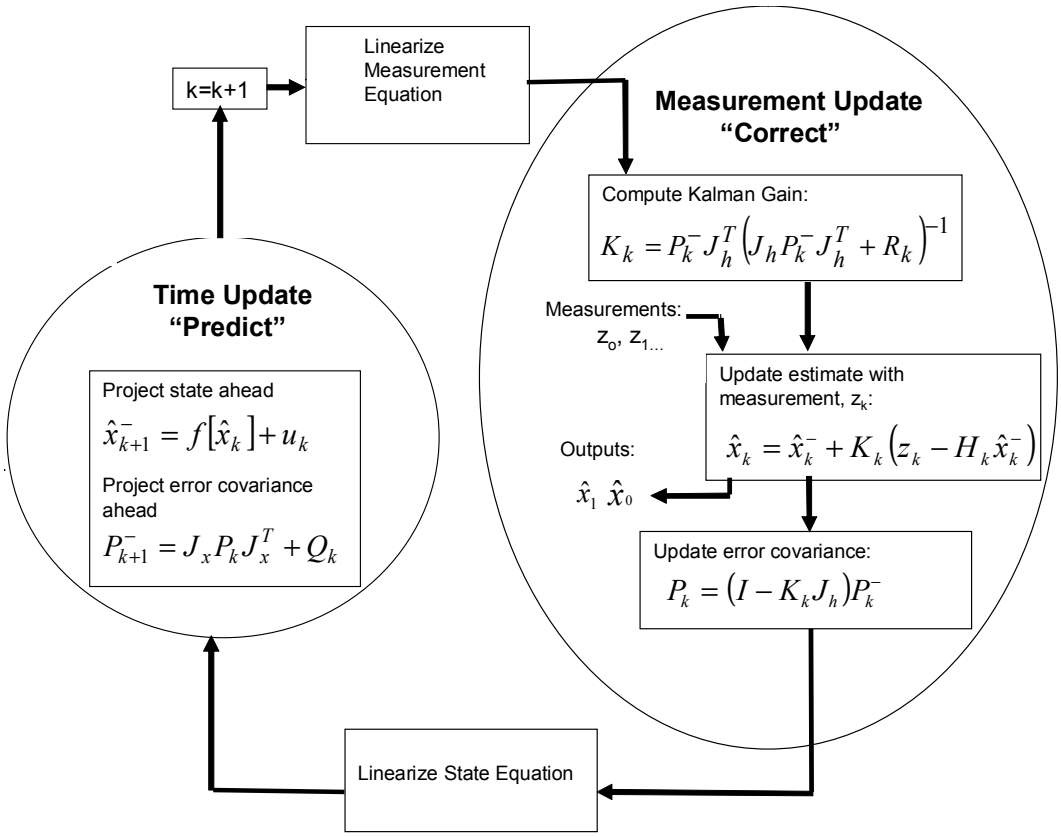

Figure 3.2: Extended Kalman Filter Recursive Loop

3.4 Observability Considerations

"Observability is the issue of whether the state of a dynamic system is uniquely determinable from its inputs and outputs, given a model for the dynamic system" (Grewal and Andrews, 1993). In other words a system model is observable if its states can be determined from the model's inputs and outputs. If a system is unobservable the measurements are not providing enough information to estimate all the state variables of the system. In order to determine if a system model is completely observable, an observability test has been developed (Franklin et.al. 1990).

Suppose that an N dimension system is represented by:

$$X(k+1) = AX(k) \qquad [3.43]$$

$$Z(k) = HX(k) \qquad [3.44]$$

The system (A, H) is observable if for any $X(0)$, there is a finite N such that $X(0)$ can be computed from observation of $Z(0), Z(1), Z(2)....Z(N-1)$ (Franklin et al., 1990). The outputs, in matrix form, from $k = 0$ to $k = N-1$ are:

$$\begin{bmatrix} Z(0) \\ Z(1) \\ Z(2) \\ ..... \\ Z(N-1)) \end{bmatrix} = \begin{bmatrix} H \\ HA \\ HA^2 \\ ..... \\ HA^{N-1} \end{bmatrix} X(0) = O_N \times X(0), \qquad [3.45]$$

where:

$$O_N = \begin{bmatrix} H \\ HA \\ HA^2 \\ ..... \\ HA^{N-1} \end{bmatrix} \qquad [3.46]$$

The system model is observable when the rank of the observability matrix (O_N) is equal to the system length. The system length of the matrix is the number of states being estimated.

Chapter 4 discusses the system model and the implementation of the Extended Kalman Filter for leak detection in an underground water pipeline.

Chapter 4: Distribution Line Model and Applying the EKF for Leak Detection

This chapter discusses the water distribution model that was developed. The water distribution line was modeled to represent any pipeline that has either a large reservoir or a pump at the beginning of the pipeline and a valve at the end of the pipe. The equations developed for the water distribution model are transformed into state space form and the transition matrix is determined in order to apply the Extended Kalman Filter (EKF). The observability issue is also discussed and the necessary changes are performed to ensure the model is observable. The number of significant figures for this thesis was chosen to be five. A discussion of significant figures is included in Chapter 6 in which a sensitivity study is considered. In actuality, the number of significant figures which should be used would be dependent on the significant figures (resolution) associated with the sensors and the inputs.

4.1 Melfort Pipeline Characteristics and Notation

The model developed below roughly represents the raw water pipeline for Melfort, Saskatchewan with a few notable exceptions. In order for consistency throughout this thesis, all of the numerical inputs and outputs are recorded to five significant figures. This number of significant figures is chosen in order to provide enough information to compare different simulations. The model consists of a pipeline with a length of 52382m with an upstream reservoir of 139.00m of head and a downstream reservoir of 135.50m of head. The actual upstream head in the Melfort line is produced by a constant pressure pump. There is an elevation difference of 107.00m between the upstream and downstream reservoir, which is represented in the two head measurements above. The pipeline has an internal diameter of 0.50800 m and has a valve located just before the downstream reservoir. The model of the Melfort pipeline is simplified and does not include any elbow, bends etc. in order to determine if the EKF technique is a viable option for leak detection in underground water distribution lines. A graphical representation is shown in Figure 4.1.

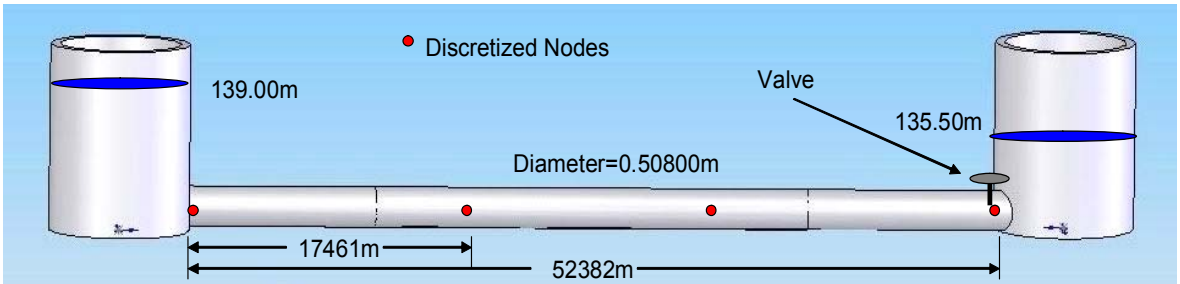

Figure 4.1: Melfort Raw Water Line

The modeled pipeline is discretized into three segments. Each segment is 17461m and there are four nodes along the pipeline. One node is located at the supply reservoir, two interior nodes located at 17461m and 34921m, respectively, and one node at the valve and downstream reservoir. The model assumes that leakage occurs at both interior nodes (17461m and 34921m). The two "fictitious" leakage locations are known; however their magnitudes are the unknown states. The model assumes no leakage occurs at the upstream or downstream nodes. The actual leakage location is determined from an interpolation of the two "fictitious leakages" at the interior nodes. A completed derivation of the leakage magnitude and location is explained in Section 5.1

The equations to model the pipeline are taken from Lesyshen, 2005. In order to be consistent with the notation presented in Lesyshen the following notation is adopted for this thesis.

The nodes are labelled 1 to 4 with node 1 occurring at the upstream reservoir, node 2 and 3 are the interior nodes and node 4 is located at the valve and downstream reservoir. The notation for head is $H_{2,k}$, which represents the head at node 2 at the current time step, k. The notation for flow is $Q_{21,k}$, which represents the flow at the beginning of the second segment at the current time step, whereas the notation, $Q_{12,k-1}$, represents the flow at the end of the first segment at the previous time step and $Q_{L2,k}$ is the leakage at node 2. The first two-digit subscript numbers represent the spatial position and the second subscript represents the temporal position. This notation is explained in Figure 4.2.

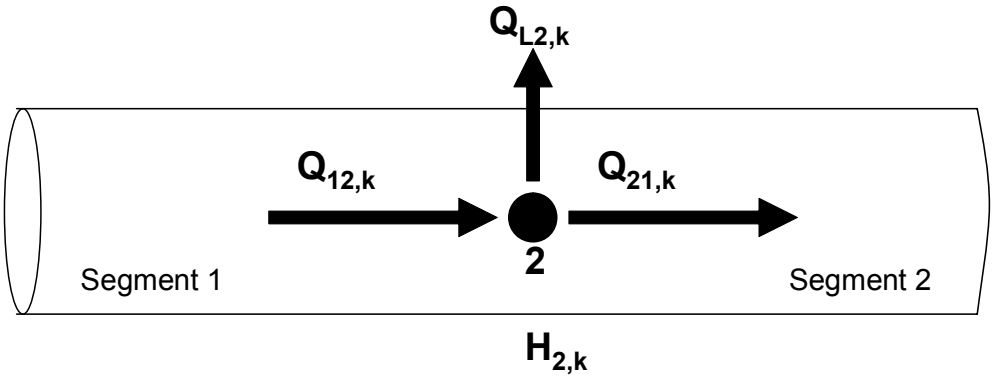

Figure 4.2: Notation for Flow and Pressure at step k

4.2 Boundary Condition

In order to solve the transient pipe equations developed in Chapter 2, special boundary conditions are required. For this thesis, the boundary condition for the upstream reservoir is obtained from the energy equation. The boundary condition for the downstream reservoir is obtained from the valve equation. Both boundary conditions are explained in detail in the following section and the equations necessary for calculating the pressure and flow throughout the pipeline are presented.

4.2.1 The Upstream Reservoir and/or Pump (Node 1)

There is a pump at the beginning of the Melfort raw water distribution line. However, since the pump was "online", determining the pump characteristic curve was not possible. Therefore, the upstream boundary condition was modeled as a constant-level upstream reservoir (see Appendix A for derivation and justification). The water level in the reservoir was assumed to remain constant because the upstream reservoir was assumed to have a large cross-sectional area. Therefore transient conditions do not affect the level of the water in the reservoir. The energy equation gives a relationship between the head and the flow at the upstream reservoir. To obtain the boundary condition the energy equation and the C^- Equation [2.56] must be solved. These equations are stated below:

$$H_{1,k} = H_{R1} - (1+k)\frac{{Q_{11,k}}^2}{2gA^2}, \qquad [4.1]$$

$$Q_{11,k} = \frac{\left(H_{1,k} - H_{2,k-1} + BQ_{12,k-1} - RQ_{12,k-1}\left|Q_{12,k-1}\right|\right)}{B} \qquad [4.2]$$

H_{R1} represents the head at the supply reservoir and k is the entrance loss coefficient. The entrance loss is chosen to be 0.5. This value is taken from (Watters, 1984) and represents the constant for minor losses at a sharp-edged pipe entrance. The pipe wave velocity constant, $B = a/gA$, and the pipe friction constant, $R = f\Delta x/\left(2gDA^2\right)$, were derived in Chapter 2. When Equation [4.1] is substituted into Equation [4.2] for $H_{1,k}$, and neglecting the negative sign from the radical term, Equation [4.3] is obtained.

$$Q_{11,k} = \frac{-B + \sqrt{B^2 - 4\left(\frac{1+K}{2gA^2}\right)\left(H_{2,k-1} - BQ_{12,k-1} + RQ_{12,k-1}\left|Q_{12,k-1}\right| - H_{R1}\right)}}{1 + K/gA^2} \qquad [4.3]$$

Now $H_{1,k}$ may be solved using Equation [4.1].

4.2.2 The Downstream Reservoir and Valve (Node 4)

The boundary condition for the downstream reservoir-valve is modeled using the valve equation. The development of this boundary condition is taken from Chaudhry, 1987 and Wylie/Streeter, 1983. Since the datum for elevation is taken at the downstream valve, the orifice equation for steady state flow through a valve is:

$$Q_{32,0} = (C_d A)_0 \sqrt{2g(H_{4,0} - H_{R2})} \qquad [4.4]$$

where $Q_{32,0}$ is the steady state flow through the valve, $(H_{4,0} - H_{R2})$ is the steady state head loss across the valve, H_{R2} is the downstream reservoir head, C_d is the steady state discharge coefficient and A is the valve opening area. The Reynolds number is around

100,000 therefore it is assumed that the flow is turbulent through the valve. For a general opening the flow may be described as:

$$Q_{32,k} = (C_d A)\sqrt{2g(H_{4,k} - H_{R2})} \qquad [4.5]$$

where $(H_{4,k} - H_{R2})$ is the instantaneous drop across the valve.

Dividing Equation [4.4] by [4.5] results in:

$$Q_{32,k} = \frac{Q_{32,0}}{\sqrt{(H_{4,0} - H_{R2})}}\tau\sqrt{(H_{4,k} - H_{R2})} \qquad [4.6]$$

where τ is the dimensionless valve opening given as:

$$\tau = \frac{(C_d A)}{(C_d A)_0}. \qquad [4.7]$$

For a fully open valve, as in this case, $\tau = 1$.

Using the C^+ Equation [2.55], (from the method of characteristics) to obtain $H_{4,k}$ and substituting $H_{4,k}$ into Equation [4.6] results in:

$$Q_{32,k}{}^2 + C_v B Q_{32,k} - C_v\left(H_{3,k-1} + BQ_{31,k-1} - RQ_{31,k-1}\left|Q_{31,k-1}\right| - H_{R2}\right) = 0 \qquad [4.8]$$

where C_v is a valve opening coefficient given by:

$$C_v = \frac{(Q_{32,0}\tau)^2}{(H_{4,0} - H_{R2})}. \qquad [4.9]$$

Using the quadratic equation to solve Equation [4.8] for $Q_{32,k}$ results in:

$$Q_{32,k} = \frac{-C_v B + \sqrt{(C_v B)^2 + 4C_v\left(H_{3,k-1} + Q_{31,k-1}\left(B - R\left|Q_{31,k-1}\right|\right) - H_{R2}\right)}}{2}. \qquad [4.10]$$

Therefore, the hydraulic head at node 4, $H_{4,k}$, may now be calculated from the C^+ equation from the Method of Characteristics as:

$$H_{4,k} = H_{3,k-1} + BQ_{31,k-1} - RQ_{31,k-1}\left|Q_{31,k-1}\right| - BQ_{32,k}\,. \qquad [4.11]$$

4.3 Interior Nodes (Nodes 2 and 3)

The model of the pipeline is developed with leakage occurring at both interior nodes, node 2 and node 3. Flow can be modeled at the interior nodes using the C^+ and C^- equations for incoming and outgoing flow. The leakage at each interior node is modeled using an orifice equation. The heads on either side of the interior nodes are assumed to be equivalent. Therefore for node 2:

$$Q_{12,k} = \left(-H_{2,k} + H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right|\right)\Big/B\,, \qquad [4.12]$$

$$-Q_{21,k} = \left(-H_{2,k} + H_{3,k-1} - BQ_{22,k-1} + RQ_{22,k-1}\left|Q_{22,k-1}\right|\right)\Big/B\,. \qquad [4.13]$$

Ground water pressure is considered to be negligible, hence:

$$-Q_{L2} = -\lambda_2\sqrt{H_{2,k}}\,, \qquad [4.14]$$

where λ_2 is the unknown leakage area constant for node 2. $-Q_{21,k}$ and $-Q_{L2}$ are negative to represent flow leaving the node 2. In order to solve for $H_{2,k}$, a summation of the flow for node 2 is shown as:

$$\sum Q = 0 = \frac{-2H_{2,k} + H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| + H_{3,k-1} - BQ_{22,k-1} + RQ_{22,k-1}\left|Q_{22,k-1}\right|}{B} - \lambda_2\sqrt{H_{2,k}} \quad [4.15]$$

Squaring Equation [4.15] and using the quadratic equation yields $H_{2,k}$ as:

$$H_{2,k} = \frac{1}{8}\left(B^2{\lambda_2}^2 \pm \left(\sqrt{B^4{\lambda_2}^4 - 16\left(H^2{}_{1,k-1} + B^2Q^2{}_{11,k-1} - R^2Q^2{}_{11,k-1}\left|Q^2{}_{11,k-1}\right| + H^2{}_{3,k-1} - B^2Q^2{}_{22,k-1} + R^2Q^2{}_{22,k-1}\left|Q^2{}_{22,k-1}\right|\right)}\right)\right) \quad [4.16]$$

Similarly flow and pressure at node 3 are determined as:

$$Q_{22,k} = \left(-H_{3,k} + H_{2,k-1} + BQ_{21,k-1} - RQ_{21,k-1}\left|Q_{21,k-1}\right|\right)\Big/B, \quad [4.17]$$

$$-Q_{31,k} = \left(-H_{3,k} + H_{4,k-1} - BQ_{32,k-1} + RQ_{32,k-1}\left|Q_{32,k-1}\right|\right)\Big/B, \quad [4.18]$$

$$-Q_{L3} = -\lambda_3\sqrt{H_{3,k}}, \quad [4.19]$$

$$H_{3,k} = \frac{1}{8}\left(B^2{\lambda_3}^2 \pm \left(\sqrt{\left(B^4{\lambda_3}^4 - 16\left(H^2{}_{2,k-1} + B^2Q^2{}_{22,k-1} - R^2Q^2{}_{22,k-1}\left|Q^2{}_{22,k-1}\right| + H^2{}_{4,k-1} - B^2Q^2{}_{32,k-1} + R^2Q^2{}_{32,k-1}\left|Q^2{}_{32,k-1}\right|\right)\right)}\right)\right) \quad [4.20]$$

4.4 A Simplified Prediction Model and State Space Representation

In order to implement the Extended Kalman Filter the equations must be in state-space representation. Since the leakage is expressed by the square root valve equation it makes Equation [4.12], [4.13] and [4.16] very long and complex when transformed into state space as can be seen below:

$$H_{2,k} = \frac{1}{8}\left(B^2{\lambda_{2,k-1}}^2 \pm \left(\sqrt{B^4{\lambda_{2,k-1}}^4 - 16\begin{pmatrix} H^2{}_{1,k-1} + B^2Q^2{}_{11,k-1} - R^2Q^2{}_{11,k-1}\left|Q^2{}_{11,k-1}\right| \\ + H^2{}_{3,k-1} - B^2Q^2{}_{22,k-1} + R^2Q^2{}_{22,k-1}\left|Q^2{}_{22,k-1}\right| \end{pmatrix}}\right)\right) \quad [4.21]$$

$$Q_{12,k} = \left(\begin{array}{l} -\frac{1}{8}\left(B^2 {\lambda_{2,k-1}}^2 \pm \left(\sqrt{B^4 {\lambda_{2,k-1}}^4 - 16\left(\begin{array}{l} {H^2}_{1,k-1} + B^2 {Q^2}_{11,k-1} - R^2 {Q^2}_{11,k-1}\left|{Q^2}_{11,k-1}\right| \\ + {H^2}_{3,k-1} - B^2 {Q^2}_{22,k-1} + R^2 {Q^2}_{22,k-1}\left|{Q^2}_{22,k-1}\right| \end{array} \right)} \right) \right) \\ + H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| \end{array} \right) \Big/ B \quad [4.22]$$

$$Q_{21,k} = \left(\begin{array}{l} \frac{1}{8}\left(B^2 {\lambda_{2,k-1}}^2 \pm \left(\sqrt{B^4 {\lambda_{2,k-1}}^4 - 16\left(\begin{array}{l} {H^2}_{1,k-1} + B^2 {Q^2}_{11,k-1} - R^2 {Q^2}_{11,k-1}\left|{Q^2}_{11,k-1}\right| \\ + {H^2}_{3,k-1} - B^2 {Q^2}_{22,k-1} + R^2 {Q^2}_{22,k-1}\left|{Q^2}_{22,k-1}\right| \end{array} \right)} \right) \right) \\ - H_{3,k-1} + BQ_{22,k-1} - RQ_{22,k-1}\left|Q_{22,k-1}\right| \end{array} \right) \Big/ B \quad [4.23]$$

Programming these equations is prone to mistakes and implementing the Extended Kalman Filter requires partial derivatives for linearization of each equation which adds even more complexity and length. Equations [4.21], [4.22] and [4.23] are based on Equation [4.14], ($-Q_{L2} = -\lambda_2\sqrt{H_{2,k}}$) in which the leakage flow is replaced by the product of the area constant and pressure drop. Because of the complexity of Equation [4.21] to [4.23], it was decided not make the leak substitution as was done above. Instead the leak is simply determined from the previous leak. Performing this simplification replaces Equation [4.14] with:

$$-Q_{L2,k} = -Q_{L2,k-1} \quad [4.24]$$

Leakage is now a constant flow rate and is to be estimated from the previous time step. Now performing the summation of the flows on Node 2 results in:

$$\sum Q = 0 = \frac{-2H_{2,k} + H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| + H_{3,k-1} - BQ_{22,k-1} + RQ_{22,k-1}\left|Q_{22,k-1}\right|}{B} - Q_{L2,k-1} \quad [4.25]$$

Rearranging Equation [4.25] results in:

$$H_{2,k} = \frac{1}{2}\left(H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| + H_{3,k-1} - BQ_{22,k-1} + RQ_{22,k-1}\left|Q_{22,k-1}\right| - BQ_{L2,k-1}\right) \quad [4.26]$$

Substituting Equation [4.26] into Equation [4.12] results in:

$$Q_{12,k} = \frac{1}{2}\left(H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| - H_{3,k-1} + BQ_{22,k-1} - RQ_{22,k-1}\left|Q_{22,k-1}\right| + BQ_{L2,k-1}\right)/B \quad [4.27]$$

Similarly, substituting Equation [4.26] into Equation [4.13] results in:

$$Q_{21,k} = \frac{1}{2}\left(H_{1,k-1} + BQ_{11,k-1} - RQ_{11,k-1}\left|Q_{11,k-1}\right| - H_{3,k-1} + BQ_{22,k-1} - RQ_{22,k-1}\left|Q_{22,k-1}\right| - BQ_{L2,k-1}\right)/B \quad [4.28]$$

Equations [4.26-4.28] represent the flow and head at node 2 in state space representation. These equations are much simpler to differentiate because of their reduced length and complexity.

The state vector may be described as:

$$x_j = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \\ x_{12} \end{bmatrix} = \begin{bmatrix} H_{1,k} \\ H_{2,k} \\ H_{3,k} \\ H_{4,k} \\ Q_{11,k} \\ Q_{12,k} \\ Q_{21,k} \\ Q_{22,k} \\ Q_{31,k} \\ Q_{32,k} \\ Q_{L2,k} \\ Q_{L3,k} \end{bmatrix} \quad [4.29]$$

The inputs into the model are the upstream and downstream head and the valve coefficient.

$$u_j = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} H_{R1} \\ H_{R2} \\ C_v \end{bmatrix} \quad [4.30]$$

A transient may be introduced into the pipeline through either input. Changes in the reservoir heads may represent waves on the reservoir or the fluctuating pressure of a pump.

