A Probabilistic Approach for the Design of an Early Warning Source Water Monitoring Station

by

Heather Patricia Mustard

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Applied Science

in

Civil Engineering

Waterloo, Ontario, Canada, 2007

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis involves the design of an early warning source water monitoring station for a riverine source of drinking water. These stations provide downstream water utilities with advanced notification of contamination events so they have time in which to implement a response, such as closing their intakes.

Many threats facing riverine water supplies, such as accidental spills, are uncertain in nature. Therefore, designing a monitoring station for the detection of these events requires a probabilistic modelling approach. Sources of uncertainty considered in this research include the location, mass and duration of a spill event as well as the flow at the time of the spill and the water quality model parameters. Probability distributions for each of these uncertainties were defined and a Monte Carlo experiment was conducted.

The design objectives include maximizing the probability of detection and maximizing the probability of having a threshold amount of warning time. These objectives are in conflict with each other because the probability of detection improves as the station moves closer to the intake and the amount of warning time increases as the station is located further upstream. Values for the competing objectives were calculated for a number of potential monitoring station locations at multiple sample intervals and the tradeoff solutions were analyzed.

This methodology was applied to the Hidden Valley Intake which services the Regional Municipality of Waterloo's Mannheim Water Treatment Plant. The Hidden Valley Intake is located in Kitchener, Ontario and withdraws up to 72 ML of water per day from the Grand River.

Based on an analysis of the Monte Carlo simulation results for the case study application, it was found that locating the monitoring station near the Victoria Street Bridge, approximately 11 km upstream of the intake, represents the best tradeoff in the design objectives. Sampling at least once per hour is recommended to increase the amount of warning time.

The impact of various sources of uncertainty was also explored in this thesis. It was found that the flow at the time of a spill and the spill location are the only sources of uncertainty that significantly impact the probability distributions of relevant model results.

Acknowledgements

I would like to thank my supervisors, Dr. Donald Burn and Dr. Bryan Tolson, for their guidance and support throughout the completion of my Masters degree. I would also like to thank them for the valuable suggestions and feedback they provided for this thesis.

I would like to acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) for funding this research through a Canada Graduate Scholarship.

Some of the graphs and maps presented in this thesis were produced using information under license with the Grand River Conservation Authority (Copyright © Grand River Conservation Authority, 2007). I would like to thank staff at the Grand River Conservation Authority for providing data and assistance, particularly, Mark Anderson, Water Quality Engineer.

I would also like to thank the staff at the Regional Municipality of Waterloo for providing water and wastewater treatment plant data and answering my questions. I would particularly like to thank Amy Domaratzki, Hydrogeologist, for her assistance.

Finally, I would like to thank my family, friends, and fellow grad students who have supported me throughout the completion of my Masters degree. I would especially like to thank my parents, Paul and Jackie, for their support over the years, and my fiancé, John, for his patience and encouragement.

Table of Contents

List of Tables

List of Figures

1 Introduction

1.1 Background

In 1854, London, England was hit by its third epidemic of cholera in 25 years. Physician John Snow's hypothesis that the disease was being spread by the city's drinking water supply and his subsequent recommendation to shut down a contaminated well saved thousands of lives. In addition to beginning the science of epidemiology, Snow's discovery represents one of the modern world's first recognitions of the importance of source water protection.

Today, the ability of water to carry disease is undisputed. The use of disinfection has significantly reduced health risks associated with microbial pathogens (Davies and Mazumder, 2003). As a result, cholera and many other waterborne diseases no longer pose a significant threat in developed countries such as Canada. Sophisticated drinking water treatment technologies, such as membrane filtration and ultraviolet disinfection, are becoming increasingly effective at removing some of the most persistent chemical and microbial contaminants in a water supply. In addition, stringent regulations governing drinking water quality now exist in most developed parts of the world. As a result of these advances, drinking water quality has improved substantially since 1854. However, most of these efforts have been focused on treating source water rather than protecting it. Not until 150 years after Snow's breakthrough was source water protection officially defined by Ontario's Ministry of the Environment as "protecting current and future sources of drinking water from potential contamination and depletion" (Technical Experts Committee, 2004).

The fear of a depleted water supply is far from the minds of most Canadians because Canada is bestowed with nearly 7% of the world's renewable freshwater supply (Environment Canada, 2004). As a result, Canadians are fortunate to have the second most inexpensive water supply in the world (National Utility Service, 2006). However, Canada's abundant and inexpensive freshwater supply is grossly overused. In a 2007 survey of water consumption rates amongst 32 nations, Canada ranked a disappointing 31st with an annual per capita water consumption rate of more than double the average of all other participating nations (Organisation for Economic Co-operation and Development, 2007).

In addition to taking for granted that there will always be enough water, Canadians also take for granted that their water supply will always be safe. Gord Miller, Ontario's Environmental Commissioner, has described the myth of detachment as society's belief that we are not part of the ecosystem and are therefore subject to different rules. However, societal impacts such as wastewater discharges, industrial effluents, agricultural runoff, and negligence continue to pollute the ecosystem. This was clearly evidenced in May 2000 when seven people died and over 2000 became ill in the Town of Walkerton, Ontario. Their drinking water had been contaminated with *E.coli 0157:H7* and *campylobacter* originating from a nearby farm. This event highlighted that Canadians are not immune to adverse drinking water quality.

In response to the tragic events in Walkerton, the Ontario government mandated an inquiry led by Commissioner Dennis O'Connor. Commissioner O'Connor released a report in 2002 in which he identified the need for watershed-based source protection planning. As a result, The Clean Water Act (Ontario Ministry of the Environment, 2006b) was created for the purpose of "protecting existing and future sources of drinking water". According to Ontario's Environment Minister, Laurel Broten, this new legislation is an integral part of the Province's multi-barrier approach for protecting drinking water from source to tap (Ontario Ministry of the Environment, 2006c). The legislation requires that every watershed in the province develop a local, science-driven source protection plan that identifies and mitigates risks to drinking water quality and quantity.

Investing in source water protection results in a lower risk of acute and chronic health problems as a result of adverse water quality, in addition to decreased treatment requirements, and fewer treatment residuals and by-products (Gostin, 2000; Davies and Mazumder, 2003). For these reasons, source water protection is a prudent management decision not only from an environmental and public health perspective but also from a financial one (Davies and Mazumder, 2003). With alarming rates of water consumption and a growing number of threats to drinking water quality, source water protection is essential if Canadians want to ensure an adequate supply of safe drinking water for future generations.

1.2 Scope of Research

Source water protection involves issues relating to the quantity and quality of drinking water supplies. The scope of this thesis is limited to surface water quality, and more specifically to riverine sources of drinking water. Most water quality threats facing riverine water supplies are due to unintentional spills or discharges such as wastewater bypasses, transportation accidents, and agricultural runoff. All of these threats can cause an immediate deterioration in water quality and may arrive at downstream water treatment plant intakes within minutes or hours.

Although the purpose of source water protection is to identify and mitigate potential threats to prevent them from ever entering a source water supply, accidental spills cannot be entirely avoided. Early warning source water monitoring stations can provide an additional barrier in the multi-barrier approach for producing clean drinking water. The positioning of such a station upstream of a drinking water intake can provide downstream water utilities with advanced notice of contamination events, allowing them sufficient time to implement a response. Even though contamination cannot always be prevented from entering a source water body, it can be prevented from entering a drinking water treatment plant with the use of an early warning source water monitoring station.

The purpose of this research is to design a source water monitoring station upstream of a riverine drinking water treatment plant intake. Since spills are inherently uncertain in nature, a large part of this research involved probabilistic modelling of spill scenarios using Monte Carlo simulations. The results of these simulations were analyzed at a number of potential monitoring station locations and sampling intervals. This research also involved an examination of the sources of uncertainty impacting the design of a source water monitoring station. The resulting methodology was applied to a case study example.

1.1 Organization of Thesis

This thesis is organized into the following chapters:

Chapter 1: This chapter introduces background information about source water protection and describes the scope of this research.

Chapter 2: This chapter provides a review of literature relevant to this research.

Chapter 3: This chapter presents an overview of the methodology developed for designing an early warning source water monitoring station.

Chapter 4: This chapter discusses background information about the Case Study application, the Hidden Valley Intake, which is located in the Grand River Watershed and services the Mannheim Water Treatment Plant in Kitchener, Ontario.

Chapter 5: This chapter presents and analyzes the results for locating a source water monitoring station upstream of the Hidden Valley Intake. Graphs and discussion of the results at potential monitoring stations are presented. A discussion of the most significant sources of uncertainty affecting the decision making process is also included.

Chapter 6: This chapter provides general conclusions and recommendations regarding the importance of source water monitoring and protection. Specific conclusions and recommendations related to the case study example are also presented.

2 Literature Review

The literature relevant to this thesis is organized into the following five topics, which are discussed in Sections 2.1 to 2.5:

- Water quality monitoring;
- Hydrodynamic modelling;
- Water quality modelling;
- Uncertainty; and
- Multi-objective optimization.

2.1 Water Quality Monitoring

In the following section, various objectives of water quality monitoring are discussed. Early warning source water monitoring is described in further detail along with a description of general and probabilistic design considerations.

2.1.1 Purposes of Water Quality Monitoring

Prior to designing a water quality monitoring program, its purposes must be clearly identified (Palmer and MacKenzie, 1985; Harmancioglu and Alpaslan, 1992; Dixon and Chiswell, 1996). Since water quality monitoring can be expensive and time consuming, failing to clearly identify the goals of a monitoring program can result in the collection of sub-optimal data that are of little use for decision making. This represents a waste of money and resources that could be better allocated in order to achieve monitoring objectives (Palmer and MacKenzie, 1985).

The United States Environmental Protection Agency (2007) identified the following purposes of water quality monitoring:

1. Characterization of waters and identification of changes or trends in water quality over time;
2. Identification of existing or emerging water quality problems;
3. Collection of information to design pollution prevention or remediation programs;

4. Determination of whether program goals, such as compliance with pollution regulations, are being met; and
5. Response to emergencies, such as spills and floods.

The purpose of an early warning source water monitoring station is similar to Objective 5 listed above. Early warning stations serve to detect sudden changes in water quality due to low probability but high impact events, such as spills (Grayman et al., 2001). This type of monitoring is described further in the following section.

2.1.2 Early Warning Source Water Monitoring

Grayman and Males (2002) define an early warning monitoring system as "a mechanism for detecting, characterizing and providing notification of a source water contamination event." These systems typically consist of one or more monitoring stations upstream of a drinking water treatment plant intake and provide advanced warning of contamination caused by events such as industrial spills, wastewater bypasses, and transportation accidents (International Life Sciences Institute, 1999). Benefits of early warning monitoring stations include improved decision making, reduced risks of adverse drinking water quality, increased public confidence in a water supply, and increased motivation for dischargers to follow reporting regulations (International Life Sciences Institute, 1999; Gullick et al., 2003; Mikol et al., 2007).

2.1.2.1 Monitoring Technologies

A wide range of monitoring technologies can be used at early warning monitoring stations. Basic online monitors usually operate continuously and measure common parameters such as temperature, conductivity, pH, dissolved oxygen, and turbidity (Grayman et al., 2001). However, basic monitors are unable to detect many types of spill events (Grayman et al., 2001). Advanced analytical methods, such as gas chromatography/mass spectrometry (GCMS) and liquid chromatography/mass spectrometry (LCMS), can be used to detect a wide array of contaminants such as organics, fluorescence for oils, and immunoassays for herbicides (International Life Sciences Institute, 1999). Video surveillance at road and railway crossings can also be used for early warning monitoring stations on large rivers (International Life Sciences Institute, 1999).

Biomonitors are another option for source water quality monitoring. A biomonitor consists of a population of bivalves, fish, zooplankton, or algae that are exposed to the source water. Their behaviours, such as swimming patterns, ventilation rates, and avoidance patterns, can be continuously monitored to provide notification of adverse source water quality (Grayman et al., 2001; Mikol et al., 2007). Biomonitors are commonly used in European and Asian nations but their use in North America is more limited (International Life Sciences Institute, 1999; Grayman et al., 2001). Biomonitors are sensitive to a wide array of organic and inorganic chemicals; however, they are unable to identify the specific type of contamination (Mikol et al., 2007). Biomonitors may also react to substances that are not harmful to humans. Another drawback is that biomonitors may not identify some types of contamination events fast enough, particularly for chemicals with low acute toxicity (Mikol et al., 2007). Other considerations related to biomonitors include cost, care, and feeding (Grayman et al., 2001).

2.1.2.2 Existing Early Warning Systems

Grayman et al. (2001) provide an excellent review of existing early warning monitoring stations around the world that use a diversity of technologies. For example, monitoring stations on the Danube River in Europe use conventional analyzers (e.g., pH, turbidity) and stations on the Mississippi and Ohio Rivers utilize advanced methods such as gas chromatographs. Both the Moselle River in France and the River Han in Korea make use of biomonitors (Grayman et al., 2001; Gullick et al., 2003). Some monitoring systems are multi-jurisdictional and have many monitoring stations (e.g., the Danube River) and others involve a single station upstream of a specific intake (e.g., the River Trent in the United Kingdom). All of these applications and others are described in great detail by Grayman et al. (2001).

2.1.3 Water Quality Monitoring Design Considerations

Sanders et al. (1983) presented a general methodology for the design of a water quality monitoring system from the identification of its purpose through to its operating and reporting procedures. This methodology is illustrated in Figure 2-1.

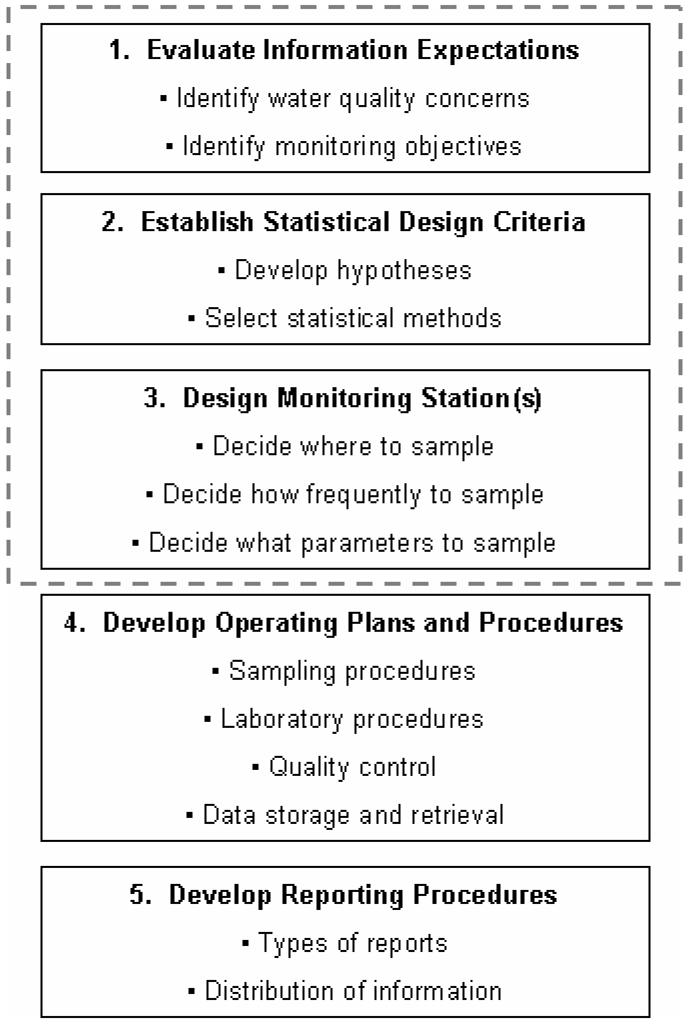

Figure 2-1. The Design of a Water Quality Monitoring System
(Adapted from Sanders et al., 1983)

The first, second, and third steps of the method presented by Sanders et al. are addressed as part of this thesis. The first step involves establishing expectations and objectives; in the case of the current research, the objective is to create a monitoring station or network of stations to provide notice of contamination events upstream of a drinking water intake. The second step involves establishing statistical design criteria. The focus of this thesis is the third step which involves determining how many monitoring stations should be implemented, where the station(s) should be located, the frequency of monitoring, and the parameters for which the station(s) will monitor. Literature relating to these decisions is presented in the following sections.

2.1.3.1 Location

Selection of a location is the most important decision in the design of a monitoring station. If samples are collected from a non-representative location all other design decisions, such as sample frequency, are irrelevant (Sanders et al., 1983).

In the past, monitoring locations were selected based upon personal experience, intuition, and expert judgment of the local conditions (Reinelt et al., 1988; Dixon and Chiswell, 1996; Ning and Chang, 2004; Ning and Chang, 2005). These sites generally coincided with streamflow gauges and were frequently located near known industrial and wastewater discharges (Harmancioglu and Alpaslan, 1994). Convenience was also a consideration and locations with easy access to required facilities, such as laboratories, were often selected (Harmancioglu and Alpaslan, 1994).

Most recent literature on monitoring network design involves the use of statistical and optimization methods (MacKenzie et al., 1987; Reinelt et al., 1988; Ning and Chang, 2004). Examples include the use of fuzzy set theory (Ning and Chang, 2004), gradient search algorithms (Palmer and MacKenzie, 1985), and the entropy principle of information (Harmancioglu and Alpaslan, 1992). Design objectives usually involve maximizing some measure of the amount of information (e.g., capability to detect changes, capability to sufficiently represent a given area, statistical power of the network, capability to meet regulatory requirements) (Palmer and MacKenzie, 1985) and minimizing costs or applying a budget constraint (Ning and Chang, 2004).

For the purpose of locating an early warning monitoring station, Gullick et al. (2003) recommend consideration of the following factors:

- The location of contaminant sources;
- The time of travel from these sources to the intake;
- The amount of mixing and dilution that occur;
- The response time of the monitoring instrument;
- The type of treatment process and its capability to handle various contaminants;
- Security to protect the monitoring station; and

- Access to electricity and potentially telephone lines.

The first two considerations are an important part of this thesis. The location of contaminant sources must be considered in an attempt to minimize the risk of failing to detect events originating between the station and a downstream water treatment plant intake (International Life Sciences Institute, 1999). However, travel times are also an important consideration so that sufficient warning is provided in which to implement a response at the water treatment plant (International Life Sciences Institute, 1999).

In addition to determining the monitoring station's location along the river's length, its vertical and lateral positioning should also be considered (Gullick et al., 2003). For small or well mixed rivers, one station located at the river's center or at one of the banks may be adequate. In some applications, however, multiple intakes across the river's width may be necessary to sufficiently characterize the river's water quality (Gullick et al., 2003). The vertical positioning of a monitor may depend upon the type of analysis it performs. For example, a station monitoring for substances with surface slicks, such as oil, will be sensitive to depth (Gullick et al., 2003).

Access to electrical lines is also a consideration for selecting potential station locations because most monitoring technologies require electricity. Depending on the type of communication between the station and the response team, telephone wires may also be required. Furthermore, the station must be accessible to vehicles. As a result, early warning monitoring stations are usually located in close proximity to a bridge (International Life Sciences Institute, 1999).

2.1.3.2 Sample Frequency

Selection of the frequency at which a monitoring station will sample is another important design decision. The chosen sample frequency can significantly impact both operating costs as well as the usefulness of the data collected (Sanders et al., 1983).

In the past, sample frequencies were often selected using professional judgment with consideration for cost constraints (Sanders et al., 1983; Harmancioglu and Alpaslan, 1994). Although this is often still the case, the literature discusses the use of statistical methods for the optimal selection of monitoring frequency. Some of these statistical methods are discussed further in Sanders et al. (1983) and Harmancioglu and Alpaslan (1994); however, most of these methods are relevant for long-term monitoring and trend detection networks. Therefore, these methods will not be discussed further as they are not as relevant for the current research which focuses on the detection of events that cause sudden deterioration in water quality.

In the case of an early warning source water monitoring station, defining a sample frequency should involve consideration of the duration of typical contamination events and their travel time to the intake. Less frequent sampling leads to decreased response time and could also result in the failure to detect events of short duration (Grayman et al., 2001). As cited by Grayman et al. (2001), Waldon et al. (1989) found that a daily frequency was not appropriate for an early warning monitoring station as it failed to detect many spill events. Continuous or near-continuous monitoring is often preferred, particularly for rivers with high velocities, rivers with low dispersion, or for cases where the monitoring station is located at the water treatment plant intake (International Life Sciences Institute, 1999; Grayman et al., 2001).

2.1.3.3 Parameters to Sample

Early warning source water monitoring stations usually monitor for chemical and radioactive threats in a riverine water supply. Microbiological threats are not usually monitored because their analysis requires hours or days (Gullick et al., 2003). Selecting parameters for which to monitor is a challenging task because a vast array of contaminant events are possible. This decision should be largely driven by a threats and vulnerability assessment of the water supply (Grayman et al., 2001).

Similar to other design decisions, the selection of parameters must also give consideration to the monitoring purpose. For example, if the monitoring program exists for the detection of pesticide transport in a water supply, then sampling dissolved oxygen and turbidity may not be sufficient (International Life Sciences Institute, 1999).

The selection of parameters for which to monitor is often based upon budget and technology constraints. Many parameters are quite costly to sample, and indicator variables may be more appropriate. Sanders (1983) suggests investigating relationships between water quality parameters to see if there are correlations which can be used to reduce the number of constituents to be monitored. Another option is the use of biomonitors which can be very effective at identifying adverse water quality associated with a wide range of contaminants (Grayman et al., 2001; Mikol et al., 2007).

2.1.4 Probabilistic Design of Early Warning Monitoring Stations

This section discusses the use of probabilistic modelling in order to assess various monitoring station designs. The use of probabilistic modelling in the design of a source water monitoring station was demonstrated by Waldon et al. (1989) (as cited by Grayman et al. (2001)), who suggested that future research could involve conducting Monte Carlo simulations.

The American Water Works Association Research Foundation (AwwARF) published a report discussing early warning monitoring in detail (Grayman et al., 2001). This report includes a description of a probabilistic model that uses Monte Carlo simulations to analyze various monitoring station designs. The model is called Spill Risk and it includes uncertainties associated with the fact that spills are random in time, location, duration, and quantity. The model also includes uncertainties associated with whether or not the spill is reported by the public or spill generator.

Spill Risk is a one-dimensional advection-dispersion model with the following assumptions:

- Only a single reach is being modelled with constant flow throughout the reach and no tributaries;
- A one-dimensional model is appropriate as concentrations are vertically and laterally averaged;
- The system can be represented with a simple advection-dispersion model;
- Flows are seasonal (up to 12 seasons per year);
- The flow during a spill simulation is constant;
- Each spill is of a single constituent;

- Spills are not seasonal; and
- There is no interaction between spills (Grayman et al., 2001).

Spill events modelled in Spill Risk are defined by the location of the spill generator, the contaminant spilled, the probability of occurrence (e.g., a 500-year event has a probability of 1/500), as well as a distribution for the spill magnitude, duration, and probability of public or agency report. The spill magnitude is assumed to have a triangular distribution defined by the minimum, maximum, and most likely spill amount, and the spill duration is defined as a uniform distribution ranging from minimum to maximum expected durations.

Spill Risk defines the flow at the time of the spill through the use of a shifted log-normal distribution for each season. Within a season the flow is probabilistic but it is constant within a reach at a given time (Grayman et al., 2001).

Potential monitoring designs tested in the Spill Risk model are defined by their location, contaminants they measure, frequency of sampling, and method detection limit (MDL). The monitors initiate a detection when the concentration of the contaminant is above the MDL. The minimum frequency that can be tested with the model is one sample every hour, to simulate near-continuous monitoring (Grayman and Males, 2002).