In order to implement the Extended Kalman Filter the equations must be in state space form which are a function of prior states, present inputs, and process noise. In order to put these equations into state space form, one state equation has to be substituted into another so that the current states are only a function of prior states, inputs or constants. Transforming the model equations into state space gives:

$$x_{1,k} = u_{1,k} - \frac{1+K}{2gA^2}\left(\frac{-B+\sqrt{B^2-4\left(\frac{1+K}{2gA^2}\right)\left(x_{2,k-1}-Bx_{6,k-1}+Rx_{6,k-1}\left|x_{6,k-1}\right|-u_{1,k}\right)}}{\frac{1+K}{2gA^2}}\right)^2 + w_{1,k-1}, \quad [4.31]$$

$$x_{2,k} = \frac{1}{2}\left(x_{1,k-1}+Bx_{5,k-1}-Rx_{5,k-1}\left|x_{5,k-1}\right|+x_{3,k-1}-Bx_{8,k-1}+Rx_{8,k-1}\left|x_{8,k-1}\right|-Bx_{11,k-1}\right)+w_{2,k-1}, \quad [4.32]$$

$$x_{3,k} = \frac{1}{2}\left(x_{2,k-1}+Bx_{7,k-1}-Rx_{7,k-1}\left|x_{7,k-1}\right|+x_{4,k-1}-Bx_{10,k-1}+Rx_{10,k-1}\left|x_{10,k-1}\right|-Bx_{12,k-1}\right)+w_{3,k-1}, \quad [4.33]$$

$$x_{4,k} = x_{3,k-1}+Bx_{9,k-1}-Rx_{9,k-1}\left|x_{9,k-1}\right|-B\frac{u_{3,k}B+\sqrt{\left(u_{3,k}B\right)^2+4u_{3,k}\left(x_{3,k-1}+x_{9,k-1}\left(B-R\left|x_{9,k-1}\right|\right)-u_{2,k}\right)}}{2}+w_{4,k-1}, \quad [4.34]$$

$$x_{5,k} = \frac{-B - \sqrt{B^2 - 4\left(\frac{1+K}{2gA^2}\right)\left(x_{2,k-1} - Bx_{6,k-1} + Rx_{6,k-1}\left|x_{6,k-1}\right| - u_{1,k}\right)}}{1 + K \diagup gA^2} + w_{5,k-1}, \quad [4.35]$$

$$x_{6,k} = \frac{1}{2}\left(x_{1,k-1} + Bx_{5,k-1} - Rx_{5,k-1}\left|x_{5,k-1}\right| - x_{3,k-1} + Bx_{8,k-1} - Rx_{8,k-1}\left|x_{8,k-1}\right| + Bx_{11,k-1}\right)/B + w_{6,k-1},$$

[4.36]

$$x_{7,k} = \frac{1}{2}\left(x_{1,k-1} + Bx_{5,k-1} - Rx_{5,k-1}\left|x_{5,k-1}\right| - x_{3,k-1} + Bx_{8,k-1} - Rx_{8,k-1}\left|x_{8,k-1}\right| - Bx_{11,k-1}\right)/B + w_{7,k-1},$$

[4.37]

$$x_{8,k} = \frac{1}{2}\left(x_{2,k-1} + Bx_{7,k-1} - Rx_{7,k-1}\left|x_{7,k-1}\right| - x_{4,k-1} + Bx_{10,k-1} - Rx_{10,k-1}\left|x_{10,k-1}\right| + Bx_{12,k-1}\right)/B + w_{8,k-1},$$

[4.38]

$$x_{9,k} = \frac{1}{2}\left(x_{2,k-1} + Bx_{7,k-1} - Rx_{7,k-1}\left|x_{7,k-1}\right| - x_{4,k-1} + Bx_{10,k-1} - Rx_{10,k-1}\left|x_{10,k-1}\right| - Bx_{12,k-1}\right)/B + w_{9,k-1},$$

[4.39]

$$x_{10,k} = \frac{u_{3,k}B + \sqrt{\left(u_{3,k}B\right)^2 + 4u_{3,k}\left(x_{3,k-1} + x_{9,k-1}\left(B - R\left|x_{9,k-1}\right|\right) - u_{2,k}\right)}}{2} + w_{10,k-1}, \quad [4.40]$$

$$x_{11,k} = x_{11,k-1} + w_{11,k-1}, \quad [4.41]$$

$$x_{12,k} = x_{12,k-1} + w_{12,k-1}. \quad [4.42]$$

The output equation is given below:

$$z_k = \begin{bmatrix} x_{1,k} \\ x_{4,k} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} x_k + v_k, \quad [4.36]$$

where v_k represents measurement noise. The measurements taken from the system and the outputs of the model are the upstream and downstream pipeline heads. The EKF requires the state and output equations to be represented in state space representation and therefore these equations above are in the correct form to implement the EKF. However, the Jacobian matrix still needs to be determined and is derived in Section 4.5.

4.5 Jacobian Matrix Equations

Since the model equations are non-linear, the Extended Kalman Filter is implemented (Welsch and Bishop, 2001). This requires the equations to be linearized around the current estimate and can be determined by computing the Jacobian matrix. The Jacobian, or partial derivatives of the state equations, is what "drives" the filtering process. The Jacobian matrix is the rate of change within the state vector, with respect to each state, and is needed for determining the *a priori* covariance matrix given by Equation [3.35]. The Jacobian is given by Equation [3.33] as:

$$J_x = \left.\frac{\partial f[x_k]}{\partial x}\right|_{\hat{x}_k} = \begin{bmatrix} \frac{\partial f[x_{1k}]}{\partial x_1} & \frac{\partial f[x_{1k}]}{\partial x_2} & \cdots \\ \frac{\partial f[x_{2k}]}{\partial x_1} & \frac{\partial f[x_{2k}]}{\partial x_2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}_{\hat{x}_k} \quad [3.33]$$

It is crucial to explain the notation adopted from Lesyshen, 2005, in order to understand the development of the Jacobian matrix. $J_{2,1}$ is the partial derivative of state equation x_2 with respect to x_1. Therefore, the partial derivative is taken of the first subscript with respect to the second subscript. A list of the non-zero elements of J_x is given below. The state space representation, x_k and u_k, are dropped and the head and flow variables are used in order to more easily understand these equations. The entire equation set, collectively defined as Equation [4.37], is given as:

$$J_{1,2} = \frac{-B + \sqrt{B^2 - 3\left(H_{2,k-1} - BQ_{12,k-1} + RQ_{12,k-1}^2 - H_{R1}\right)/\left(gA^2\right)}}{\sqrt{B^2 - 3\left(H_{2,k-1} - BQ_{12,k-1} + RQ_{12,k-1}^2 - H_{R1}\right)/\left(gA^2\right)}},$$

$$J_{1,6} = \frac{\left(-B + \sqrt{B^2 - 3\left(H_{2,k-1} - BQ_{12,k-1} + RQ_{12,k-1}^2 - H_{R1}\right)/\left(gA^2\right)}\right)\left(-B + 2RQ_{12,k-1}\right)}{\sqrt{B^2 - 3\left(H_{2,k-1} - BQ_{12,k-1} + RQ_{12,k-1}^2 - H_{R1}\right)/\left(gA^2\right)}},$$

$$J_{2,1} = J_{2,3} = J_{3,2} = J_{3,4} = J_{6,11} = -J_{7,11} = J_{8,12} = -J_{9,12} = \frac{1}{2},$$

$$J_{2,5} = J_{3,7} = \frac{1}{2}\left(B - 2RQ_{(i)1,k-1}\right) \ where i = 1:2,$$

$$J_{i,8} = J_{i,10} = \frac{1}{2}\left(-B + 2RQ_{(i)2,k-1}\right) \ where i = 2:3,$$

$$J_{2,11} = J_{3,12} = -\frac{1}{2}B,$$

$$J_{4,3} = 1 - \frac{BCv}{\sqrt{(BCv)^2 + 4Cv\left(H_{3,k-1} + BQ_{31,k-1} - RQ_{31,k-1}^2 - H_{R2}\right)}},$$

$$J_{4,9} = B - 2RQ_{31,k-1} - \frac{1}{4}\frac{4Cv\left(B - 2RQ_{31,k-1}\right)}{\sqrt{(BCv)^2 + 4Cv\left(H_{3,k-1} + BQ_{31,k-1} - RQ_{31,k-1}^2 - H_{R2}\right)}},$$

$$J_{5,2} = \frac{-1}{\sqrt{B^2 - 3\left(H_{2,k-1} - BQ_{12,k-1} + RQ_{12,k-1}^2 - H_{R1}\right)/\left(gA^2\right)}},$$

$$J_{5,6} = \frac{-B + 2RQ_{12,k-1}}{\sqrt{B^2 - 3\left(H_{2,k-1} - BQ_{12,k-1} + RQ_{12,k-1}^2 - H_{R1}\right)/\left(gA^2\right)}},$$

$$J_{6,1} = -J_{6,3} = J_{7,1} = -J_{7,3} = J_{8,1} = -J_{8,3} = J_{9,1} = -J_{9,3} = \frac{1}{2B},$$

$$J_{6,5} = J_{7,5} = \frac{1}{2B}\left(B - 2RQ_{11,k-1}\right),$$

$$J_{6,8} = J_{7,8} = \frac{1}{2B}\left(B - 2RQ_{22,k-1}\right),$$

$$J_{8,7} = J_{9,7} = \frac{1}{2B}\left(B - 2RQ_{21,k-1}\right),$$

$$J_{8,10} = J_{9,10} = \frac{1}{2B}\left(B - 2RQ_{32,k-1}\right),$$

$$J_{10,3} = \frac{Cv}{\sqrt{(BCv)^2 + 4Cv\left(H_{3,k-1} + BQ_{31,k-1} - RQ_{31,k-1}^2 - H_{R2}\right)}},$$

$$J_{10,9} = \frac{1}{4}\frac{4Cv\left(B - 2RQ_{31,k-1}\right)}{\sqrt{(BCv)^2 + 4Cv\left(H_{3,k-1} + BQ_{31,k-1} - RQ_{31,k-1}^2 - H_{R2}\right)}},$$

$$J_{11,11} = J_{12,12} = 1. \qquad [4.37]$$

The pressure heads at the ends of the pipeline are the measurements taken from the system. Therefore the measurement Jacobian J_h is given as:

$$J_h = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}'. \qquad [4.38]$$

4.6 Issues with Observability

As discussed in Section 3.4 a system is unobservable if one or more of the state variables cannot be uniquely determined from the measurements of the system. In many other studies dealing with the Extended Kalman Filter the issue of observability is overlooked. "A system is known to be completely observable if the rank of the observability matrix is equal to the number of state variable", (Ogata, 2002). The rank is the number of linearly independent rows or columns within a matrix. In this work it was believed that there may be an issue with observability since initially there was limited success in getting the EKF to converge to the correct states. However, other studies have been able to get the EKF to converge on the correct answer through extensive tuning (adjusting the error covariance matrices, Q_k and R_k) even if the system was unobservable (Lesyshen, 2005).

The observability test (as discussed in Section 3.4) was performed on the model to determine if the model was observable. The number of states determined by the model was 12 and therefore the observability matrix had to have a rank of 12 in order to ensure the model was observable. With only the upstream and downstream head taken as measurements from the system the observability matrix had a rank of 10. This shows that the current model was unobservable. It was necessary to investigate if the model could be made observable by using more measurements from the system. Indeed, it was determined that if four head measurements (one measurement from each of the four nodes) were inputted into the model then the rank of the observability matrix was 12. Therefore the model could be made observable and the output equation was changed to:

$$z_k = \begin{bmatrix} x_{1,k} \\ x_{2,k} \\ x_{3,k} \\ x_{4,k} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} x_k + v_k \qquad [4.37]$$

Also the measurement Jacobian J_h was changed to:

$$J_h = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}'. \qquad [4.38]$$

Once the model was observable it was determined that the model now produced repeatable results because the model could be computed, with the same inputs, and the same results were calculated. The model could now be used for different systems with different parameters without extensive tuning.

4.7 The Addition of Noise to the Simulation

The plant and the measurements were corrupted with zero mean noise in the simulation model. Noise was added to the plant by adding noise to the upstream head input. This noise was intended to simulate small disturbances at the upstream head, which travel down the pipeline. In actuality, these disturbances would be produced by the pump dynamics or waves on a reservoir. The measurement noise was added to simulate the uncertainty associated with the head measurements.

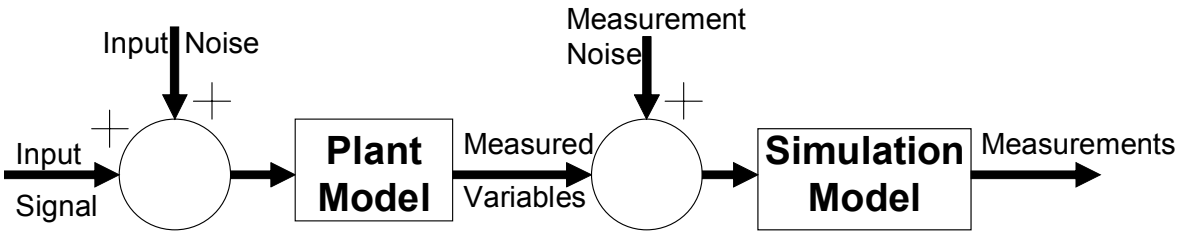

Figure 4.3: Noise Added to Plant and Simulation Model

4.8 Initial Conditions and Covariance

Initial conditions are required for the Extended Kalman filter to begin the estimation process. These initial conditions include the initial state estimates, $\hat{x}_0^-$, the *a priori* error covariance P_0^-, and the process and measurements noise, Q_k and R_k. The initial state estimates were determined from a steady state analysis, assuming no leakage in the

pipeline. The initial value for P_0^- must also be determined. If it was known that the initial state estimates were correct then P_0^- would be set to 0. However since the initial state estimates are determined by assuming zero leakage, it is known that the initial states are not correct. Therefore, P_0^- must be chosen to be something larger than 0. One could choose almost any P_0^- and the filter would *eventually* converge if modeling uncertainties are not too large and reflected by the magnitude of the noise covariance matrices (Welsch and Bishop, 2001). The initial error covariance, P_0^-, must be a square matrix of the same length as the number of estimated states. Therefore, for this thesis, $P_0^- = IC$, where I is 12x12 identity matrix and C is some large constant. As stated above, the choice of C is not crucial; the rate of convergence will be largely affected. The filter adjusts the error covariance matrix for every iteration and therefore the final value will be similar. For this thesis the value of 10^5 was chosen for C.

The initial conditions for the process and measurement noise, Q_k and R_k, are required for the filter to begin the estimation process. The measurement noise is set to resemble the amount of uncertainty in the sensor (i.e the precision of the sensor). If R_k was set to 0 then it would be assumed that the sensor had no noise and the state estimates would be based completely on the measurements. Setting the process noise, Q_k, is generally much more difficult because it is the difference between what the model is predicting and what is occurring in the real system. However, usually it is not known what is occurring in the real system and that is why the model was produced. In many situations, one or both of these covariance matrices are unknown; however acceptable results may be found through trial and error (Brown and Hwang, 1997).

The measurement error covariance, R_k, for this thesis was set by taking into consideration the noise added to the measurements. This is similar to setting the measurement error covariance to the resolution of the sensors. As stated above the process error covariance may be set by trial and error. A starting point is to set the measurement and process noise to equal values, in order for the filter to weigh the

measurements and the model equally. The process noise was then adjusted until acceptable results were achieved. For this work the error covariance matrices were set to:

$$R_k = \begin{bmatrix} 0.02 & 0 & 0 & 0 \\ 0 & 0.02 & 0 & 0 \\ 0 & 0 & 0.02 & 0 \\ 0 & 0 & 0 & 0.02 \end{bmatrix}, \quad Q_k = \begin{bmatrix} 1e^{-5} I_4 & 0 & 0 \\ 0 & 5e^{-6} I_6 & 0 \\ 0 & 0 & 5e^{-7} I_2 \end{bmatrix}, \qquad [4.39]$$

where I_4 is a 4x4 identity matrix, I_6 is a 6x6 identity matrix, and I_2 is a 2x2 identity matrix. The three different identity matrices are the size of the head, flow and leakage flow states.

Chapter 5 will show how the model is able to predict the leak location and magnitude from the two fictitious leakage states that are estimated from the Extended Kalman Filter.

Chapter 5: Fictitious Leaks Approach and Simulated Results

The Extended Kalman Filter (EKF) technique is used for estimating the magnitude of two "fictitious" leaks along the pipeline at "known" or "specified" locations from the four pressure measurements. The four pressure measurements must be equally spaced and occur at the spatial discretization. The reason the leaks are termed as "fictitious" is because of the Method of Characteristics technique; leaks can only be modeled at nodes along the pipeline. The "fictitious" leak positions do not necessarily coincide with the actual leakage point along the pipeline. The fictitious leaks can, however, be used by the proposed method in this Chapter for determining leaks between the nodes. The sum of the two fictitious leaks equals the actual leakage in the pipeline. If there is no leakage in the pipeline then the two fictitious leaks will oscillate around zero and their sum will be zero. The location of the leak is determined from an interpolation based on the magnitude and location of the fictitious leaks. A complete explanation and derivation is discussed in Section 5.1.

5.1 Fictitious Leaks Approach

The concept and approach of the two fictitious leaks is taken directly from Lesyshen, 2005. The idea of the fictitious estimates comes from the concept of equivalent systems. Figure 5.1 displays the head measurements for two different pipelines. Both pipelines have the same overall leak magnitude; however, one pipeline has one leak and the other pipeline has two leaks. Considering only a steady state analysis, these two pipelines are equivalent with respect to their boundaries.

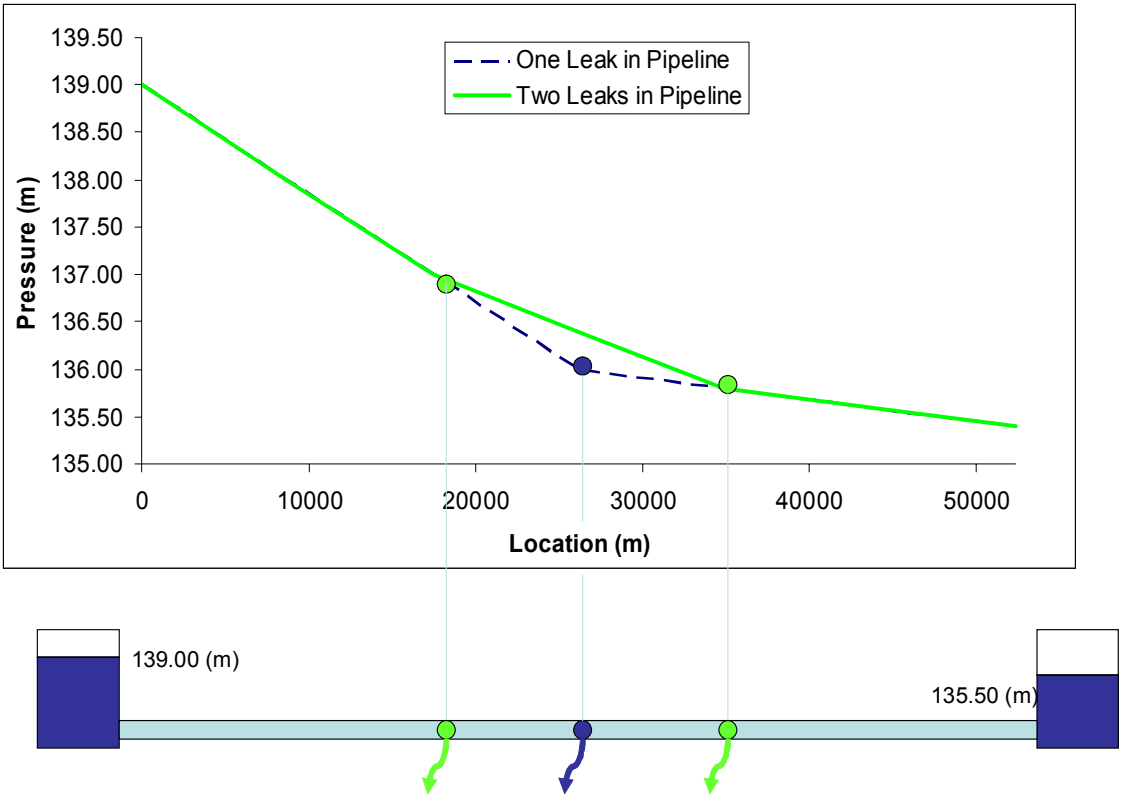

Figure 5.1: Head Measurements for Two Separate Pipelines. The sum of the two leak flows equals the one leak flow scenario.

The equations necessary for locating one non-discrete leak, which includes the actual leakage location and magnitude, given the two fixed fictitious filter leakages are derived below. Figure 5.2 shows two separate pipelines, *pipe a* and *pipe b*, with leakage occurring along the pipelines. *Pipeline a* is the real line with leakage, Q_L, occurring at any location, x_L, along the pipeline. *Pipeline b* is the prediction model (used by the EKF) with "fictitious" leakages, Q_{L1}, and Q_{L2}, occurring at known locations x_{L1} and x_{L2}.

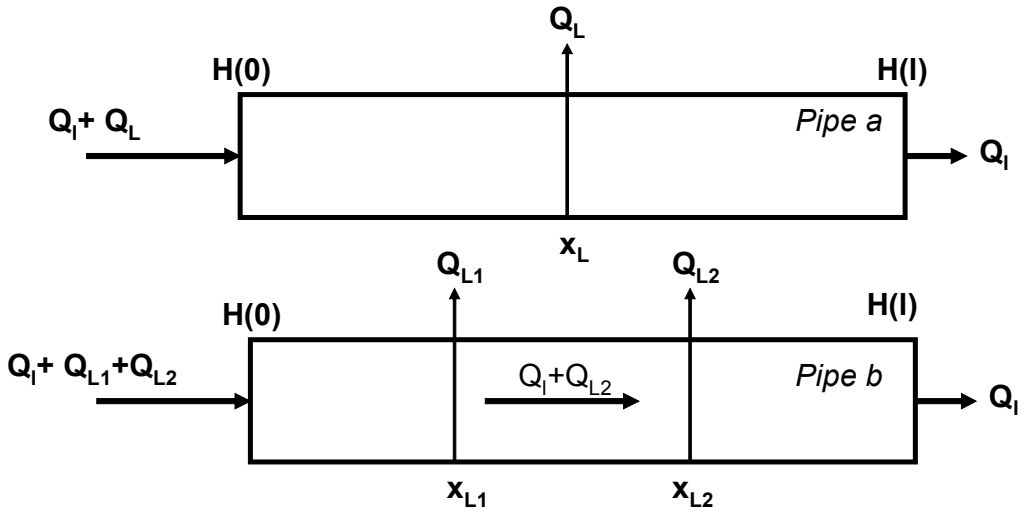

Figure 5.2: Flow within two Identical Pipelines (Lesyshen, 2005)

The momentum and continuity equations, which are derived in Chapter 2, are repeated here as they represent flow dynamics within a pipeline.

$$\frac{\partial Q}{\partial t} + gA\frac{\partial H}{\partial x} + \frac{fQ|Q|}{2DA} = 0\,, \qquad [5.4]$$

$$\frac{\partial H}{\partial t} + \frac{a^2}{gA}\frac{\partial Q}{\partial x} = 0\,. \qquad [5.5]$$

Assuming steady state conditions, the temporal terms within Equations [5.4] and [5.5] disappear and the equations become:

$$\frac{\partial H}{\partial x} = \frac{fQ|Q|}{2gDA^2}\,, \qquad [5.6]$$

$$\frac{\partial Q}{\partial x} = 0\,. \qquad [5.7]$$

At steady state Equation [5.7] states that the flow rate is constant and not dependent on time or position. Therefore the flow is constant throughout the pipeline and is given by the boundary condition at $x = l$. This is given as:

$$Q = Q_l . \qquad [5.8]$$

Integrating Equation [5.6] and substituting $Q = Q_l$ gives the Darcy Weisbach equation for steady state flow as:

$$H(x) - H(0) = \frac{fQ_l^2}{2gDA^2} x , \qquad [5.9]$$

where *H(x)* denotes the steady state head at distance x from the upstream boundary. In order to ensure both pipelines *a* and *b* are equivalent it is necessary to find Q_L and x_L so that the steady state conditions are the same within both pipes. If the flow is to be the same in both pipelines, continuity states that:

$$Q_L = Q_{L1} + Q_{L2} . \qquad [5.10]$$

Applying Equation [5.9] to both pipelines *a* and *b* gives:

$$H(l) - H(0) = \frac{f}{2gDA^2}\left((Q_L + Q_l)^2 x_L + Q_l^2 (l - x_L)\right), \quad \text{pipe a} \qquad [5.11]$$

$$H(l) - H(0) = \frac{f}{2gDA^2}\left((Q_l + Q_{L1} + Q_{L2})^2 x_{L1} + (Q_l + Q_{L2})^2 (x_{L2} - x_{L1}) + Q_l^2 (l - x_{L2})\right), \quad \text{pipe b.} \qquad [5.12]$$

In order for the two pipes to be considered equivalent, the steady state head loss across the two pipelines must be the same. Therefore equating Equation [5.11] and Equation [5.12] results in:

$$(Q_L + Q_l)^2 x_L + Q_l^2 (l - x_L) = (Q_l + Q_{L1} + Q_{L2})^2 x_{L1} + (Q_l + Q_{L2})^2 (x_{L2} - x_{L1}) + Q_l^2 (l - x_{L2})$$

[5.13]

Simplifying Equation [5.13] gives:

$$Q_L{}^2 x_L + 2Q_L Q_l x_L = Q_{L1}^2 x_{L1} + Q_{L2}^2 x_{L2} + 2Q_{L1}Q_{L2}x_{L1} + 2Q_{L1}Q_l x_{L1} + 2Q_{L2}Q_l x_{L2}. \qquad [5.14]$$

Multiplying both sides by $\frac{1}{Q_l^2}$:

$$\left(\frac{Q_L}{Q_l}\right)^2 x_L + 2\frac{Q_L}{Q_l}x_L = \left(\frac{Q_{L1}}{Q_l}\right)^2 x_{L1} + \left(\frac{Q_{L2}}{Q_l}\right)^2 x_{L2} + 2\frac{Q_{L1}Q_{L2}}{Q_l^2}x_{L1} + 2\frac{Q_{L1}}{Q_l}x_{L1} + 2\frac{Q_{L2}}{Q_l}x_{L2},$$

[5.15]

The second order terms in Equation [5.15] may be neglected if it is assumed that the leakage is much smaller than the main flow, Q_l, because their magnitude is small compared to the first order terms (Benkherouf and Allidina, 1988). Equation [5.15] then becomes:

$$Q_L x_L \approx Q_{L1}x_{L1} + Q_{L2}x_{L2}. \qquad [5.16]$$

The leak magnitude therefore may be determined by Equation [5.10] and its position is given by:

$$x_L \approx \frac{Q_{L1}x_{L1} + Q_{L2}x_{L2}}{Q_L}. \qquad [5.17]$$

Thus, since x_{L1} and x_{L2} are specified, Q_{L1} and Q_{L2} are estimated, and Q_L is a summation of Q_{L1} and Q_{L2}, x_L can be readily calculated.