Once the response is determined by the model, the resulting treated water concentrations are calculated based on user defined treatment efficiencies for the various contaminants (under both normal and enhanced treatment conditions). A set of spill simulations are performed and the output statistics are accumulated. Using all the accumulated results from a set of simulations, the impact of a given monitoring station design is calculated as a function of the duration a contaminant was present in the treated water above the maximum contaminant level (MCL) and the population that was exposed. An effectiveness index is used to compare different monitoring station designs, which is calculated as the impact reduction divided by the impacts associated with taking no action (Grayman et al., 2001).

The design of a station for the purpose of early warning source water monitoring is a relatively new area of research. No further literature relating specifically to the probabilistic design of such stations is currently available.

2.1.4.1 Similarities with Distribution System Monitoring

Early warning monitoring in distribution systems has obvious parallels to the current research as both involve locating monitoring stations for the purpose of detecting highly uncertain events. Similar probabilistic approaches have been applied in the literature for the purpose of monitoring for intentional contamination events in a distribution system.

Cozzolino et al. (2006) presented a Monte Carlo-based method for locating monitoring stations within a distribution network. Their model considers uncertainties associated with user demand on the system, as well as the uncertainty associated with the node at which the contamination enters the system. Equally probable time-varying hydraulic situations are generated and modelled in order to determine an optimal monitoring location.

Ostfeld and Salomons (2005) discuss a similar approach for the design of an early warning distribution system monitoring program. Their method includes uncertainties associated with the volume and location of a deliberate injection, the demand on the system, as well as possible delays in response and the sensitivity of the monitoring equipment. This methodology is further described by Ostfeld and Salomons (2004) and involves the use of a genetic algorithm to optimize station locations in order to identify deliberate terrorist acts in the distribution system. The tradeoff between the number of stations and the probability of detection is also explored in their work.

2.2 Hydrodynamic Modelling

Information about the pathway, volume, and velocity of water is essential in order to determine how contaminants move and behave in it (Martin and McCutcheon, 1999). Therefore, the use of a hydrodynamic model is required for the current research.

2.2.1 Fundamental Equations

Continuity and momentum equations are used in hydrodynamic models to describe variations in flow. The equation of continuity represents the fact that the change in storage (*S*) over time is equal to the difference between inflows (*I*) and outflows (*O*) (Martin and McCutcheon, 1999). Mathematically this can be expressed as:

$$\frac{dS}{dt} = I - O \qquad \textbf{2-1}$$

If river depths and velocities change over time at a given location, as they do for most natural open channels, the flow is classified as unsteady. For unsteady, gradually varied flow in open channels, the momentum equation can be expressed as (Martin and McCutcheon, 1999):

$$\frac{\partial U}{\partial t} + U\frac{\partial U}{\partial x} = -g\frac{\partial h}{\partial x} - \frac{gn^2}{\delta^2 R^{4/3}}U^2 \qquad \textbf{2-2}$$

where U represents the average longitudinal velocity,
t represents time,
x represents longitudinal distance,
g represents acceleration due to gravity,
n represents Manning's roughness coefficient,
δ represents a unit conversion (1 for SI units and 1.49 for English units),
h represents the water surface elevation above a given datum, and
R represents the hydraulic radius.

Equations 2-1 and 2-2 are called the Saint-Venant equations and are used by all models that simulate dynamic water movement in rivers (Chapra, 1997). Unfortunately, there is no closed form solution for these equations so they must be solved numerically.

2.2.2 Solution Techniques

The Saint-Venant equations are commonly solved using finite difference methods. Finite difference methods involve the use of finite quantities to approximate derivatives that cannot be solved analytically. For example, a river can be divided into finite segments over space (Δx) and time (Δt).

Finite difference methods can be classified as explicit or implicit. Explicit methods only involve one unknown (the value of a given segment at the next time step) because they do not consider adjacent segments at the next time step, which will have an impact on the solution (Chapra, 1997). Implicit solution methods do not have this limitation; however, they are more computationally expensive because the spatial derivatives for the next time step for all points along the x-axis must be solved simultaneously (Chapra, 1997; Martin and McCutcheon, 1999).

The weighted four-point implicit method is one of the most commonly used and efficient solution techniques for solving the Saint-Venant equations (Martin and Wool, 2002). It is considered the standard solution method used in many one-dimensional hydraulic models including the U.S. Army Corps of Engineers CE-QUAL-RIV1 and the National Weather Service models DWOPER, DAMBRK and FLDWAV (Martin and McCutcheon, 1999).

To solve for a value at point Z, as shown in Figure 2-2, the values at all four corners of the surrounding box are required.

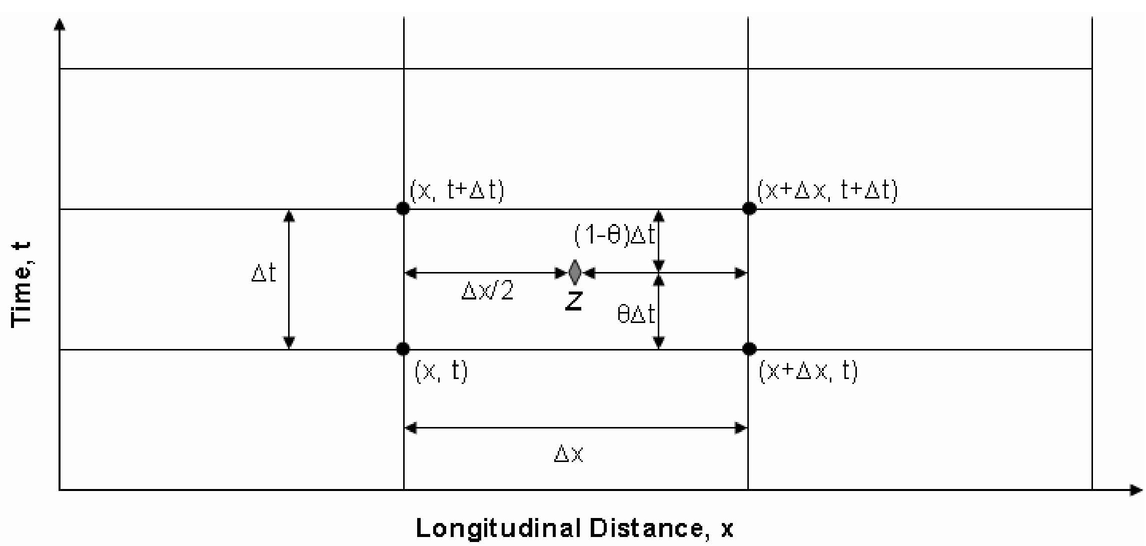

Figure 2-2. Schematic of the Weighted Four Point Implicit Finite Difference Scheme
(Adapted from Martin and McCutcheon, 1999)

A weighting factor, θ, is used to determine the position between two time steps (t and $t+\Delta t$) and its choice is important for the stability of the solution. If θ is set to zero, the method becomes explicit as no values at the future time step are used. In this case, the time step will be restricted by the Courant number which represents the ratio between the distance moved during one time step (Δt) and the segment length (Δx). The Courant number (γ) can be calculated using the following equation (Chapra, 1997):

$$\gamma = \frac{U \Delta t}{\Delta x} \tag{2-3}$$

To ensure stability, the Courant number must be less than one so the water cannot move more than one segment length at a given velocity (U) within one time step.

The solution is stable for 0.5< θ <1, though stability increases as θ approaches values of 1. If θ is set to 1, the solution becomes fully implicit. Martin and McCutcheon (1999) recommend a θ value of 0.6 to improve accuracy while avoiding possible instabilities associated with values closer to 0.5.

The Newton Raphson algorithm is commonly used to solve simultaneous non-linear flow equations, such as those produced using the four-point implicit method. The Newton Raphson method is an optimization technique that can be used to find roots, minimums and maximums of real-valued functions. Martin and McCutcheon (1999) provide an excellent discussion regarding the application of the four-point implicit scheme and use of the Newton-Raphson method to solve the resulting equations.

2.3 Water Quality Modelling

The Law of Conservation of mass states that mass can neither be created nor destroyed, only transferred or transformed. This law is a fundamental basis for most mechanistic water quality models, which are commonly called mass balance models (Chapra, 1997) and can be expressed as (Martin and Wool, 2002):

$$\frac{\partial \alpha}{\partial t} + U\frac{\partial \alpha}{\partial x} = D\frac{\partial^2 \alpha}{\partial x^2} + \frac{q}{A}(\gamma - \alpha) - K_s \alpha + SINKS \tag{2-4}$$

where α represents the water quality constituent of interest,

U represents the average longitudinal velocity,

q represents lateral inflow rate,

A represents cross-sectional area,

γ represents the concentration of runoff input to the channel by distributed flow q,

K_s represents biochemical decay and growth rates, and

SINKS represents biochemical sources or sinks.

Equation 2-4 is known as the one-dimensional advection-dispersion equation and is used in many water quality modelling applications (Kashefipour and Falconer, 2002). The first term in equation 2-4 represents the change in constituent concentration over time, the second term represents advection, the third term represents diffusion, the fourth lateral inflows or withdrawals, the fifth reactions, and the sixth sources or sinks (Martin and Wool, 2002).

2.3.1 One-Dimensional Models

One-dimensional water quality models assume that velocities and concentrations are reasonably represented by cross sectional averages (Martin and McCutcheon, 1999). Although complete cross sectional mixing rarely occurs, it is often a good engineering approximation (Martin and McCutcheon, 1999; Grayman et al., 2001).

Prior to selecting a one-dimensional model, its limitations must be understood. One-dimensional models assume that a contaminant is completely and instantaneously mixed across a given cross section. However, pseudo-complete cross sectional mixing does not occur until some distance downstream from a release. The distance to complete mixing is called the mixing length (Martin and McCutcheon, 1999). Complete mixing is attained when the ratio of minimum to maximum concentration at a given cross section is close to 1; values of 0.95 or 0.98 are commonly used (Rutherford, 1994). For a ratio of 0.98, the mixing length for a mid-channel injection is calculated as:

$$L_m = 0.134 \frac{v_x b^2}{D_y} \qquad \textbf{2-5}$$

For releases on either bank, the mixing length is calculated as (Rutherford, 1994):

$$L_m = 0.536 \frac{v_x b^2}{D_y} \qquad \textbf{2-6}$$

where v_x represents the average velocity (m/s),

b represents the channel width (m), and

D_y is the longitudinal dispersion coefficient (m^2/s).

In general, for a mid-channel release, vertical mixing is complete at a distance of approximately 50 river depths downstream and lateral mixing is complete a distance of 100 to 300 river widths downstream (Rutherford, 1994). For bankside releases, the mixing lengths are four times as long (Rutherford, 1994).

The dominant processes affecting a soluble, conservative spill as it moves downstream are illustrated in Figure 2-3.

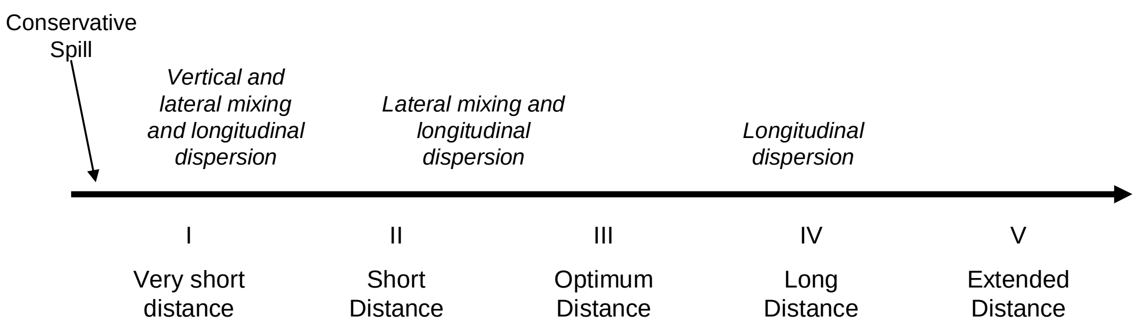

Figure 2-3. Stages of River Mixing
(Adapted from Kilpatrick, 1993)

Prior to reaching section I on Figure 2-3, mixing exists in all three dimensions. At section I, vertical mixing is complete and lateral mixing continues. At some optimum distance (section III), the response curves at various points throughout the section all have equal areas. At this point, dispersion is almost all in one dimension. However, the peak concentration in the center of the river may still be considerably larger than that at the banks. At far distances (section IV), the areas of the response curves are the same and the peak concentration in the center is similar to that at the banks. At this point the dispersion is almost entirely in the longitudinal direction, which continues indefinitely downstream in the absence of any boundaries (Kilpatrick, 1993). Downstream of section III, it is usually reasonable to assume one-dimensional mixing because longitudinal dispersion is the dominant process (Jobson, 1997).

In the case of the current research, the implication of using a one-dimensional model is that more detections will be simulated than would occur in reality. For example, spills that are less than one mixing length upstream of a monitoring station will be detected by the model, but in reality may pass by the monitoring station on one side or the other and go undetected.

2.3.2 Longitudinal Dispersion

Longitudinal dispersion includes the effects of molecular diffusion, turbulent mixing, and mixing due to shear in both the transverse and vertical direction (Singh and Beck, 2003). It represents a measurement of the degree of pollutant mixing in a natural stream and is one of the most important parameters in one-dimensional river water quality modelling (Iwasa and Aya, 1991; Deng et al., 2002). Longitudinal dispersion coefficients can be estimated through the use of tracer experiments, in which a known quantity of a conservative, soluble dye is injected into a river to simulate the movement of soluble contaminants (Kilpatrick, 1993). Concentration profiles of the dye must be measured at two locations downstream of the injection, at a distance far enough such that the cross sectional concentration is approximately uniform (Jobson, 1996; Singh and Beck, 2003). In other words, the concentration profiles should be measured at least as far downstream as section III (Figure 2-3).

Using measured tracer concentration profiles at two downstream locations (x_1 and x_2), the travel time ($\bar{t}$) and temporal variance (s_t^2) can be calculated to determine the mean velocity (U) using the following equations and n measured concentrations (Chapra, 1997):

$$\bar{t} = \frac{\sum_{i=0}^{n-1}(c_i t_i + c_{i+1} t_{i+1})(t_{i+1} - t_i)}{\sum_{i=0}^{n-1}(c_i + c_{i+1})(t_{i+1} - t_i)} \qquad \textbf{2-7}$$

$$s_t^2 = \frac{\sum_{i=0}^{n-1}(c_i t_i^2 + c_{i+1}t_{i+1}^2)(t_{i+1} - t_i)}{\sum_{i=0}^{n-1}(c_i + c_{i+1})(t_{i+1} - t_i)} - (\bar{t})^2 \quad \textbf{2-8}$$

$$U = \frac{x_2 - x_1}{\bar{t}_2 - \bar{t}_1} \quad \textbf{2-9}$$

The longitudinal dispersion coefficient (D) can then be calculated as:

$$D = \frac{U^2(s_{t2}^2 - s_{t1}^2)}{2(\bar{t}_2 - \bar{t}_1)} \quad \textbf{2-10}$$

If tracer testing data are not available, dispersion coefficients can be estimated using a variety of empirical equations. These equations are often relied upon as many projects lack the resources to conduct the field work necessary to collect tracer concentration profiles (Wallis and Manson, 2004). Most empirical equations that have been developed rely on hydraulic properties of the river such as the cross sectional average longitudinal velocity (U), the cross-sectional average shear velocity ($U*$), the top width of the river (W), and the hydraulic depth (H). These equations are simpler to apply than collecting tracer testing data, although they may be insufficient for capturing the complex mixing behaviour that occurs in some rivers (Wallis and Manson, 2004).

The first empirical equation for predicting longitudinal dispersion in natural streams was proposed by Fischer in 1975 and has been extensively used since that time (Seo and Baek, 2005). Many researches have continued with Fischer's work and a host of empirical equations have been developed and shown excellent performance for various case study applications. Some common empirical equations available in the literature are shown in Table 2-1. Corresponding references are also shown and the reader is directed to the original sources for more information on the development of these equations. Wallis and Manson (2004) provide a good discussion and comparison of the equations shown.

Table 2-1. Empirical Equations for Estimating Longitudinal Dispersion in Rivers

Equation	Reference	Equation Number
$\frac{D}{HU_*} = 0.011\left(\frac{U}{U_*}\right)^2\left(\frac{W}{H}\right)^2$	(Fischer, 1975)	2-11
$\frac{D}{HU_*} = 0.18\left(\frac{U}{U_*}\right)^{0.5}\left(\frac{W}{H}\right)^2$	(Liu, 1977)	2-12
$\frac{D}{HU_*} = 2\left(\frac{W}{H}\right)^{1.5}$	(Iwasa and Aya, 1991)	2-13
$\frac{D}{HU_*} = 5.915\left(\frac{U}{U_*}\right)^{1.428}\left(\frac{W}{H}\right)^{0.620}$	(Seo and Cheong, 1998)	2-14
$\frac{D}{HU_*} = \frac{0.15}{8k_1}\left(\frac{U}{U_*}\right)^2\left(\frac{W}{H}\right)^{5/3}$ where $k_1 = 0.145 + \left(\frac{1}{3520}\right)\left(\frac{U}{U_*}\right)\left(\frac{W}{H}\right)^{1.38}$	(Deng et al., 2002)	2-15
$\frac{D}{HU_*} = 10.612\left(\frac{U}{U_*}\right)^2$	(Kashefipour and Falconer, 2002)	2-16

2.4 Uncertainty

As presented in Section 2.1.4, the design of an early warning source water monitoring station involves a number of uncertainties that should be considered using methods such as Monte Carlo simulations. The following section deals more specifically with the role of uncertainty in risk analysis.

2.4.1 The Role of Uncertainty in Risk Analyses

The Random House Dictionary defines risk as an "exposure to the chance of injury or loss." A common engineering definition is the probability of an accident multiplied by the losses per accident. In the case of water resources engineering losses can be defined on a public health perspective (e.g., how many people will become ill). Since risk by definition involves the aspect of chance, consideration of uncertainty is imperative when conducting risk analyses

(Morgan and Henrion, 1990). Uncertainty analyses allow decision makers the opportunity to consider the reliability of model predictions and collect more data if necessary in order to make more defensible decisions (Morgan and Henrion, 1990; Reckhow, 1994). For these reasons, uncertainty analyses are an essential part of decision and policy making.

2.4.2 Water Quality Model Uncertainty

The use of water quality models has become increasingly important in water resources decision making. Deterministic water quality models produce a set of output values for a given set of model inputs that are assumed to be perfectly known (Portielje et al., 2000; Boano et al., 2006). However, it has been found that many deterministic models fail to produce reasonable results for many basic biological constituents (McIntyre et al., 2003b). Although most water quality model applications involve some degree of calibration, models are still simplified representations of complex and dynamic environmental processes and will always have a degree of uncertainty associated with their results (McIntyre and Wheater, 2004).

In the case of the current research, uncertainties exist regarding the nature of spill events, as described in Section 2.1.4. When using a water quality model, the following sources of uncertainty may also be significant:

- Model input uncertainty (i.e., due to errors in analytical methods);
- Parameter uncertainty; and
- Model structure uncertainty (Lindenschmidt et al., 2005; Zheng and Keller, 2006).

Investigating model structural uncertainty requires significant human and computational resources, so it is often assumed to be adequately represented as parameter uncertainty (McIntyre et al., 2003b). For the current research, therefore, the probabilistic nature of spill events is considered in addition to the uncertainty associated with water quality model parameters.

2.4.3 Methods to Quantify Uncertainty

Monte Carlo simulations are commonly used to propagate uncertainty in water quality modelling studies. This involves defining a probability distribution for all uncertain inputs or model parameters and then sampling a random value from each of these distributions. The sampled values are then used in the model to compute a corresponding output value. This process is repeated *m* times to generate *m* output values, which can be used to form a probability distribution of the model output (Morgan and Henrion, 1990).

Before such a procedure can be undertaken, probability distributions of the model parameters and inputs must be defined. Typical ranges for most water quality model parameters are well documented in the literature (Perera and Ng, 2001; Martin and Wool, 2002; McIntyre et al., 2003a; Cox and Whitehead, 2005; Lindenschmidt, 2006; Osidele et al., 2006). However, the use of overly conservative parameter distributions can result in unnecessarily high levels of model prediction uncertainty. Parameter uncertainty estimates can be refined by using Monte Carlo based methods such as Generalized Likelihood Uncertainty Estimation (GLUE) (Beven and Binley, 1992), Regional Sensitivity Analysis (RSA) (Spear and Hornberger, 1980), Markov Chain Monte Carlo (MCMC) (Brooks, 1998), or Uniform Covering by Probabilistic Rejection (UCPR) (Klepper and Hendrix, 1994). All of these methods involve conditioning the parameters using a set of calibration data. Conditioning experiments are computationally expensive as they require conducting an entire set of simulations before the model can be used for prediction. In addition, a set of good calibration data is required with which to assess the performance of each parameter set.

2.5 Multi-objective Optimization

Many environmental management problems involve multiple competing objectives (Singh et al., 2003; Muleta and Nicklow, 2005). For example, ground water problems may involve minimizing the risk of contamination while maximizing pumping rates to meet a specified demand. Similarly, water distribution network applications may involve minimizing costs while ensuring water pressure is maximized and water resources problems could entail maximizing agricultural profit while minimizing sediment yield (Muleta and Nicklow, 2005). In each of these examples, the objectives are in conflict because improving one generally leads to a decline in the other.

Two of the major considerations for locating a monitoring station identified in Section 2.1.3.1 include the amount of threats coverage and warning time provided. These considerations are in competition because improving performance with respect to one results in worse performance with respect to the other (this is discussed further in Section 3.1).

There are many different methods with which to solve multi-objective optimization problems. Multiple objectives can be converted to a single objective by applying a weighting factor to each objective. For example, if one wants to minimize two objectives, F_1 and F_2, the following single objective function, F_3, may be used:

$$\textit{Minimize } F_3 = w_1 F_1 + w_2 F_2$$

where w_1 and w_2 are weightings assigned to objectives 1 and 2, respectively. The use of this method allows the user to convert a multi-objective problem to a single objective problem and then apply traditional optimization methods such as linear, dynamic, or non-linear programming (Reddy and Kumar, 2007).

An alternative method is to convert all but one of the multiple objectives to constraints. For a two objective problem, the second objective (F_2) may be converted to a constraint as shown below.

$$\textit{Minimize } F_1$$

$$\textit{Subject to } F_2 \leq t.$$

Once again, this allows the use of traditional single objective solution methods.

If the problem is solved using a single objective, some solutions that represent a tradeoff between the original objectives may fail to be identified (Reddy and Kumar, 2007). As a result, the full information value is not captured when one or more objectives are limited to subjectively defined weightings or constraints (Fenicia et al., 2007). To overcome this limitation and explore a range of tradeoff solutions, the concept of Pareto optimality can be applied.

A Pareto optimal front is defined by all solutions for a given problem that are not dominated by another solution. A solution (S_1) dominates another solution (S_2) if it is better in at least one objective and not worse than S_2 in any other objective. The Pareto optimal set of solutions is defined as all of the solutions that are not dominated by any other feasible solution. For a two objective problem, these solutions can be plotted as a curve in objective space, called the Pareto front.

A Pareto front is shown in Figure 2-4, with each axis representing an objective to be minimized (i.e., solutions towards the origin are preferred). It can be seen from this figure that solution C is dominated because solutions A and B are preferable with respect to both objectives. Neither A nor B are dominated by any other solutions, and it can be seen that they are two solutions contained on the Pareto front.

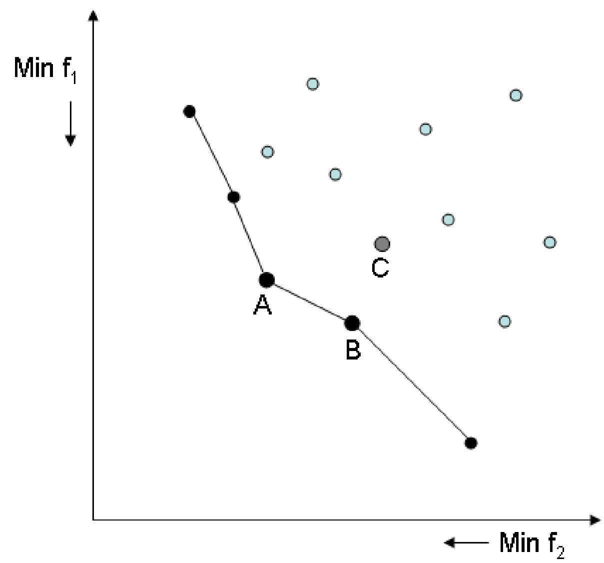

Figure 2-4. Pareto Optimal Curve

Presenting a complete or approximate representation of the Pareto front enables decision makers to consider an array of feasible solutions without pre-defining preferences between objectives or restricting objectives to be represented as constraints.

The main limitation of analyzing multiple objectives using a Pareto optimal curve is that solutions that perform poorly for one or more objectives may be identified as feasible options. For example, solutions that would be identified if each objective is optimized individually lie on the Pareto curve even though they may have unacceptable performance with respect to the other objectives. However, this limitation is easy to overcome as decision makers need not consider these solutions if they do not represent appropriate or desired tradeoffs. The benefit of constructing a Pareto curve is to provide decision makers with a range of tradeoff solutions for consideration.

2.6 Summary of Relevant Literature

In the preceding discussion it was found that a probabilistic approach for designing an early warning monitoring station is required in order to capture the uncertainties associated with the nature of spill events. The research described in this thesis builds on that of Grayman et al. (2001) by considering multiple objectives in locating the stations and using a different method for incorporating uncertain flows. This methodology is described in further detail in Chapter 3.

To conduct a probabilistic design of an early warning monitoring station, the use of both a hydrodynamic model and water quality model is required. One-dimensional models have been used for similar applications and are widely accepted as good engineering approximations. Monte Carlo simulations are commonly used to propagate uncertainty which, in the case of the current research, is associated with both the spill scenarios modelled as well as the water quality model parameters. Common Monte Carlo based methods for the inclusion of parameter uncertainty were also identified.

Two important considerations for locating a source water monitoring station were identified to be the amount of coverage and the amount of warning time provided by the station. Instead of calculating a population exposure as the metric with which to compare station designs, this thesis involves the use of a multi-objective design problem so that tradeoff solutions can be analyzed by decision makers.

3 Methodology

The methodology developed to design an early warning source water monitoring station involved the following tasks:

- Specification of design objectives;
- Identification and prioritization of potential threats;
- Identification of potential monitoring station locations;
- Modelling of probable spill scenarios; and
- Analysis of model results to determine an optimal monitoring station design.

Each of the above steps is described in Sections 3.1 through 3.6 and a specific application of the methodology is discussed in Chapter 4.

3.1 Specification of Design Objectives

The following decisions are required for the design of an early warning source water monitoring system:

- Number of monitoring stations;
- Location of monitoring station(s);
- Sampling frequency;
- Parameters to monitor;
- Monitoring technology;
- Communication linkages; and
- Response protocol.

The purpose of this research is to design a single early warning monitoring station upstream of a specific drinking water treatment plant intake. Decisions with regards to communication linkages, monitoring technology, and response protocols are left to the local decision makers.

Therefore, the scope of this methodology is limited to determining the optimal location and sample frequency of an early warning source water monitoring station.

In Grayman et al. (2001), various monitoring station designs were assessed based on an effectiveness index (see Section 2.1.4). The methodology applied in this research involves a multi-objective approach for assessing potential designs, as described below.

The monitoring station should be located where the probability of detecting a contamination event is maximized. However, it must also provide sufficient warning to allow downstream water treatment plants time to implement a response. This response could include initiating further sampling to characterize the event, conducting predictive modelling, implementing advanced treatment, or shutting down the plant intake (International Life Sciences Institute, 1999). Therefore, in addition to maximizing the probability of detection, the station should be located where the probability of achieving a threshold amount of warning time is maximized.

These two objectives cannot be optimized simultaneously. Warning time is maximized when the station is as far as possible upstream and the probability of detection is maximized when the station is as close as possible to the intake in order to provide greater coverage of potential threats. Therefore, locating the monitoring station required consideration of the tradeoffs between these competing objectives.

More frequent sampling increases the probability of detecting a spill event. However, cost was also a consideration so determining the tradeoff between increased sampling and the improvement in the objective values was investigated. For example, if sampling every hour achieves only a minor improvement in the objective values over sampling every two hours, then it may not be worth the added expense of sampling twice as often.

3.2 *Threats Inventory*

The probabilistic modelling that was undertaken as part of this research was based on threats located within the study area. These threats were identified and prioritized using the procedure described below.

3.2.1 Identification

In the province of Ontario, any person who causes or permits a spill to be released to the natural environment, or any public sector employee who has knowledge of such a release, is required by law to report to the Ontario Ministry of the Environment's Spills Action Centre (SAC) (Ontario Ministry of the Environment, 1990). The SAC maintains records of all reported spills including the medium to which the spill occurred (e.g., air, land or water) as well as the type and amount of substance release, if known. Each spill is classified as oil (e.g. crude, gasoline, petroleum), waste (e.g. industrial, hazardous liquids, sewage), gas or particulate, chemical (e.g. acids, bases, pesticides, solvents), or other.

Historical spill data were obtained from the SAC to help identify threats within the case study area. However, due to the low probability of most spill events, these data were very limited so province-wide data were reviewed. In 2006, 4541 spills were reported in Ontario and are classified in Table 3-1 (Ontario Spills Action Centre, 2006b).

Table 3-1. Spills Reported in Ontario in 2006

Type	Number	Percentage of Total
Oils	2516	55.4%
Wastes	778	17.1%
Gases and particulates	593	13.1%
Chemicals	459	10.1%
Other	195	4.3%

From Table 3-1, it can be seen that oils and wastes are the most abundant type of spill in Ontario. The SAC also reported that the most common spill source was transportation-related spills, which represented 22% of all spills reported in 2006 (Ontario Spills Action Centre, 2006b). Therefore, identifying sources of oil and waste spills and the nature of potential transportation-related spills provided a good starting point for conducting the threats inventory.

In addition to considering previous spill data, the Ministry of the Environment's Draft Guidance Modules for Source Water Protection Planning were also reviewed. These modules list a number of drinking water threats of provincial concern, which are shown in Table 3-2.

Table 3-2. Drinking Water Threats of Provincial Concern
(Ontario Ministry of the Environment, 2006a)

Direct Introduction	Landscape Activities	Storage of Potential Contaminants
• Water treatment plant waste water discharge • Sewage treatment plant effluent • Sewage treatment plant bypasses • Industrial effluents	• Road salt application • De-icing activities • Snow storage • Stormwater management systems • Cemeteries • Landfills • Organic soil-conditioning • Septage application • Hazardous waste disposal • Liquid Industrial waste • Mine tailings • Biosolids application • Manure application • Fertilizer application • Pesticide/herbicide application • Historical activities – contaminated lands	• Fuels/hydrocarbons • DNAPLs • Organic Solvents • Pesticides • Fertilizers • Manure

The information presented was used as a starting point to direct the identification of threats within the study area. The results of this inventory for the case study application are discussed in Section 4.2.2.

3.2.2 Prioritization

After the threats were identified, they had to be prioritized according to their level of risk for the purpose of defining probability distributions from which to sample during the Monte Carlo simulations (see Section 3.5.1.3). Risk estimates require detailed knowledge and understanding of potential threats in order to define their probabilities of occurrence and the magnitude of their effects. This is a subjective exercise often conducted by a panel of experts based on professional experience, and may involve site visits and discussions with local stakeholders.

Due to limitations related to the case study application, a comprehensive prioritization of threats could not be completed. However, in the Province of Ontario, the Clean Water Act requires source protection committees to identify drinking water threats and perform a semi-quantitative risk assessment in the near future (Ontario Ministry of the Environment, 2006b). For the purposes of this research, best estimates of relative risk levels have been used. The results can be easily updated when a more comprehensive risk assessment has been completed for the case study area.

3.3 Identification of Potential Monitoring Locations

As described in Section 2.1.3.1, early warning monitoring stations are commonly located near bridges for access to electricity (International Life Sciences Institute, 1999). Multiple locations that coincided with bridges and were well spaced throughout the case study area were selected as described in Section 4.2.3. These locations represent a set of discrete solutions for this optimization problem.

3.4 Modelling

Once the threats and potential station locations were identified, probabilistic modelling was undertaken to generate results upon which to base design decisions. Setting up the model to perform these simulations represented a significant component of the methodology. The following sections describe this procedure in more detail.

3.4.1 Model Selection

As reviewed in Section 2.2, water quality models require information about the pathway, volume, and velocity of the water to determine how contaminants move and behave (Martin and McCutcheon, 1999). Therefore, both a hydrodynamic model and a water quality model were required for this research.

For many modelling applications, rivers are assumed to have adequate lateral and vertical mixing so that homogeneous cross sectional concentrations and average cross sectional velocities can be used (Martin and McCutcheon, 1999). Although complete mixing never occurs, it is often assumed to be a good engineering approximation (Martin and McCutcheon, 1999). The complete mixing assumption has been used in similar modelling studies (e.g. Grayman and Males (2002)), and was selected as a reasonable choice for the current research.

Brief descriptions of the one-dimensional models considered for this research are listed in Table 3-3. This is not an exhaustive list and the reader is referred to Grayman et al. (2001) for more comprehensive descriptions of these and other models.

All of the models listed in Table 3-3 are well documented and widely accepted for modelling one-dimensional rivers (Jobson, 1996; Grayman et al., 2001). Furthermore, all of the water quality models are based on the conservation of mass equation, so their differences are thought to be less important than the quality of input data with which they are provided (Jobson, 1996).

Since all the models considered are technically appropriate for the current research, the criteria for selecting a hydrodynamic and water quality model included:

- The water quality model must be able to run from the DOS prompt or have the capacity to run repeatedly without user intervention (so that Monte Carlo simulations can be performed);
- The hydrodynamic and water quality models must easily link together;
- The models must have well documented User Manuals; and
- The models must be easily obtainable and inexpensive.

Table 3-3. List of One-Dimensional Hydrodynamic and/or Water Quality Models

Model	**Type**	**Description**
RIVMODH (Dames and Moore, 1994)	Hydraulic	RIVMODH is a hydraulic model used for rivers, estuaries and other one-dimensional water bodies with unsteady flow. It can be linked with a water quality model such as WASP.
BRANCH (Schaffranek, 1987)	Flow	BRANCH can simulate steady and unsteady flows in a single river branch or network of branches
DYNHYD (Ambrose et al., 1993)	Hydrodynamic	DYNHYD is a hydrodynamic model that can be linked with WASP. It is used for well mixed rivers and estuaries.
CE-QUAL-RIV1 (U.S. Army Corps of Engineers-WES, 1990)	Hydrodynamic and Water Quality	CE-QUAL-RIV1 has both hydrodynamic and water quality modules. It can be used for highly unsteady as well as steady conditions.
WASP (Ambrose et al., 1993)	Water Quality	WASP has been extensively used due to its effectiveness in modelling a wide variety of pollutants. It can be used in one, two, or three dimensions. WASP is commonly linked with CE-QUAL-RIV1H or DYNHYD. However, WASP may be prone to numerical dispersion problems (Grayman et al., 2001).
BLTM (Jobson and Schoellhamer, 1987)	Water Quality	BLTM is a one-dimensional water quality model that has been widely applied on small streams to large rivers. It can be easily linked with DAFLOW.

With the considerations listed above in mind, the model EPD-RIV1 was selected. EPD-RIV1 is a one dimensional hydrodynamic and water quality model that is based upon the U.S. Army Corps of Engineers Waterways Experiment Station's CE-QUAL-RIV1 model (Martin and Wool, 2002). In 1993, the Georgia Environmental Protection Division (EPD) identified CE-QUAL-RIV1 as the most appropriate model to use for a large modelling project for the Chattahoochee River. However, some limitations were noted which resulted in extensive updates to improve the model's performance and ease of use. The updated model is referred to as EPD-RIV1 (Martin and Wool, 2002).

EPD-RIV1 can be downloaded free of charge from the internet and can be run from both the Windows interface and DOS prompt. EPD-RIV1 consists of two modules, RIV1H and RIV1Q, which simulate hydrodynamics and water quality, respectively. RIV1H can be used for both steady and highly unsteady flow conditions and uses the four-point implicit finite difference numerical scheme to estimate flows, velocities and water surface elevations. As discussed in Section 2.2.2, the four-point implicit finite difference scheme is one of the most accurate methods for solving the Saint-Venant equations (Martin and McCutcheon, 1999; Martin and Wool, 2002). RIV1H writes a hydrodynamic linkage file upon completion to provide transport information to RIV1Q. The water quality model can simulate a number of constituents including organic nitrogen, ammonia nitrogen, nitrate nitrogen, organic phosphorus, orthophosphate, biochemical oxygen demand, dissolved oxygen, algae, iron, manganese, coliforms, temperature, as well as two arbitrary constituents (Martin and Wool, 2002).

3.4.1.1 Model Limitations

Although EPD-RIV1 was identified as the preferred model for this research, there are some limitations that should be noted.

By definition, concentrations in a one-dimensional model are vertically and laterally averaged. This means that the model will provide no results on which to base decisions regarding the lateral and vertical positioning of a monitoring station. As a result, unless there is compelling evidence to do otherwise (e.g. a vast majority of threats are along the same bank as the intake), the station should be located near the center of the river.

Another limitation of a one-dimensional model is that discharges located a short distance upstream of potential monitoring locations may be detected by the model more than they would be in reality. For example, the model will detect any discharge located upstream (of sufficient quantity to reach the monitoring station) since it assumes complete and instantaneous mixing. However, prior to theoretically achieving complete mixing the contaminant must travel at least one mixing length. This means that some discharges may pass the monitoring station without being detected.

Another limitation of the model is the fact it can only simulate soluble constituents that can be approximated with first-order decay. Although oil spills were identified as one of the most common spill sources, all of the models considered for this study are unable to simulate oil spills. These models are not designed to handle some of the transport and fate processes of sparingly soluble, buoyant substances, such as the spreading of the surface slick and the interaction of the slick with shorelines (Hibbs and Gulliver, 1999; Grayman et al., 2001). Although oil spills are not modelled as part of this research, decision makers may still select a monitoring technology that can detect hydrocarbons or volatile organic carbons. This methodology could also be adopted for use with a model that can simulate riverine oil spills. Yapa and Shen (1994) provide an excellent review of some models available for this purpose.

3.4.2 Model Setup and Preprocessing

Prior to conducting simulations, a significant amount of data collection and processing was required to setup the model. These efforts are described in more detail below.

3.4.2.1 Model Extent and Spatial Grid

The first step in setting up the model was to establish its spatial extent. The upstream model boundary was located at a flow gauging station since flow or stage measurements were required. A gauging station was selected that was sufficiently far upstream so that most major threats could be included in the model. The downstream extent of the model was selected as the closest gauging station below the drinking water treatment plant intake. Tributaries and abstractions that impact the river flow within the study area were identified so that necessary input data could be collected.

Once the model extent was established it was discretized into nodes, or cross sections. The placement of nodes was limited to those areas along the river that had surveyed cross sectional geometry available.

The surveyed cross sectional geometries required some processing before being entered into the model. The model requires geometries to be entered as (x, y) pairs, where x represents lateral position in the cross section and y represents depth, both in units of feet. The origin must be at the top of the left stream bank, where the left bank is defined for an observer looking downstream. Conversion of the cross sectional survey data from elevations to depths was required. This was accomplished by subtracting the bed elevation from the elevation of the origin (left bank), as illustrated in Figure 3-1.

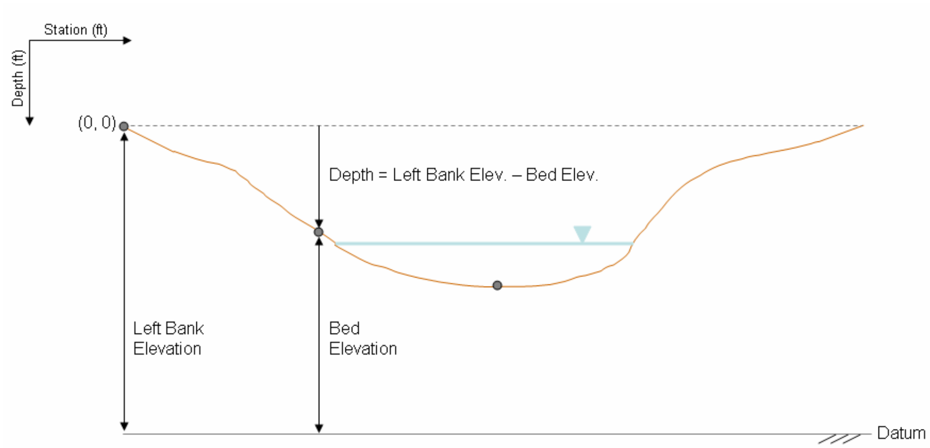

Figure 3-1. Cross Sectional Geometry Depth Calculation

3.4.2.2 Hydrodynamic Input File

The hydrodynamic input file required the following information for each node:

- Node name;
- Distance to downstream node;
- Slope between node and downstream node;
- Manning's roughness coefficient;
- Initial water surface elevation;
- Initial flow; and
- Invert elevation (elevation of lowest point of the cross section).

Distances, slopes, and invert elevations were easily calculable from the cross sectional data that were collected. Manning's values used in previous modelling studies by the local conservation authority were applied and initial water surface elevations and flows were approximated for various starting conditions. Since rivers tend to "wash-out" the initial conditions quickly, their exact specification is usually not necessary (Martin and McCutcheon, 1999). However, reasonable values were estimated and adjusted by trial and error, where necessary, to ensure that stability problems were not encountered on startup of the model.

The hydrodynamic input file also required the specification of some hydraulic parameters. A tolerance parameter was required to determine when the model converges. This parameter was set to 0.05, which was large enough to decrease computation time but small enough that the model's performance was not impacted (Martin and Wool, 2002). A theta value of 0.6 was adopted to improve the model's accuracy while avoiding instabilities, as recommended by Martin and McCutcheon (1999) (see Section 2.2.2 for a description of theta).

Prior to running RIV1H, the computational time step also required specification. The selection of the time step was limited by the Courant Number, γ, which represents the ratio between the distance moved during one time step (Δt) and the segment length (Δx). It can be calculated using equation 2-3 as shown in Section 2.2.2. To ensure stability, the Courant number must be less than one so the water cannot move more than one segment length at a given velocity (U) within one time step. A computational time step of 60 seconds was used to satisfy the Courant condition

3.4.2.3 Water Quality Input File

RIV1Q uses the same model discretization as RIV1H which simplified the initial setup process. Initial concentrations for all water quality constituents being simulated were required. Dispersion coefficients for each cross section also required specification, as did water quality parameters, such as decay rates of the various constituents. Since these parameters were adjusted as part of the Monte Carlo simulations, their specification is further described in Section 3.5.1.1.

The water quality model output time step needed to be relatively frequent to ensure peak concentrations were recorded. A set of initial simulations using time steps of three and six minutes were conducted. For each simulation, the peak concentration recorded with the larger time step was less than one percent different than that recorded with the smaller time step.

Therefore, an output time step of six minutes was selected to minimize the size of output files while maintaining enough information to produce accurate concentration profiles.

3.4.2.4 Boundary Conditions

At the upstream boundary, a time series of hydrodynamic data (flow or stage) was required. Hourly flow data from a gauging station were available for this purpose. The downstream boundary condition was defined by a rating curve of stage versus discharge.

3.4.2.5 Inflows and Withdrawals

Input files to define inflows and withdrawals to the river were also required. Separate lateral inflow files were used for RIV1H and RIV1Q as the flow data were at a different frequency than the quality data. However, within each input file values had to be entered at the same frequency. As a result, the flows from the wastewater treatment plant had to be converted to an hourly frequency, using linear interpolation, to coincide with the hourly flow data available for the tributaries.

Withdrawal rates for the water treatment plant and a permitted abstraction were available at a daily time step and were entered into a withdrawal file.

3.4.3 Hydrodynamic Calibration

Once the hydrodynamic model was set up, it was calibrated to ensure that it modelled the system accurately. The only parameter usually calibrated in RIV1H is Manning's roughness value. Default Manning's values used in previous modelling studies were entered as a starting point for each cross section. However, calibration data (stage measurements) were only available at the downstream boundary and one other gauging station within the modelled reach. As a result, the river was divided into two segments and the Manning's values within each segment were calibrated together (i.e. all were adjusted by the same factor). The Nash Sutcliffe coefficient was used as the metric with which to compare the observed and modelled elevations and was calculated using equation 3-1. A coefficient value of one indicates a perfect match.

$$E = 1 - \frac{\sum_{t=1}^{T} (Q_o^t - Q_m^t)^2}{\sum_{t=1}^{T} (Q_o^t - \overline{Q}_o)^2} \qquad \textbf{3-1}$$

where E = Nash Sutcliffe coefficient,

Q_o = observed value (stage in this case),

Q_m = modelled value (stage in this case), and

$\overline{Q}_o$ = average of all observed values.

Adjustments were made manually until the modelled elevations matched the measured elevations within a desired level of accuracy. The results of the hydrodynamic calibration for the case study are presented in Section 4.3.

3.4.4 Water Quality Calibration

RIV1Q was not calibrated for reasons discussed in Section 3.5.1.1.

3.5 Monte Carlo Simulations

As discussed in Section 2.4.3, Monte Carlo simulations are commonly used to quantify uncertainty in water quality modelling studies. The procedure used to conduct the Monte Carlo Simulations as part of this research is described in Section 3.5.1 to Section 3.5.3.

3.5.1 Sources of Uncertainty

The design of an early warning source water monitoring station involves many sources of uncertainty. These sources include the model parameters, timing of the spill, and nature of the spill, as illustrated in Figure 3-2 and described in the following sections.

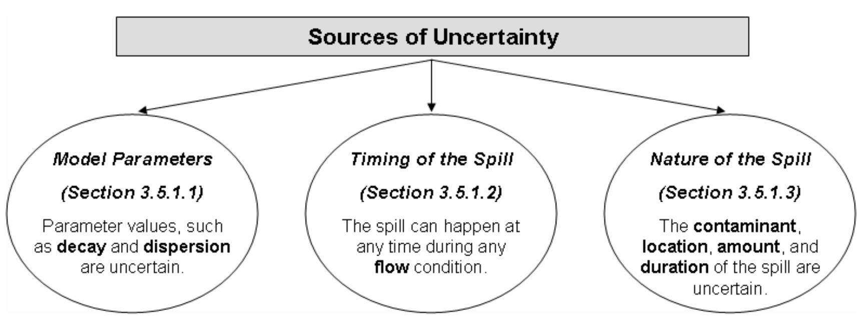

Figure 3-2. Sources of Uncertainty

3.5.1.1 Model Parameter Uncertainty

As discussed in Section 2.4.2, consideration of the uncertainty associated with model parameters is essential in order to establish the reliability of model predictions. Before an uncertainty analysis could be undertaken, probability distributions of the model parameters and uncertain inputs were required. Typical ranges suggested by the literature were used to define uniform probability distributions for each parameter (Perera and Ng, 2001; Martin and Wool, 2002; McIntyre et al., 2003a; Cox and Whitehead, 2005; Lindenschmidt, 2006; Osidele et al., 2006), as shown in Appendix A. This meant that all parameter values had an equal likelihood of being sampled so the model could be tested over a wide range of parameter values (Cox and Whitehead, 2005).