5.2 Simulated Results

The pipeline discussed in Chapter 4 was modeled using the commercial software MATLAB © (See Appendix C for code). The equations used to model the pipeline were developed in Chapter 2. The pipeline is referred to as "the plant" and mathematically represents the actual pipeline. The word "actual" refers to the leak location or magnitude in the plant model. The word "estimate" refers to the leak location or magnitude from the EKF technique. In order to ensure the estimate of the leak location and magnitude reached steady state, the simulation results for the EKF technique were calculated from the $5{,}000^{th}$ time step until the end of the simulation.

For this work, when referring to leak location, "accurate" is referred to when the difference between the actual leak location and estimated leak location is within 5% of the length of the pipeline. Also for this work, when referring to leak magnitude, "accurate" is referred to when the difference between the actual leak magnitude and the estimated leak magnitude is 5% of the actual magnitude. This percent error was selected because currently Colt Engineering has a model developed called LineGuard^{TM} in which they can detect a leak in a pipeline within 3 kilometers in a 53 kilometer pipeline. This represents approximately 5.7% error in their leak detection model (Colt Technologies, 2006)

All of the results presented in this chapter are simulated results. The number of significant figures in the results would depend on the resolution of the pressure transducers. For this thesis, except for the sensitivity analyses, the number of significant figure reported does not necessarily represent the practical number of significant figures. The number of significant figures for this thesis was chosen to be five in order to determine the accuracy of the simulation model.

The plant pipeline was 52382m long and was discretized into twenty-four equal lengths of approximately 2182.6m. A leak was placed at 26191m with a magnitude of $4.1000 \times 10^{-3} m^3/s$. Since the flow rate leaving the pump was $4.1000 \times 10^{-2} m^3/s$, the magnitude of the leakage was approximately 10% of the total flow. The four head

measurements from the plant nodes were corrupted with white noise to simulate the actual sensor measurements. The four head measurements (given in meters) were used as inputs into the model and are shown in Figure 5.3.

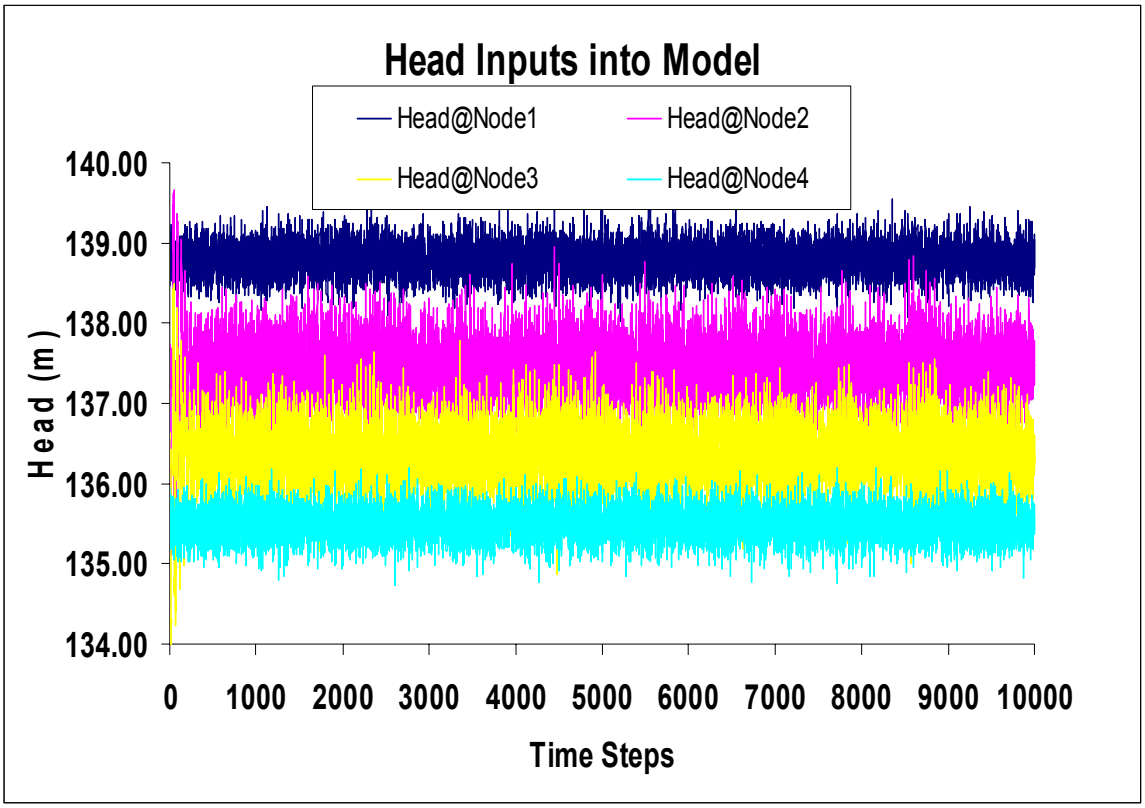

Figure 5.3: Head inputs into the model

Following the procedure outlined in Section 5.1 the model was simulated with a pipeline length of 52382m. The pipeline was discretized into 3 equal sections of 17461m. Fictitious leakage estimates were placed at 17461m and 34921m from the upstream reservoir. Since the leakage was located in the middle of the pipeline, the two fictitious leakages should result in similar magnitudes. Figure 5.4 displays the magnitudes of the two fictitious leaks Q_{L1} and Q_{L2}.

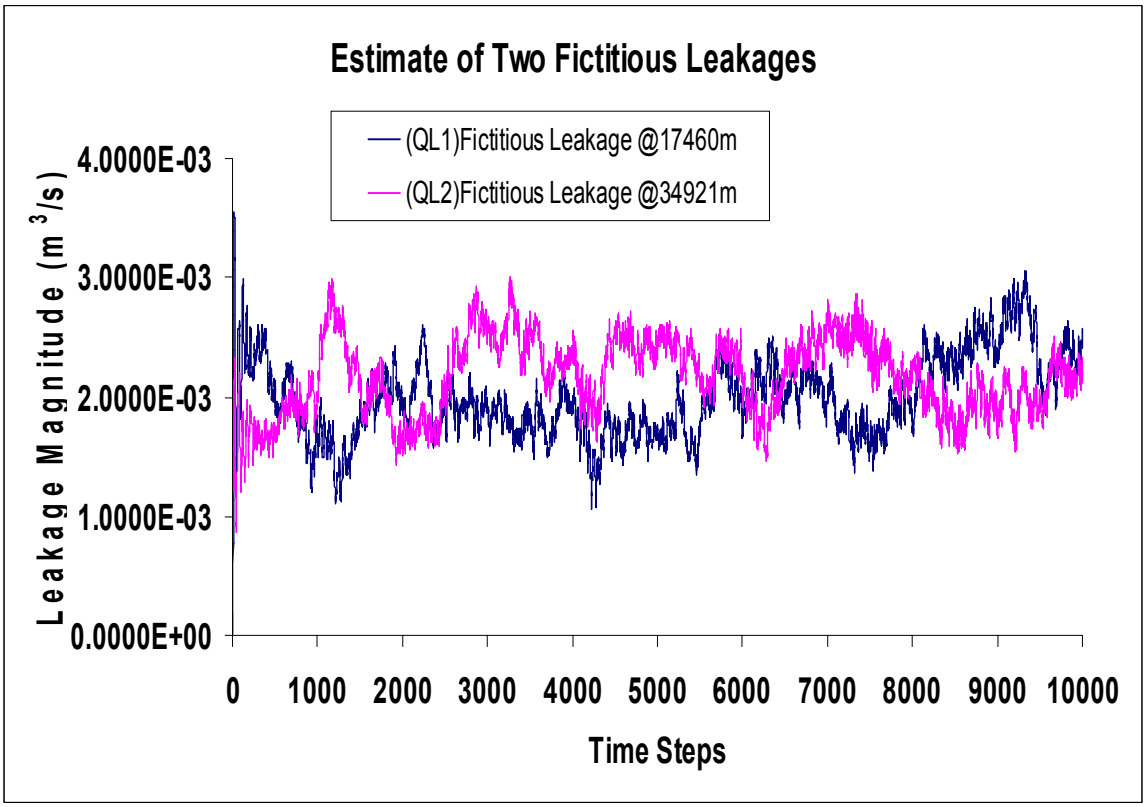

Figure 5.4: Estimates of Two Fictitious Leakages
(Mean Values: Q_{L1}=1.9920x10^{-3}m^3/s and Q_{L2}=2.1680x10^{-3}m^3/s)

Figure 5.4 shows the two estimated fictitious leakages. Q_{L1} had a mean value of 1.9920x10^{-3}m^3/s and Q_{L2} had a mean value of 2.1680x10^{-3}m^3/s. Since the leakage was located in the centre of the pipeline the two fictitious leakages should have been exactly equal in order to give the exact leak location and magnitude. Once the fictitious leakages were determined by the EKF technique they were used in Equation [5.10] and Equation [5.17] to determine the estimated magnitude and location of the leak. Since Q_{L2} is larger than Q_{L1} the estimated leakage should be just over the centre of the pipeline. Figures 5.5 and 5.6 display the estimated and actual leak location and magnitude, respectively. The mean values were calculated from the 5,000th time step until the end of the simulation in order to ensure stability of the leak location and magnitude estimates were achieved. The standard deviation is calculated from the difference between the estimate value and the actual value.

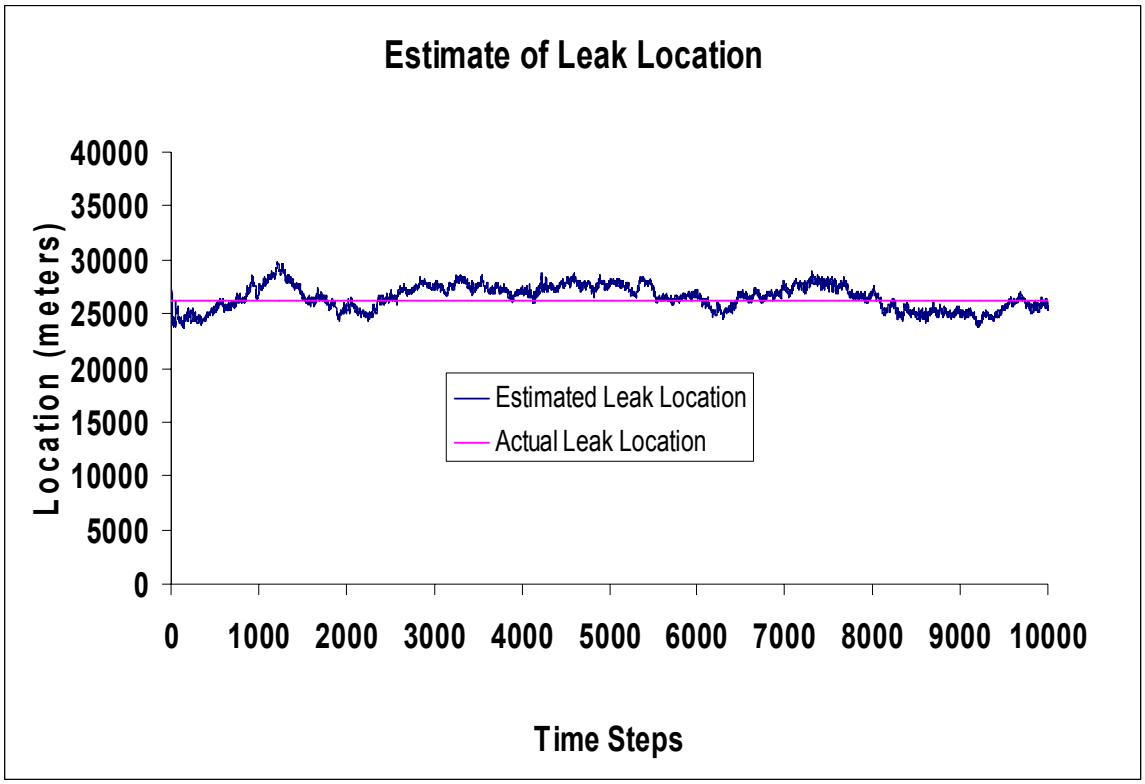

Figure 5.5:Actual Leak Location of 26191m with Mean Estimated Leak Location of 26070m and a Standard Deviation of 573.50m

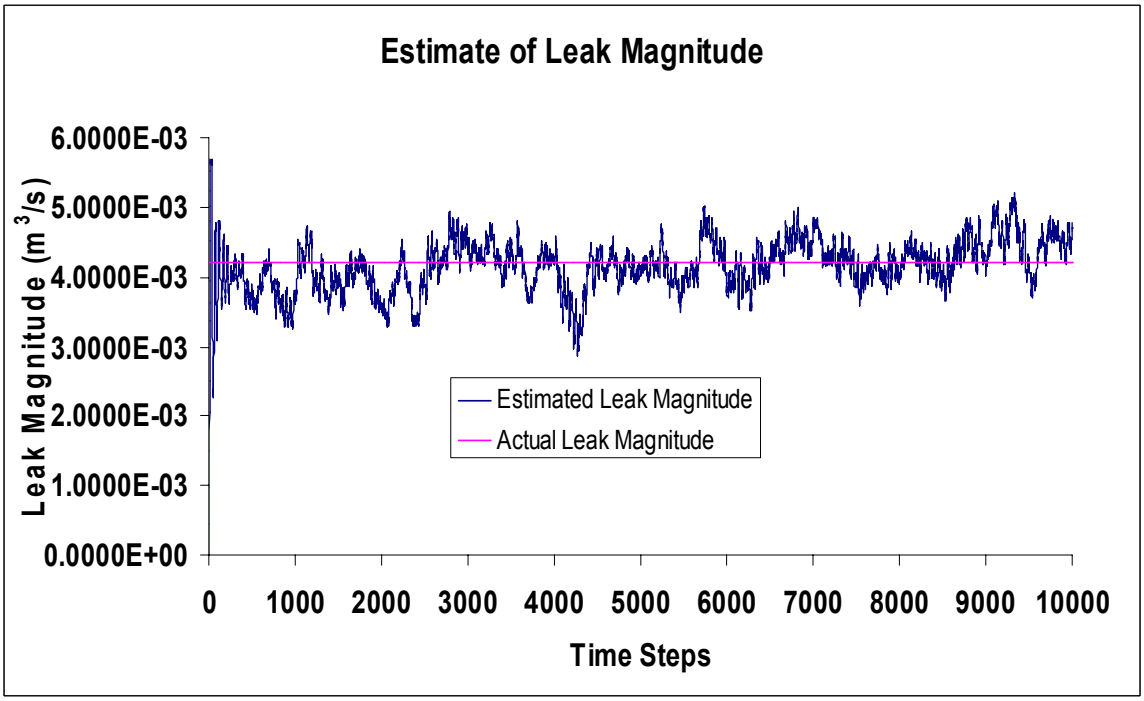

Figure 5.6: Actual Leak Magnitude of $4.1000 \times 10^{-3} m^3/s$ with Mean Estimated Leak Magnitude of $4.0760 \times 10^{-3} m^3/s$ and a Standard Deviation of $5.6689 \times 10^{-5} m^3/s$

The mean estimate of the leak location and magnitude were calculated to be X_L= 26070m and Q_L = 4.0760x10^{-3}m^3/s. The mean error and standard deviation in the leak location estimate for a 10% leak located at 26191m was 0.90% and 573.50m, respectively. The mean error and standard deviation in the leak magnitude estimate for a 10% leak located at 26191m was 0.95% and 5.6689x10^{-5}m^3/s, respectively.

5.2.1 Leak Location Variation

In order to determine the usability of the EKF technique, a leak location variation test was performed. This involved introducing a 10% leak into the pipeline at varied locations along the pipeline. In total seven different locations were tested in order to determine the effectiveness of the EKF technique with respect to leak location. Table 5.1 shows the results from the seven different leak locations. The actual leak is determined using Equation [4.14]; therefore as the location in the pipeline changes the head changes and the actual leak is slightly different from one location to another.

Table 5.1: Leak Estimates for Varied Locations

Actual Leak Location (m)	Mean Estimated Location (m)	Standard Deviation in Location	%Error In Location	Actual flow Magnitude (m^3/sec)	Mean Estimated flow Magnitude (m^3/sec)	Standard Deviation in Flow (m^3/sec)	%Error in Flow
6547.8	15548	9092.3	17.18	4.7170E-03	1.7382E-03	2.9797E-03	63.15
13095	17074	4217.8	7.60	4.1060E-03	3.0363E-03	1.0731E-03	26.05
19643	19569	254.07	0.44	4.1040E-03	4.0877E-03	8.8729E-05	1.92
26190	26070	573.50	0.90	4.1000E-03	4.0760E-03	5.6689E-05	0.95
32739	32462	332.36	0.53	4.0920E-03	4.0625E-03	8.5770E-05	1.38
39286	34296	5005.0	9.53	4.0912E-03	3.0291E-03	1.0650E-03	25.96
45834	35903	10207	18.96	4.0870E-03	1.5308E-03	2.5575E-03	62.54
	Average Results	**4240.3**	**7.88**		**Average Results**	**1.1295E-03**	**25.99**

Table 5.1 shows that the EKF technique is able to accurately detect the location and magnitude of a leak in certain sections of the pipeline. When the leak was located at 19643m the standard deviation in the leak location and magnitude was 254.07m (0.44% error) and 8.8729x10^{-5}m^3/s (1.92% error), respectively. When the leak was located at 45834m the standard deviation in the leak location and magnitude was 10207m

(18.96% error) and $2.5575 \times 10^{-3} m^3/s$ (62.54% error), respectively. Overall, for the seven test location, the EKF technique had an average standard deviation, for the leak location and magnitude, of 4240.3m (7.88% error) and $1.1295 \times 10^{-3} m^3/s$ (25.99% error), respectively.

When the leakage is located in between the first (17460m) and second (34921m) fictitious leakage the EKF technique is able to predict the leakage location and magnitude to a high level of accuracy. However, when the leak is not located between the two fictitious leakages the EKF technique is unable to accurately estimate the leak location or the leak magnitude. The reason this occurs is because when the leak is not located between the fictitious leakages then either Q_{L1} or Q_{L2} must be larger than the actual leakage in order to produce a leakage estimate that is outside the two fictitious leakages (Refer to Equation [5.17]). However, this is not possible because Q_{L1} or Q_{L2} cannot be larger than the actual leakage because the information transmitted in the head measurements doesn't contain that information.

It was decided that the EKF technique needed to be improved in order to increase the accuracy of the model along the entire pipeline. Chapter 6 explains the improvements that were made to the EKF technique and discusses the results.

Chapter 6: Theoretical and Experimental Results using the EBA-EKF Technique on a Large Pipeline

The original EKF technique required four equally-spaced pressure measurements along the pipeline in order to ensure the model was observable and produced repeatable results. However, since the EKF technique did not produce accurate results along the entire pipeline, an improved technique was developed. This improved technique is referred to as the "Extended Boundary Approach (EBA)-Extended Kalman Filter (EKF) technique" because it extended the two interior nodes to the outside of the actual pipeline. This allows the actual leakage to always occur between the two fictitious leakages. Since the original EKF technique produced accurate results when the actual leakage was located in between the two fictitious leakages, it was believed that the EBA-EKF technique would produce accurate results along the entire pipeline. In the EBA-EKF technique, the two internal pressure readings are measured from the plant and the two external pressure measurements are calculated. Therefore, only two pressure measurements are required instead of four pressure measurements as in the original EKF technique. Two internal flow measurements are also required in order to facilitate the calculations of the two external pressure measurements. Chapter 6 includes an explanation of the EBA and the simulated and experimental results from the EBA-EKF technique.

6.1 Extended Boundary Approach (EBA)

The concept of fictitious estimates, as discussed in Section 5.1, was also used in the Extended Boundary Approach model. However, the EBA technique extended the interior nodes in order to improve the estimates of the leakage location and magnitude along the entire pipeline. The following is an explanation and derivation of the Extended Boundary Approach (here after referred to as EBA).

It is important to re-state certain definitions that are critical to the understanding of the EBA. The pipeline is referred to as "the plant" and mathematically represents the actual pipeline. The word "actual" refers to the leak location or magnitude in the plant model. The word "estimate" refers to the leak location or magnitude from the EBA-EKF

technique. The points where head and flow measurements were made are labelled "H_2, H_3 and Q_2, Q_3", respectively, and are represented by black dots in subsequent figures. The points where head measurements were calculated are labelled "H_1 and H_4" and are represented by white dots

Figure 6.1 shows two pipelines, pipe A and pipe B. Pipe A has a length of X_L and pipe B has a length of $3X_L$. Pipe A is the pipeline representing the plant and Pipe B is the pipeline used by the EBA. The middle pipe section of the Pipe B corresponds to the plant Pipe A and the outer two pipes are "fictitious" lines that are "extended" beyond the bounds of Pipe A, hence the name Extended Boundary Approach (EBA). The symbols H_2, H_3 and Q_2, Q_3 represent pressure and flow measurements at the beginning and end of the plant pipeline.

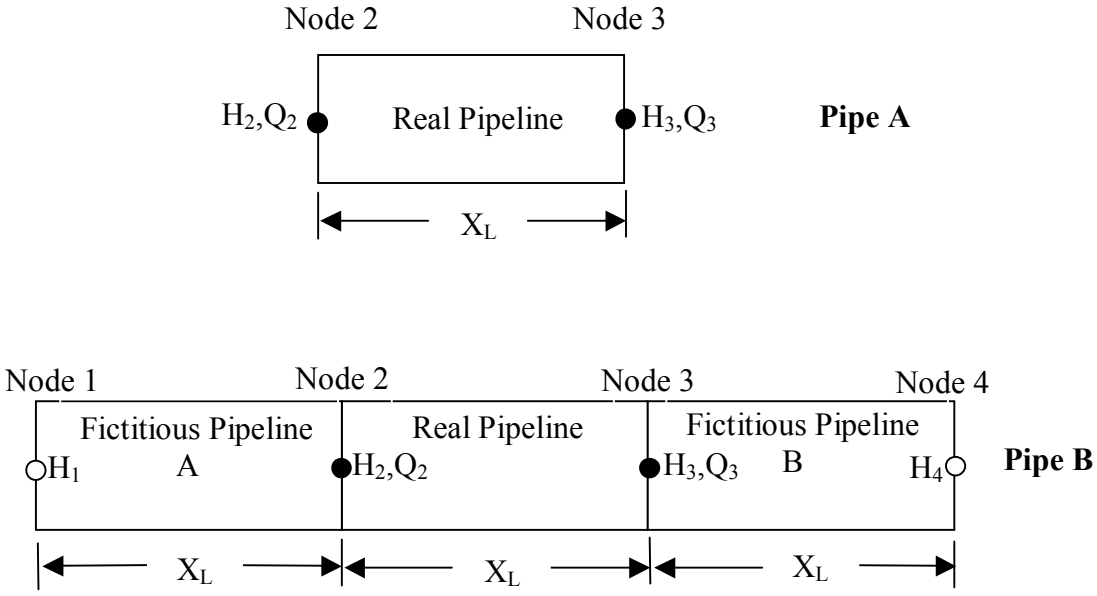

Figure 6.1: Diagram of Plant Pipeline (Pipe A) and EBA model Pipeline (Pipe B)

As explained in Section 4.6, it is necessary to have four pressure measurements in order to ensure the EKF and EBA-EKF technique are observable. In the original approach (EKF technique), two internal pressure measurements in Pipe A were required along with the pressures at the two boundaries. In the EBA technique, H_2, H_3 and Q_2, Q_3 measurements are collected from the boundaries of the plant and H_1 and H_4 are estimated

at the exterior boundaries of the two fictitious pipes as indicated in Figure 6.1. These pressures (H_1 and H_4) are estimated in the following manner.

In Chapter 2, the momentum and continuity equation were presented and are repeated here for convenience.

$$\frac{\partial Q}{\partial t} + gA\frac{\partial H}{\partial x} + \frac{fQ|Q|}{2DA} = 0 \qquad [6.1]$$

$$\frac{\partial H}{\partial t} + \frac{a^2}{gA}\frac{\partial Q}{\partial x} = 0 \qquad [6.2]$$

If the flow through these fictitious pipelines is assumed to be steady state and has no leakage, then:

$$\frac{\partial H}{\partial t} = 0 \qquad [6.3]$$

$$\frac{\partial Q}{\partial t} = 0 \qquad [6.4]$$

The continuity and momentum equations then become:

$$\frac{\partial H}{\partial x} = \frac{fQ|Q|}{2gDA^2}, \qquad [6.5]$$

$$\frac{\partial Q}{\partial x} = 0 \qquad [6.6]$$

At steady state, Equation [6.6] shows that the flow rate is constant and not dependent on time or position and therefore is constant throughout the extended sections of the pipeline. Integrating Equation [6.5] with respect to position (x) results in the Darcy–Weisbach friction formula.

$$\Delta H = \frac{f(\Delta x)Q^2}{2gDA^2} \quad . \qquad [6.7]$$

ΔH represents the head loss in length Δx for a flow of Q. Therefore, H_1 in the extended Pipe B section can be estimated by using Equation [6.7] and adding the ΔH to the measured H_2, being that H_1 is upstream of H_2. Similarly H_4 can be determined using Equation [6.7] by subtracting ΔH from the measured H_3 because H_4 is downstream of H_3. This allows the transient pressure information to be captured at H_2 and H_3 and then extended to H_1 and H_4. An inherent assumption here is that the extended pressure down stream will still be positive.