The probability distributions used for dispersion coefficients were determined differently than those used for the other parameters because they are dependent on the hydraulics of the river, not the constituents being modelled. When tracer data are not available to calibrate dispersion coefficients, empirical equations based on top widths, velocities, and hydraulic depths can be used (Jobson, 1997). For this methodology six empirical equations were selected to provide a range of possible dispersion values at each cross section. Equations 2-11 to 2-16, contained in Table 2-1, were used for this purpose and are described in detail by Wallis and Manson (2004).

The six resulting dispersion coefficients for each cross section were utilized to establish a triangular probability distribution. A triangular distribution is defined by a parameter's minimum value (*a*), maximum value (*b*), and most frequently occurring value (*c*). In this case *a* represents the minimum calculated dispersion coefficient, *b* represents the maximum calculated dispersion coefficient and *c* was set as the average of the six calculated dispersion coefficients.

The cumulative triangular distribution function can be defined using the following equation:

$$F(x|a,b,c)=\begin{cases}\dfrac{(x-a)^2}{(b-a)(c-a)} & \text{for } a\le x\le c\\[2ex] 1-\dfrac{(b-x)^2}{(b-a)(b-c)} & \text{for } c< x\le b\end{cases} \qquad \textbf{3-2}$$

A random sample (*x_test*) from the cumulative triangular distribution function was determined using the following equation:

$$x\_test = \begin{cases} a + \sqrt{u(b-a)(c-a)} & \text{for } u \le \dfrac{c-a}{b-a} \\ b - \sqrt{(1-u)(b-a)(b-c)} & \text{for } u > \dfrac{c-a}{b-a} \end{cases} \qquad \textbf{3-3}$$

where *u* represents a uniform random number between 0 and 1.

Since dispersion at all cross sections is related (i.e. at a given time, dispersion would not be at its maximum value at one cross section and its minimum value immediately downstream), the same uniform random number was used to sample from each triangular distribution. This allowed dispersion to vary from one cross section to the next based on differing hydraulic properties, but the percentile of the sampled coefficient values was the same for each cross section.

The parameter distributions identified from the literature and those calculated for the dispersion coefficients were overly conservative and had the potential to result in unnecessarily high levels of model prediction uncertainty. For this reason, conditioning experiments are traditionally conducted to refine parameter ranges using a method such as Generalized Likelihood Uncertainty Estimation (GLUE), as discussed in Section 2.4.3. Methods such as GLUE require a set of good calibration data with which to assess the performance of each parameter set.

At this point it is important to recognize a major difference between typical modelling studies and the modelling conducted as part of this research. Typically, models are used predictively; for example, a model may be used to predict the time of arrival of a contaminant during a real-time spill scenario. Suitable model parameter ranges will have been pre-defined as part of a conditioning experiment so that model parameter uncertainty can be propagated through the model to describe prediction uncertainty. Other sources of uncertainty *may* be included in the assessment of this arbitrary example, such as the amount spilled (e.g. about 100 to 150 kg), the duration of the spill (e.g. 1 to 1.5 hours), the flow at the time of the spill (e.g. measured as 100 cms ±2%), and the location of the spill relative to the intake (e.g. 1 to 1.25 km upstream). In situations such as this, the uncertainty associated with water quality parameters often leads to significant model prediction uncertainty. This arbitrary example is labeled as "Predictive Modelling of a Specific Situation" in Figure 3-3.

In the case of the current research, a specific situation is not being modelled, but rather a series of probabilistic situations. In addition to model parameter uncertainty, great amounts of uncertainty exist due to the timing and nature of the spill. Depending on the time of the spill, the river flow could range from a trickle to a flood condition. Similarly, the duration, mass, and location of the spill also have large ranges. Figure 3-3 compares the relative magnitudes of the different uncertainties in each of these arbitrary scenarios.

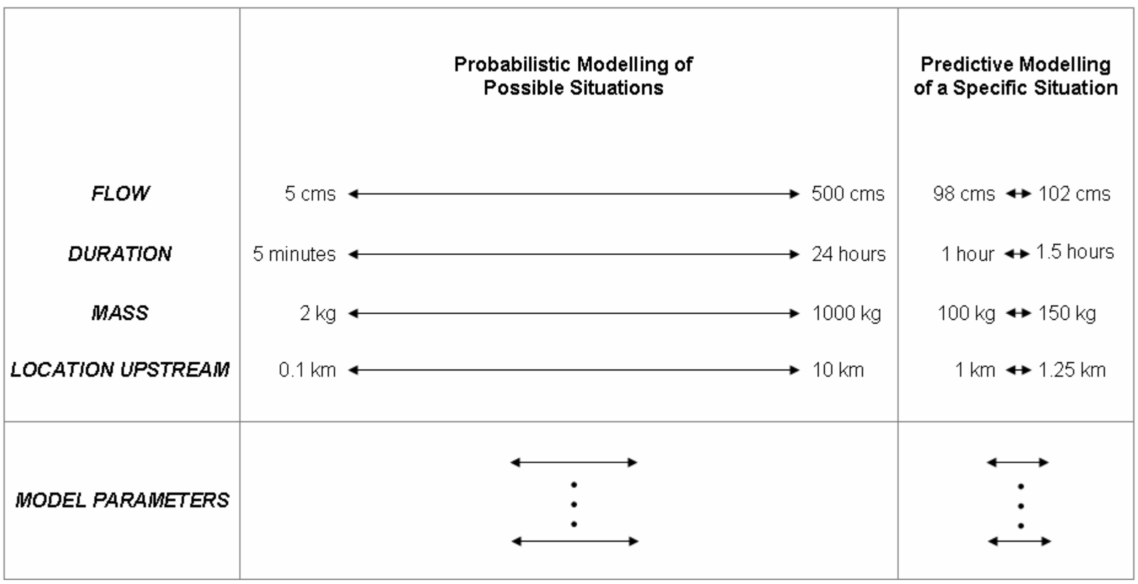

Figure 3-3. Comparison of Different Sources of Uncertainty

Note: Ranges are arbitrary and are shown as linear for simplicity. Not all sources of uncertainty are uniformly sampled from within their ranges. The distribution of each source of uncertainty is described below.

It was hypothesized that model parameter uncertainty may not be significant compared with the uncertain nature of the spills being modelled as part of this research. Therefore, prior to investing in computationally expensive and data intensive procedures to estimate parameter uncertainty, its significance on the decision making process was tested. A set of scenarios were executed with all sources of uncertainty and another set with only the flow as uncertain (with the parameters set to the midpoint of their ranges). For this preliminary analysis, the mass and duration were held constant. If the relevant results produced using the conservative parameter ranges are similar to the results produced when only flow is uncertain, no calibration is required. However, if the uncertainty associated with the parameters is shown to impact the results, then calibration using a method such as GLUE is required. The results of this experiment are described further in Section 5.1.1.

Note that the method described above was only used for the water quality model parameters. The only parameters in the hydrodynamic model are the Manning's coefficients. Manning's equation can be represented as

$$V = \frac{1}{n} R^{2/3} S^{1/2} \qquad \textbf{3-4}$$

where V represents the cross sectional average velocity (m/s);
N represents Manning's coefficient;
R represents the hydraulic radius (m); and
S represents the energy gradient.

Therefore, changing Manning's coefficient also changes the average velocity. However, as described further in Section 3.5.1.2, the flow distribution used for the Monte Carlo simulations ranged multiple orders of magnitude. Depending on the type of channel, Manning's coefficients typically do not range by more than a factor of 1.5 (Martin and McCutcheon, 1999). Therefore, the impact of the Manning's coefficient on the resulting water velocities is minimal compared with the impact of the uncertain flow at the time of a spill.

3.5.1.2 Time of Spill

The uncertain nature of spills means that they can occur at any time into any river flow. Uncertain flows include the upstream boundary flow, tributary inflows, wastewater treatment plant effluent flows, and any withdrawal flows. In previous work, Grayman et al. (2001) assumed the flow at the time of a spill is constant and that no tributaries exist. However, for many case study applications major tributaries may be encountered that can significantly impact the river's flow. Therefore, an alternate method was developed so that flows could vary throughout a spill simulation and tributary impacts could be included.

Generating a time series of synthetic flows and preserving the correlations between each of the flow inputs is a complex task. Therefore, instead of defining a distribution to synthetically generate a time series of flows, uniform random sampling was conducted to select a historical point in time. The time series of flows recorded at that time for all flow inputs were then used in the model. Using this method meant that realistic flow relationships (between the main channel and its tributaries) were automatically preserved.

Since this method limits the flows to past conditions, it was important to use a long record of data. Fifteen years of hourly flow data for the upstream boundary condition and all tributary inflows were collected for this purpose. Since water and wastewater treatment plant flows are not correlated with river flows, they were not sampled in the same way. Treatment plant flows recorded many years ago are no longer representative as a result of a growing population, so average flow values for recent years were used in the model.

3.5.1.3 Nature of the Spill

For the purpose of this research, the nature of a spill was defined to include:

- The substance spilled;
- Location of the spill;
- Quantity spilled; and
- Duration of the spill.

A spill scenario is defined as a combination of a location and contaminant. Based on the threats inventory that was completed, a number of likely spill scenarios were identified. Since a comprehensive threats inventory could not be completed, arbitrary contaminants were used for all scenarios except for one which involved a wastewater spill. The decay rate for the arbitrary spills was set to a range of 0 to 1 d^{-1} to simulate a range of potential contaminants.

Instead of pre-defining a spill scenario distribution, a set of Monte Carlo simulations were run for each identified spill scenario and weighting factors, representing best estimates of each scenario's relative level of risk, were applied after the simulations were complete. Although the weighting factors applied for this research may not be ideal, they can easily be updated once the complete threats inventory and risk assessments are completed as required by the Clean Water Act (Ontario Ministry of the Environment, 2006a). The sensitivity of the chosen weightings was explored and is discussed with the results in Section 5.4.2.

3.5.1.4 Uncertainties Not Considered

The probability that a spill event is reported by the spill generator or a member of the public was not considered. It was assumed that only the monitoring station can alert a downstream water treatment plant of a contamination event. This was a conservative, simplifying assumption because in many cases spills may be reported before the contamination even reaches the monitoring station.

It was also assumed that the monitor performs perfectly. The probability that the monitor fails due to power interruptions, technical problems, or incomplete river mixing was not considered.

3.5.2 Number of Simulations Required for Each Spill Scenario

A sufficient number of Monte Carlo simulations should be conducted such that the average and standard deviation of the accumulated model outputs stabilize. 2000 simulations were executed and the cumulative average and standard deviation of the following outputs were calculated for each potential monitoring station:

- Duration the contaminant was detected;
- Travel time to the intake; and
- Peak concentration.

The results of this analysis are described in Section 5.1.2.

3.5.3 Monte Carlo Procedure

Once the above steps were completed, the required number of Monte Carlo simulations (*n*) were performed for each of the spill scenarios identified. Code was written in MatLab in order to sample all uncertain parameters and update the relevant model input files. The procedure used to perform these simulations is described below and shown graphically in Figure 3-4.

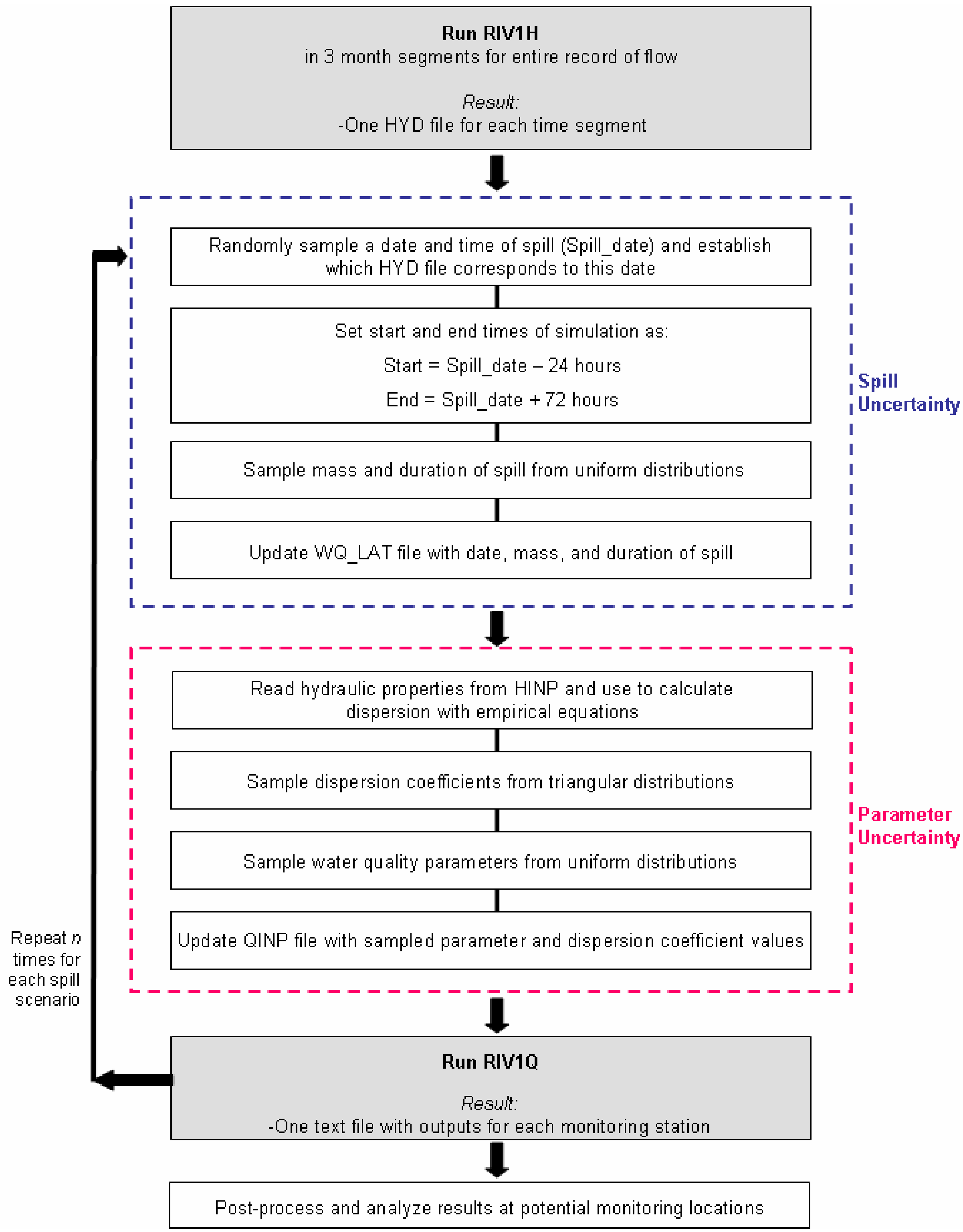

Figure 3-4. Monte Carlo Flow Chart

To reduce the computation time and simplify the procedure, it was decided not to run RIV1H for every simulation. This decision meant that a hydrodynamic linkage file was required for all 15 years of flow data prior to conducting the Monte Carlo simulations. A single hydrodynamic linkage file for 15 years of flow would require approximately 10.8 GB and would need to be opened and searched for each simulation. Instead, 60 linkage files were created at three month intervals, which reduced the file size to approximately 180 MB each and increased the speed at which RIV1Q could execute.

Inherent in the decision not to run RIV1H for each simulation was the assumption that the spills would not have a significant flow contribution. Of the spills reported in Ontario in 2006, less than 3% were of volumes greater than 10,000 L (Ontario Spills Action Centre, 2006b). A spill of 10,000 L released over one minute results in a flow rate of only 0.17 m^3/s. The lowest hourly flow recorded for the case study river in 2006 was 7.1 m^3/s. Therefore, it was reasonable to assume that the spills do not contribute significantly to the river's flow.

The simulated spills required some volume of flow in which to enter the river, so a negligible amount of flow was modelled at each spill location. As a result of the low flows conveying each spill, some of the spill concentrations were overly high. However, the model assumes complete and instantaneous mixing, so the resultant mass loading is all that is important.

Once all 60 linkage files were created, a series of Monte Carlo simulations were conducted for each of the spill scenarios. The first step of each simulation was to sample a uniform random date and time from the 15 years of hourly flow data. Code was written to establish which hydrodynamic linkage file (HYD) corresponded to the sampled date. The start and end time of the simulation were then written to the input file. The start time was set as 24 hours prior to the spill time and the end time was set as 72 hours after the spill time, which was a sufficient amount of time for all spills to arrive at the intake.

In addition to the time of the spill, the mass and duration of the spill scenario were also uncertain. These inputs were both randomly sampled from uniform distributions. The mass and the duration were then written to the water quality lateral inflow file (WQ_LAT).

The model parameters were sampled next. As discussed, six empirical equations were used to calculate dispersion and define its distribution. In order to perform these calculations, hydraulic data were read from the appropriate hydrodynamic linkage file (HYD). Once the calculations were complete, a value for each dispersion coefficient was sampled from the corresponding triangular distribution as described in Section 3.5.1.1. The remaining water quality parameters were then randomly sampled from their uniform distributions. All dispersion coefficients and parameters were then written to the water quality input file (QINP).

After sampling of the spill conditions and parameters was complete, RIV1Q was executed. Each simulation took approximately 15 seconds to complete. The outputs for each monitoring station were written to a text file after each simulation and accumulated for subsequent analyses. This procedure was repeated *n* times for each of the identified spill scenarios. All spill events were assumed to be independent and only one event occurred for each simulation.

3.6 Post Processing and Analysis of Results

In order to assess the performance of potential monitoring stations, the values of the design objectives required calculation, as is described in Section 3.6.1. Once the objective values were calculated, they were used to create plots of objective space to determine if any solutions were dominated and provide a visual representation of the tradeoffs between different solutions. Empirical cumulative distribution functions (CDFs) and concentration profiles were also created to provide more information to the decision makers, as described in Sections 3.6.3 to 3.6.4. Section 3.6.5 discusses the analysis of spills that had background concentrations in the river, such as wastewater constituents.

3.6.1 Calculation of Objective Function Values

As identified previously, the design objectives for this thesis included

- Maximizing the probability of detecting a contamination event; and
- Maximizing the probability that a minimum amount of warning time is achieved, *given* that a detection has occurred.

In order to analyze the results, the values of both design objectives had to be calculated for each potential station location, at each sample interval considered. Specification of a method detection limit (MDL) was required prior to calculating the objective values. An MDL of 0.01 mg/L was selected because many common analyzers have an MDL of 0.01 mg/L or lower (Grayman et al., 2001). The impact of the chosen MDL on the final results is described further in Section 5.4.3. Prior to calculating objective values, the sample interval also required specification. For this thesis, sample intervals of 0.1, 1, 2, 6, and 12 hours were considered.

3.6.1.1 Probability of Detection

To calculate the probability of detection, the time that a detection occurs was conservatively calculated to be the sum of the time that the concentration first exceeds the MDL at a given location and the sample interval. This calculation results in worst case scenario performance for all simulations because it assumes that a detection occurs at the last possible instant. Although the performance of a monitoring station will be better in reality, this assumption ensured that no station could randomly perform better than it should relative to another. The calculation of detection time is illustrated in Figure 3-5.

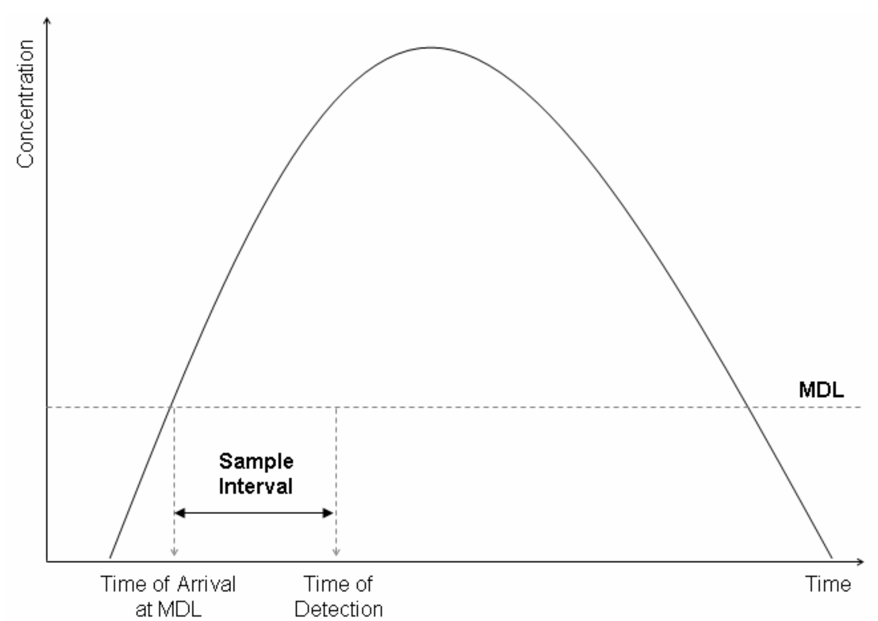

Figure 3-5. Use of MDL and Sample Interval to Calculate Time of Detection

A successful detection occurs at a monitoring station if concentrations above the MDL exist at the station for a duration greater than the sample interval. For example, in Figure 3-6 the monitoring station fails to detect the spill event because the sample frequency is greater than the duration the contaminant is present above the MDL. Once again, this is due to the conservative way in which the time of detection is calculated; in reality, short events such as this will be detected with some probability.

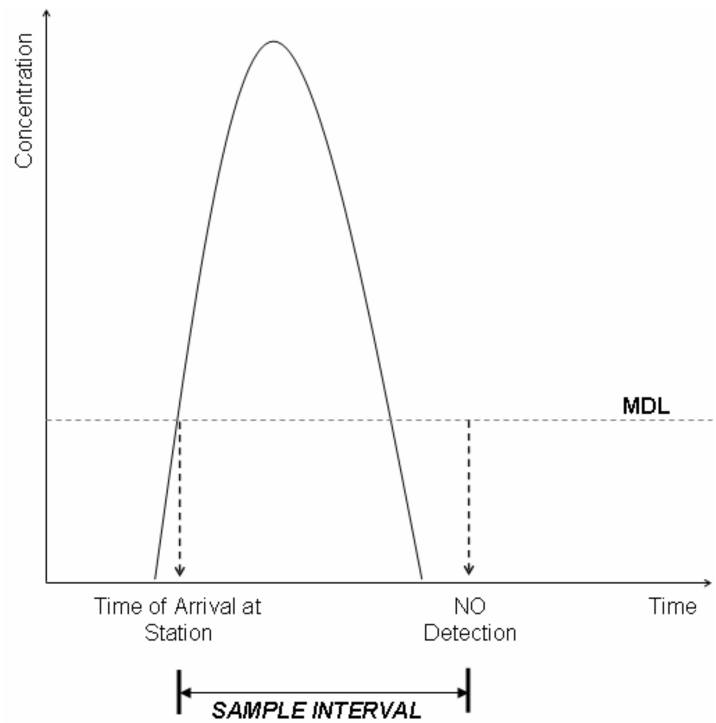

Figure 3-6. Example of a Failure to Detect in Time

MatLab code was written to cycle through all the output files for each monitoring station to determine the number of detections. The probability of detection was then calculated as the number of simulations that had a successful detection divided by the total number of simulations.

Therefore, the calculation of the probability of detection for a given spill scenario is a function of the specified MDL and sample interval, as well as the simulation results for the arrival time at the MDL and the duration a contaminant is above the MDL at a given station.