The Extended Boundary Approach (EBA) uses the fact that the pressure measurements at the ends of the real pipeline can be extended to the outer ends of the fictitious pipelines. Using these new pressures as inputs to the EBA, fictitious leakage could be estimated at the two interior nodes (nodes 2 and 3). It is important to recall from the initial feasibility studies, that the EKF technique could estimate leakage accurately between the two interior nodes. Thus it was believed that by making the plant pipeline "A" as the interior section of an extended pipeline, the same level of leakage prediction accuracy could be achieved. Implementation of this approach (via simulation) is now considered.

6.2 Simulated Results for a Large Pipeline

The pipeline discussed in Chapter 4 was implemented using the commercial software MATLAB © (See Appendix C for Code). The equations used to model the pipeline were developed in Chapter 2. In order to avoid ambiguity the following definition for accuracy is restated. For this work, when referring to leak location, "accurate" is referred to when the difference between the actual leak location and estimated leak location is within 5% of the length of the pipeline. Also for this work, when referring to leak magnitude, "accurate" is referred to when the difference between the actual leak magnitude and the estimated leak magnitude is 5% of the actual magnitude. As stated before, this particular percent error was selected based on the results that other leakage detection approaches have produced (Colt Technologies, 2006).

To test the EBA-EKF technique, the plant pipeline (pipeline A, 52382m) was simulated by discretizing the length into twenty-four equal lengths of 2182.6m. A leak was placed at 26191m with a magnitude of $4.2000 \times 10^{-3} m^3/s$. Since the flow rate leaving the pump was 4.2000×10^{-2} m^3/s, the magnitude of the leakage was approximately 10% of the total flow. The two head measurements from the plant nodes were corrupted with white noise to simulate the actual sensor measurements. The two head measurements (given in meters) at nodes 2 and 3 (obtained from the plant model) and the calculated head measurements at node 1 and 4 (calculated from Equation [6.7]) were used as inputs into the EBA-EKF technique and are shown in Figure 6.2.

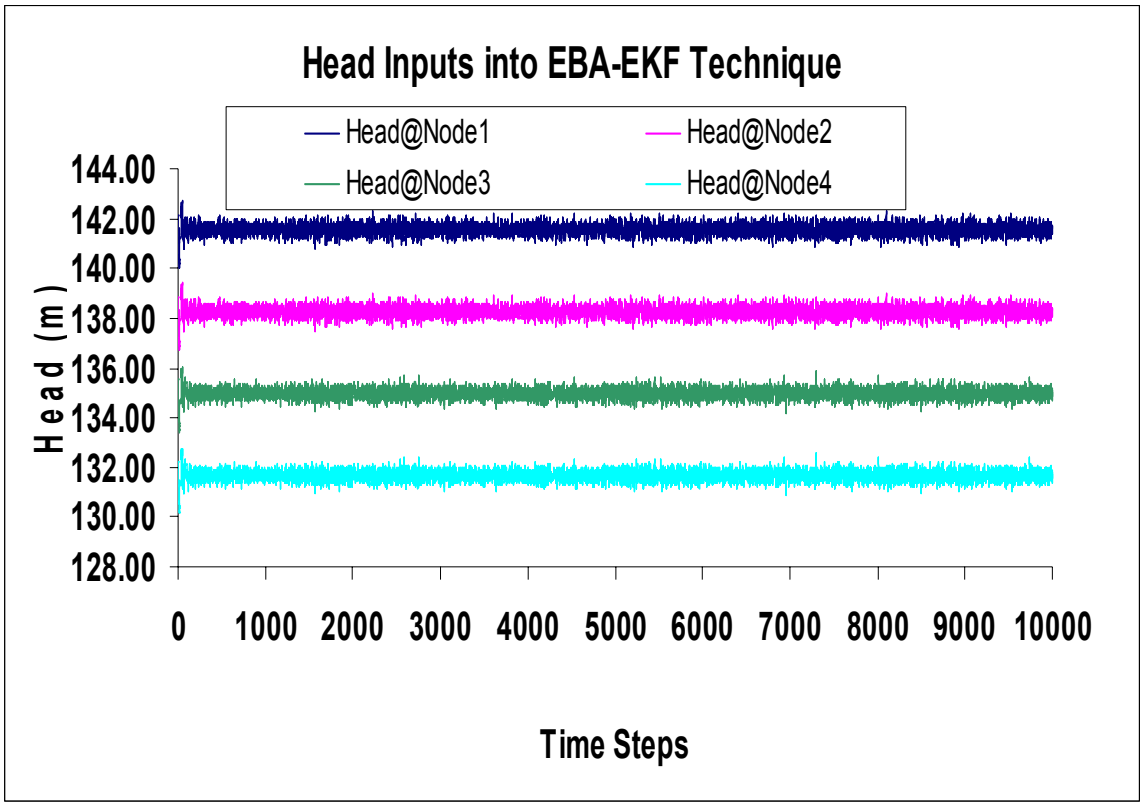

Figure 6.2: Head Inputs into EBA-EKF technique

The two head measurements at the extended boundaries (H_1 and H_4) were estimated from H_2, Q_2 (node 2) and H_3, Q_3 (node 3) measurements using Equation (6.7). Following the procedure outlined in Sections 5.1 and 6.1 the EBA-EKF technique was implemented with a pipeline length of 157146m (three times the original length of 52382m). The

pipeline was discretized into 3 equal sections of 52382m. Fictitious leakage estimates were placed at 52382m and 104764m from the upstream reservoir. Since the actual leakage was located in the middle of the pipeline, the two estimated fictitious leakages should result in similar magnitudes. Figure 6.3 displays the magnitudes of the two fictitious leaks Q_{L1} and Q_{L2}.

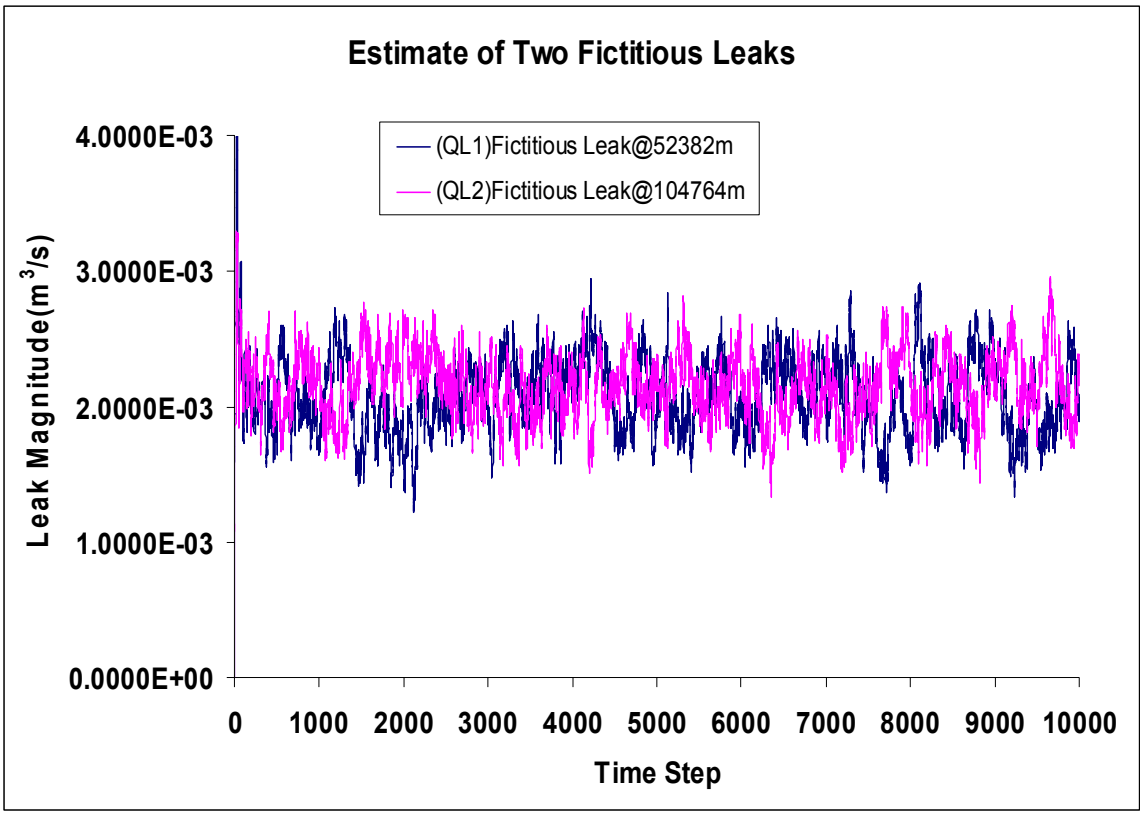

Figure 6.3: Estimates of Two Fictitious Leaks
(Mean Values: $Q_{L1}=2.0832x10^{-3}$ m^3/s and $Q_{L2}=2.1554x10^{-3}m^3/s$)

Figure 6.3 shows that both fictitious leakage estimates had noise associated with them. Q_{L1} had a mean value of $2.0832x10^{-3}m^3/s$ and Q_{L2} had a mean value of $2.1554x10^{-3}m^3/s$ (approximately a 4% difference). Once the fictitious leakages were determined by the EBA-EKF technique they were inputted into Equation [5.10] and Equation [5.17] to give the estimated magnitude and location of the leak. Based on the discussion in Section 5.1 this would translate into the leakage estimate being further downstream then the centre of the pipeline because Q_{L2} is about 4% larger than Q_{L1}. Figures 6.4 and 6.5 display the estimated and actual leak location and magnitude,

respectively. It should be pointed out that the mean values were calculated from 5,000th time step until the end of the simulation execution time in order to ensure stability of the leak location and magnitude estimates had been achieved.

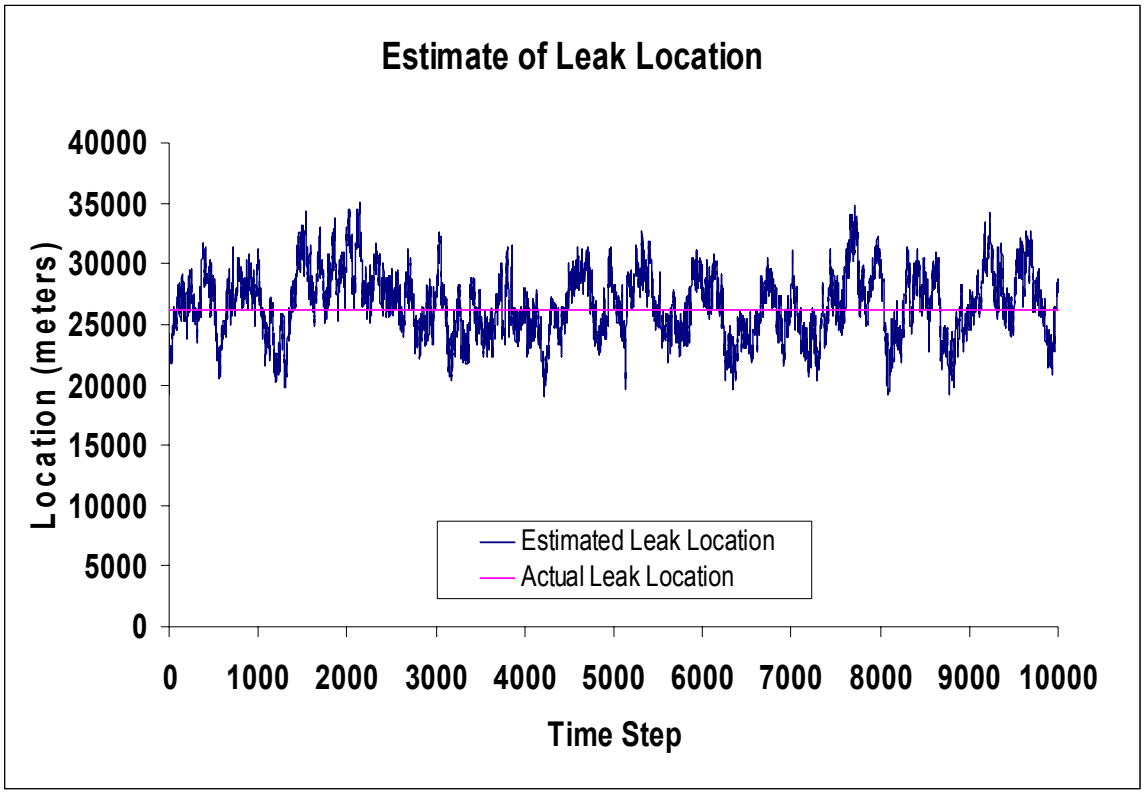

Figure 6.4: Actual Leak Location of 26191m with Mean Estimated Leak Location of 26737m and a Standard Deviation of 689.86m

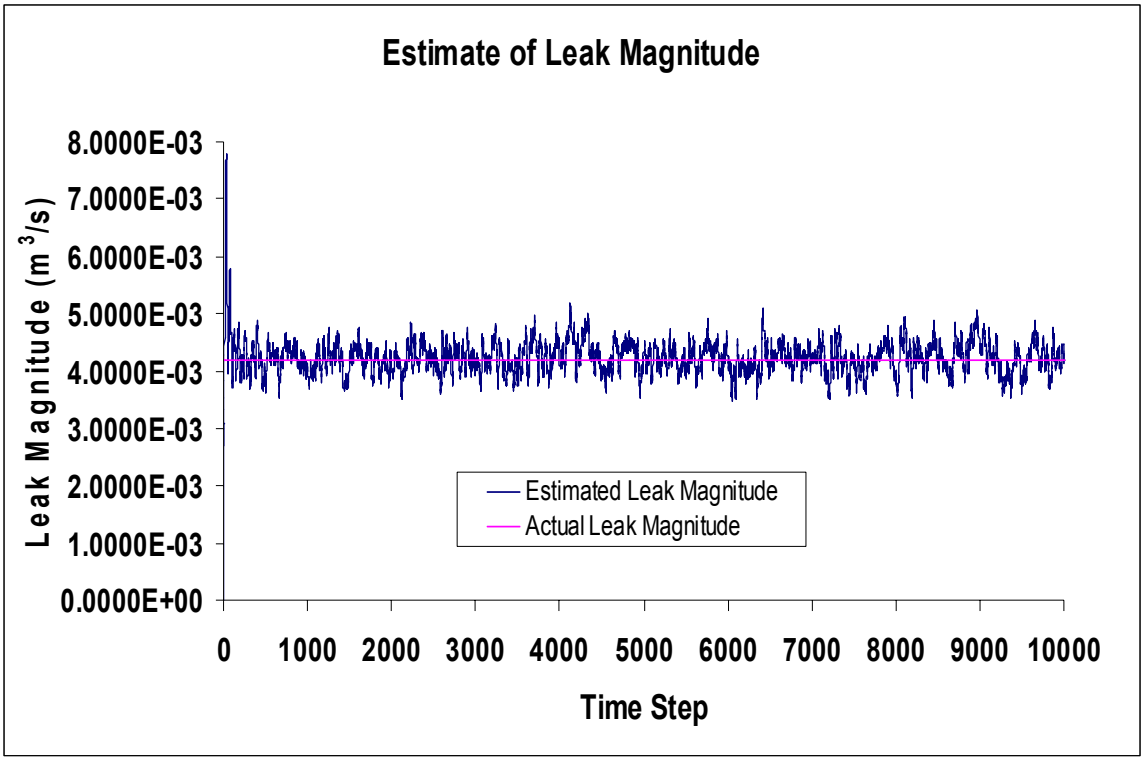

Figure 6.5: Actual Leak Magnitude of $4.2000 \times 10^{-3} m^3/s$ with Mean Estimated Leak Magnitude of $4.2060 \times 10^{-3} m^3/s$ and a Standard Deviation of $1.2286 \times 10^{-5} m^3/s$

The mean estimates of the leak location and magnitude were calculated to be $X_L = 26737m$ and $Q_L = 4.2060 \times 10^{-3} m^3/s$. The mean error and standard deviation in the leak location estimate for a 10% leak located at 26191m were 1.10% and 689.86m, respectively. The mean error and standard deviation in the leak magnitude estimate for a 10% leak located at 26191m were 0.28% and $1.2286 \times 10^{-5} m^3/s$, respectively.

6.2.1 Leak Location Variation

In order to determine the effectiveness of the model, with respect to leak location, a 20% leak was placed at varying locations along the pipeline. Table 6.1 shows the mean, error and standard deviation of the estimated leak locations and magnitudes for a 20% leak. The actual leak is determined using Equation [4.14]; therefore as the location in the pipeline changes the head changes and the actual leak is slightly different from one location to another.

Table 6.1: Leak Estimates for Varied Locations

Actual Leak Location (m)	Mean Estimated Location (m)	Standard Deviation in Location (m)	%Error in Location	Actual flow Magnitude (m^3/sec)	Mean Estimated flow Magnitude (m^3/sec)	Standard Deviation in Flow (m^3/sec)	%Error in Flow
6547.8	7338	818.22	1.51	9.3770E-03	9.4059E-03	3.4204E-05	0.31
13095	14125	1034.5	1.97	9.0060E-03	9.0503E-03	5.0169E-05	0.49
19643	20893	1263.9	2.39	8.9970E-03	9.0390E-03	4.5548E-05	0.47
26191	27503	1319.4	2.50	9.3270E-03	9.3839E-03	5.9054E-05	0.61
32739	17465	1108.3	2.09	8.8490E-03	9.2198E-03	4.5733E-05	0.47
39286	40270	994.53	1.88	8.8320E-03	8.8555E-03	2.5184E-05	0.27
45834	46598	801.19	1.46	9.2970E-03	9.3148E-03	2.4280E-05	0.19
	Average Results	**1048.6**	**1.97**		**Average Results**	**4.0596E-05**	**0.40**

Table 6.1 shows that the EBA-EKF technique is able to accurately detect the location of a leak at the seven different test locations in the pipeline. The accuracy of the technique is the highest when the leak occurs near the boundaries of the pipeline. When the leak was located at 45834m the standard deviation in the leak location and magnitude was 801.19m (1.46% error) and $2.4280 \times 10^{-5} m^3/s$ (0.19% error), respectively. The accuracy of the technique was reduced when the leak was located closer to the center of the pipeline. When the leak was located at 26191m the standard deviation in the leak location and magnitude was 1319.4m (2.50% error) and $5.9054 \times 10^{-5} m^3/s$ (0.61% error). The overall standard deviation in the estimated leak locations was 1048.6m (1.97% error) and the overall standard deviation in the estimated leak magnitude was $4.0596 \times 10^{-5} m^3/s$ (0.40% error). Figure 6.6 shows a graphical representation of the estimated leak locations.

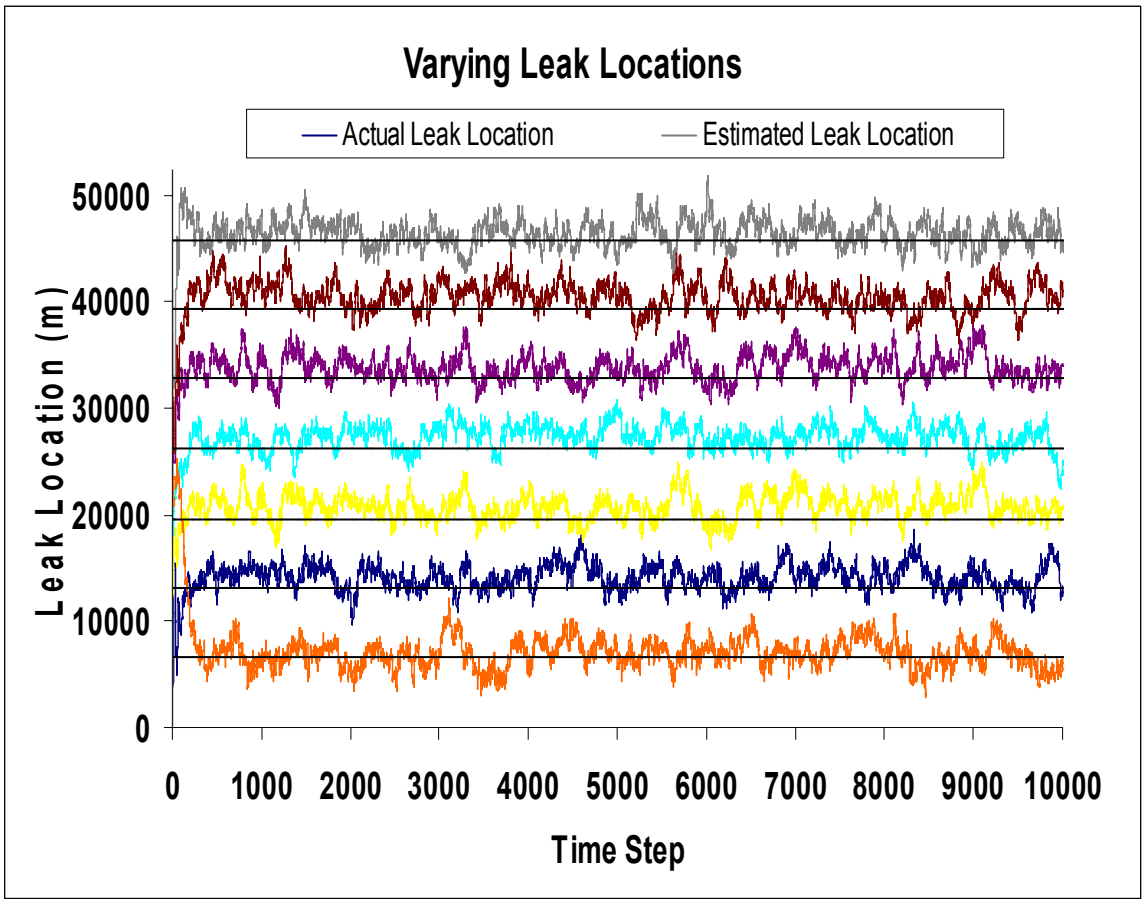

Figure 6.6: Estimated Leak Location at Varied Location

6.2.2 Sensitivity to Leak Magnitude

The ability of the model to detect leaks in a pipeline depends on the magnitude of the leakage. A further limiting factor in the estimation of the leakage is the resolution and DC drift of the pressure and flow sensors that are located on either end of the pipeline. In general, the less noise associated with the sensors the more accurate the EBA-EKF technique will be in predicting the location and magnitude of the leakage.

In order to determine the threshold of the leak magnitude that can be accurately detected a series of simulations were performed. A leak was placed at 26191m and the magnitudes of the leakage were varied from 1, 2, 5, 10, 15 and 20% of the total flow. Table 6.2 shows the results of these simulations.

Table 6.2: Leak Estimates for Varied Magnitudes

% Leakage	Mean Estimated Location (m)	Standard Deviation in Location	%Error In Location	Actual flow Magnitude (m^3/sec)	Mean Estimated flow Magnitude (m^3/sec)	Standard Deviation in Flow (m^3/sec)	%Error in Flow
1	50338	40934	55.95	4.2700E-04	4.5817E-04	4.8411E-05	9.49
2	28888	3514.8	6.31	8.9100E-04	8.4875E-04	4.5079E-05	4.74
5	26232	805.77	1.22	2.3370E-03	2.3405E-03	2.4173E-05	0.89
10	26737	689.86	1.10	4.2060E-03	4.2060E-03	1.2286E-05	0.28
15	26973	794.08	1.49	6.4340E-03	6.4758E-03	4.4269E-05	0.65
20	27503	1320.4	2.51	9.3270E-03	9.3839E-03	5.9054E-05	0.61

Table 6.2 shows the standard deviation and the percent errors of the different magnitudes of leakage. The mean location estimate for the 10% leakage had a standard deviation of 689.86m (1.10% error) and the magnitude estimate had a standard deviation of 1.2286×10^{-5} m^3/s (0.28% error), respectively. When the leakage was 1% of the total flow the estimated leak location had a standard deviation of 40934m (55.95% error), and the estimated magnitude had a standard deviation of $4.8411 \times 10^{-5} m^3/s$ (9.49% error). Therefore according to the definition of accuracy for this thesis, the model was able to accurately predict the leak location when the percent leakage was 5% of the total flow or larger. However when the leakage decreased to 2% or lower the accuracy of the model decreased to an unacceptable level. Figures 6.7 to 6.9 show a graphical representation of the leak location based on the percent leakage. It can be seen that as the percent leakage decreased the noise in the estimates increased.

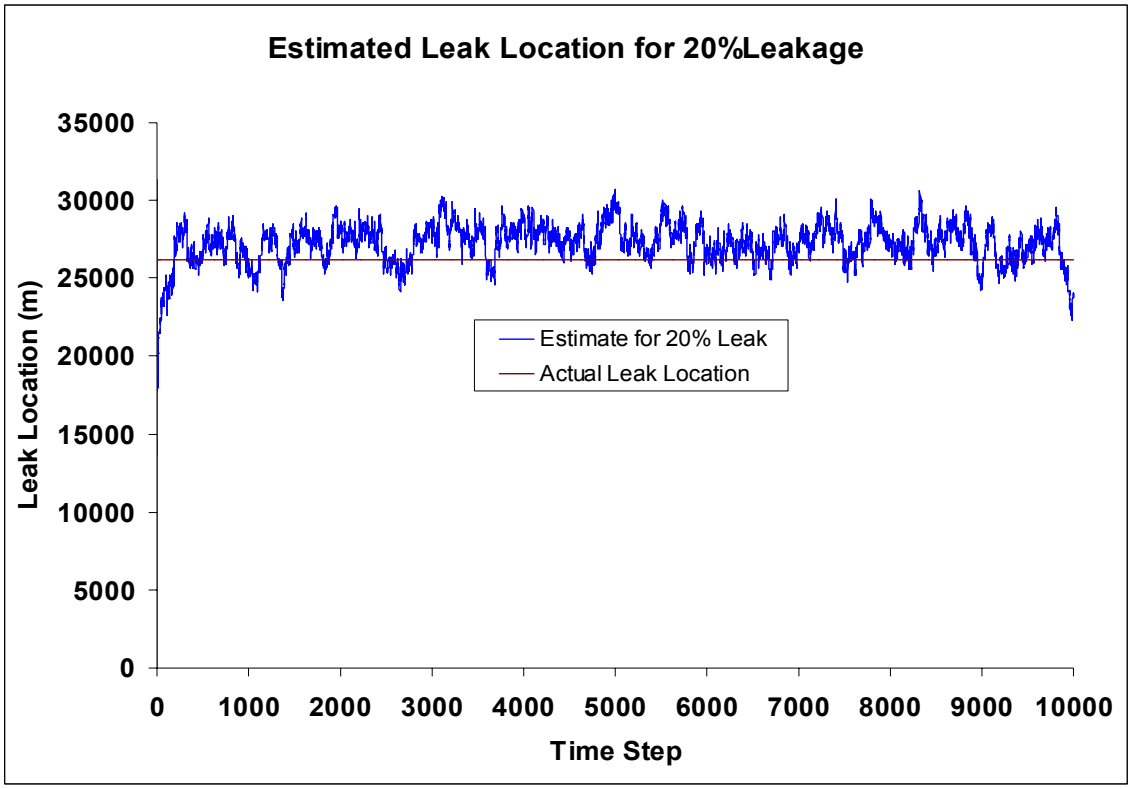

Figure 6.7: Actual Leak Location of 26191m with Mean Estimated Leak Location of 27503m and a Standard Deviation of 1320.4m

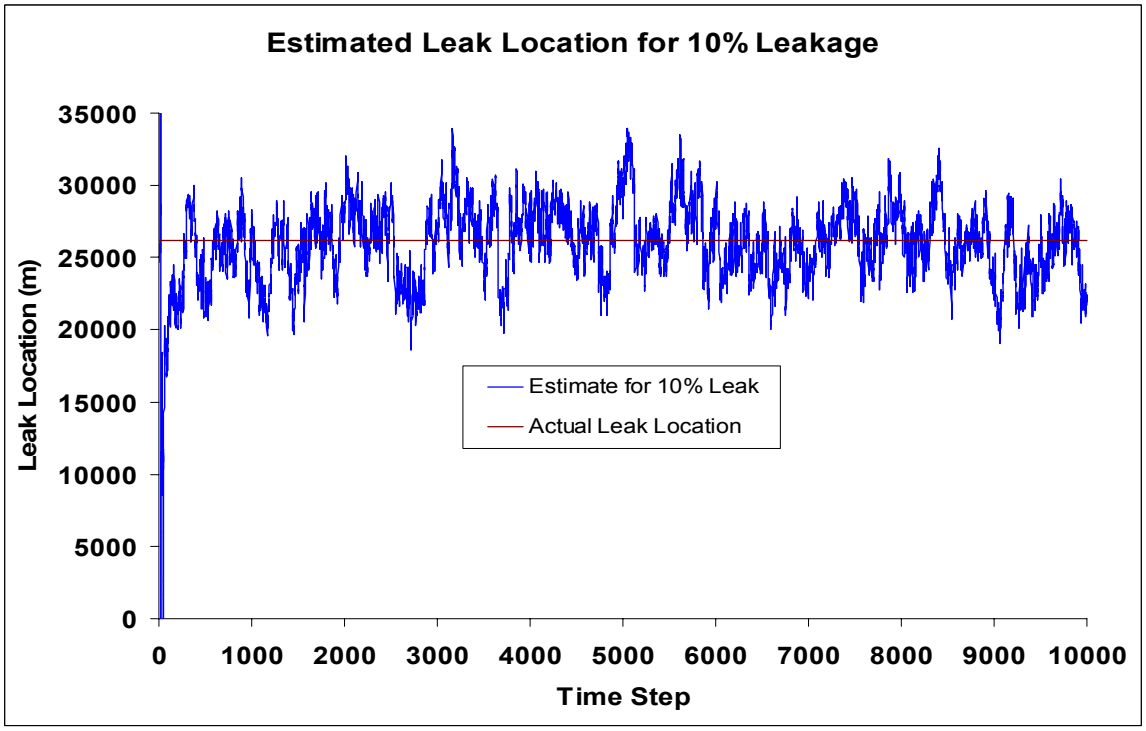

Figure 6.8: Actual Leak Location of 26191m with Mean Estimated Leak Location of 26737m and a Standard Deviation of 689.86m

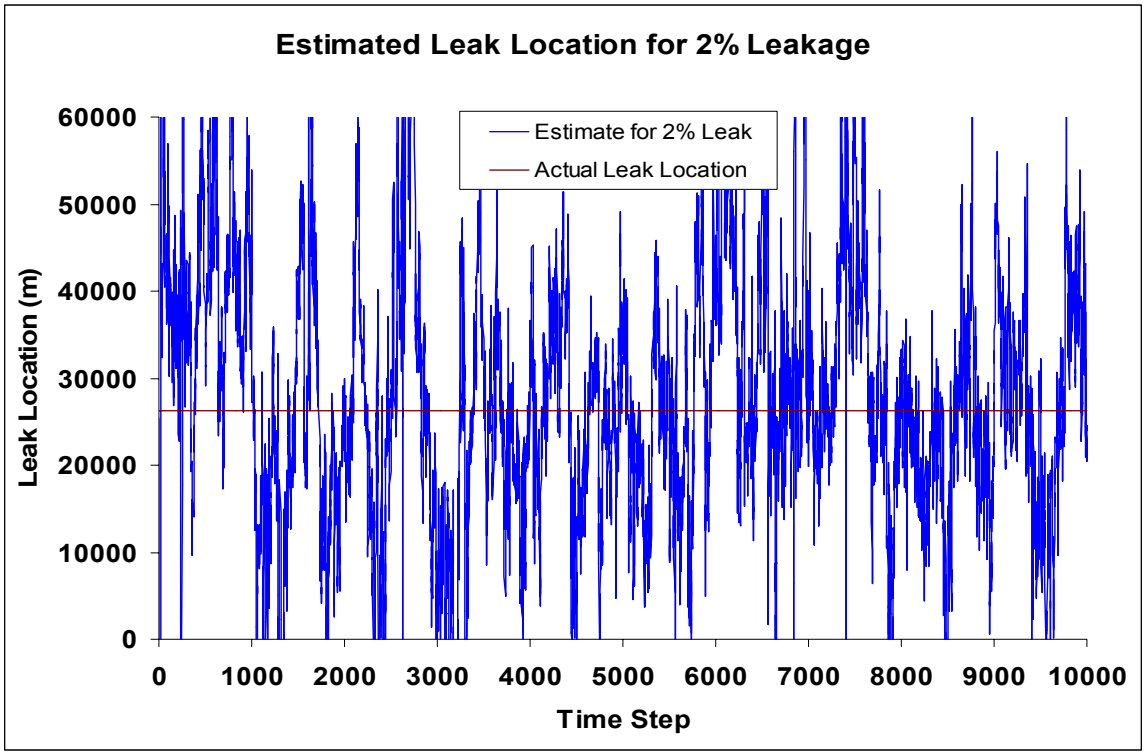

Figure 6.9:Actual Leak Location of 26191m with Mean Estimated Leak Location of 28888m and a Standard Deviation of 3514.8m

The EBA-EKF technique is able to predict the leakage magnitude accurately for any leak that is equal to or greater than 5% of the total flow.

Figure 6.10 shows the estimated magnitudes based on the percent leakage. The EBA-EKF technique is able to accurately detect the magnitude of a leak when the percent leakage is 2% or higher. When the percent leakage is reduced to 1% the EBA-EKF technique is unable to accurately predict the magnitude of the leak.

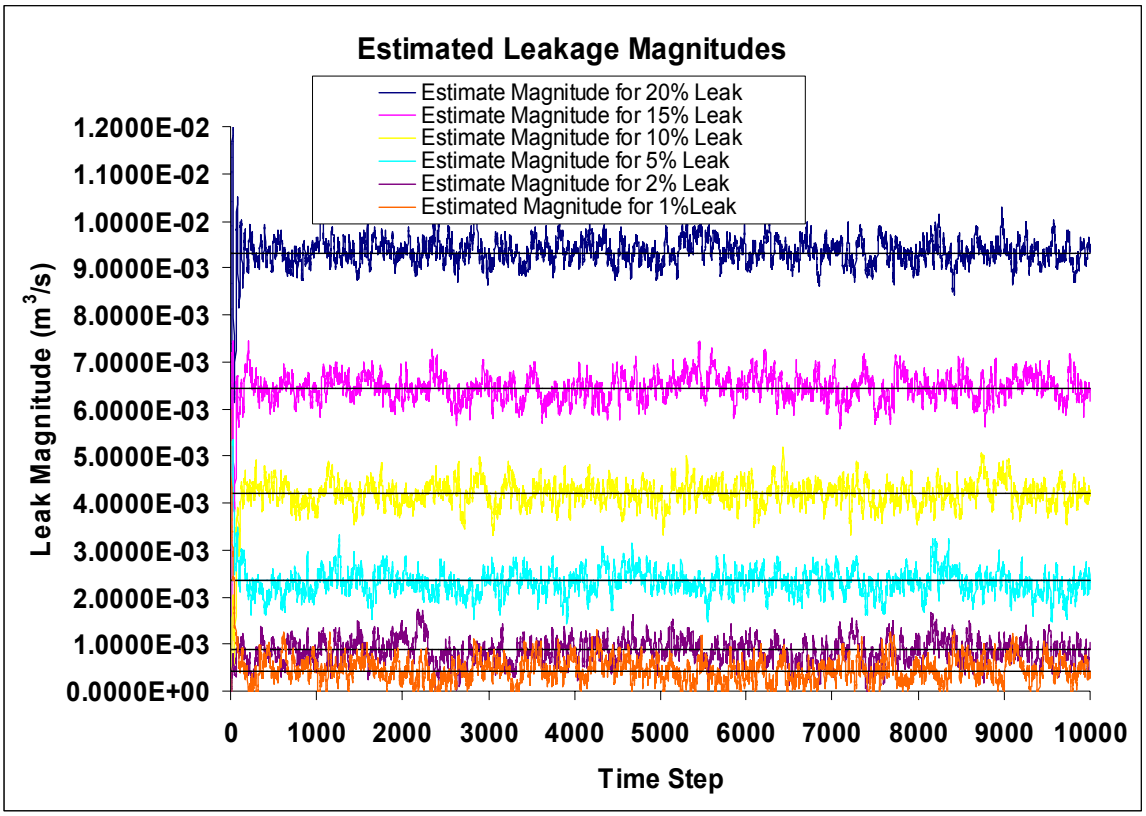

Figure 6.10: Estimated Leakage Magnitude for 1-20% Leaks

The 20% leakage resulted in the worst percent error (2.51%) and the highest standard deviation, while still being "accurate." The 1% and 2% leakages had a higher percent error and standard deviation; however they were not accurate because they had a percent error that was larger than 5%. Therefore the percent leakage that had the worst results, while still being accurate was the 20% leakage. This leakage was then selected to further investigate the effectiveness of the model with respect to leak location as was shown in Section 6.2.1

6.3 Simulated Melfort Plant Pipeline Results

The pipeline discussed in Chapter 4 represents the Melfort raw water distribution pipeline. The simulation was again performed using the Melfort line as a basis for simulation parameters with the individual simulated leaks located at approximately the same location as the physical access points along the Melfort line. The leak magnitudes

were approximately 20% of the total flow leaving the pump. Table 6.3 shows the results of these simulations.

Table 6.3: Leaks Located at Access Points along Melfort Pipeline (simulated plant results only)

Actual Leak Location (m)	Mean Estimated Location (m)	Standard Deviation in Location (m)	%Error in Location	Actual leak Magnitude (m^3/sec)	Mean Estimated flow Magnitude (m^3/sec)	Standard Deviation in Flow (m^3/sec)	%Error in Flow
2182	2429.8	262.44	0.47	9.7620E-03	9.7809E-03	2.1862E-05	0.20
15520	16664	1148.4	2.18	9.6930E-03	9.7353E-03	4.4919E-05	0.44
25143	26430	1309.4	2.46	9.7480E-03	9.6992E-03	5.9974E-05	0.50
	Average Results	**906.74**	**1.70**		**Average Results**	**4.2252E-05**	**0.38**

Table 6.3 shows the mean estimated leak locations have an average standard deviation 906.74m (1.70% error), respectively. The mean estimated magnitudes have an average standard deviation $4.2252 \times 10^{-5} m^3/s$ (0.38% error). As before the overall accuracy of the model increases as the leak approaches the boundaries of the pipeline.

6.3.1 Experimental Results and Discussion from Melfort Pipeline

Experimental tests were performed on the Melfort pipeline in order to test the EBA-EKF technique on a real pipeline. Pressure and flow measurements were collected at the beginning and end of the pipeline. Two separate experiments were conducted. One allowed water to "leak" out of a valve located 2170.0m from the upstream pumping station. The other experiment allowed water to "leak" out of a valve located 25240m from the upstream pumping station. During both experiments the leak magnitudes were varied from 20% to 2% of the total flow. The pressure and flow measurements were collected for 25 minutes to ensure the rich transient information traveled down the pipeline and were obtained in the pressure and flow measurements. The pressure measurements were inputted into the model; however the model would not converge to the correct answer. This implied that the EBA-EKF technique was not using the transient pressure information from the pressure measurements. This implied a significant change in the understanding of how the EBA-EKF was actually performing in the prediction of

the leakage in simulation studies. To re-examine this issue, simulation studies were again considered.

6.3.1.1 Re-examination of pressure transient input information to the EBA-EKF approach

Simulations were performed on the simulated plant in order to determine if the EBA-EKF technique was using the transient pressure information. A leak was located at 17461m and was turned "on" when the time step reached 5000. Figure 6.11 shows the effects the leak had on the pressure measurement at 17461 m.

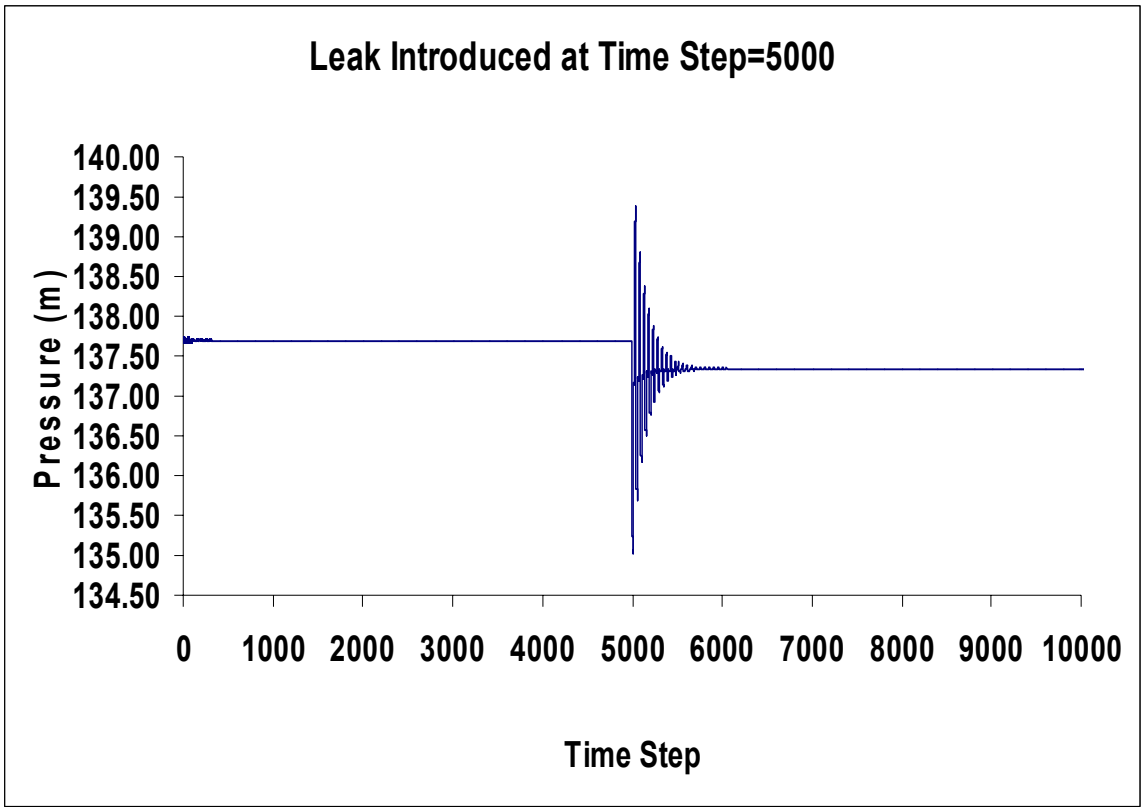

Figure 6.11: Pressure Measurement at 17461m when the leak was turned on at the 5000th Time Step

Figure 6.11 shows that the pressure oscillates when the leak is introduced into the pipeline and becomes stable at approximately 6000 time steps. The four pressure measurements are inputted into the model and Figure 6.12 and 6.13 shows the estimated leak magnitude and location, respectively.

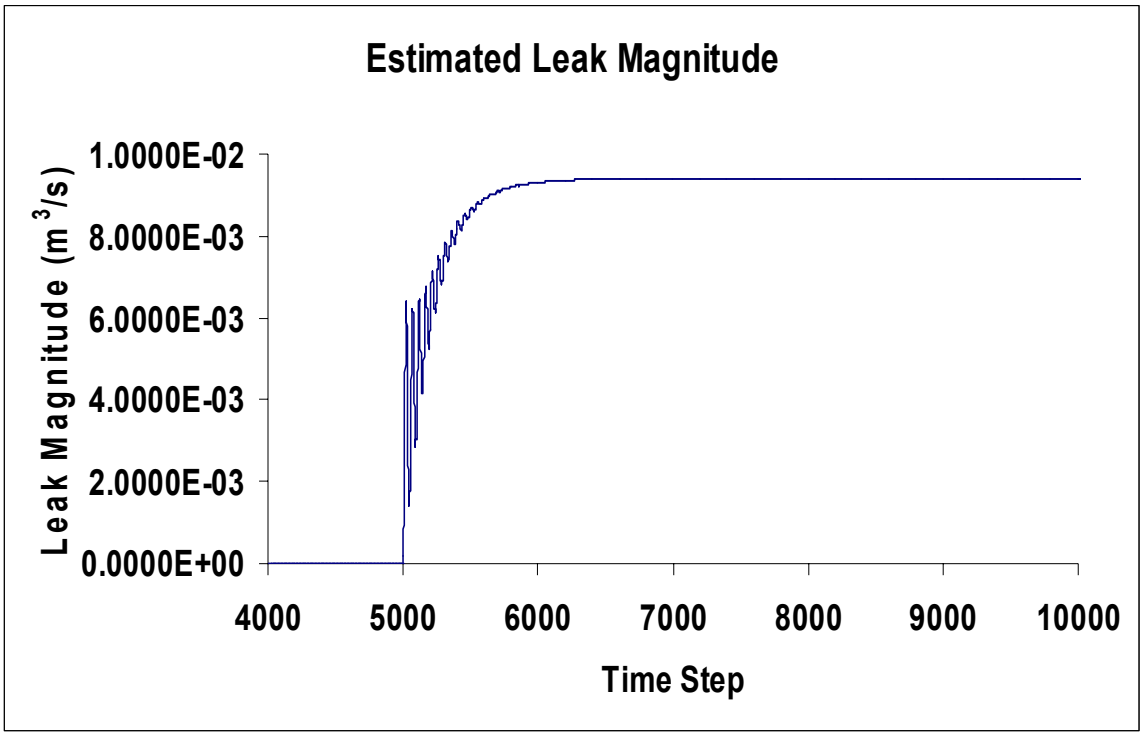

Figure 6.12: Estimated Leak Magnitude when the Leak was turned on at the 5000th Time Step

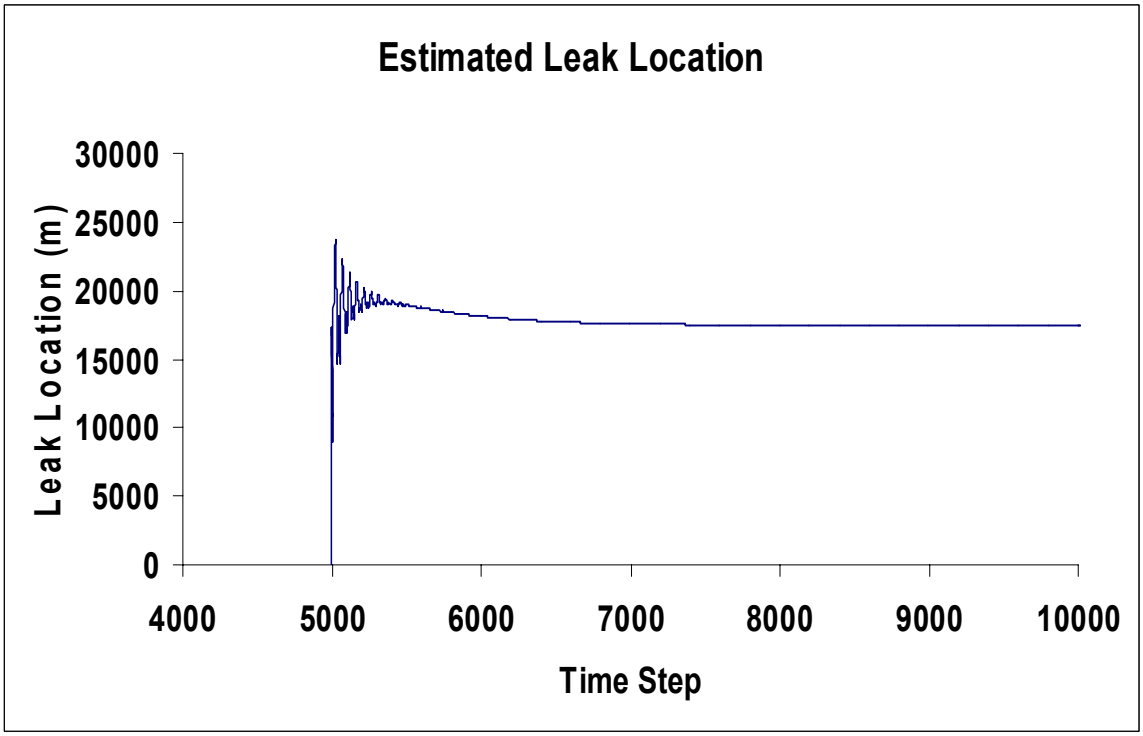

Figure 6.13: Estimated Leak Location when the Leak was turned on at the 5000th Time Step

From Figure 6.12 it can be seen that the estimated magnitude becomes steady at approximately 6000 time steps. This is the same time that the four input pressures also become steady. Figure 6.13 shows that the estimated leak location also becomes steady once the pressure measurements become steady state values. Therefore the EBA-EKF technique was not using the transient information; rather it converged when the pressure measurements became stable.

Pressure transients were added at the upstream head to allow a continuous perturbation to travel along the pipeline. The most accurate predictions from the EBA-EKF technique occurred when there were no perturbations added to the upstream head. The accuracy of the leak location and magnitude estimates decreased as the amplitude and/ or frequency of the perturbations increased. This is because the EBA-EKF technique was not able to use the dynamic information associated with the perturbations. Instead the EBA-EKF technique was able to filter out the perturbation (i.e. noise), but as the perturbations increased the EBA-EKF technique results were less accurate because the perturbations disguised the steady state values. Since the EBA-EKF technique required only the steady state measurements, a significant figure analysis and sensitivity analysis were performed to determine the level of significant figures required for each input, in order to achieve accurate simulated results.

6.3.2 Significant Figures Analysis

Since it was determined that the EBA-EKF technique only requires steady state measurements, it was necessary to investigate the significant figures required to produce accurate results. All of the inputs (i.e. density, gravity, modulus of elasticity, upstream/downstream head, head measurements at four nodes etc.) had five significant figures in the plant and the simulation. The results of all the outputs were recorded (i.e head measurements at the four nodes, upstream/downstream flow, fictitious leaks etc.). All of the inputs were reduced to four significant figures and then to three significant figures and the outputs were recorded for three and four significant figures. The results did not change when the number of significant figures were reduced from five to four;

however, when the significant figures were reduced from four to three the accuracy of the model was significantly reduced.