The probabilities of detection for each set of spill scenarios were combined to determine a single value for each potential monitoring station (at a given sample interval) using the following equation:

$$P_{\det} = \sum_{i=1}^{n} w_i P_i \qquad \textbf{3-5}$$

where P_{det} represents the probability of detection at a given monitoring station;

P_i represents the probability of detection at a given monitoring station for the i^{th} spill scenario;

w_i represents a weighting factor describing the relative risk level associated with the i^{th} spill scenario; and

n represents the number of identified spill scenarios.

For small sample intervals (i.e., when the sample interval is shorter than the duration a contaminant is above the MDL at a given location), the probability of detection at a monitoring location can be simplified to:

$$P_{\det} = \sum_{i=1}^{n} w_i \qquad \textbf{3-6}$$

3.6.1.2 Probability of Minimum Warning Time Given a Detection

In order to determine the probability that a minimum amount of warning time is achieved, the minimum amount of warning time required for the case study had to be specified. The Ontario Ministry of the Environment suggests a minimum travel time of two hours for defining intake protection zones (Ontario Ministry of the Environment, 2006a). This travel time represents the minimum amount of time required for water treatment plant operators to respond to an adverse event. Therefore, a minimum amount of warning time of two hours was also adopted for this research.

Post processing code was written to calculate the amount of warning time as the time of the detection subtracted from the time of arrival at the intake, as illustrated in Figure 3-7. For the purpose of this research, the time of arrival at the intake was calculated as the time the concentration reaches the MDL. This method of calculation assumes that concentrations below the MDL are not of interest; if they are, a technology with a lower MDL should be selected. In reality the concentration of interest at the intake may be even higher than the MDL, but this conservative approach was adopted for this research.

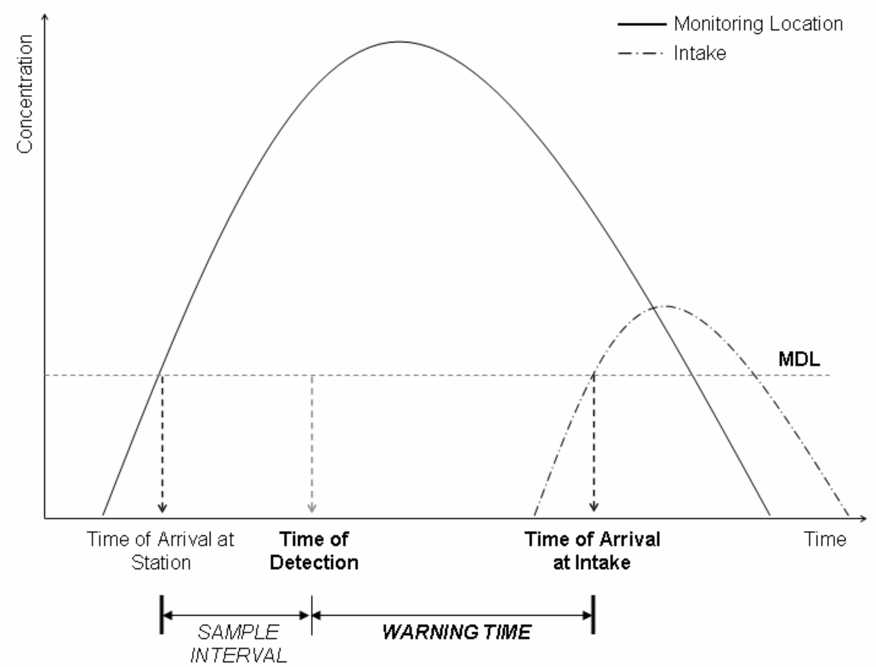

Figure 3-7. Warning Time Calculation

The probability that there are at least two hours of warning *given that a detection occurred* was then calculated as the number of simulations with a warning time greater than two hours divided by the number of simulations *that had a successful detection.*

Therefore, the probability of having at least two hours of warning at a given monitoring location for a specific spill scenario is a function of the specified MDL and sample frequency, as well as the arrival time at a given monitoring station, the duration a contaminant is present above the MDL, and the arrival time above the MDL at the intake.

The probabilities of having at least two hours of warning time given a detection occurs for each set of spill scenarios were combined to determine a single value for each potential monitoring station (at a given sample interval) using the following equation:

$$P_2 = \sum_{i=1}^{n} w_i P_{2i} \qquad \textbf{3-7}$$

where P_2 represents the probability of having two hours of warning time (given a detection occurs) at a given monitoring station;

P_{2i} represents the probability of having two hours of warning time (given a detection occurs) at a given monitoring station for the i^{th} spill scenario;

w_i represents a weighting factor describing the relative risk level associated with the i^{th} spill scenario; and

n represents the number of identified spill scenarios.

3.6.2 Pareto Optimal Curves

As discussed in Section 2.5, Pareto optimal curves can be used to analyze multi-objective problems. Therefore, the calculated objective function values were plotted in objective space for each potential station, with each axis of the graph representing one of the objectives. The non-dominated points form the Pareto optimal curve and represent the best set of feasible monitoring station locations. Deciding between the tradeoffs of the two objectives is further discussed in Chapter 5.

3.6.3 Empirical Cumulative Probability Distributions

Code was also written to calculate empirical cumulative probability distributions for the arrival time of a contaminant above the MDL and the duration a contaminant is above the MDL at a given monitoring station. The CDFs provided further information about the distribution of results used to calculate the objective values.

To create a CDF, the relevant outputs (i.e., duration or arrival time), were sorted in ascending order. The probability of each value was then calculated using the Weibull plotting position formula:

$$Probability = \frac{Rank}{n+1} \qquad \textbf{3-8}$$

where n represents the total number of simulations (McCuen, 1998).

3.6.4 Concentration Profiles

Although the monitoring station design is based on the performance of each potential station with respect to both design objectives, the creation of concentration profiles was of interest to provide more information to the decision makers. Code was written to calculate average, 95th percentile, and 5th percentile concentration profiles for a given monitoring station. The code cycles through the results of each simulation and sorts the concentration values at each time step. The average, 95th and 5th percentiles are then calculated for each time step. An example plot created for one of the monitoring stations is shown in Figure 3-8.

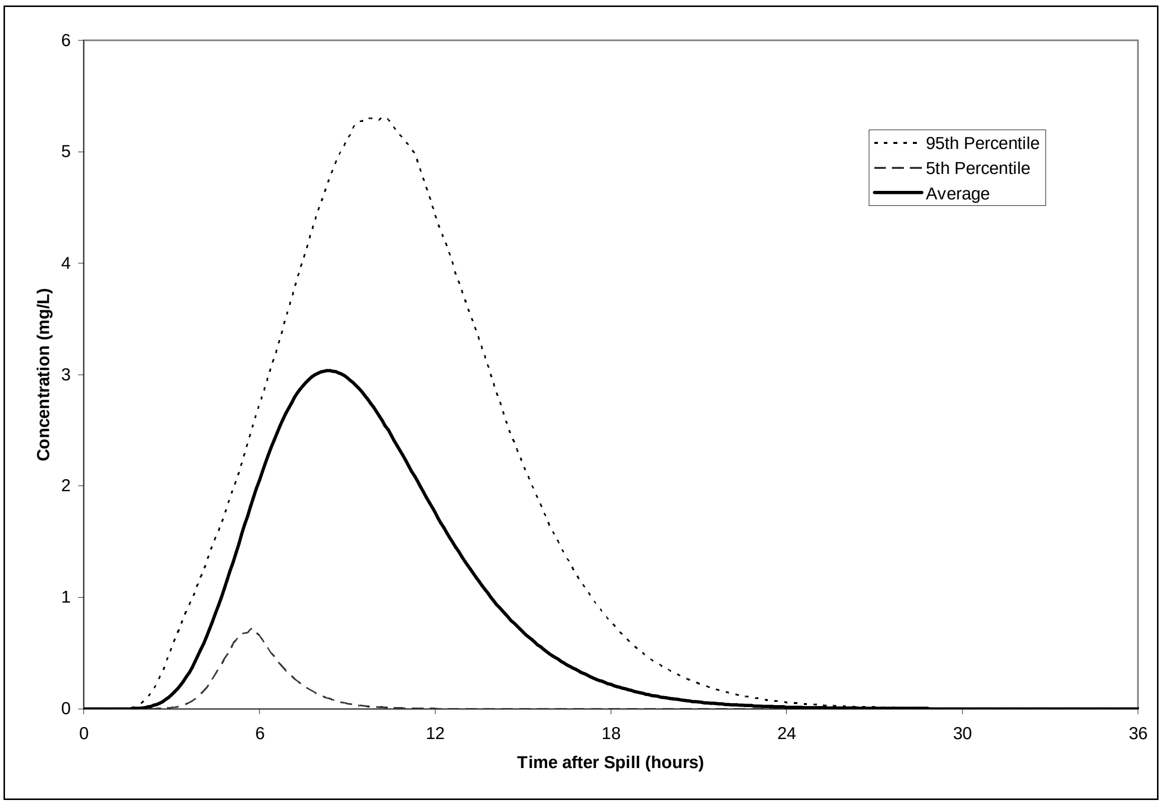

Figure 3-8. Example of 95th Percentile, 5th Percentile, and Average Concentration Profiles

3.6.5 Contaminants with Background Concentrations

Defining when a detection occurs was more complex for constituents that already existed in the water, such as nitrogen and phosphourus. Before a spill occurs, these constituents are already at concentrations above the MDL, so an alternate flag was required in order to generate a detection. Two options were considered, both of which require a comprehensive set of historical data.

The first option was to examine a long record of historical data and determine the maximum recorded concentration or some percentile value. This value could then be used as the threshold above which a detection occurs.

The second option was to initiate detection if the rate of change of the concentration is greater than some threshold value. This requires analysis of historical data recorded at a frequent interval to determine typical rates of change. The threshold rate of change can then be set to some value near the recorded maximum. This method should also have a maximum concentration specified to ensure that slow rate of change cannot mask dangerously high concentrations.

Although both methods require a long record of water quality data, the rate of change method requires more frequent data so that maximum rates can be accurately calculated. Since water quality data in the study area were very limited, the rate of change method was not feasible. Therefore, the maximum concentration recorded in the previous five years was used as the threshold above which a detection is initiated. Since water quality data were collected so infrequently, it was assumed that none of the measurements occurred during a previous spill event.

4 Case Study Application

4.1 Background

The 6,800 km^2 Grand River Watershed is home to southern Ontario's largest inland river system and is illustrated in Figure 4-1. The Grand River begins in the Village of Dundalk and empties 300 km downstream into Lake Erie at Port Maitland (Grand River Conservation Authority, 2006a).

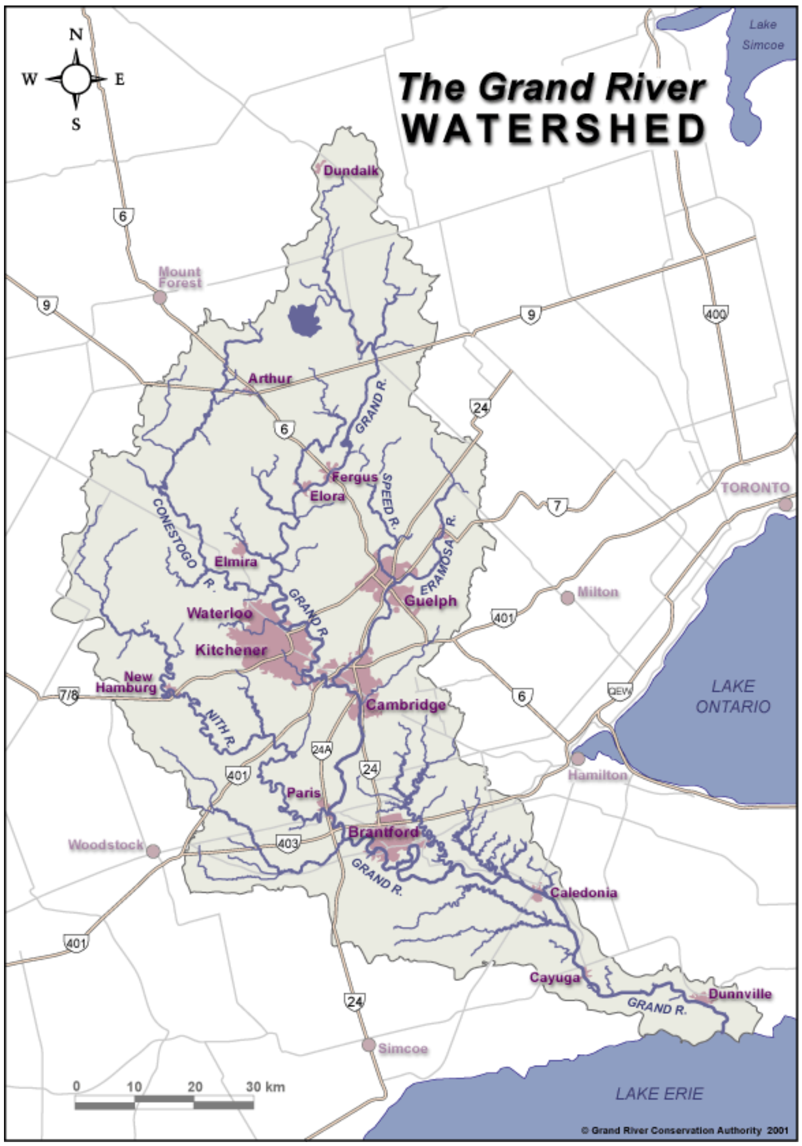

Figure 4-1. The Grand River Watershed
(Grand River Conservation Authority, 2006a)

Approximately 29% of Grand River Watershed residents obtain their drinking water from the Grand River, many of whom are serviced by the Mannheim Water Treatment Plant (WTP) located in Kitchener, Ontario. The Mannheim WTP is operated by the Regional Municipality of Waterloo and receives Grand River water via the Hidden Valley Intake. Up to 72 ML of water are withdrawn daily and stored in the Hidden Valley Reservoir prior to being pumped 10 km to the Mannheim WTP (Walton, 2006). The reservoir represents an additional level of source protection as it has multiple cells that can be isolated and diverted in the event of contamination (Stantec Consulting Ltd., 2006). The location of the Hidden Valley Intake within the Grand River Watershed is shown in Figure 4-2.

The Ministry of the Environment's Spills Action Centre (SAC) notifies operations staff at the Mannheim WTP when a spill occurs upstream of the Hidden Valley Intake. Although mandatory spill reporting exists in Ontario, some incidents go unreported if spill generators are unaware they have caused a spill or try to circumvent the law. Unreported spills may be detected by analyzers at the intake, which continuously monitor for dissolved oxygen, temperature, conductivity, ammonia, turbidity, and pH. However, most conventional online monitors are unable to detect many types of spill events (Grayman et al., 2001). Although the plant is capable of treating many spills through the use of ozonation, chlorine disinfection, and granulated activated carbon, the conservative response of shutting down the intake is typically chosen in the interest of public perception (Walton, 2006).

Real-time water quality monitoring stations exist at some locations throughout the Grand River Watershed. However, within the study area there is only one real-time station located approximately 17 km upstream of the Hidden Valley Intake. This station samples for temperature, pH, conductivity, and dissolved oxygen. Additional sampling sites that are part of the Provincial Water Quality Monitoring Network are also within the study area, but these sites are only sampled eight or nine times per year.

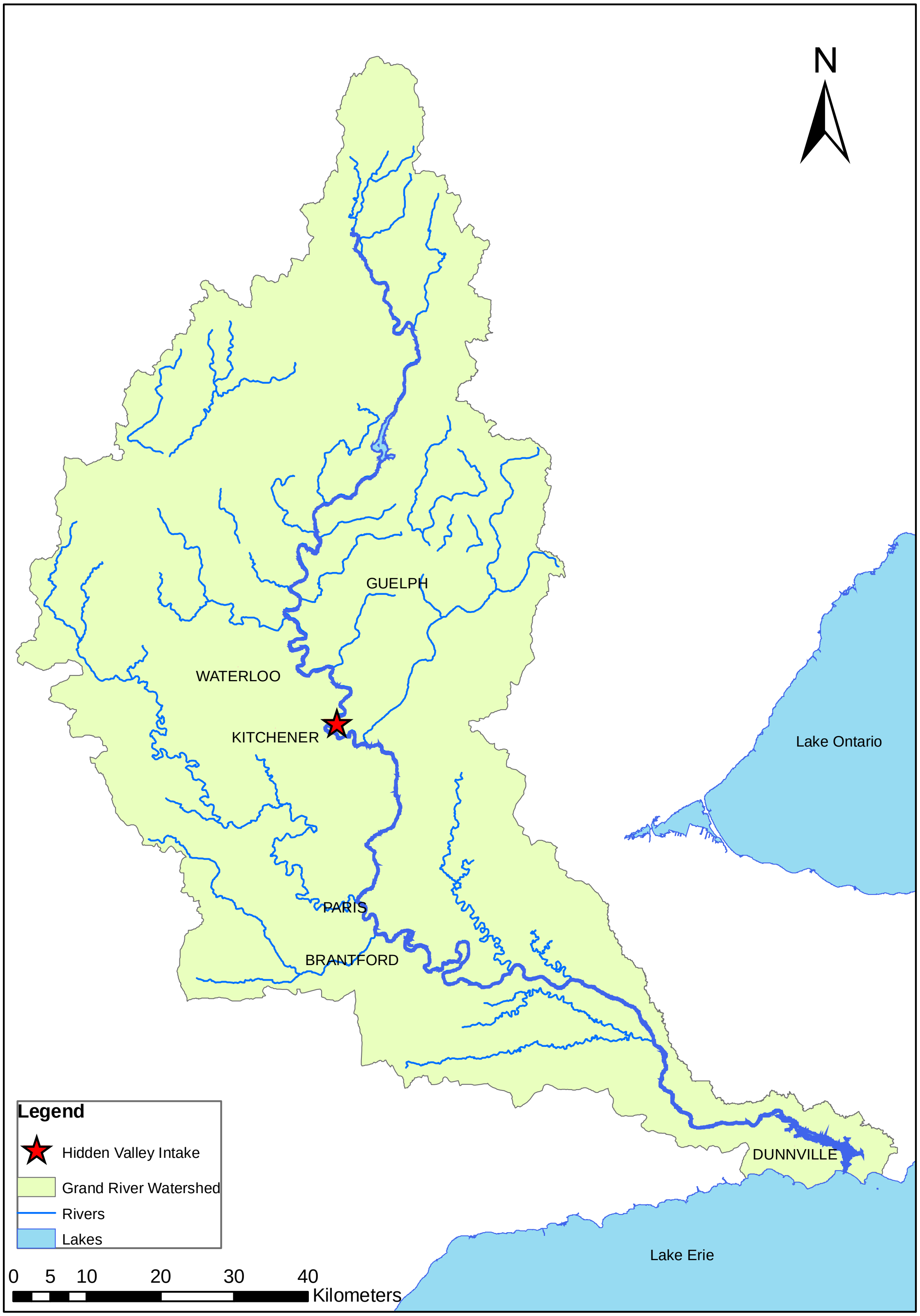

Figure 4-2. Location of the Hidden Valley Intake Within the Grand River Watershed

(Produced using data under license with the Grand River Conservation Authority, 2007)

Hazards associated with land use, spills, and wastewater treatment plant (WWTP) bypasses have been identified as significant threats to the Hidden Valley Intake because they cause sudden deterioration of water quality and can impair the drinking water treatment process (Cooke, 2006). The implementation of an early warning source water monitoring station could detect spill events that routine monitors fail to identify. In addition, positioning an early warning monitoring station upstream of the intake would provide more response time in which to close the intake and prevent contaminated water from entering the reservoir, eliminating the need to isolate and divert it after the fact. For these reasons, the need for an early warning monitoring station was identified to complement the Regional Municipality of Waterloo's source water protection planning.

4.2 Study Site

The probabilistic modelling procedure described in Chapter 3 was applied to design an early warning monitoring station for the Hidden Valley Intake. The spatial extent of the model, threats identified within the study area, and potential monitoring locations considered are illustrated in Figure 4-3.

4.2.1 Spatial Extent

As discussed in Section 3.4.2.1, the upstream and downstream boundary conditions were set at flow gauging stations. The West Montrose gauging station was selected as an appropriate boundary as it is located 38 km upstream of the intake, allowing coverage of a number of threats. It was assumed that there is minimal benefit in extending the model upstream of West Montrose so it was not worth the added computational cost.

A weir located immediately downstream of the intake served as an ideal downstream boundary condition. A rating curve was used to define the relationship between water surface elevation and flow at the downstream boundary and is illustrated in Figure 4-4 (Grand River Conservation Authority, 2006b).