The significant figures of each input were then reduced individually. This was completed to determine which input's significant figures the model relied upon to produce accurate results. All of the inputs could have their significant figures reduced from five to three without affecting the accuracy of the model, with the exception of the four head measurement that were inputted into the Extended Kalman Filter. When the four head measurements were reduced from five significant figures to three the accuracy of the model was greatly reduced. Section 6.3.3 shows the results of this simulation

6.3.3 Sensitivity Analysis for the Simulated Melfort Pipeline

Since it was determined that the number of significant figures of the head measurements greatly affect the accuracy of the model, a sensitivity analysis was conducted on the head measurements. In order to determine if the EBA-EKF technique could be implemented on the Melfort pipeline it was necessary to determine the level of sensitivity required from the two pressure transducers. Numerous simulations were performed on the Melfort pipeline model in order to determine the level of accuracy required from the pressure transducers. The transients from the pump and the sensor noise were removed from the simulation to eliminate any uncertainty in the sensitivity analysis. A 20% leak was placed at 6547.8m and Table 6.4 shows the results from the sensitivity analysis on the Melfort model pipeline.

Table 6.4: Sensitivity Analysis for Melfort Model Pipeline

H1 (m)	H_2 (m)	H_3 (m)	H_4 (m)	Actual Location (m)	Estimated Location (m)	% Error	Actual Magnitude (m^3/s)	Estimated Magnitude (m^3/s)	% Error
142.0962	137.8723	134.9444	132.2017	6547.8	7217.3	**1.28**	8.7900E-03	8.7898E-03	0.002
142.096	137.872	134.944	132.202	6547.8	7233.1	**1.31**	8.7900E-03	8.7951E-03	0.058
142.10	137.87	134.94	132.20	6547.8	7438.6	**1.70**	8.7900E-03	8.8313E-03	0.470
142.1	137.9	134.9	132.2	6547.8	11538	**9.53**	8.7900E-03	8.9375E-03	1.678

H_2 and H_3 are the measured pressure measurements and H_1 and H_4 are the calculated pressure measurements using the EBA. When the pressure measurements were accurate to four decimal places the percent error in the estimated leak location was 1.28% and the percent error in the estimated leak magnitude was 0.002%. However when the accuracy of the pressure transducer was reduced (by eliminating the last decimal place) the percent error of the leak location and leak magnitude increased. When the pressure measurements were accurate to only one decimal place the percent error in the leak location and magnitude was 9.53% and 1.678%, respectively. Therefore, for a pipeline with parameters similar to the Melfort pipeline the pressure transducers must be accurate to ± 0.01m. This translates into ± 96Pa or ± 0.014psi. This level of accuracy is unobtainable with the current level of technology available for pressure transducers.

A physical interpretation of this result can be seen by observing the difference between the heads at the inlet and outlet (which include elevation differences). The pressure gradient along the lines with respect to nodes 2 and 3 are very small for the Melfort pipeline. Thus any small slight "bias" in the slope at either end would result in a substantial change in the intersection of the pressure gradients. Thus the pressure gradients at the end points appear to have a significant effect on the ability of the EBA-EKF technique to predict an accurate leakage location.

It was decided to investigate a much smaller pipeline to determine if using the EBA-EKF technique could be feasible on a pipeline with smaller parameter in which a much larger pressure gradient would exist. This is considered in Chapter 7.

Chapter 7: Theoretical and Experimental Results using the EBA-EKF Technique on a Small Pipeline

The sensitivity analysis presented in the last Chapter showed that implementing the EBA-EKF technique on a large pipeline is limited by the level of accuracy of the pressure transducers. It was believed that the small pressure gradients at the ends of the pipeline due to elevation differences could be one reason for this sensitivity. Therefore, a smaller pipeline with much larger pressure gradients was investigated to determine if the EBA-EKF technique would produce accurate results at least from a theoretical modeling and experimental point of view. The pipeline consisted of 19.150m of steel tubing. The pipeline had an outside diameter of 1.2700×10^{-2}m (0.5 inch) and an internal diameter of 8.8900×10^{-3}m (0.35 inch). The hydraulic flow rate with no leakage was set to $1.8900 \times 10^{-4} m^3/s$ (3gpm). Theoretical simulations were completed to determine if the EBA-EKF technique would be capable of determining the leak location and magnitude in a smaller pipeline. Experimental tests were also performed on the smaller pipeline to determine if a leak location and magnitude could be experimentally determined using the EBA-EKF technique.

7.1 Results for a Simulated Small Pipeline

The pipeline discussed above was simulated using the commercial software MATLAB ©. The equations used to model the pipeline were developed in Chapter 2. In order to avoid ambiguity the following definition for accuracy is restated; "accurate" is referred to when the estimated value is within 5% of the actual value.

A 20% leak was placed at three different locations along the pipeline. Table 7.1 shows the results from these simulations.

Table 7.1: Simulated Results for Small Pipeline

Actual Leak Location (m)	Mean Estimated Location (m)	Standard Deviation in Location	%Error In Location	Actual flow Magnitude (m^3/sec)	Mean Estimated flow Magnitude (m^3/sec)	Standard Deviation in Flow (m^3/sec)	%Error in Flow
2.3920	2.6340	0.24405	**1.27**	4.2720E-05	4.2740E-05	2.0189E-08	0.047
9.5810	10.077	0.49740	**2.60**	4.1430E-05	4.1429E-05	2.0000E-09	0.004
16.760	16.975	0.21461	**1.12**	4.3230E-05	4.3244E-05	1.4255E-08	0.032
	Average Results	0.31869	**1.66**		Average Results	1.2148E-08	0.028

Table 7.1 shows very similar results to the theoretical results of the large pipeline in Table 6.2. When the leak was located near the boundaries of the pipeline the percent error in the leak location was lower than when the leak was located in the centre of the pipeline. When the leak was located at 16.760m the standard deviation in the leak location and magnitude was 0.21461m (1.12% error) and $1.4255 \times 10^{-8} m^3/s$ (0.032% error), respectively. The overall standard deviation in the estimated leak locations was 0.31869m (1.66% error) and the overall standard deviation in the estimated leak magnitude was $1.2148 \times 10^{-8} m^3/s$ (0.028% error). Therefore, these simulations show that the EBA-EKF technique can accurately detect the leak location and magnitude in a small hydraulic pipeline given the assumed significant figures in the pressure head measurements.

These simulations included pump transients and sensor noise. However, when the pump transients and sensor noise were eliminated the percent error in the leak location and magnitude estimates decreased. The most accurate leak location and magnitude estimates resulted when steady state pressure measurements were inputted into the model.

Also, as more perturbations were added to the pump, the accuracy of the estimated leak location and magnitude were reduced. This is very similar to the result in Chapter 6 and shows that the EBA-EKF technique was not able to use the dynamic information associated with the perturbation in the small pipeline. Rather it appears that the EBA-EKF technique uses the steady-state information and in fact has to filter out the noise associated with perturbations of the pump and the sensor noise. As a consequence of this discovery, a sensitivity analysis was performed to determine the level of accuracy

required by the pressure sensors in order to accurately detect the location and magnitude of a leak in a small pipeline as described above.

7.1.1 Sensitivity Analysis for a Simulated Small Pipeline

In order to determine if the EBA-EKF technique could be implemented on a small experimental hydraulic pipeline, a sensitivity analysis was performed. Once again the pump transients and the sensor noise were removed from the plant model simulation to determine the level of accuracy required by the pressure transducers to produce accurate results. A 20% leak was placed at 2.3900m and Table 7.2 shows the results of the sensitivity analysis for the small pipeline.

Table 7.2: Sensitivity Analysis for Small Pipeline (Simulated Results)

H_1 (m)	H_2 (m)	H_3 (m)	H_4 (m)	Actual Location (m)	Estimated Location (m)	% Error	Actual Magnitude (m^3/s)	Estimated Magnitude (m^3/s)	% Error
668.1923	458.2039	312.6928	176.3928	2.3900	2.5247	**0.70**	4.1570E-05	4.1522E-05	0.12
668.192	458.204	312.693	176.393	2.3900	2.5244	**0.70**	4.1570E-05	4.1519E-05	0.12
668.19	458.20	312.69	176.39	2.3900	2.5234	**0.70**	4.1570E-05	4.1517E-05	0.13
668.2	458.2	312.7	176.4	2.3900	2.5194	**0.68**	4.1570E-05	4.1515E-05	0.13
668	458	313	176	2.3900	2.2962	**0.49**	4.1570E-05	4.1429E-05	0.34
670	460	310	180	2.3900	4.4797	**10.91**	4.1570E-05	4.2302E-05	1.76

H_2 and H_3 are the pressure measurements from the plant and H_1 and H_4 are the calculated pressure measurements using the EBA approach (Equation 6.7). When the pressure measurements were accurate to four decimal places the percent error in the estimated leak location was 0.70% and the percent error in the estimated leak magnitude was 0.12%. However when the accuracy of the pressure transducer was reduced (by eliminating the last decimal place) the percent error of the leak location and leak magnitude increased once it reached the "ten" position. When the pressure measurements were rounded off at the "ten" position the percent error in the leak location and magnitude increase significantly to 10.91% and 1.76%, respectively. Therefore, for a small pipeline with

parameters similar to the pipeline described in the beginning of Chapter 7, the pressure transducers must be accurate to ± 1.0m. This translates into ± 8142Pa or ± 1.2Psi. This level of accuracy is still very high; however it is much lower than the level of accuracy required for a pipeline with head traces similar to the Melfort pipeline.

The reason the sensitivity tolerance is higher for the small pipeline is because there is a much larger head loss from the beginning to the end of the pipeline. The small pipeline had a head loss of approximately 145m of head, whereas the large pipeline only had a head loss of approximately 3m of head (due to elevation changes). With a larger head loss the level of accuracy required by the pressure sensors reduces because the pressure gradients are much larger and a small bias does not have such a huge effect on the point of intersection between the two slopes.

Even though the accuracy of the sensor would have to be very high for a small pipeline, an experiment was set up in the laboratory to determine if the current EBA-EKF technique could accurately detect the leak location and magnitude of a real pipeline.

7.1.2 Experimental Results for Small Pipeline

A small pipeline was set up to determine if the EBA-EKF technique could accurately determine the location and magnitude of a leak. As discussed in the beginning of Chapter 7, the pipeline was 19.l50m long and had an internal diameter of $8.8900\text{x}10^{-3}\text{m}^3/\text{s}$
(0.35 inch). The fluid was hydraulic oil and the pump was a pressure compensated type. A needle valve was located at the end of the pipeline (similar to the pipeline discussed in Chapter 4) and was set so that the pump was deadheading. Therefore, when a leak was introduced into the pipeline the pressure at the pump decreased and the flow rate increased which is similar to the situation in the Melfort line. One pressure transducer and flow meter were located at the beginning of the pipeline and one pressure transducer was located at the end of the pipeline. Flow at the end of the line was determined by subtracting the leakage flow (known) from the inlet flow. The inlet and outlet flow rates were needed to apply the EBA technique. Since the small pipeline was much shorter than

the Melfort pipeline the steady state information could travel the length of the pipeline and be seen at the endpoints very quickly.

A valve was located at three different locations to allow fluid to "leak" out of the pipeline. The three leak locations were 3.0300m, 9.3000m and 15.780m. Varying percent leakages were placed at 9.3000m to test the EBA-EKF technique with varying leakage magnitude. The magnitude of the leak was determined by recording the amount of time required for a measured volume amount of fluid to "leak" out of a valve. H_2 and H_3 were the pressure measurements made on the experimental system. H_1 and H_4 were the pressure measurements calculated using Equation [6.7]. Table 7.3 shows the results from the experimental data.

Table 7.3: Experimental Results for Small Pipeline

% Leakage	Experiment Data (m)		Calculated Data (m)		Actual Location (m)	Estimated Location (m)	% Error	Actual Magnitude(m^3/s)	Estimated Magnitude (m^3/s)	% Error
	H_2	H_3	H_1	H_4						
28	494.5	320.9	814.8	156.9	9.3000	1.4662	**40.91**	7.5266E-05	7.4840E-05	**0.57**
15	497.8	334.2	722.2	170.2	9.3000	0.026989	**48.42**	3.2113E-05	3.1879E-05	**0.73**
11	500.4	335.4	709.4	171.4	9.3000	0.67399	**45.04**	2.4416E-05	2.4174E-05	**0.99**
6	499.2	347.7	685.3	183.7	9.3000	-11.260	**107.36**	1.2370E-05	1.2171E-05	**1.61**
2	500.7	348.3	671.9	184.3	9.3000	-31.592	**213.54**	4.1892E-06	3.9596E-06	**5.48**
23	495.1	334.4	771.1	170.4	3.0300	-0.55984	**18.75**	5.6260E-05	5.5934E-05	**0.58**
25	498.4	304.7	794.1	140.7	15.780	4.9510	**56.55**	6.4857E-05	6.4515E-05	**0.53**

Table 7.3 shows the experimental results for the three different leak locations with varying leak magnitudes. The estimated leak locations have significant error for all the leak locations. Therefore, for all intent and purposes, the EBA-EKF technique was unable to predict the leak location in the experimental design. It is interesting to note that the EBA-EKF technique was able to accurately estimate the leak magnitude (not position) when the percent leakage was greater than or equal to 6% of the total flow.

It was postulated that perhaps a "bias" in the pressure transducer calibration existed. If this bias could be determined and removed, then perhaps the accuracy of the leak position estimate could be improved. The pressure measurements for the 11% leakage at 9.3000m were "adjusted" until the percent error in the leak location was just below 5% error. This

translated into adjusting H_3 and H_4 by 17.400m (0.14200 MPa) or 5.5% of the measured pressure. The other H_3 and H_4 pressure measurements that were collected during the experiment at other leakage rates and leak locations were adjusted by 5.5% to determine if the "adjusted" data would accurately detect the leak location and magnitude. Table 7.4 shows the "adjusted" experimental results.

Table 7.4: Adjusted Experimental Results for Small Pipeline

% Leakage	Adjusted Experiment Data (m)		Adjusted Calculated Data (m)		Actual Location (m)	Estimated Location (m)	% Error in Location	Actual Flow Magnitude (m^3/s)	Estimated Flow Magnitude (m^3/s)	% Error in Flow
	H2	H3	H1	H4						
28	494.5	303.3	814.8	139.3	9.3000	3.9300	**28.04**	7.5266E-05	7.4840E-05	**0.57**
15	497.8	316.8	722.2	152.8	9.3000	5.9054	**17.73**	3.2113E-05	3.1879E-05	**0.73**
11	500.4	318.0	709.4	154.0	9.3000	8.3947	**4.73**	2.4416E-05	2.4174E-05	**0.99**
6	499.2	329.6	685.3	165.6	9.3000	5.4681	**20.01**	1.2370E-05	1.2171E-05	**1.61**
2	500.7	330.2	671.9	166.2	9.3000	19.270	**52.06**	4.1892E-06	3.9596E-06	**5.48**
23	495.1	317.0	771.1	153.0	3.0300	2.8191	**1.10**	5.6260E-05	5.5934E-05	**0.58**
25	498.4	794.1	288.9	124.9	15.780	7.4000	**43.76**	6.4857E-05	6.4515E-05	**0.53**

When the data was adjusted, the percent error in the estimated leak location did decrease but not to an "accurate" percent error (less than 5%). The EBA-EKF technique required the pressure sensors to be ± 1.18Psi; however it was determined that the adjusted pressure sensors were approximately ± 10.25Psi. Therefore the pressure measurements were not accurate enough for the model to be able to accurately predict the leak location. Also the sensors may have had some drift which would significantly affect the experimental results. It can be concluded that the current EBA-EKF technique relies heavily on accurate pressure measurements and therefore field implementation is not a feasible option.

Chapter 8: Conclusions and Recommendations

8.1 Project Summary

The objective of this project was to determine if the leak detection model proposed by R. Lesyshen could be feasibly implemented on a real water distribution system. As the human population continues to grow and expand throughout the world, higher demands are placed on the water supply system. These demands are increasing every day and are threatening the quantity and quality of the water throughout the world. One area of specific concern is the amount of water loss through underground water distribution systems. It is not uncommon for a water distribution system to loose 20-30% of the pumped water through leakages. Water managers have many different options for leak detection including steady-state mass volume analysis, acoustic analysis and ground penetrating radar to name a few. Each of these techniques have positive and negative aspects and must be analyzed to determine if they are suitable for a given application.

The leak detection algorithm developed in this research is only applicable to single pipelines. The model of the pipeline was developed from the momentum and continuity equations and the Method of Characteristics was used to determine the pressure and flow at each node along the pipeline. The Extended Kalman Filter was applied to determine the magnitude of two "fictitious" leaks at two known locations along the pipeline. In order to ensure the model was observable four head measurements were required from four equally spaced locations along the pipeline. However, for the new model (the Extended Boundary Approach) only the head and flow measurements were required at the actual upstream and downstream locations because the pipeline could be "Extended" and the head measurements could be calculated at the "fictitious" upstream and downstream locations. Once the two "fictitious" leaks were determined, the magnitude and location of the actual leak was determined from the concept of equivalent systems. The magnitude of the actual leak is simply the sum of the two leakage estimates and the location of the actual leak is determined from a linear interpolation.

Simulated data was used to determine if the model could accurately detect the leak magnitude and location in an underground water pipeline. Using only the Extended Kalman Filter technique, the model could only accurately detect the magnitude and location of a simulated leak in the centre one third of the pipe. Therefore the EBA-EKF technique was developed and the model could accurately detect a leak, as small as 5% of the total flow, along the entire pipeline. When the leak was 10% of the total flow the standard deviation in the leak location was 689.86m (1.10% error) compared to 53556m (87.97% error) when the leak was 1% of the total flow. When the 20% leak was located at varying locations along the pipeline the average standard deviation was 1048.6m (1.97% error).

Real data was collected from the pipeline located in Melfort; however, the model was not able to accurately detect the location and magnitude of the leak. Since the pressure measurements were collected for about 25 minutes only the transient information was collected in the measurements. It was believed that the model was not using the transient information and therefore simulations were performed to verify this assumption.

It was determined that the model was not able to use the transient information provided from the upstream perturbations. The most accurate estimates of the leak location and magnitude occurred when there were no perturbations added to the upstream reservoir. Since only the steady state information was required, a sensitivity analysis was done to determine the level of accuracy required for the head measurements in order to produce accurate results. It was determined that for the large pipeline described in this thesis the pressure measurements must be ± 0.01m which translates into ± 96Pa (± 0.014Psi). Since this level of accuracy is impractical, a smaller pipeline was investigated.

For the small pipeline, the theoretical average standard deviation of the leak location for a 20% leak was 0.31869m (1.66% error). A sensitivity analysis was conducted for the small pipeline, and it was determined that the accuracy required by the pressure measurements was ± 1.0m which translates into ± 8142Pa (± 1.2Psi). A small pipeline was constructed to determine if the problem with the accuracy of the sensor could be

overcome by "adjusting" the measurements. It was determined that any drift along with inaccurate measurements caused the model to produce inaccurate estimates and could not be overcome. Therefore implementing this leak detection method is not practical on a real water distribution system.

8.2 Conclusions

Water is becoming more precious as the human population grows. Water distribution systems typically loose over 20% of the pumped water through pipeline leakage. Also contamination may enter the pipeline through leakages and therefore the need to reduce leakages in water distribution systems is extremely important for the well being of everyone.

Based on the simulated and limited experimental work, the following conclusions are drawn:

1. It can be concluded, that for this leak detection algorithm, it is crucial that the model must be observable. Once four pressure measurements were inputted into the EKF (making the procedure fully observable), the model was able to accurately detect the location and magnitude of a leak in a pipeline. In Lesyshen's research, observability was not satisfied and a solution could only be found with extensive tuning. When applied to a wide variety of leakage locations, Lesyshen's algorithm was very limited in what it could do, even under simulated conditions. The issue of observability should not be overlooked if the Kalman Filter or Extended Kalman Filter are being implemented.

2. It is concluded that the proposed EBA-EKA could accurately predict the location of leaks in both simulated large and small pipelines provided the level of significant figures in the head values were greater than three.

3. It is concluded that the proposed EBA-EKA in its existing form is not feasible for practical implementation for leak detection. The leak detection algorithm in this work proved to be able to theoretically determine the leak location and magnitude for any leak

in a simulated pipeline that was 5% of the total flow or larger. However, practical implementation of the leak detection algorithm is limited due to the level of accuracy required from the sensors. For a large pipeline as described in this thesis, the pressure sensors must be accurate to ± 96Pa (± 0.014Psi). Also for a small pipeline as described in this thesis, the pressure sensors must be accurate to ± 8142Pa (± 1.2Psi). Since these levels of accuracy are not practical for today's technology the implementation of this leak detection algorithm is not recommended.

4. It was concluded that the EBA-EKF approach did not make use of the transient information input into it but relied on steady state information. This was a result that was contrary to what was expected based on other studies using the EKF. This is an area that needs much more in-depth study.

8.3 Recommendations for Future Work

Although the leak detection methods in this thesis did not allow practical implementation it has provided a launching pad for future work.

A leak detection model that is capable of using the dynamic transient information may be a feasible option. This may eliminate the issue with the accuracy of the sensors and allow the model to be implemented on a real pipeline. It was not fully understood why the transient information could not be used successfully by the EBA-EKF because the perturbations were continuously present both in the simulated and experimental work. It is recommended that this problem be investigated in greater detail.

The algorithm in this thesis may be used to detect drift or inaccuracies in sensors. A pipeline that is known to have zero leakages (i.e. visible line in a chemical plant) and stable conditions may be instrumented with sensors. The necessary sensor measurements can be inputted into the model and when the estimates from model begin to change it can be determined that the sensors are drifting or becoming inaccurate.

A leak detection method that is capable of locating multiple leaks would be more practical. Possibly a technique that uses multiple filters would allow the model to differentiate between multiple leaks.

References List

1.) American Water Works Association (AWWA) Water Loss Control Committee "Applying worldwide BMPs in water loss control." J. Am. Water Works Assoc., 95(8), 65-79. (2003).

2.) Benkherouf, A. and Allidina, A.Y., "Leak detection and location in a gas pipeline" Institution of Electrical Engineers (IEE) Proceedings. (1988). Vol 135. Pg 142-148.

3.) Bishop G., Welch G., An Introduction to the Kalman Filter. (2001). http://www.cs.unc.edu/~tracker/media/pdf/SIGGRAPH2001_CoursePack_08.pdf

4.) Brown R.G. and Hwang P.Y.C., Introduction to Random Signal and Applied Kalman Filtering. John Wiley & Sons, Inc., Toronto (1997).

5.) Brunone B., and Ferrante M., "Detecting leaks in pressurized pipes by means of transients." Journal of Hydraulic Research. Vol 39. 539-547. (2001).

6.) Brunone B., "Transient Test-Based Techniques for Leak Detection in Outfall Pipes" Journal of Water Resources Planning and Management Vol. 125, 302-306 (1999).

7.) Buchberger S. G. and Nadimpalli G., "Leak Estimation in Water Distribution Systems by Statistical Analysis of Flow Readings." Journal of Water Resources Planning And Management, Vol 130, 321-329. (2004).

8.) Cao H., Parameter Estimation Using Extended Kalman Filter for the Swash Plate Assembly and the Control Piston in a Load Sensing Pump. (2001). University of Saskatchewan.

9.) Carpentier P. and Cohen G., "State estimation and leak detection in water distribution networks," Civil Engineering Systems Vol. 8, 247-257 (1991).

10.) Cesario L., "Modeling, Analysis, and Design of Water Distribution Systems." American Water Works Association, Denver Colorado. (1995).

11.) Chaudhry M.H., Applied Hydraulic Transients (Van Nostrand Reinhold Company Inc., New York 1987).

12.) Cheong, L.C. Unaccounted for water and the economics of leak detection. Proceedings of the 18th International Water Supply Congress and Exhibition, 15-31 May 1991, Copenhagen, published in *Water Supply,* 9:3&4:IR1.1, (1991).

13.) Clark R.M., "Urban Drinking Water Distribution systems: AUS Perspective" Proc. Conf. Water Conservation, Water Supply, and Systems Integration, Valencia, Spain. (1998).

14.) Cykowski, E.E., “Pipeline Hydraulic Transient and Leak Detection Analysis”. American Society of Civil Engineers. (1984). Pg 42-59.

15.) Environment Canada. (2006). http://www.ec.gc.ca/water/en/info/facts/e_intro.htm

16.) European Commission. World Water Day 2002. http://ec.europa.eu/research/dossier/do220302/index_en.html.

17.) Franklin G., Powell D., and Workman M., Digital Control of Dynamic Systems, Addison Wesley Publishing, (1990).

18.) Gleick P.H., “Human Population and Water: To the Limits in the 21st Century” *Pacific Institute for Studies in Development, Environment, and Security, Oakland, California.* (1995). http://www.aaas.org/international/ehn/fisheries/gleick.htm

19.) Grewal M.S., Andrews A.P., *Kalman Filtering-Theory and Practice* Prentice Hall, Englewood Cliffs, New Jersey, 1993

20.) Guru, M.V., Horne, J.E., “The Ogallala Aquifer.” The Kerr Center for Sustainable Agriculture Inc.(2000). http://www.kerrcenter.com/publications/ogallala_aquifer.pdf

21.) Hargesheimer E. E., “Identifying Water Main Leaks With Trihalomethane Tracers” American Water Works Association Journal. Vol 77, Pg 71-75. November (1985).