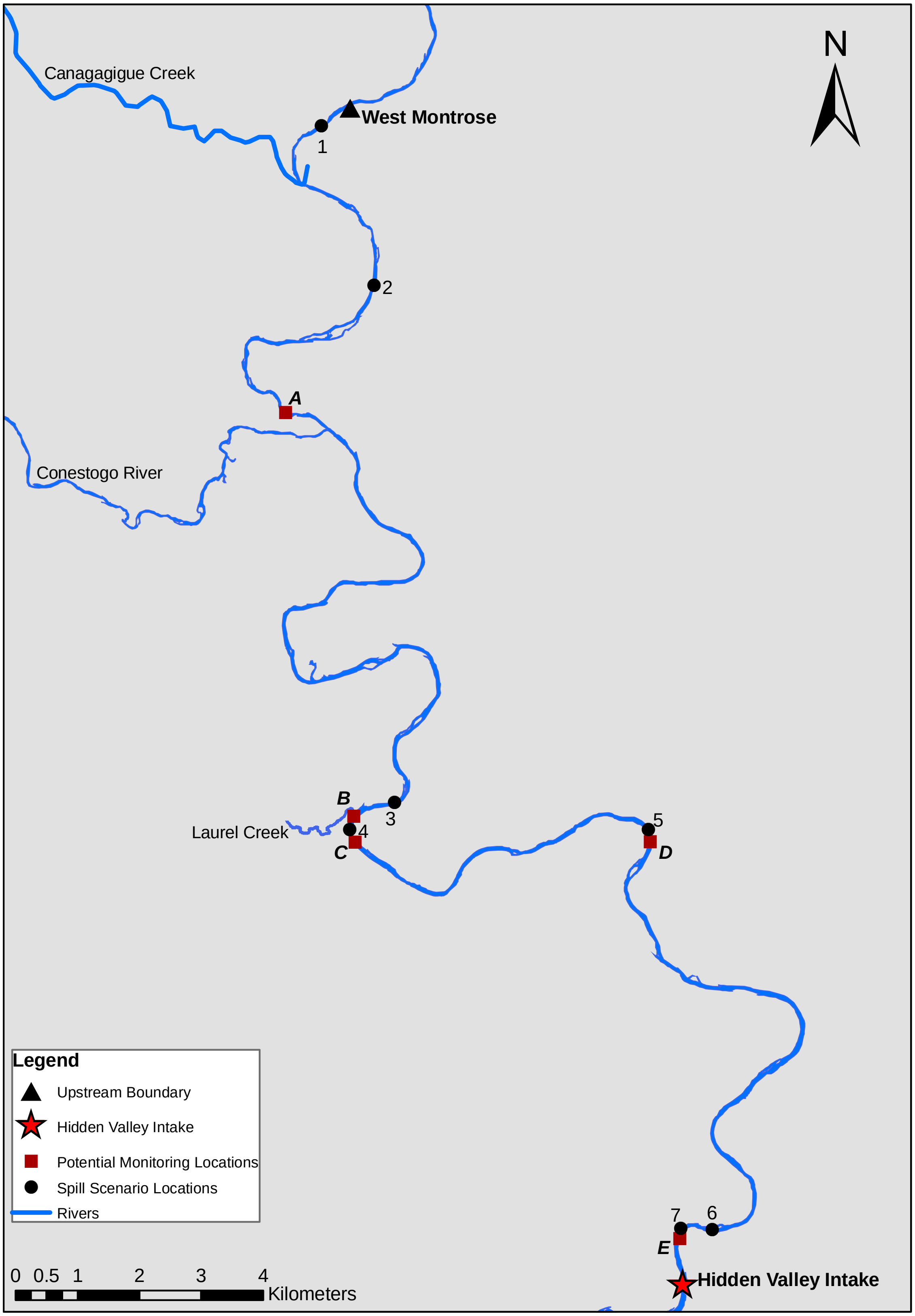

Figure 4-3. Study Site

(Produced using data under license with the Grand River Conservation Authority, 2007)

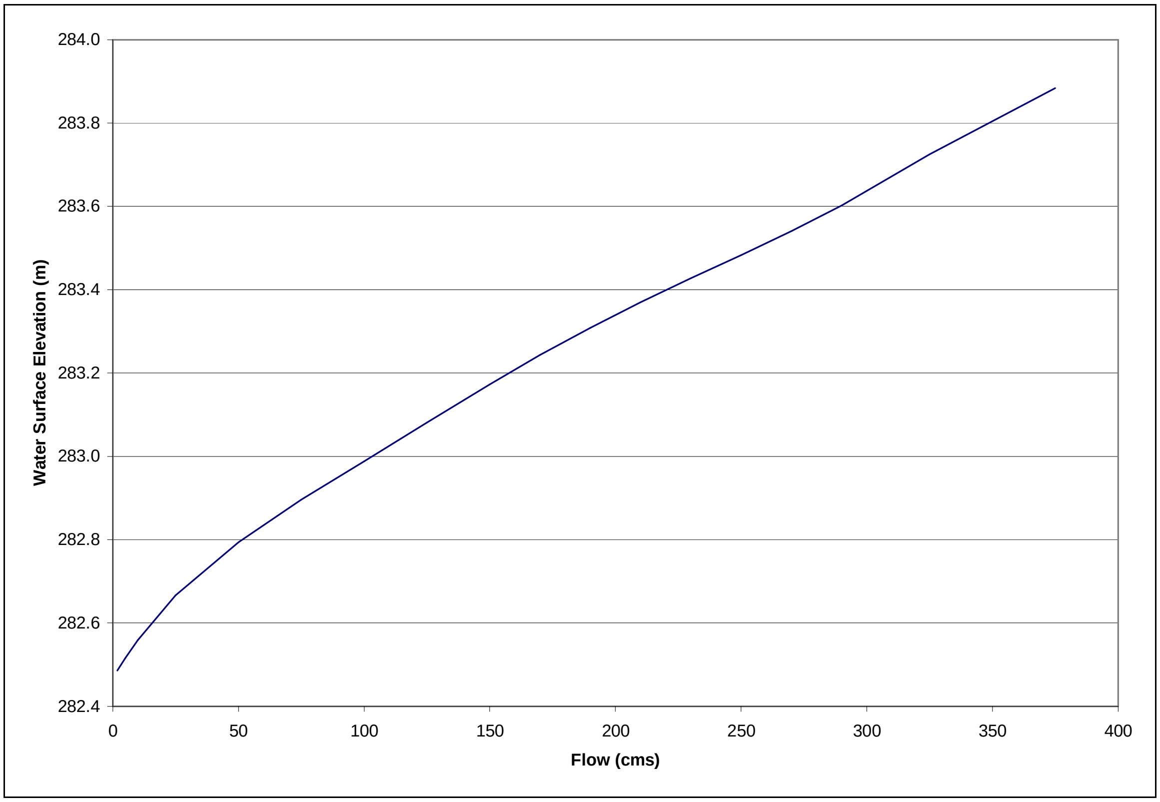

Figure 4-4. Hidden Valley Weir Rating Curve

Three major tributaries exist within the model bounds including Canagagigue Creek, Conestogo River, and Laurel Creek. Smaller tributaries, such as Cox Creek and Hopewell Creek, were not included in the model since they do not represent significant flow contributions.

Fifteen years of hourly flow data recorded at the West Montrose gauging station and stations on each of the three modelled tributaries were obtained from the Grand River Conservation Authority (Grand River Conservation Authority, 2006c). These data were subject to the following qualifications:

- The flow data are not corrected for backwater due to ice, debris or for the effects of aquatic vegetation, which may cause the flow estimates to be larger than the actual river flows.
- Sediment accumulation can cause the intakes to become partially plugged. Best efforts were made to identify plugged intakes and remove this data from the record.

- Best efforts were made to identify periods when float tapes malfunctioned and remove this data from the record.
- The data provided by the GRCA is provisional and is subject to change.

The Waterloo Wastewater Treatment Plant (WWTP) also discharges its effluent within the study area. Two years of daily flow data for the Waterloo WWTP, as well as withdrawals from the Mannheim WTP, were provided by staff at the Regional Municipality of Waterloo (2006).

Water quality data were required for the upstream boundary and for all tributary inflows. These data were available from the Ministry of the Environment as part of their Provincial Water Quality Monitoring Network. As mentioned, the water quality data were very sparse as only eight or nine data samples were taken per year (Ontario Ministry of the Environment, 2006d). As a result, seasonal average values were determined for each water quality input. Water quality data for the Waterloo WWTP were difficult to obtain, but typical values were provided for both treated effluent and raw wastewater quality (Andrews, 2007).

The study site was discretized into 51 cross sections with surveyed geometry provided by the GRCA (2006d). The cross sectional geometries were surveyed in 1975 and were resurveyed at a few locations near the intake in 2006 by Stantec Consulting. A comparison of the cross sections that were surveyed in both years was conducted to determine if the 1975 data are representative of current conditions. Two examples of this comparison are shown in Figures 4-5 and 4-6 at locations 2.3 km and 3.0 km upstream of the Hidden Valley Intake, respectively. The geometry appears to have changed only minimally at the stations compared and no major floods have occurred in the study area since 1975, so the available survey data were assumed to be a reasonable representation of the current river condition.

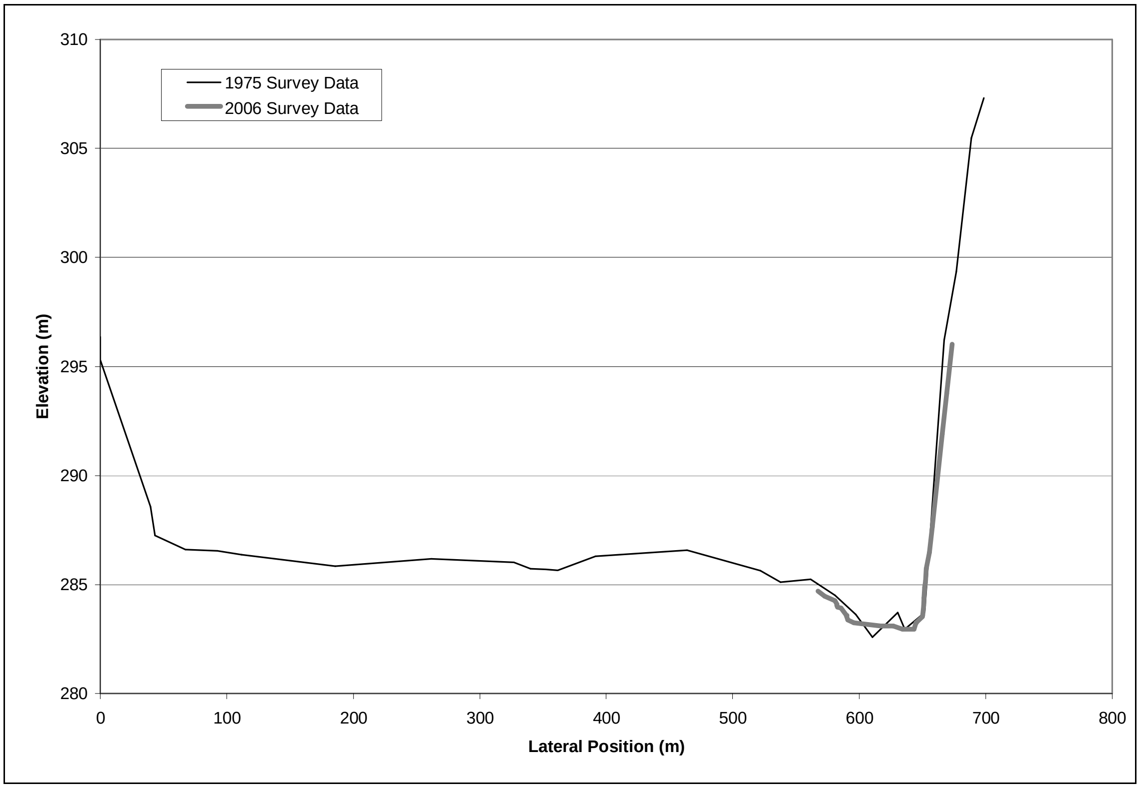

Figure 4-5. 1975 and 2006 Surveyed Cross Sections (2.3 km upstream of intake)

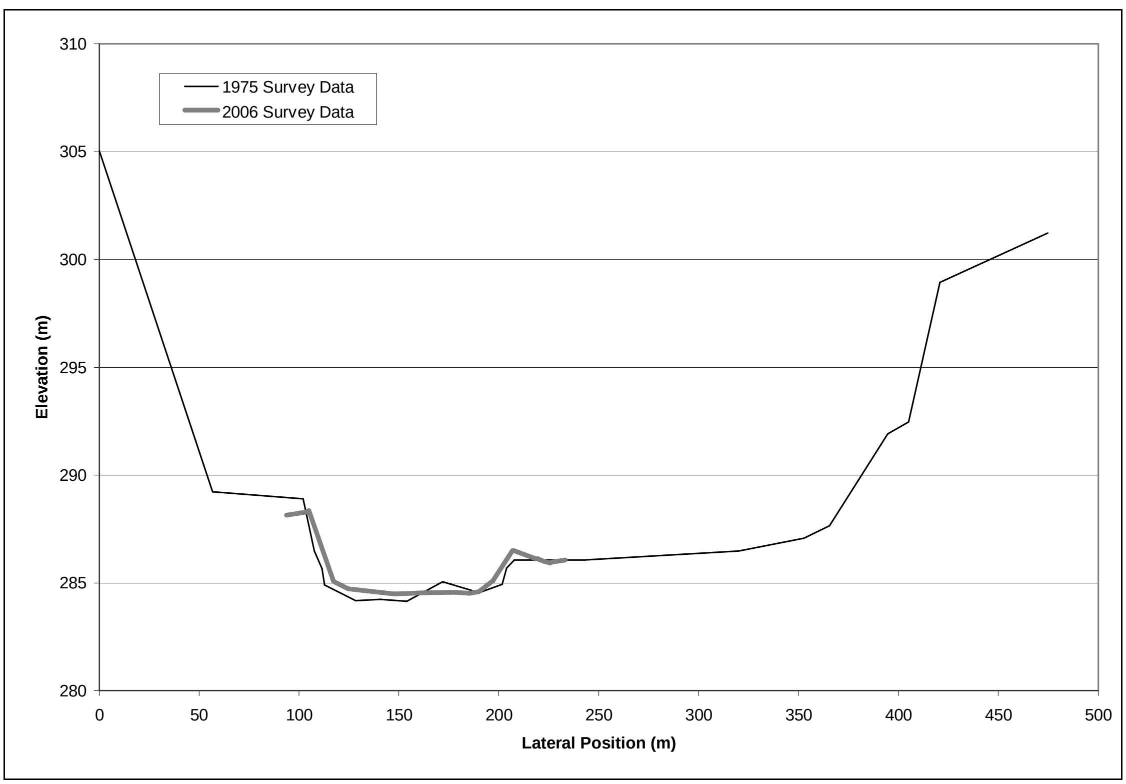

Figure 4-6. 1975 and 2006 Surveyed Cross Sections (3.0 km upstream of intake)

4.2.2 Threats Upstream of the Hidden Valley Intake

To begin the threats identification process, three years of historical spill data were obtained from Ontario's Spills Action Centre (SAC). Within this time period only 14 spills were reported in the study area. Their sources are identified in Figure 4-7.

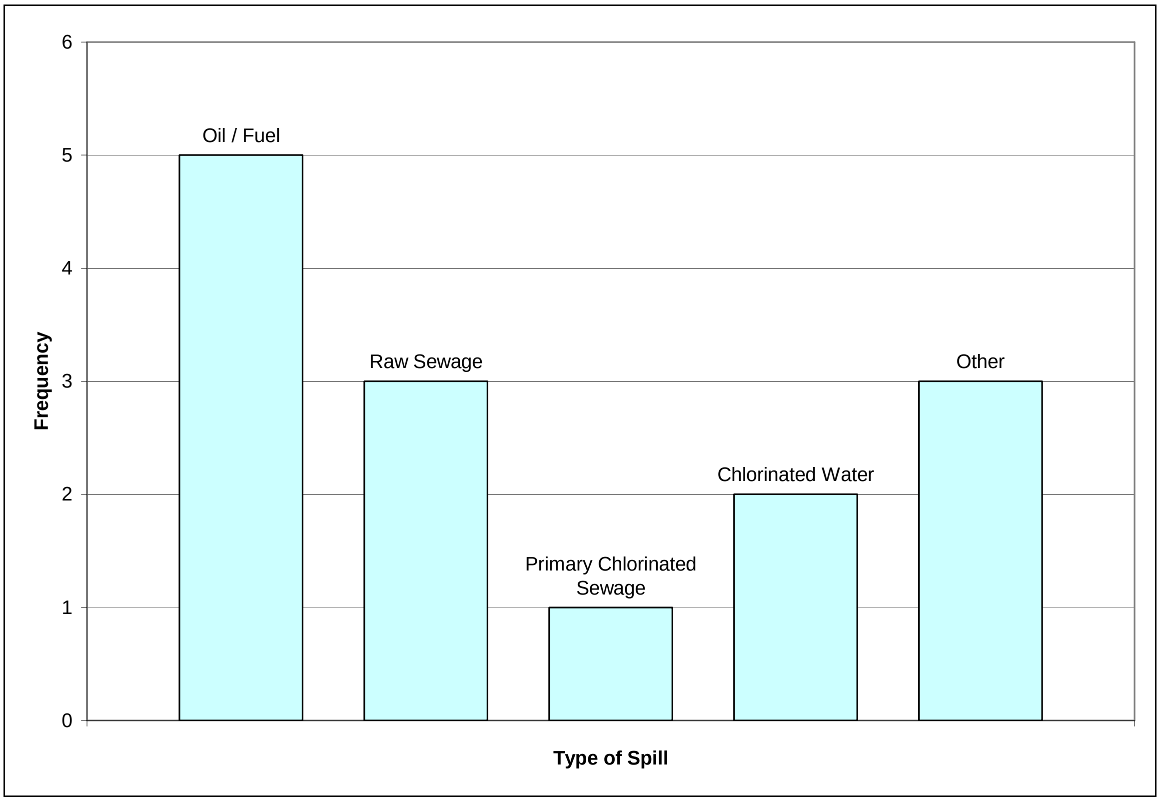

Figure 4-7. Spills Reported in Study Area, 2003-2005
(Ontario Spills Action Centre, 2006a)

Examination of the limited historical data and the province-wide data provided in Section 3.2.1 led to the identification of some threats within the study area. As previously discussed, a complete threats inventory could not be completed due to data access limitations associated with the case study.

Oil and fuel spills were identified as a definite threat within the study area and have been spilled in the past. For example, an estimated 4000 to 10000 L of gasoline were spilled adjacent to the Grand River in Cambridge in July of 2004 (Ontario Spills Action Centre, 2006a). However, for reasons described in Section 3.4.1.1, oil spills were not included as part of this research.

Cooke (2006) identified wastewater bypasses as a significant threat to drinking water treatment plants along the Grand River. The Waterloo WWTP is less than 17 km upstream of the Hidden Valley Intake and has the potential to release partially treated wastewater in the event of a bypass or raw sewage in the unlikely event of a plant failure.

Approximately 80% of the Grand River watershed is agricultural, some of which is located within the study extent. Although urban areas account for only 5% of the watershed area, much of the urban land is concentrated in the Cities of Kitchener and Waterloo which are located just upstream of the Hidden Valley Intake (Cooke, 2006). Therefore, threats associated with both urban and agricultural land uses exist within the study area.

Transportation-related spills were the most common spill source in Ontario in 2006 according to the SAC (Ontario Spills Action Centre, 2006b). Seven bridges cross the Grand River within the study area. Highway 8 has the highest traffic volume of all the bridges and is located less than 1 km upstream of the intake, representing a significant threat. The next highest volume bridges are King Street and Victoria Street, which are located 1.5 km and 11.1 km upstream of the intake, respectively. Rail bridges also exist at both King and Victoria Streets. Four other bridges within the study site are located further upstream and have much lower traffic volumes.

Based on the threats inventory that was performed, seven probabilistic spill scenarios were identified and modelled as part of this research. Information about each of these scenarios, including their location, associated contaminants, and approximate distance upstream of the Hidden Valley Intake, is listed in Table 4-1. The locations of each of the spill scenarios is shown on Figure 4-3.

Table 4-1. Spill Scenarios Modelled

Spill Scenario #	Location	Contaminant(s) Modelled	Approximate Distance U/S of Intake (km)
1	Agricultural Area Near Canagagigue Creek	Arbitrary contaminant	37.1
2	Peel Street Bridge	Arbitrary contaminant	33.7
3	Urban/residential Area in Waterloo	Arbitrary contaminant	17.1
4	Waterloo WWTP	Nitrogen and phosphorus species	16.8
5	Victoria Street Bridge	Arbitrary contaminant	11.1
6	King Street Bridge	Arbitrary contaminant	1.5
7	Highway 8 Bridge	Arbitrary contaminant	0.8

The first scenario was chosen to represent an agricultural spill occurring near the upstream extent of the model. The Peel Street Bridge was selected as the second spill scenario to represent a transportation-related spill. Spills located this far upstream are out of the public eye and may go unreported. The third spill scenario was selected at a location where the land use begins to change from agricultural to urban. The fourth spill scenario was identified as a spill of raw sewage from the Waterloo WWTP and three major bridges were identified as the fifth through seventh spill scenarios.

As shown in Table 4-1, six of the seven spills were modelled as arbitrary contaminants. Arbitrary contaminants were simulated using decay values uniformly sampled between 0 and 1 d^{-1} to represent a broad range of potential contaminants. The other spill scenario represented a spill of raw wastewater. Raw wastewater involves a number of constituents (e.g., nitrogen species, phosphorous species, dissolved oxygen), but only one was required for subsequent analyses. Total Kjehldahl nitrogen (TKN), which represents the sum of organic nitrogen and ammonia-nitrogen, was used for this purpose. A detection of a wastewater spill was generated using the method described in Section 3.6.5. A review of historical TKN data revealed that the maximum recorded concentration in the previous five years was 1.62 mg/L. Therefore, a threshold value of 1.62 mg/L was used in order to initiate a detection for the raw wastewater spill scenario.

A uniform distribution for the mass and duration of each spill was used to simulate a large range of possible scenarios with equal likelihood. The spill duration was uniformly sampled between 1 and 12 hours for each of the seven spill scenarios. One hour was selected as the lower end to simulate near instantaneous releases. Twelve hours was chosen as the upper limit to represent spills that may be discharged for a period of time before corrective action is taken. The uniform mass distribution was selected to range from 0.1 to 35 tonnes, which represent typical spill masses that have been reported in the past (Environment Canada, 2005).

4.2.3 Potential Monitoring Station Locations

Five potential monitoring station locations were identified within the study area. These stations are spaced throughout the study site and are all located at or near a bridge. Information about each of these stations is contained in Table 4-2 and their corresponding locations are indicated on Figure 4-3.

Table 4-2. Information about Potential Monitoring Locations

Station	Distance Upstream of Intake (km)	Bridge	Access Point	Major Threats Upstream
A	29.8	Sawmill Road	Dr. George Priddle Park	Canagagigue Creek Conestogo River Agricultural land use
B	16.9	Bridge Street	Economical Insurance Trailway	Laurel Creek Agricultural land use
C	16.5	Downstream of Bridge Street	Economical Insurance Trailway	Bridge Street Bridge Laurel Creek Waterloo WWTP
D	11.0	Victoria Street	Peter Hallman Family Trailway	Hopewell Creek Urban land use
E	0.8	Highway 8	Schneider Park	King Street Bridge Urban land use

Since Station A provides adequate warning time during the highest flow events, there was no need to assess any additional upstream monitoring locations (which would unnecessarily sacrifice threats coverage).   Although it provides minimal warning time, Station E was selected as the most downstream location to provide coverage of the high traffic volume King

Street and Highway 8 bridges. The attractiveness of Station E as a monitoring location will depend upon the consideration given to the retention time provided by the Hidden Valley Reservoir. As can be seen in Table 4-2, Stations B and C are very close together. Station C would normally not be considered as it is not located at a bridge; however, it was desired to include a station immediately downstream of the Waterloo WWTP in the analysis.

4.3 Hydrodynamic Calibration

Prior to conducting the hydrodynamic calibration, it was of interest to compare modelled and observed flows to determine the accuracy of the flow balance. This comparison was conducted using one year of flows from both the Bridgeport and Hidden Valley gauging stations. The results of this comparison at the Bridgeport station are illustrated in Figure 4-8.

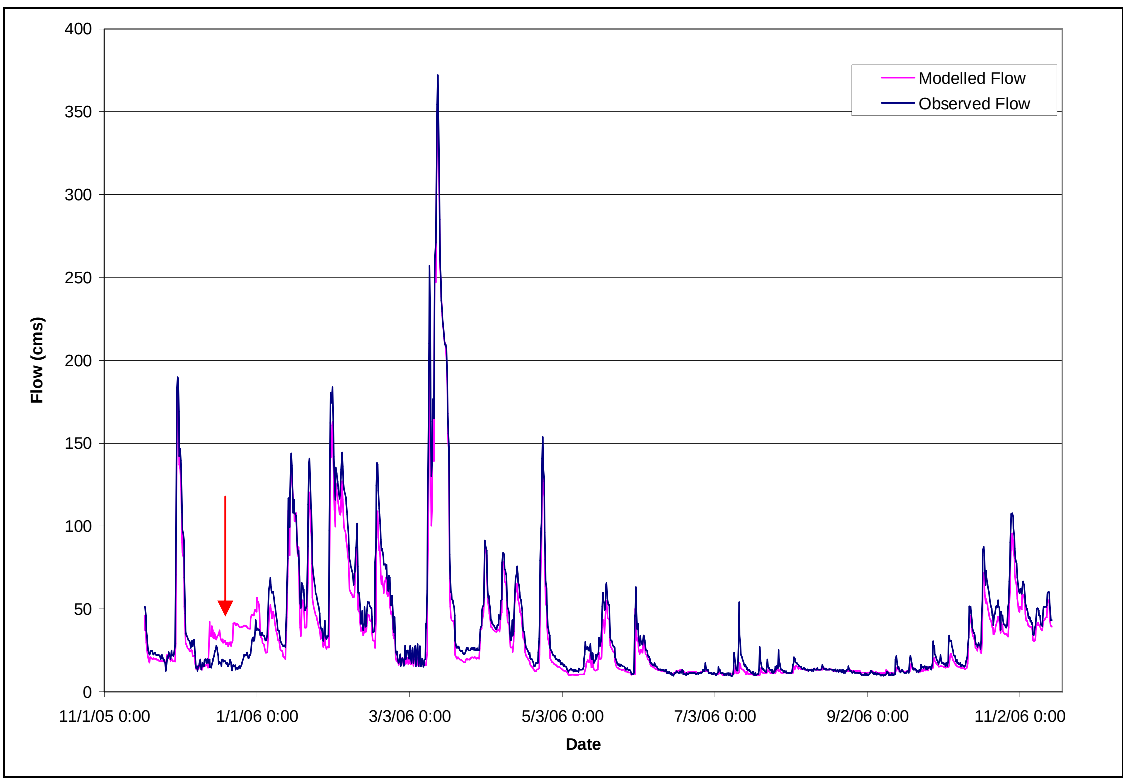

Figure 4-8. Flow Balance at Bridgeport

With the exception of the period highlighted with an arrow on Figure 4-8, the flow balance was excellent. The period where the discrepancy occurs is during the winter and may be due to backwater effects as a result of ice, as the observed data were not corrected for this. If the flow inputs from the tributaries were artificially high due to backwater effects, the resulting modelled flow would also be higher than what was observed.

Once the flow balance was determined to be satisfactory, the Manning's roughness coefficients were calibrated. One year of observed and modelled water surface elevations at a six hour frequency for both the Bridgeport and Hidden Valley gauging stations were used for this purpose. The resulting Nash Sutcliffe coefficients at the two calibration locations were 0.96 and 0.97. The excellent match between predicted and measured elevations is illustrated in Figure 4-9 and Figure 4-10, at both the Bridgeport and Hidden Valley locations, respectively. The resulting Manning's coefficients ranged from 0.03 to 0.06.

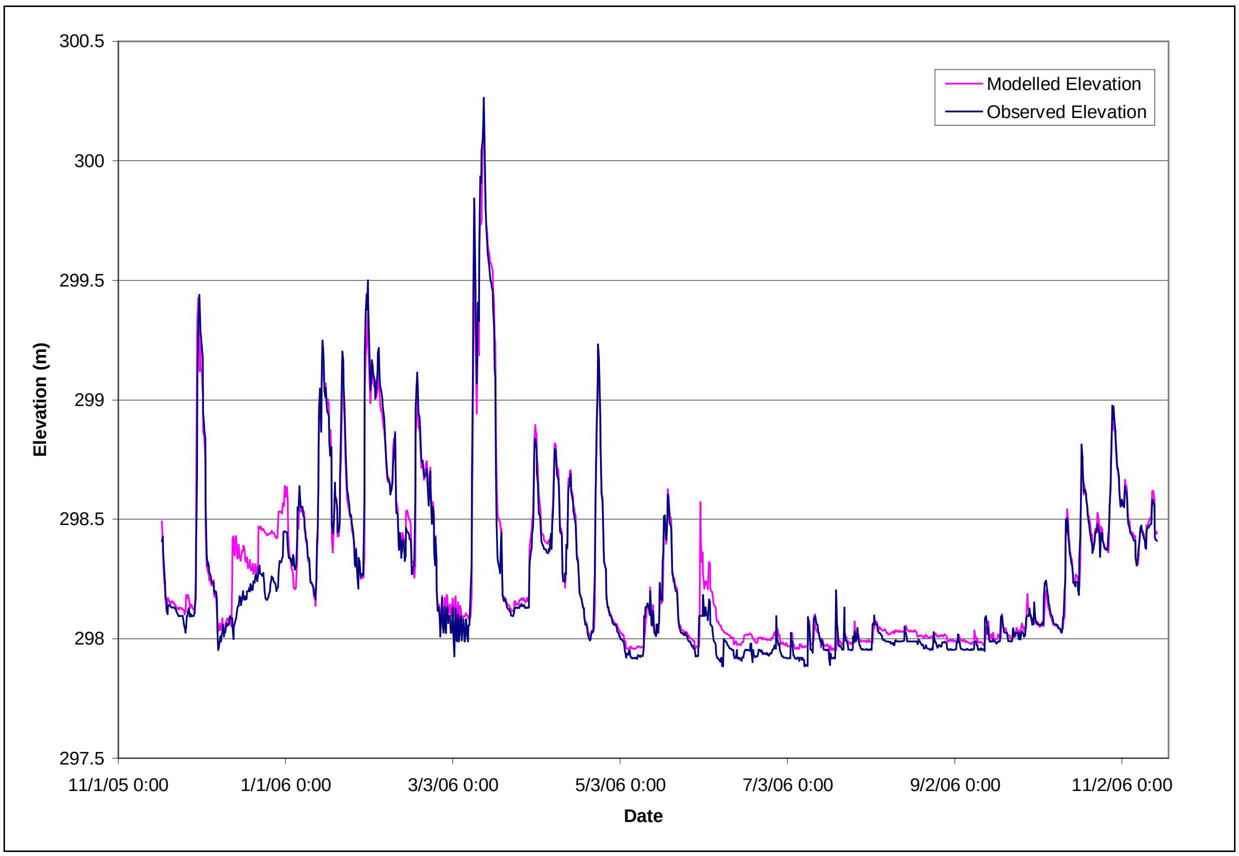

Figure 4-9. Elevation Calibration at Bridgeport

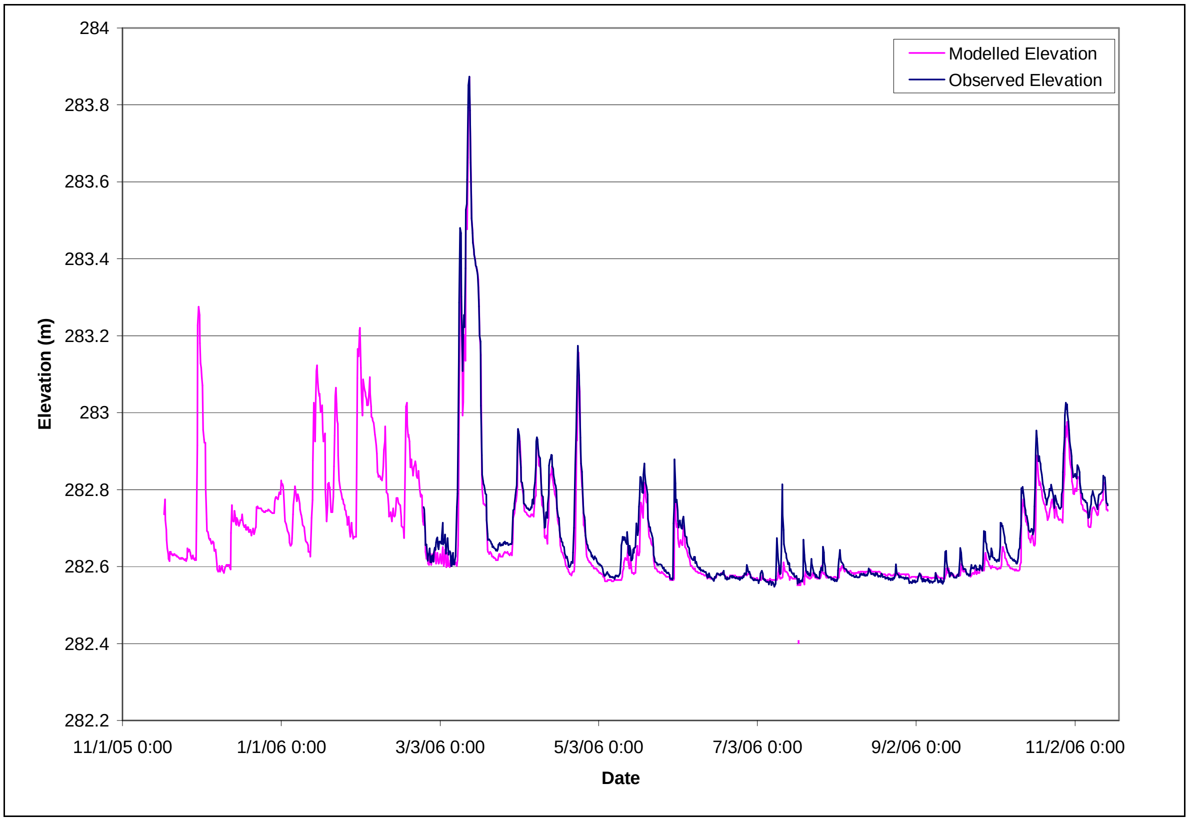

Figure 4-10. Elevation Calibration at Hidden Valley

5 Results and Analysis

This chapter presents the results obtained for the case study application. The performance of potential monitoring station locations and various sample frequencies are analyzed. A discussion of the effect of the spill scenario weightings and the method detection limit (MDL) is also presented. The chapter concludes with an analysis of the sources of uncertainty that impact the design of a source water monitoring station.

5.1 Results of Initial Experiments

Prior to conducting all of the simulations, some initial computational experiments were performed to determine the effect of parameter uncertainty, the number of simulations required, and the sample frequencies to consider in subsequent analysis. These results are presented in Sections 5.1.1 to 5.1.3, respectively.

5.1.1 Parameter Uncertainty

As discussed in Chapter 3, it was hypothesized that model parameter uncertainty may not be significant compared with the uncertain nature of the spills being modelled as part of this research. Therefore, prior to investing in computationally expensive and data intensive procedures to estimate parameter uncertainty, its significance on the decision making process was tested. A set of simulations of an arbitrary spill were executed with both flow and the parameters set as uncertain, and another set with only flow as uncertain (with the parameters set to average values). For this preliminary analysis mass and duration were held constant.

Empirical cumulative distribution functions (CDFs) were generated for the duration a contaminant is above the MDL and the time of arrival at the MDL at a given monitoring station for each set of simulations. The resulting graph for Station C is presented in Figure 5-1.

As can be seen in Figure 5-1, the pair of distributions for both the detection duration and arrival time are very similar. Since these are the only model outputs affecting the calculation of objective function values (see Figure 5-1), this indicates that parameter uncertainty contributes minimally to the overall model prediction uncertainty. Similar results were obtained for all other monitoring stations.

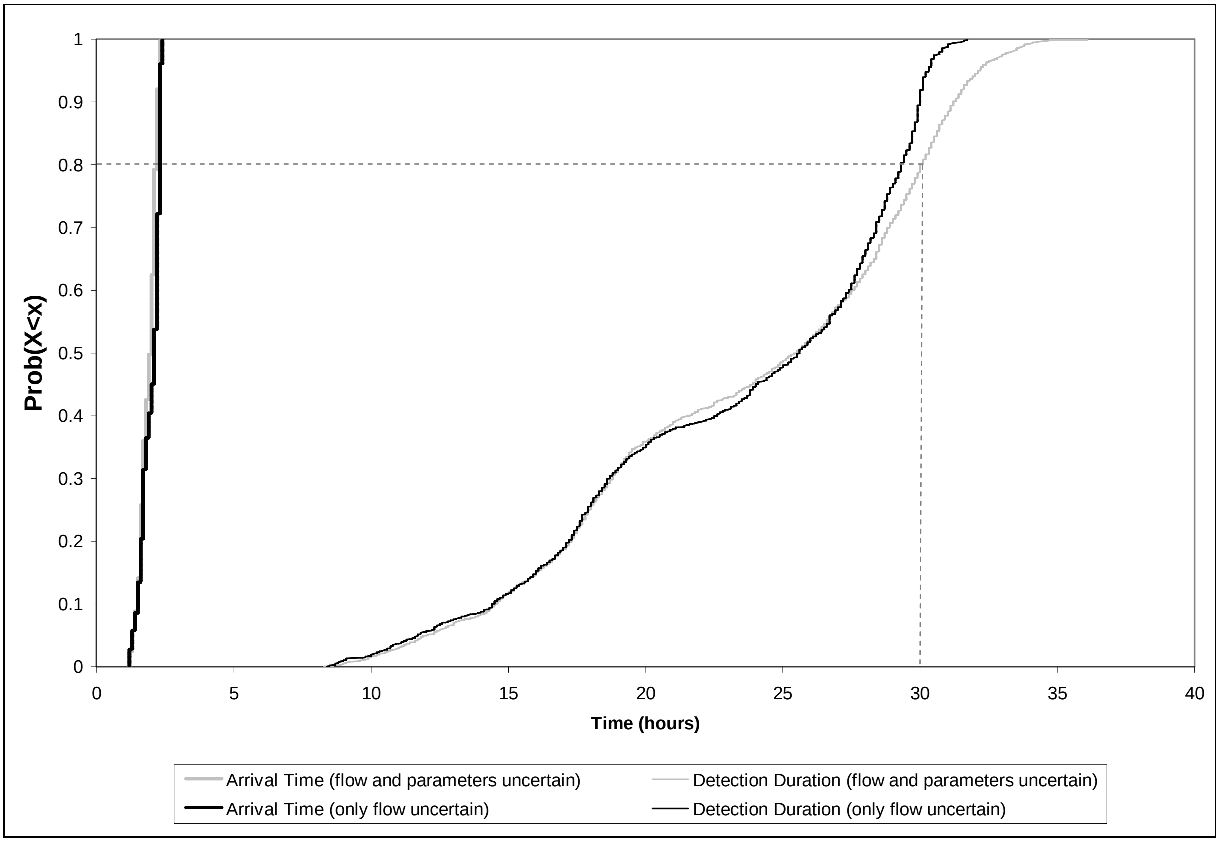

Figure 5-1. Cumulative Distribution Functions for Arrival Time and Detection Duration

The only noticeable impact of parameter uncertainty occurred for detection duration values greater than approximately 27 hours (for the specific example shown in Figure 5-1), which corresponded with low flow conditions. This minor impact, which was observed at each potential monitoring location, was attributed to the increased importance of dispersion during low flows when there is more time for contaminant plumes to spread. This slight impact on results was not a concern because design decisions were not based on low flow conditions. For example, a sample interval of 30 hours at this station is not justifiable because this would result in a failure to detect a spill event more than 80% of the time, as shown by the dashed line in Figure 5-1.

Based on the preceding results, it was concluded that parameter uncertainty is not significant for the decision making process. As a result, no conditioning experiments were required to refine the a priori parameter ranges. Since it did not significantly impact the results one way or the other, the parameters were considered uncertain in subsequent simulations rather than setting them to arbitrarily chosen discrete values.

It is important to note that as a result of not having calibrated the water quality parameters, the model is not optimally configured for use as a real-time water quality prediction tool in its current state.

5.1.2 Number of Simulations Required

In order to determine the number of simulations required, a preliminary experiment involving 2000 simulations was conducted. The average and standard deviation of the following outputs were calculated for each potential monitoring station:

- Duration the contaminant was detected;
- Warning time; and
- Peak concentration.

Figures 5-2 and 5-3 illustrate the results at one of the monitoring stations for the cumulative average of the output values and cumulative standard deviation of the output values. Based on these results, it was concluded that 1000 simulations were sufficient for subsequent model runs because the average and standard deviation values after 1000 simulations were within 1.5% of their values after 2000 simulations.

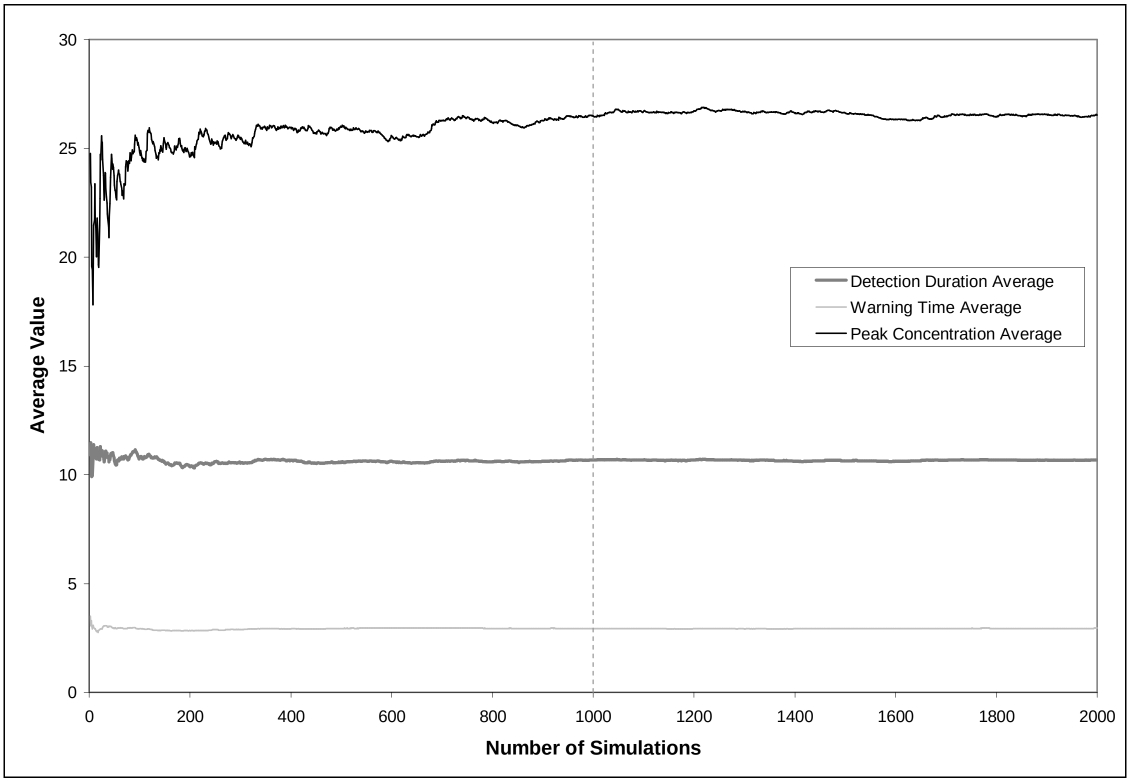

Figure 5-2. Cumulative Average of Output Values

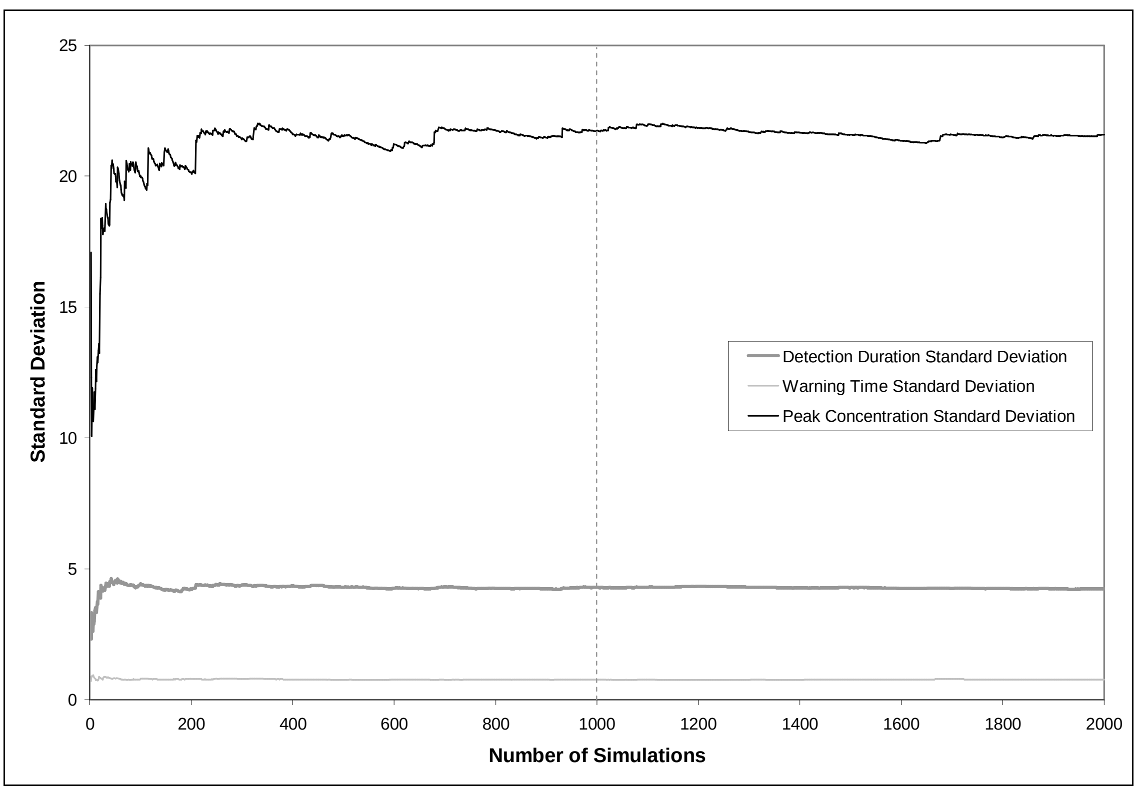

Figure 5-3. Cumulative Standard Deviation of Output Values

5.1.3 Sample Frequencies to Analyze

As discussed in Chapter 3, the time of detection was conservatively calculated as the sum of the time a contaminant arrives at a given station above the MDL and the sample interval (see Figure 5-4). As the sample interval increases, the amount of warning time decreases. If samples are collected too infrequently, a monitoring station may fail to detect an event altogether. Therefore, more frequent sampling can lead to improved values for both of the design objectives. In some cases, however, minor improvements in objective values may not justify the cost of increased sampling.

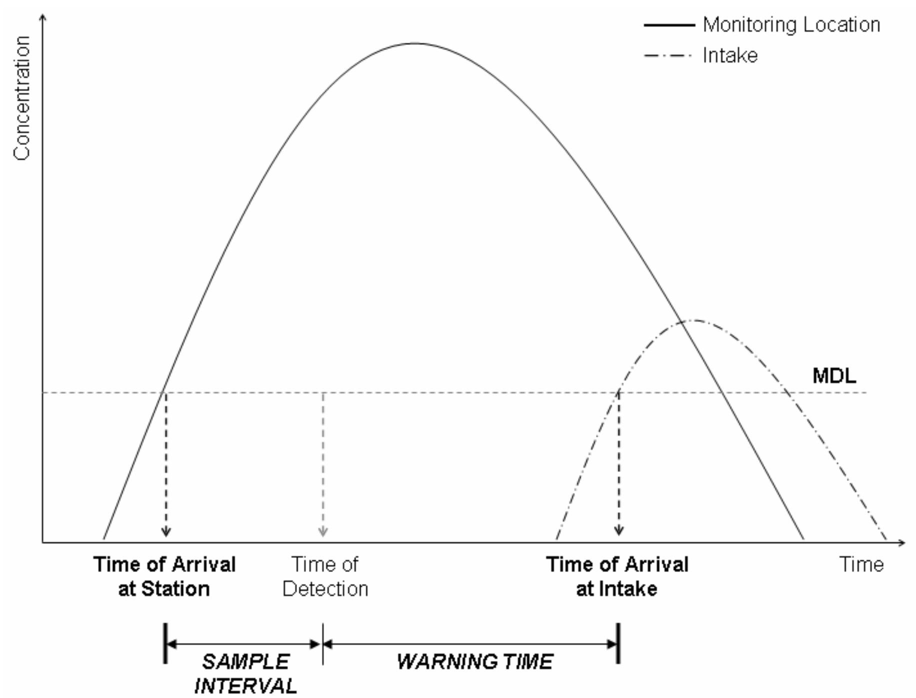

Figure 5-4. Sample Interval and Warning Time

At the outset of this research, it was desired to consider sample intervals of 0.1, 1, 2, 6, 12, and 24 hours. A preliminary analysis was performed to determine if these frequencies produced reasonable values for both of the design objectives. CDFs for the duration a contaminant is present above the MDL at the furthest upstream location (Station A) and the closest station to the intake (Station E) are presented in Figures 5-5 and 5-6. The results from each of the seven spill scenarios were equally weighted to create the CDFs for this preliminary analysis.