22.) Hunaidi O. and Giamou P. “Ground-Penetrating Radar for Detection of Leaks in Buried Plastic Water Distribution Pipes.” 1998. Lawrence, Kansas, USA, Seventh International Conference on Ground-Penetrating Radar.(1998).

23.) International Water Management Institute. “*World water supply and demand*” Colombo, Sri Lanka. (2000). http://www.iwmi.cgiar.org/pubs/WWVisn/WWSDHtml.htm

24.) Isermann R., "Process Fault Detection Based on Modeling and Estimation Methods - A Survey," Automatica Vol. 20, 387-404 (1983).

25.) Jowitt P. W., and Xu C., “Optimal valve control in water distribution networks.” J. Hydraul. Eng., 120(8), 934-955.(1990).

26.) Kalman R.E., "A New Approach to Linear Filtering and Prediction Problems," Transaction of the ASME - Journal of Basic Engineering 82(Series D), 35-45 (1960).

27.) Lee P.J., Vitkovsky J.P., Lambert M.F., Simpson A.R., Ligett J.A., “ Frequency Domain Analysis for Detecting Pipeline Leaks” Journal of Hydraulic Engineering. (2005). Vol 131. 596-604.

28.) Lennart J. L., “Computer and Laboratory studies of Leak Detection using Hydraulic Transients.” Proceedings of the EWRA 95 Symposium: Nicosia, Cyprus, March 14-18 (1995).

29.) Lesyshen R., “Water Distribution Line Leak Detection using Extended Kalman Filtering.” (2005). Thesis. University of Saskatchewan

30.) Liggett J.A., and Chung C.L., "Inverse Transient Analysis in Pipe Networks," Journal of Hydraulic Engineering Vol. 120, 934-955 (1994).

31.) Maybeck P. S., Stochastic models, estimation, and control. Vol1. Academic Press. United States of America. (1979).

32.) Misiunas D., Vitkovsky J., Olsson G., Simpson A., Lambert M., “Pipeline Break Detection Using Pressure Transient Monitoring.” Journal of Water Resources Planning and Management. Vol 131, 316-325. (2005).

33.) Mpesha W., Gassman S.L., and Chaudhry M.H., "Leak Detection in Pipes by Frequency Response Method," Journal of Hydraulic Engineering Vol. 127, 134-147 (2001).

34.) Ogata K., Modern Contro, Engineering, Fourth Ed. Prentice Hall, (2002)

35.) O’Brien G.G., Hyman M.A., and Kaplan S. “A Study of the Numerical Solution of Partial Differential Equations” J. Mathematical Physics, Vol 29, (1951).

36.) Pelletier G., Maihot A., and Villeneuve J., “Modeling Water Pipe Breaks – Three Case Studies,” Journal of Water Resource Planning and Management. Vol. 129, 115-123. (2003).

37.) Prime Kermit. U.S. “Water Infrastructure AT RISK” Heating/Piping/Air Conditioning Engineering. Vol 73, Iss 2, Pg 9. (July-Dec 2001).

38.) Pudar R.S., and Liggett, J.A., "Leaks in Pipe Networks," Journal of Hydraulic Engineering Vol. 118, 1031-1046 (1992).

39.) Rollins K., Freh J., Tate D., Zachariah O., “Resource Valuation and Public Policy: Consumers’ Willingness to Improving Water Service Infrastructure” Canadian Water Resources Journal. Vol 22, Pg 185-195. April (1997).

40.) Smith L.A., Fields K. A., Chen, A. S. C., and Tafuri A. N., “Options for leak and break detection and repair of drinking water systems,” Batelle, Columbus, Ohio. (2000).

41.) Tafuri A.N., “Locating Leaks with acoustic technology” American Water Works Association Journal. Vol 92, Pg 57-66. July (2000).

42.) Thompson W.C.Jr.,“Leak Detection on Petroleum Pipelines.” Proceedings of the International School of Hydrocarbon Measurement, 85-91 (1985).

43.) U.S. Environmental Protection Agency (USEPA). “1999 drinking water infrastructure needs survey,” EPA Fact sheet-816-F-01-001, Washington, D.C. (2001).

44.) Verde C., “Minimal Order Nonlinear Observer for Leak Detection,” Journal of Dynamic Systems, Measurement and Control. Vol. 126, 467-472 (2004).

45.) Verde, C. “Multi-leak detection and isolation in fluid pipelines,” Control Engineering Practice. Vol. 9, 673-682 (2000).

46.) Vitkovsky J.P., Simpson A.R., and Lambert M.F., "Leak Detection and Calibration Using Transients and Genetic Algorithms," Journal of Water Resources Planning and Management Vol. 126, 262-265 (2000).

47.) Wang X., Lambert M., Simpson A., “Detection and Location of a Partial Blockage in a Pipeline Using Damping of Fluid Transients.” Journal of Water Resources Planning and Management. Vol 131, 244-249. (2005).

48.) Wang X., Lambert M.F., Simpson A.R., Liggett J.A., and Vitkovsky J.P., "Leak Detection in Pipelines using the Damping of Fluid Transients," Journal of Hydraulic Engineering Vol. 128, 697-711 (2002).

49.) Water Quality and Health Council. (2006). http://www.waterandhealth.org/drinkingwater/safewater.html

50.) Watters G.Z., Analysis and Control of Unsteady Flow in Pipelines (Butterworth Publishers, Stoneham, MA 1984).

51.) Williams G.S., and Hazen A., Hydraulic Tables. 3rd Edition. New York. Wiley. (1947).

52.) Wikipedia. The Free Encyclopedia. (2006). http://en.wikipedia.org/wiki/Water_resources

47.) Willsky A.S., "A Survey of Design Methods for Failure Detection in Dynamic Systems," Automatica Vol. 12, 601-611 (1976).

52.) Wylie E.B. and Streeter V.L., Fluid Transients (Feb Press, Ann Arbor, Michigan 1983).

Appendix A: Upstream Boundary Condition

The upstream boundary condition is modeled as a constant-level upstream reservoir. In actuality there is a pump located at the upstream boundary; however the pump is "online" and determining the pump characteristic curve is not possible. The following equations verify the assumptions of modeling the upstream boundary condition as a constant-level reservoir instead of a pump.

A.1 Equations for a Constant-Level Upstream Reservoir

The equations for the constant-level upstream reservoir come from Chaundhry, 1987. The head at the entrance of the pipe is given by Equation [A.1].

$$H_{1,k} = H_{R1} - (1+K)\frac{Q_{11,k}{}^2}{2gA^2} \qquad [A.1]$$

Equation [A.1] and the negative characteristic Equation [2.56] are solved simultaneously to eliminate $H_{1,k}$ and the resulting equation for flow is:

$$Q_{11,k} = \frac{-1+\sqrt{1+4k_1\left(C_m + C_a H_{R1}\right)}}{2k_1} \qquad [A.2]$$

where:

$$k_1 = \frac{C_a\left(1+K\right)}{2gA^2} \qquad [A.3]$$

H_{R1} represents the head at the supply reservoir

K is the coefficient entrance loss

C_m = Constant from Method of Characteristics (Equation 2.60)

$$C_a = \frac{gA}{a} \qquad [A.4]$$

A.2 Equations for a Constant Pressure Pump at Upstream End

It is assumed that the pump is pressure compensated to provide a constant head. This is a reasonable assumption given the actual pump unit on the Melfort line. The equations for the constant pressure pump come from Chaundhry, 1987. The head at the entrance of the pump is given by Equation [A.5].

$$H_{1,k} = H_{SH} - C_8 Q_{11,k}{}^2 \qquad [A.5]$$

where:

H_{SH} = shut-off head (head when there is no discharge),

$$C_8 = \frac{\left(H_{SH} - H_r\right)}{Q_r^2} \qquad [A.6]$$

H_r = Head at maximum pump efficiency
Q_r = Flow at maximum pump efficiency

Solving Equation [A.5] and the negative characteristic Equation [2.56] simultaneously results in :

$$Q_{11,k} = \frac{-1 + \sqrt{1 + 4C_a C_8 \left(C_m + C_a H_{SH}\right)}}{2C_a C_8} \qquad [A.7]$$

Therefore from Equations [A.1] and [A.5]

$$C_8 = \frac{1+K}{2gA^2}, \text{ assuming that } H_{R1} = H_{SH} \qquad [A.8]$$

Also from Equation [A.2] and [A.7]

$$k_1 = C_a C_8 \qquad [A.9]$$

Equating Equation [A.3] and [A.9] results in

$$C_a C_8 = \frac{C_a\left(1+K\right)}{2gA^2} \qquad [A.10]$$

which is the same as Equation [A.8], that is

$$C_8 = \frac{1+K}{2gA^2} \qquad [A.8]$$

Therefore with the assumption that the head reservoir is equal to the shut-off head, a constant pressure pump may be modeled as a constant level reservoir.

Appendix B: Probability and Statistics

This section includes a basic introduction to probability and random numbers. "*Random signals cannot be described with explicit mathematical functions like sine waves or step functions*" (Brown, 1997). Therefore random numbers must be explained through probability. Information is extracted from noisy signals by using the Kalman Filter. Noise is treated as a random signal and therefore a basic understanding of probability is required. This section will cover expectation, averages, variance and covariance.

B.1 Probability

Probability can be explained as the percent chance that one event will occur over another assuming that the outcome is random. Mathematically probability can be explained as:

$$p(A) = \frac{Possible\ outcomes\ favoring\ event\ A}{Total\ number\ of\ \ possible\ outcomes} \qquad [B.1]$$

B.2 Probability with Random Variables

A random variable may be describes as a function that assigns real numbers to all points within the sample space. For a continuous random variable the probability of a single event occurring is $p(A) = 0$. This is because there are infinite number of outcomes and the chance of one event occurring is zero. Probability is used to define the event of encountering a range of numbers and is referred to as the cumulative distribution.

The function, $F_X(x)$, is a cumulative function because it is the probability of all events up to and including x occurring, or:

$$F_X(x) = p(-\infty, x], \qquad [B.2]$$

where,

$$F_X(x) \to 0 \ \ as\, x \to -\infty,$$

$$F_X(x) \to 1 \ \ as\, x \to \infty.$$

The probability density function is defined as the derivative of the cumulative distribution function and is shown below as:

$$f_X(x) = \frac{d}{dx} F_X(x), \qquad \text{[B.3]}$$

where

1. $f_X(x)$ is a non-negative function
2. $\int_{-\infty}^{\infty} f_X(x)dx = 1.$

The probability over a defined interval is determined by integrating the probability density function over the given interval. This is given as:

$$p_X[a,b] = \int_a^b f_X(x)dx. \qquad \text{[B.4]}$$

B.3 Mean and Variance

For N samples the mean or average value of a discrete sample space is given as:

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_N}{N}, \qquad \text{[B.5]}$$

$\overline{X}$ denotes the average of the sample X_1, X_2,….The expected value of X may be described as:

$$Expected\,value\,of\ X = E(X) = \sum_{i=1}^{n} p_i x_i\ , \qquad \text{[B.6]}$$

where the probability p_i essentially weights the importance of each discrete outcome x_i within the summation of all possible realizations n. Similarly for the continuous random variable the expected value is given as:

$$Expected\ value\ of\ X = E(X) = \int_{-\infty}^{\infty} x f_X(x)dx\,. \qquad [B.7]$$

Similarly functions of the random variable X are defined as:

$$E\big(g(x)\big) = \int_{-\infty}^{\infty} g(x) f_X(x)dx\,. \qquad [B.8]$$

Equation [B.8] is known as the first statistical moment. Variance is derived from the second statistical moment which is shown below:

$$E\big(X^2\big) = \int_{-\infty}^{\infty} x^2 f_X(x)dx\,. \qquad [B.9]$$

B.4 Variance

Variance is the second moment of a function that is the difference between the random variable and the expected value. Mathematically the function is defined as:

$$g(x) = X - E(X) \qquad [B.10]$$

Therefore the variance is defined as:

$$\begin{aligned} Variance\ of\ X = E\big(g(x)^2\big) &= E\big[(X - E(X))^2\big] \\ &= E(X^2) - E(X)^2 \end{aligned}. \qquad [B.11]$$

Variance is a very useful statistical property for random signals. The magnitude of the variance is the "noise" or "jitter" and is described as the dispersion that exists in a random signal. Another useful statistical expression is known as the standard deviation and is the square root of the variance.

$$\text{Standard deviation of } X = \sigma_X = \sqrt{Variance\ of\ X}\,. \qquad [B.12]$$

B.5 Covariance

The covariance of two random processes, X and Y is a measure of correlation between the two variables. The covariance of X and Y is given as:

$$Cov\,of\ X\ and\ Y = \sigma_{XY} = E\left[\left(X - \overline{X}\right)\left(Y - \overline{Y}\right)\right], \qquad [B.13]$$

where,

$\overline{X}$ is the mean of X,
$\overline{Y}$ is the mean of Y.

The covariance represents the dispersion extent of two random variables. If two functions have a covariance of 1 then they are highly correlated and if they are completely uncorrelated their covariance will be equal to zero.

Appendix C: MatLab Code

The following MatLab code is used to produce pressure measurements (the plant model) and then estimate the leak location and magnitude (using the EKF technique). The equations in the code are explained throughout the thesis and include the transient pipe flow equations developed in Chapter 2, the Extended Kalman Filter Equations developed in Chapter 3 and the pipe flow equations developed in Chapter 4.

C.1 Plant Model

```
%Set up Initial Variable Values
ro = 980; % density [kg/m^3]
K = 2.1994e9; % Fluid modulus of elasticity [Pa]
c1 = 1; % constant assuming pipe anchored with expansion joints throughout
g = 9.811; % gravity [m/sec^2]

%Set up Node Data
numnode = 25;
node = [1:numnode];
consmp = [-.04,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0.04]; %set the steady state consumption for
each node
hgl =
[138.8,138.668,138.536,138.404,138.272,138.140,138.008,137.876,137.744,137.480,137.348,137.216,137.
084,136.952,136.820,136.688,136.556,136.424,136.292,136.160,136.028,135.896,135.764,135.632,135.5];
% Hydraulic grade line
device = [1,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,2,1]; %Boundary condition settings

%Set up Pipedata
numpipe= 24;
minnumreach = 1;
pipe = 1:numpipe;
upstreamnode =
[node(1),node(2),node(3),node(4),node(5),node(6),node(7),node(8),node(9),node(10),node(11),node(12),no
de(13),node(14),node(15),node(16),node(17),node(18),node(19),node(20),node(21),node(22),node(23),nod
e(24)];  %in order [pipe1,pipe2,....pipeN]
downstreamnode =
[node(2),node(3),node(4),node(5),node(6),node(7),node(8),node(9),node(10),node(11),node(12),node(13),n
ode(14),node(15),node(16),node(17),node(18),node(19),node(20),node(21),node(22),node(23),node(24),no
de(25)]; %in order [pipe1,pipe2,....pipeN]
upstreampipe =
[pipe(1),pipe(1),pipe(2),pipe(3),pipe(4),pipe(5),pipe(6),pipe(3),pipe(4),pipe(5),pipe(6),pipe(7),pipe(8),pipe(
9),pipe(10),pipe(11),pipe(12),pipe(13),pipe(14),pipe(15),pipe(16),pipe(17),pipe(18),pipe(19),pipe(20),pipe(
21),pipe(22),pipe(23),pipe(24)]; %set the pipe to the node... in order [node1,node2,...nodeN]
lengthpipe(1:24) = [2182.58];
f(1:24) = [0.01598]; %friction factor in each pipe, assumed constant
[minlength,index] = min(lengthpipe);
diameterpipe(1:24) = [0.508];
E(1:24) = [1.965e11]; %Youngs modulus of elastisity (Steel pipe) [Pa]
e(1:24) = [0.00556];    %Pipe wall thickness [m]
for pipe = 1:numpipe
```

```
    a(pipe) = sqrt((K/ro)/(1+((K/E(pipe))*(diameterpipe(pipe)/e(pipe)))*c1)); % Wave speed within the pipe
    A(pipe) = pi/4*diameterpipe(pipe)^2;
    B(pipe) = a(pipe)/(g*A(pipe));     %Wave speed constant used frequently
end

%Set up Time Duration
dt = (minlength/minnumreach)/a(index);
t(1) = 0;
time = 1;
for kurtis = 1:300
for pipe = 1:numpipe
    dx(pipe)=a(pipe)*dt;
    R(pipe) = f(pipe)*dx(pipe)/(2*g*diameterpipe(pipe)*A(pipe)^2); %Resistance constant for each pipe
    j(pipe) = lengthpipe(pipe)/dx(pipe)+1; %Number of grid points in each pipe
    leak(1:numnode) = 0.0;

%Steady State Solver
    LQ(pipe,1:j(pipe))
=(diameterpipe(pipe)^2.5/(0.0826*f(pipe)*lengthpipe(pipe))^0.5)*sqrt(hgl(upstreamnode(pipe))-
(hgl(downstreamnode(pipe))));%-0.5*consmp(1)^2/(2*g*A(pipe)^2)
    LQ1 = LQ;
End
 %end for pipe = 1:numpipe
    for node = 1:numnode
        uppipes = find(downstreamnode==node); %determine how many pipes are connected to node in
question
        downpipes = find(upstreamnode==node);
        if device(node)==1
            consmp(node) = sign(consmp(node))*LQ(upstreampipe(node),1);
        else %if device(d)==2
            if ~isempty(uppipes)
                upQ = 0;
                upQH = 0;
            for Bc1count = 1:length(uppipes)
                upQ = upQ + LQ(uppipes(Bc1count),1);
                upQH = upQH + abs(LQ(uppipes(Bc1count),1)/(hgl(upstreamnode(uppipes(Bc1count)))-
hgl(downstreamnode(uppipes(Bc1count)))));
            end
            else
                upQ = 0;
                upQH = 0;
            end
            if ~isempty(downpipes)
                dwnQ = 0;
                dwnQH = 0;
            for Bc2count = 1:length(downpipes)
                dwnQ = dwnQ + LQ(downpipes(Bc2count),1);
                dwnQH = dwnQH +
abs(LQ(downpipes(Bc2count),1)/(hgl(upstreamnode(downpipes(Bc2count)))-
hgl(downstreamnode(downpipes(Bc2count)))));
            end
            else
                dwnQ = 0;
                dwnQH = 0;
            end
```

```
        hgl(node) = hgl(node) + ((upQ - dwnQ) - consmp(node))/(.5*(upQH + dwnQH));%this is Newton's
METHOD
      end %end if device(node)==1
  end %end for node = 1:numnode
end % for kurtis
for pipe = 1:numpipe
   for i = 1
      LH(pipe,i) = (hgl(upstreamnode(pipe)) - (LQ(pipe,1))^2);
      hgl(upstreamnode(pipe)) = LH(pipe,i);
   end
   for i = 2:j(pipe)-1
      LH(pipe,i) = (hgl(upstreamnode(pipe))-(i-1)*R(pipe)*(LQ(pipe,1))^2);
      LH1 = LH;
   end
   for i = j(pipe)
      LH(pipe,i) = hgl(downstreamnode(pipe)) + (LQ(pipe,j(pipe))^2);
      hgl(downstreamnode(pipe)) = LH(pipe,i);
   end
 end
H = LH;
Q = LQ;
Lleak = leak;

%Valve Stuff
numtimesteps = 10000;
Tau = ones(numnode,numtimesteps);
Ts = 180000; %Start of Valve Closure
leakonset = 1;
leaky(1:numnode) = 0.0;
leaky(13) = 0.00075;
Tc = floor(20/dt); %Time of valve closure (number of time steps)

for time = 1:numtimesteps

if time == leakonset
   Lleak = leaky; % Turn on leak
   leak = leaky;
end

%Calculate Integration Constants
 for pipe = 1:numpipe
   for i = 1
      Bm(pipe,i) = B(pipe);
      Cm(pipe,i) = LH(pipe,i+1) - (B(pipe) - R(pipe)*abs(LQ(pipe,i+1)))*LQ(pipe,i+1);
   end
   for i = 2:(j(pipe)-1)
      Bm(pipe,i) = B(pipe);
      Cm(pipe,i) = LH(pipe,i+1) - LQ(pipe,i+1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i+1)));
      Bp(pipe,i) = B(pipe);
      Cp(pipe,i) = LH(pipe,i-1) + LQ(pipe,i-1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i-1)));
      Q(pipe,i) = (Cp(pipe,i) - Cm(pipe,i))/(Bp(pipe,i) + Bm(pipe,i)); %Calculate MOC for Q & H @
interior elements.
      H(pipe,i) = Cp(pipe,i) - Bp(pipe,i)*Q(pipe,i);
   end
   for i = j(pipe)
      Bp(pipe,i) = B(pipe);
```

```
        Cp(pipe,i) = LH(pipe,i-1) + LQ(pipe,i-1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i-1)));
        Bm(pipe,i) = 0;
        Cm(pipe,i) = 0;
    end
end

%Simple and Ordinary One node boundary Conditions
%Calculate Bc and Cc
for node = 1:numnode
    uppipes = find(upstreamnode==node); %determine how many pipes are connected to node in question
    downpipes = find(downstreamnode==node);
    if ~isempty(uppipes)
        upBc(node) = 0;
        for Bc1count = 1:length(uppipes)
            upBc(node) = upBc(node) + 1/Bm(uppipes(Bc1count),1);
        end
        upCc(node) = 0;
        for Cc1count = 1:length(uppipes)
            upCc(node) = upCc(node) + Cm(uppipes(Cc1count),1)/Bm(uppipes(Cc1count),1);
        end %end for loop for Cc1count
    else
        upBc(node) = 0;
        upCc(node) = 0;
    end %end if uppipes
    if ~isempty(downpipes)
        downBc(node) = 0;
        for Bc2count = 1:length(downpipes)
            downBc(node) = downBc(node) + 1/Bp(downpipes(Bc2count),j(downpipes(Bc2count)));
        end
        downCc(node) = 0;
        for Cc2count = 1:length(downpipes)
            downCc(node) = downCc(node) +
Cp(downpipes(Cc2count),j(downpipes(Cc2count)))/Bp(downpipes(Cc2count),j(downpipes(Cc2count)));
        end
    else
        downBc(node) = 0;
        downCc(node) = 0;
    end %end if downpipes
    Bc(node) = (upBc(node) + downBc(node))^-1;
    Cc(node) = Bc(node)*(upCc(node) + downCc(node));

%Boundary Conditions
if device(node) == 1
    for Bc1count = 1:length(uppipes)
        HR = 138.8 + 0.2*randn(1); %Reservoir Head
if LQ(1,2) > 0
        Q(uppipes(Bc1count),1) = (-B(pipe) + sqrt(B(pipe)^2 -
4*((1+0.5)/(2*g*A(pipe)^2))*(Cm(uppipes(Bc1count),1) - HR)))/((1+0.5)/(g*A(pipe)^2));
        H(uppipes(Bc1count),1) = HR - (1 + 0.5)*Q(uppipes(Bc1count),1)^2/(2*g*A(pipe)^2); %Sharp exit
minor loss (Energy equation)
    else
        H(uppipes(Bc1count),1) = HR;
        Q(uppipes(Bc1count),1) = (H(uppipes(Bc1count),1) -
Cm(uppipes(Bc1count),1))/Bm(uppipes(Bc1count),1); % from method of characteristics
    end
    end
```

```
   for Bc2count = 1:length(downpipes)
      Cv = (consmp(node)*Tau(node,time))^2/(1*consmp(node)^2/(2*g*A(pipe)^2));%(0.5);
      Q(downpipes(Bc2count),j(downpipes(Bc2count))) = (-
Cv*Bc(node)+sqrt((Cv*Bc(node))^2+4*Cv*Cc(node)-4*Cv*(135.5)))/2;
      H(downpipes(Bc2count),j(downpipes(Bc2count))) = LH(downpipes(Bc2count),j(pipe)-1) +
B(pipe)*LQ(downpipes(Bc2count),j(pipe)-1) - R(pipe)*LQ(downpipes(Bc2count),j(pipe)-
1)*abs(LQ(downpipes(Bc2count),j(pipe)-1)) - B(pipe)*Q(downpipes(Bc2count),j(downpipes(Bc2count)));
   end

 %The Inner Nodes
   else %LEAK NODE
   for Bc2count = 1:length(downpipes)
      H(downpipes(Bc2count),j(downpipes(Bc2count))) = 1/8*B(pipe)^2*Lleak(node)^2 +
1/2*(LH(downpipes(Bc2count),j(downpipes(Bc2count))-1) +
B(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1) -
R(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-
1)*abs(LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1)) + LH(downpipes(Bc2count)+1,2) -
B(pipe)*LQ(downpipes(Bc2count)+1,2) +
R(pipe)*LQ(downpipes(Bc2count)+1,2)*abs(LQ(downpipes(Bc2count)+1,2))) -
1/8*B(pipe)*Lleak(node)*sqrt(B(pipe)^2*Lleak(node)^2 +
8*(LH(downpipes(Bc2count),j(downpipes(Bc2count))-1) +
B(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1) -
R(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-
1)*abs(LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1)) + LH(downpipes(Bc2count)+1,2) -
B(pipe)*LQ(downpipes(Bc2count)+1,2) +
R(pipe)*LQ(downpipes(Bc2count)+1,2)*abs(LQ(downpipes(Bc2count)+1,2))));
   end
   for Bc1count = 1:length(uppipes)
      H(uppipes(Bc1count),1) = 1/8*B(pipe)^2*Lleak(node)^2 + 1/2*(LH(uppipes(Bc1count)-
1,j(uppipes(Bc1count)-1)-1) + B(pipe)*LQ(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1) -
R(pipe)*LQ(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1)*abs(LQ(uppipes(Bc1count)-
1,j(uppipes(Bc1count)-1)-1)) + LH(uppipes(Bc1count),2) - B(pipe)*LQ(uppipes(Bc1count),2) +
R(pipe)*LQ(uppipes(Bc1count),2)*abs(LQ(uppipes(Bc1count),2))) -
1/8*B(pipe)*Lleak(node)*sqrt(B(pipe)^2*Lleak(node)^2 + 8*(LH(uppipes(Bc1count)-
1,j(uppipes(Bc1count)-1)-1) + B(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1) -
R(pipe)*LQ(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1)*abs(LQ(uppipes(Bc1count)-
1,j(uppipes(Bc1count)-1)-1)) + LH(uppipes(Bc1count),2) - B(pipe)*LQ(uppipes(Bc1count),2) +
R(pipe)*LQ(uppipes(Bc1count),2)*abs(LQ(uppipes(Bc1count),2))));
      Q(uppipes(Bc1count),1) = (H(uppipes(Bc1count),1) - LH(uppipes(Bc1count),2) +
B(pipe)*LQ(uppipes(Bc1count),2) -
R(pipe)*LQ(uppipes(Bc1count),2)*abs(LQ(uppipes(Bc1count),2)))/B(pipe);
      Q(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)) = Q(uppipes(Bc1count),1) +
Lleak(node)*sqrt(abs(H(uppipes(Bc1count),1)))*sign(H(uppipes(Bc1count),1));
end
end
end

XH1(time,:,:) = H; %Storage of information
XQ1(time,:,:) = Q;
XCv(time) = Cv;
LH = H;
LQ = Q;
Lleak = leak;
t(time+1) = t(time) + dt;
time = time + 1;
end %end of time stepping
```

```
pt = t(2:end); %plot timer
if numtimesteps == 10000
   Profile(:,:)=LH(:,1);
   Profile(25,1)=LH(24,2);
End
```

C.2 Filtering Code (EKF Technique)

```
%Set up Initial Variable Values
ro = 980; % density [kg/m^3]
K = 2.1994e9; % Fluid modulus of elastisity [Pa]
c1 = 1; % constant assuming pipe anchored with expansion joints throughout
g = 9.811; % gravity [m/sec^2]

%Set up Node Data
numnode = 4;
node = [1:numnode];
consmp = [-.04,0,0,0.04]; %set the steady state consumption for each node
hgl = [138.8,137.7,136.6,135.5]; % Hydraulic grade line
device = [1,2,2,1]; %Boundary condition settings

%Set up Pipedata
numpipe= 3;
minnumreach = 1;
pipe = 1:numpipe;
upstreamnode = [node(1),node(2),node(3)];  %in order [pipe1,pipe2,....pipeN]
downstreamnode = [node(2),node(3),node(4)]; %in order [pipe1,pipe2,....pipeN]
upstreampipe = [pipe(1),pipe(1),pipe(2),pipe(3)]; %set the pipe to the node... in order
[node1,node2,...nodeN]
lengthpipe(1:3) = [17460.67];
f(1:3) = [0.01598]; %friction factor in each pipe, assumed constant
[minlength,index] = min(lengthpipe);
diameterpipe(1:3) = [0.508];
E(1:3) = [1.965e11]; %Youngs modulus of elastisity (Steel pipe) [Pa]
e(1:3) = [0.00556];    %Pipe wall thickness [m]
for pipe = 1:numpipe
   a(pipe) = sqrt((K/ro)/(1+((K/E(pipe))*(diameterpipe(pipe)/e(pipe)))*c1)); % Wave speed within the pipe
   A(pipe) = pi/4*diameterpipe(pipe)^2;
   B(pipe) = a(pipe)/(g*A(pipe));     %Wave speed constant used frequently
end

%Set up Time Duration
dt = (minlength/minnumreach)/a(index);
t(1) = 0;
time = 1;

for kurtis = 1:300
for pipe = 1:numpipe
   dx(pipe)=a(pipe)*dt;
   R(pipe) = f(pipe)*dx(pipe)/(2*g*diameterpipe(pipe)*A(pipe)^2); %Resistance constant for each pipe
   j(pipe) = lengthpipe(pipe)/dx(pipe)+1; %Number of grid points in each pipe
   leak(1:numnode) = 0.0;

%Steady State Solver (NEWTONS METHOD)
   LQ(pipe,1:j(pipe))
=(diameterpipe(pipe)^2.5/(0.0826*f(pipe)*lengthpipe(pipe))^0.5)*sqrt(hgl(upstreamnode(pipe))-
(hgl(downstreamnode(pipe))));%-0.5*consmp(1)^2/(2*g*A(pipe)^2)
```

```
    LQ1 = LQ;
end %end for pipe = 1:numpipe
    for node = 1:numnode
        uppipes = find(downstreamnode==node); %determine how many pipes are connected to node in
question
        downpipes = find(upstreamnode==node);
        if device(node)==1
            consmp(node) = sign(consmp(node))*LQ(upstreampipe(node),1);
        else %if device(d)==2
            if ~isempty(uppipes)
                upQ = 0;
                upQH = 0;
            for Bc1count = 1:length(uppipes)
                upQ = upQ + LQ(uppipes(Bc1count),1);
                upQH = upQH + abs(LQ(uppipes(Bc1count),1)/(hgl(upstreamnode(uppipes(Bc1count)))-
hgl(downstreamnode(uppipes(Bc1count)))));
            end
            else
                upQ = 0;
                upQH = 0;
            end
            if ~isempty(downpipes)
                dwnQ = 0;
                dwnQH = 0;
            for Bc2count = 1:length(downpipes)
                dwnQ = dwnQ + LQ(downpipes(Bc2count),1);
                dwnQH = dwnQH +
abs(LQ(downpipes(Bc2count),1)/(hgl(upstreamnode(downpipes(Bc2count)))-
hgl(downstreamnode(downpipes(Bc2count)))));
            end
            else
                dwnQ = 0;
                dwnQH = 0;
            end
            hgl(node) = hgl(node) + ((upQ - dwnQ) - consmp(node))/(.5*(upQH + dwnQH));%this is Newton's
METHOD
        end %end if device(node)==1
    end %end for node = 1:numnode
end % for kurtis
for pipe = 1:numpipe
    for i = 1
        LH(pipe,i) = (hgl(upstreamnode(pipe)) - (LQ(pipe,1))^2);
        hgl(upstreamnode(pipe)) = LH(pipe,i);
    end
    for i = 2:j(pipe)-1
        LH(pipe,i) = (hgl(upstreamnode(pipe))-(i-1)*R(pipe)*(LQ(pipe,1))^2);
        LH1 = LH;
    end
    for i = j(pipe)
        LH(pipe,i) = hgl(downstreamnode(pipe)) + (LQ(pipe,j(pipe))^2);
        hgl(downstreamnode(pipe)) = LH(pipe,i);
    end
 end
H = LH;
Q = LQ;
Lleak = leak;
```

```
%-------------------------------------------
%                    SETUP
%-------------------------------------------
load('Melfort20%Leak@26191');

%-------------------------------------------
%                 Initial Conditions
%-------------------------------------------
Xest =
[LH(1,1);LH(2,1);LH(3,1);LH(3,2);LQ(1,1);LQ(1,2);LQ(2,1);LQ(2,2);LQ(3,1);LQ(3,2);Lleak(2);Lleak(3)]
;
P = eye(12)*1e5;            %Initial error covariance matrix

Qx = zeros(12);
for i = 1:4
   Qx(i,i) = 1e-5; %noise in head equation
end
for i = 5:10
   Qx(i,i) = 5e-6; %noise in flow equation
end
for i = 11:12
   Qx(i,i) = 5e-7; %noise in leakage estimate
end

P = P.*Qx;       %Adjust initial P matrix according to system noise
              %matrix (scaling initial error covariance to system
              %noise matrix)

Rx = eye(4);                %measurement noise covariance matrix
Rx(1,1) = 2e-1;              %noise of upstream head measurement
Rx(2,2) = 2e-1;              %noise of  head measurement @node2
Rx(3,3) = 2e-1;              %noise of  head measurement @node3
Rx(4,4) = 2e-1;              %noise of downstream head measurement

Hx = zeros(4,12);          %System output matrix
Hx(1,1) = 1;
Hx(2,2) = 1;
Hx(3,3) = 1;
Hx(4,4) = 1;

F = zeros(12);             %Set up state transistion matrix

%-------------------------------------------
%                 Parameter Estimation
%-------------------------------------------
   t(1) = 0;
      iterations =10000;
   for counter = 1:iterations
% Measurements and Inputs are entered here for each time step
      yact(1,1) = XH1(counter,1,1) + 0.2*randn;
      yact(2,1) = XH1(counter,9,1) + 0.2*randn;
      yact(3,1) = XH1(counter,17,1) + 0.2*randn;
      yact(4,1) = XH1(counter,24,2) + 0.2*randn;
      U1 = 138.8;
      U2 = 135.5;
      U3 = 0.80608;
```

```
% The block below defines the linearized state transition matrix (F)

%LH(1,1) Derivatives
for i = 1
F(i,2) = (-B(pipe)+sqrt(B(pipe)^2-3*(LH(2,1)-B(pipe)*LQ(1,2)+R(pipe)*LQ(1,2)^2-
U1)/(g*A(pipe)^2)))/(sqrt(B(pipe)^2-3*(LH(2,1)-B(pipe)*LQ(1,2)+R(pipe)*LQ(1,2)^2-
U1)/(g*A(pipe)^2)));
F(i,6) = (-B(pipe)+sqrt(B(pipe)^2-3*(LH(2,1)-B(pipe)*LQ(1,2)+R(pipe)*LQ(1,2)^2-
U1)/(g*A(pipe)^2)))*(-B(pipe)+2*R(pipe)*LQ(1,2))/(sqrt(B(pipe)^2-3*(LH(2,1)-
B(pipe)*LQ(1,2)+R(pipe)*LQ(1,2)^2-U1)/(g*A(pipe)^2)));
end
%LH(2,1) Derivatives
for i = 2
F(i,1) = 1/2;
F(i,3) = 1/2;
F(i,5) = 1/2*(B(2) - 2*R(2)*LQ(1,1));
F(i,8) = 1/2*(-B(2) + 2*R(2)*LQ(2,2));
F(i,11) = -1/2*B(2);
end
%LH(3,1) Derivatives
for i = 3
F(i,2) = 1/2;
F(i,4) = 1/2;
F(i,7) = 1/2*(B(3) - 2*R(3)*LQ(2,1));
F(i,10) = 1/2*(-B(3) + 2*R(3)*LQ(3,2));
F(i,12) = -1/2*B(3);
end
%LH(3,2) Derivatives
for i = 4
F(i,3) = 1-B(pipe)*U3/(sqrt(U3^2*B(pipe)^2+4*U3*LH(3,1)+4*U3*LQ(3,1)*B(pipe)-
4*U3*R(pipe)*LQ(3,1)^2-4*U3*U2));
F(i,9) = B(pipe)-2*R(pipe)*LQ(3,1)-1/4*B(pipe)*(4*U3*B(pipe)-
8*U3*R(pipe)*LQ(3,1))/(sqrt(U3^2*B(pipe)^2+4*U3*LH(3,1)+4*U3*LQ(3,1)*B(pipe)-
4*U3*R(pipe)*LQ(3,1)^2-4*U3*U2));
end
%LQ(1,1) Derivatives
for i = 5
F(i,2) = -1/(sqrt(B(pipe)^2-3*(LH(2,1)-B(pipe)*LQ(1,2)+R(pipe)*LQ(1,2)^2-U1)/(g*A(pipe)^2)));
F(i,6) = -(-B(pipe)+2*R(pipe)*LQ(1,2))/(sqrt(B(pipe)^2-3*(LH(2,1)-
B(pipe)*LQ(1,2)+R(pipe)*LQ(1,2)^2-U1)/(g*A(pipe)^2)));
end
%LQ(1,2) Derivatives
for i = 6
F(i,1) = 1/(2*B(1));
F(i,3) = -1/(2*B(1));
F(i,5) = 1/(2*B(1))*(B(1) - 2*R(1)*LQ(1,1));
F(i,8) = 1/(2*B(1))*(B(2) - 2*R(2)*LQ(2,2));
F(i,11) = 1/2;
end
%LQ(2,1) Derivatives
for i = 7
F(i,1) = 1/(2*B(1));
F(i,3) = -1/(2*B(1));
F(i,5) = 1/(2*B(1))*(B(1) - 2*R(1)*LQ(1,1));
F(i,8) = 1/(2*B(1))*(B(2) - 2*R(2)*LQ(2,2));
F(i,11) = -1/2;
```

```
end
%LQ(2,2) Derivatives
for i = 8
F(i,2) = 1/(2*B(2));
F(i,4) = -1/(2*B(2));
F(i,7) = 1/(2*B(2))*(B(2) - 2*R(1)*LQ(2,1));
F(i,10) = 1/(2*B(2))*(B(3) - 2*R(3)*LQ(3,2));
F(i,12) = 1/2;
end
%LQ(3,1) Derivatives
for i = 9
F(i,2) = 1/(2*B(2));
F(i,4) = -1/(2*B(2));
F(i,7) = 1/(2*B(2))*(B(2) - 2*R(1)*LQ(2,1));
F(i,10) = 1/(2*B(2))*(B(3) - 2*R(3)*LQ(3,2));
F(i,12) = -1/2;
end
%LQ(3,2) Derivatives
for i = 10
F(i,3) = U3/(sqrt(U3^2*B(pipe)^2+4*U3*LH(3,1)+4*U3*LQ(3,1)*B(pipe)-4*U3*R(pipe)*LQ(3,1)^2-
4*U3*U2));
F(i,9) = 1/4*(4*U3*B(pipe)-
8*U3*R(pipe)*LQ(3,1))/(sqrt(U3^2*B(pipe)^2+4*U3*LH(3,1)+4*U3*LQ(3,1)*B(pipe)-
4*U3*R(pipe)*LQ(3,1)^2-4*U3*U2));
end
%Lleak(2) Derivatives
for i = 11
F(i,11) = 1;
end
%Lleak(3) Derivatives
for i = 12
F(i,12) = 1;
end
%-----------------------------------------------
%   Update P and Calculate Kalman Gain K
%-----------------------------------------------

      P = F*P*F' + Qx;        %Calculate unrefined P matrix
      S = Hx*P*Hx' + Rx;
      K = P*Hx'/S;            %Calculate the Kalman Gain
      P = (eye(12) - K*Hx)*P;         %Calculate the refined P matrix

%-------------------------------------------------
% calculate states from original (nonlinear) model
%-------------------------------------------------

%Calculate Integration Constants
 for pipe = 1:numpipe
   for i = 1
      Bm(pipe,i) = B(pipe);
      Cm(pipe,i) = LH(pipe,i+1) - (B(pipe) - R(pipe)*abs(LQ(pipe,i+1)))*LQ(pipe,i+1);
   end
   for i = 2:(j(pipe)-1)
      Bm(pipe,i) = B(pipe);
      Cm(pipe,i) = LH(pipe,i+1) - LQ(pipe,i+1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i+1)));
      Bp(pipe,i) = B(pipe);
```

```
        Cp(pipe,i) = LH(pipe,i-1) + LQ(pipe,i-1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i-1)));
        Q(pipe,i) = (Cp(pipe,i) - Cm(pipe,i))/(Bp(pipe,i) + Bm(pipe,i)); %Calculate MOC for Q & H @
interior elements.
        H(pipe,i) = Cp(pipe,i) - Bp(pipe,i)*Q(pipe,i);
    end
    for i = j(pipe)
        Bp(pipe,i) = B(pipe);
        Cp(pipe,i) = LH(pipe,i-1) + LQ(pipe,i-1)*(B(pipe) - R(pipe)*abs(LQ(pipe,i-1)));
        Bm(pipe,i) = 0;
        Cm(pipe,i) = 0;
    end
end

%Simple and Ordinary One node boundary Conditions
%Calculate Bc and Cc
for node = 1:numnode
    uppipes = find(upstreamnode==node); %determine how many pipes are connected to node in question
    downpipes = find(downstreamnode==node);
    if ~isempty(uppipes)
        upBc(node) = 0;
        for Bc1count = 1:length(uppipes)
            upBc(node) = upBc(node) + 1/Bm(uppipes(Bc1count),1);
        end
        upCc(node) = 0;
        for Cc1count = 1:length(uppipes)
            upCc(node) = upCc(node) + Cm(uppipes(Cc1count),1)/Bm(uppipes(Cc1count),1);
        end %end for loop for Cc1count
    else
        upBc(node) = 0;
        upCc(node) = 0;
    end %end if uppipes
    if ~isempty(downpipes)
        downBc(node) = 0;
        for Bc2count = 1:length(downpipes)
            downBc(node) = downBc(node) + 1/Bp(downpipes(Bc2count),j(downpipes(Bc2count)));
        end
        downCc(node) = 0;
        for Cc2count = 1:length(downpipes)
            downCc(node) = downCc(node) +
Cp(downpipes(Cc2count),j(downpipes(Cc2count)))/Bp(downpipes(Cc2count),j(downpipes(Cc2count)));
        end
    else
        downBc(node) = 0;
        downCc(node) = 0;
    end %end if downpipes
    Bc(node) = (upBc(node) + downBc(node))^-1;
    Cc(node) = Bc(node)*(upCc(node) + downCc(node));

%Boundary Conditions
if device(node) == 1
    for Bc1count = 1:length(uppipes)
        U1 = 138.8;
        U2 = 135.5;
        U3 = 0.80608;
  if LQ(1,2) > 0
```

```
        Q(uppipes(Bc1count),1) = (-B(pipe) + sqrt(B(pipe)^2 -
4*((1+0.5)/(2*g*A(pipe)^2))*(Cm(uppipes(Bc1count),1) - U1)))/((1+0.5)/(g*A(pipe)^2));
        H(uppipes(Bc1count),1) = (U1 - (1 + 0.5)*Q(uppipes(Bc1count),1)^2/(2*g*A(pipe)^2)); %Sharp exit
minor loss (Energy equation)
    else
        H(uppipes(Bc1count),1) = U1;
        Q(uppipes(Bc1count),1) = (H(uppipes(Bc1count),1) -
Cm(uppipes(Bc1count),1))/Bm(uppipes(Bc1count),1); % from method of characteristics
    end
    end
    for Bc2count = 1:length(downpipes)
        Q(downpipes(Bc2count),j(downpipes(Bc2count))) = (-
U3*Bc(node)+sqrt((U3*Bc(node))^2+4*U3*Cc(node)-4*U3*(U2)))/2; %
       H(downpipes(Bc2count),j(downpipes(Bc2count))) = LH(downpipes(Bc2count),j(pipe)-1) +
B(pipe)*LQ(downpipes(Bc2count),j(pipe)-1) - R(pipe)*LQ(downpipes(Bc2count),j(pipe)-
1)*abs(LQ(downpipes(Bc2count),j(pipe)-1)) - B(pipe)*Q(downpipes(Bc2count),j(downpipes(Bc2count)));
    end

 %The Inner Nodes
    else %LEAK NODE
    for Bc2count = 1:length(downpipes)
        H(downpipes(Bc2count),j(downpipes(Bc2count))) =
1/2*(LH(downpipes(Bc2count),j(downpipes(Bc2count))-1) +
B(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1) -
R(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-
1)*abs(LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1)) + LH(downpipes(Bc2count)+1,2) -
B(pipe)*LQ(downpipes(Bc2count)+1,2) +
R(pipe)*LQ(downpipes(Bc2count)+1,2)*abs(LQ(downpipes(Bc2count)+1,2)) - B(pipe)*Lleak(node));
        %Q(downpipes(Bc2count),j(downpipes(Bc2count))) = 0.5*(LH(downpipes(Bc2count)+1,2) +
B(pipe)*LQ(downpipes(Bc2count)+1,2) -
R(pipe)*LQ(downpipes(Bc2count)+1,2)*abs(LQ(downpipes(Bc2count)+1,2)) -
LH(downpipes(Bc2count),j(downpipes(Bc2count))-1) +
B(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1) -
R(pipe)*LQ(downpipes(Bc2count),j(downpipes(Bc2count))-
1)*abs(LQ(downpipes(Bc2count),j(downpipes(Bc2count))-1)) + B(pipe)*Lleak(node))/B(pipe);
    end
    for Bc1count = 1:length(uppipes)
        H(uppipes(Bc1count),1) = 1/2*(LH(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1) +
B(pipe)*LQ(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1) - R(pipe)*LQ(uppipes(Bc1count)-
1,j(uppipes(Bc1count)-1)-1)*abs(LQ(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1)) +
LH(uppipes(Bc1count),2) - B(pipe)*LQ(uppipes(Bc1count),2) +
R(pipe)*LQ(uppipes(Bc1count),2)*abs(LQ(uppipes(Bc1count),2)) - B(pipe)*Lleak(node));
        Q(uppipes(Bc1count),1) = 1/2*(LH(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1) +
B(pipe)*LQ(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1) - R(pipe)*LQ(uppipes(Bc1count)-
1,j(uppipes(Bc1count)-1)-1)*abs(LQ(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)-1)) -
LH(uppipes(Bc1count),2) + B(pipe)*LQ(uppipes(Bc1count),2) -
R(pipe)*LQ(uppipes(Bc1count),2)*abs(LQ(uppipes(Bc1count),2)) - B(pipe)*Lleak(node))/B(pipe);
        Q(uppipes(Bc1count)-1,j(uppipes(Bc1count)-1)) = Q(uppipes(Bc1count),1) + Lleak(node);
        leak(node) = Lleak(node);

end
end
end

if counter == 30000
    Profile1(:,:)=LH(:,1);
```

```
   Profile1(4,1)=LH(3,2);
end

        LH = H;
        LQ = Q;
        Lleak = leak;
        XXest = F*Xest;
        Xest(1) = H(1,1);
        Xest(2) = H(2,1);
        Xest(3) = H(3,1);
        Xest(4) = H(3,2);
        Xest(5) = Q(1,1);
        Xest(6) = Q(1,2);
        Xest(7) = Q(2,1);
        Xest(8) = Q(2,2);
        Xest(9) = Q(3,1);
        Xest(10) = Q(3,2);
        Xest(11) = leak(2);
        Xest(12) = leak(3);
        Xest(1:4) = Xest(1:4) + K(1:4,:)*(yact - Hx*Xest);
        Xest(11:12) = Xest(11:12) + K(11:12,:)*(yact - Hx*Xest);
        LH(1,1) = Xest(1);
        LH(2,1) = Xest(2);
        LH(3,1) = Xest(3);
        LH(3,2) = Xest(4);
        LQ(1,1) = Xest(5);
        LQ(1,2) = Xest(6);
        LQ(2,1) = Xest(7);
        LQ(2,2) = Xest(8);
        LQ(3,1) = Xest(9);
        LQ(3,2) = Xest(10);
        Lleak(2) = Xest(11);
        Lleak(3) = Xest(12);


Xh(time,:,:) = H; %Storage of information
Xh1(time,:,:)=Xh(time,1,1);
Xh2(time,:,:)=Xh(time,2,1);
Xh3(time,:,:)=Xh(time,3,1);
Xh4(time,:,:)=Xh(time,3,2);

        Xhead(counter,1,:) = Xest(1:4);
        XQ(counter,1,:) = Xest(5:10);
        Xleak1(counter,:) = Xest(11);
        Xleak2(counter,:) = Xest(12);
        Xyact(counter,:,:) = yact;
        leakmag(counter) = Xest(11) + Xest(12);
        leakposition(counter) = (17460.67*Xest(11) + 34921.3*Xest(12))/leakmag(counter);
        t(counter+1) = t(counter) + dt;
        pt = t(2:end);


Z(1:4,1:12)=0;
Z(1,1)=1;
Z(2,2)=1;
Z(3,3)=1;
```

```
Z(4,4)=1;

Obs(1:36,1:12)=1;
Obs(1:4,1:12)=[Z];
Obs(5:8,1:12)=[Z*F];
Obs(9:12,1:12)=[Z*F^2];
Obs(13:16,1:12)=[Z*F^3];
Obs(17:20,1:12)=[Z*F^4];
Obs(21:24,1:12)=[Z*F^5];
Obs(25:28,1:12)=[Z*F^6];
Obs(29:32,1:12)=[Z*F^7];
Obs(33:36,1:12)=[Z*F^8];
Obs(37:40,1:12)=[Z*F^9];
Obs(41:44,1:12)=[Z*F^10];
Obs(45:48,1:12)=[Z*F^11];
Rank=rank(Obs);
Rank1(counter,1,:)=Rank;

   end

meanQL1 = mean(Xleak1(5000:10000));
meanQL2 = mean(Xleak2(5000:10000));
meanpos = mean(leakposition(5000:10000));
meanmag = mean(leakmag(5000:10000));
```