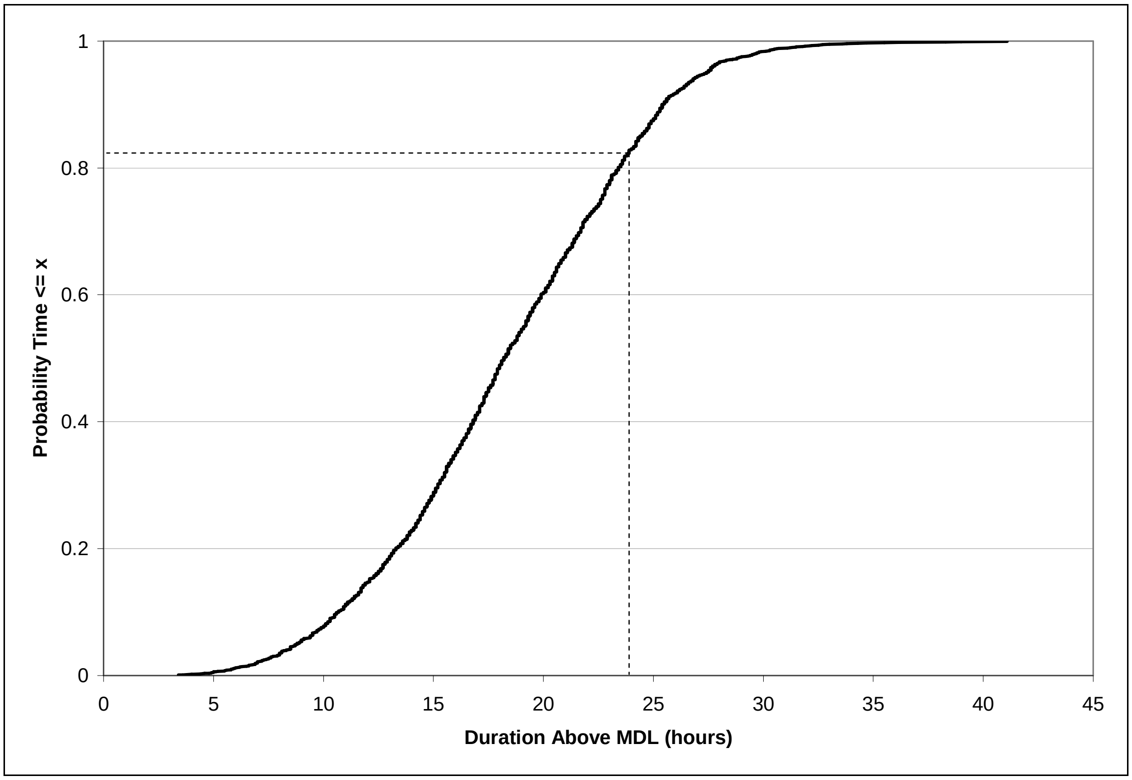

Figure 5-5. CDF of Duration Above MDL at Station A (Furthest Upstream)

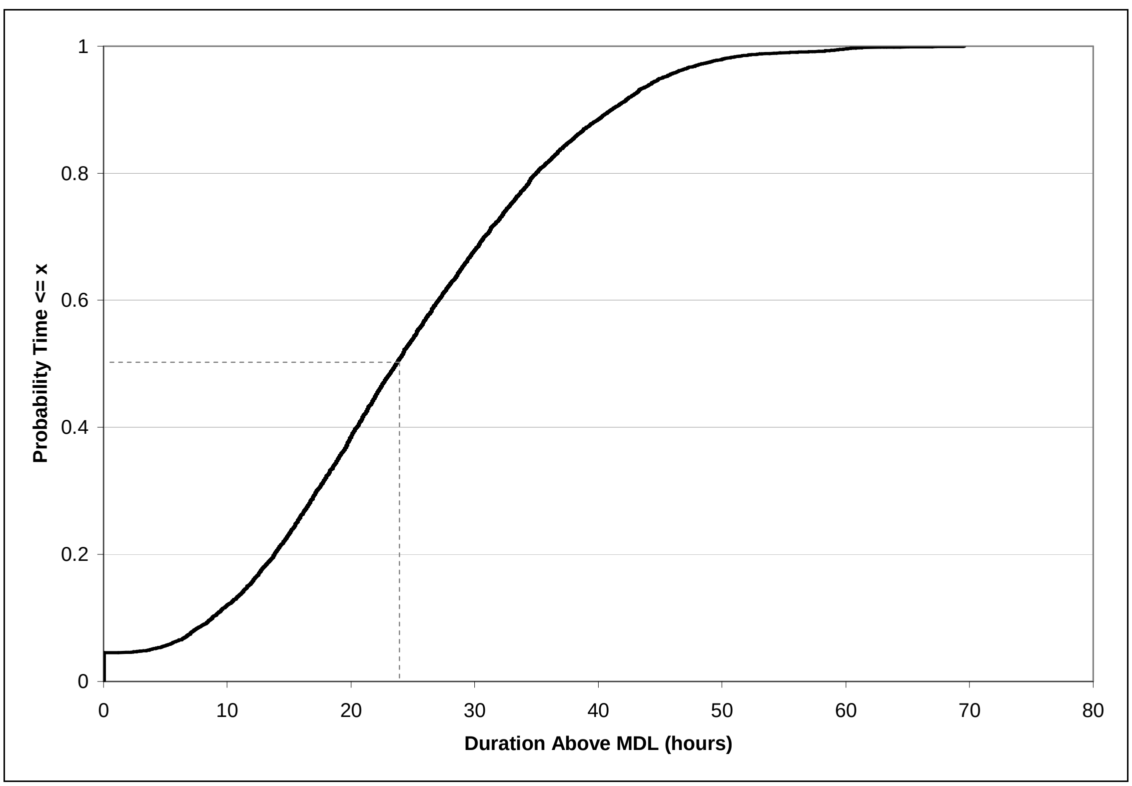

Figure 5-6. CDF of Duration Above MDL at Station E (Closest to Intake)

From Figure 5-6 it can be seen that at the most downstream station, where the concentration profile is at its widest, there is a 50% probability that contaminant concentrations will be above the MDL less than 24 hours. As a result, this station will fail to detect approximately 50% of events based on the conservative calculation of detection time used for this research. At the most upstream station where the concentration profiles are narrower, over 80% of the events will fail to be detected at a sample interval of 24 hours. This suggests that a monitoring station would be of limited value at any location using a sample interval of 24 hours.

CDFs for the difference in arrival time between each monitoring station and the intake were also plotted. The difference in arrival times is equivalent to the sum of the sample interval and warning time (see Figure 5-4). The resulting CDFs at the furthest upstream location (Station A) and the closest station to the intake (Station E) are presented in Figures 5-7 and 5-8, respectively (note the differing horizontal scales). These plots were created assuming the results from all seven spill scenarios are of equal weight. The step-wise nature of these figures is due to the fact that the model outputs at a time step of 0.1 hours.

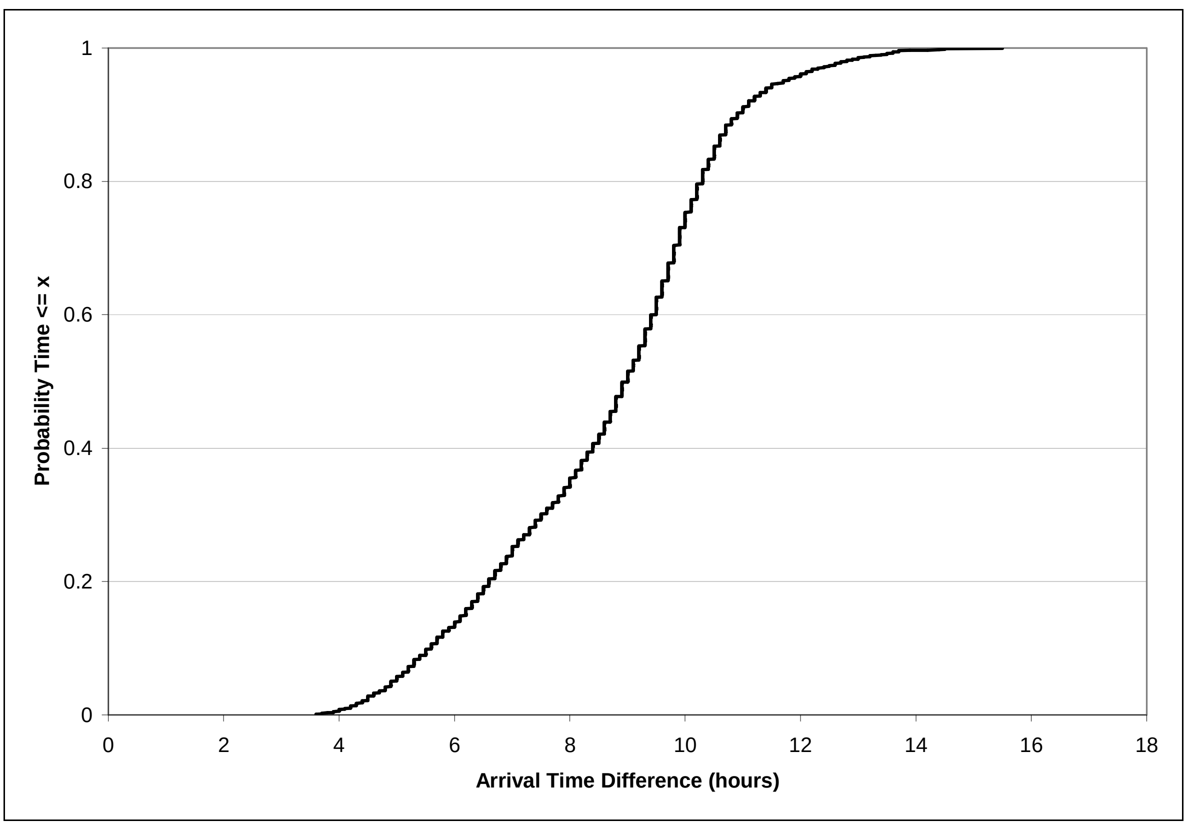

Figure 5-7. CDF of Arrival Time Difference at Station A (Furthest Upstream)

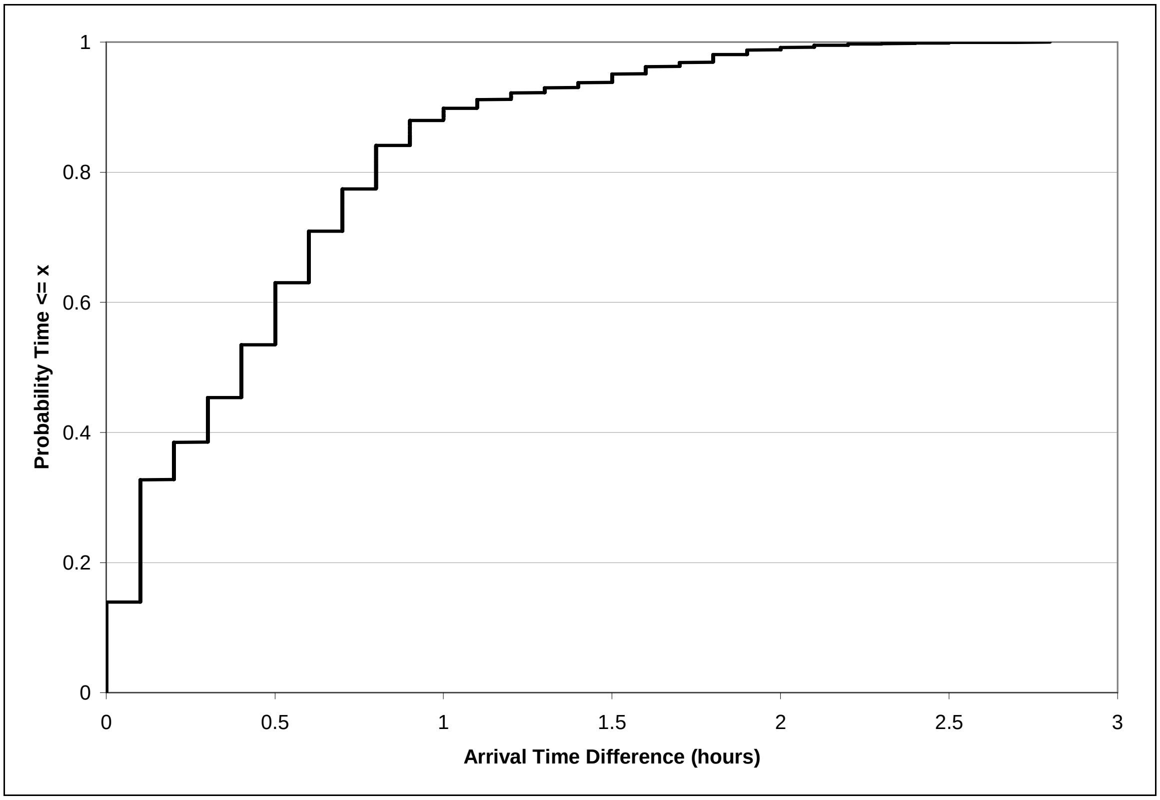

Figure 5-8. CDF of Arrival Time Difference at Station E (Closest to Intake)

In Figure 5-7 it can be seen that there is a negligible probability of having more than 16 hours between the contaminant's arrival at Station A and its arrival at the intake. Since a minimum of two hours of warning time are desired, sample intervals of more than 14 hours will almost never provide sufficient warning time at Station A or any others downstream. The few events that can be detected at sample intervals of greater than 14 hours will likely fail to provide sufficient warning time. Therefore, a 24 hour sample interval was removed from further analyses as it fails to achieve reasonable results for both design objectives.

5.2 Spill Scenario Weightings

As discussed in Chapter 4, a comprehensive threats prioritization could not be conducted due to limitations associated with the case study. As a result, defining probability distributions to represent the risk associated with different spill scenarios was difficult. Instead of limiting the analysis by attempting to define these distributions, seven discrete spill scenarios were identified and seven sets of Monte Carlo simulations were performed (as identified in Table 4-

1). This was less restricting than pre-defining a distribution because it allowed different weighting sets to be tested after the Monte Carlo simulation results were collected. The weighting sets represent relative levels of risk, which are based upon both the estimated probability of occurrence and magnitude of the effect for each spill scenario (see Section 2.4.1).

Two weighting schemes were identified for the case study; one assumed all spill scenarios were of equal risk levels and the other involved best estimates of each scenario's relative risk level using the information available. Both weighting schemes are listed in Table 5-1 and their impact on the results is discussed in Section 5.4.2.

Table 5-1. Relative Risk Weightings Applied to Each Spill Scenario

Spill Scenario	Approximate Distance U/S of Intake (km)	Weighting Scheme #1 (Equal)	Weighting Scheme #2
#1 - Agricultural Area Near Canagagigue Creek	37.1	0.142	0.1
#2 - Peel Street Bridge	33.7	0.142	0.05
#3 - Urban/residential Area in Waterloo	17.1	0.142	0.05
#4 - Waterloo WWTP	16.8	0.142	0.25
#5 - Victoria Street Bridge	11.1	0.142	0.1
#6 - King Street Bridge	1.5	0.142	0.2
#7 - Highway 8 Bridge	0.8	0.142	0.25

As can be seen in Table 5-1, the best estimate weightings (weighting scheme 2) ranged from 0.05 to 0.25 to represent different levels of risk associated with different spill scenarios. Scenarios 2 and 3 were given weightings of 0.05 as they are expected to have a low probability of occurrence. For example, the Peel Street Bridge does not have a high traffic volume so the probability of an accident is lower than it would be on King Street. Furthermore, both of these threats are relatively far upstream so their effects will be less than spills that occur immediately upstream of the intake and arrive at higher concentration levels. Scenario 1 was given a weighting of 0.1 to subjectively account for spills that may occur upstream of the model bounds. Most spills located that far upstream will be at low concentrations when they arrive at the intake. However, it is possible that a large spill could occur upstream of the model boundary so Scenario 1 was weighted higher to account for this possibility. Scenario 5 was also assigned a weighting of 0.1 because Victoria Street is closer to the intake and has more traffic than the bridges located further upstream. The King Street Bridge also has a higher traffic volume and is only 1.5 km upstream of the intake. It was therefore given a weighting of 0.2. The two most significant threats to the intake were identified as the Waterloo Wastewater Treatment Plant (WWTP) and the Highway 8 Bridge. Both of these threats were assigned weightings of 0.25.

5.3 Results

Five sample intervals and five monitoring station locations were considered, resulting in 25 discrete combinations. For each combination, the probability of detection and probability of at least two hours of warning time (given a detection) were calculated using both sets of weightings. The procedure used for these calculations is described in Section 3.6.1, and the results are contained in Tables 5-2 through 5-6 for Stations A through E, respectively.

Table 5-2. Station A Results

	Sample Interval				
	0.1 hours	**1 hour**	**2 hours**	**6 hours**	**12 hours**
Weighting #1					
Probability of Detection	0.286	0.286	0.286	0.283	0.244
Probability of $\geq$ 2 hrs Warning Time	1.000	1.000	0.995	0.666	0.003
Weighting #2					
Probability of Detection	0.150	0.150	0.150	0.149	0.130
Probability of $\geq$ 2 hrs Warning Time	1.000	1.000	0.995	0.675	0.004

Table 5-3. Station B Results

	Sample Interval				
	0.1 hours	**1 hour**	**2 hours**	**6 hours**	**12 hours**
Weighting #1					
Probability of Detection	0.429	0.429	0.428	0.417	0.350
Probability of ≥ 2 hrs Warning Time	0.994	0.928	0.774	0.067	0.000
Weighting #2					
Probability of Detection	0.200	0.200	0.200	0.196	0.171
Probability of ≥ 2 hrs Warning Time	0.995	0.944	0.806	0.081	0.000

Table 5-4. Station C Results

	Sample Interval				
	0.1 hours	**1 hour**	**2 hours**	**6 hours**	**12 hours**
Weighting #1					
Probability of Detection	0.571	0.571	0.571	0.560	0.454
Probability of ≥ 2 hrs Warning Time	0.992	0.937	0.814	0.288	0.044
Weighting #2					
Probability of Detection	0.450	0.450	0.450	0.445	0.345
Probability of ≥ 2 hrs Warning Time	0.997	0.970	0.905	0.581	0.097

Table 5-5. Station D Results

	Sample Interval				
	0.1 hours	**1 hour**	**2 hours**	**6 hours**	**12 hours**
Weighting #1					
Probability of Detection	0.714	0.714	0.714	0.701	0.585
Probability of ≥ 2 hrs Warning Time	0.912	0.703	0.516	0.153	0.005
Weighting #2					
Probability of Detection	0.550	0.550	0.550	0.541	0.425
Probability of ≥ 2 hrs Warning Time	0.930	0.771	0.662	0.347	0.011

Table 5-6. Station E Results

	Sample Interval				
	0.1 hours	**1 hour**	**2 hours**	**6 hours**	**12 hours**
Weighting #1					
Probability of Detection	1.000	1.000	1.000	0.979	0.874
Probability of ≥ 2 hrs Warning Time	0.017	0.001	0.000	0.000	0.000
Weighting #2					
Probability of Detection	1.000	1.000	0.999	0.965	0.811
Probability of ≥ 2 hrs Warning Time	0.030	0.001	0.000	0.000	0.000

Plots of objective space, with the probability of detection as the x-axis and the probability of having at least two hours of warning time (given a detection has occurred) as the y-axis, were created for each sample interval under each weighting scheme. Each plot contains five points representing the objective values for each of the potential monitoring stations. The resulting plots for a sample interval of one hour are shown in Figures 5-9 and 5-10, for weighting schemes 1 and 2, respectively. Similar plots for sample intervals of 0.1, 2, 6, and 12 hours are contained in Appendix B.

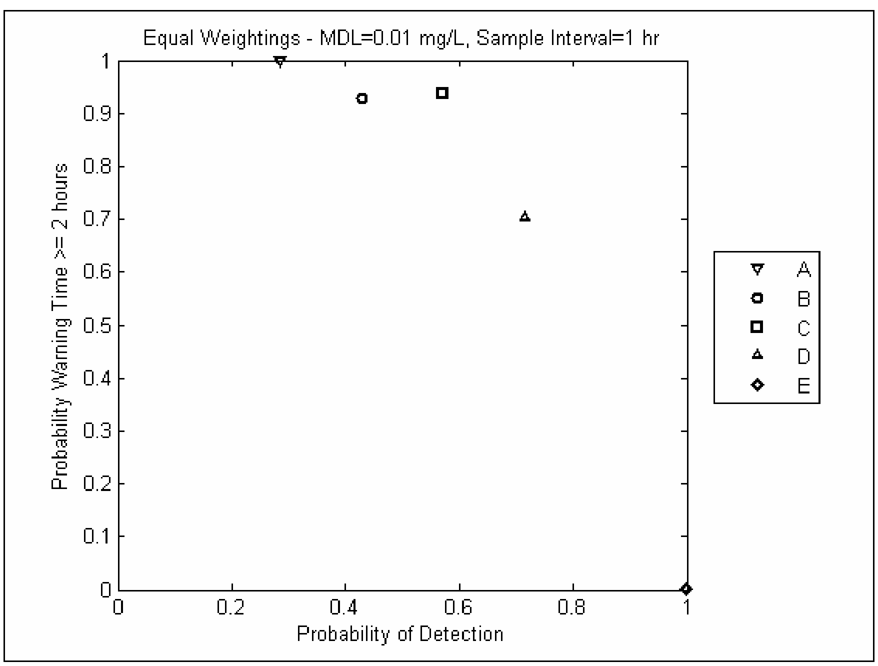

Figure 5-9. Results for 1 hour Sample Interval Using Weighting Scheme 1

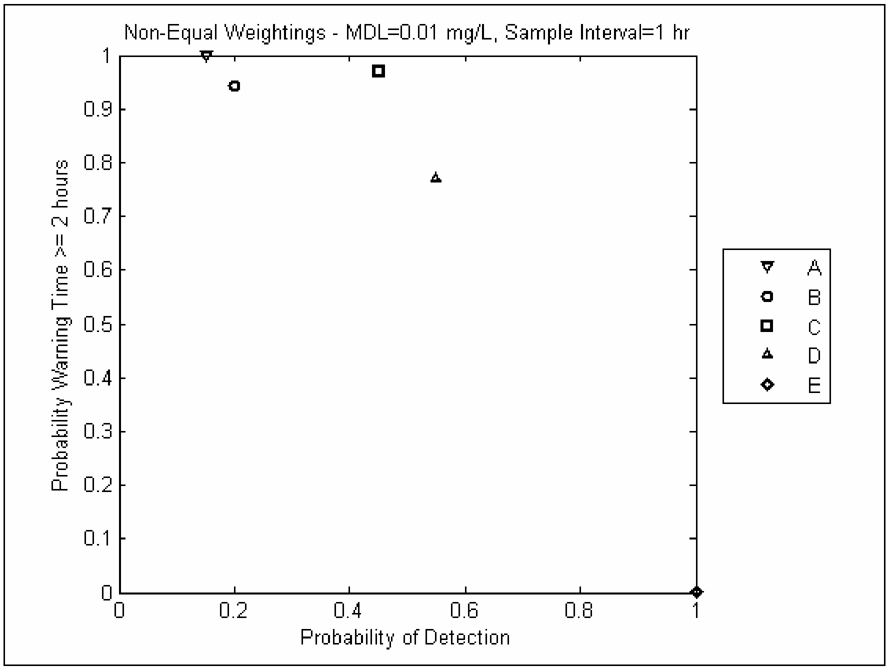

Figure 5-10. Results for 1 hour Sample Interval Using Weighting Scheme 2

5.4 Discussion and Analysis of Results

5.4.1 Inferiority of Station B

As can be seen in Figures 5-9 and 5-10, Station B is dominated by Station C because it is inferior in both objectives. This result was not expected because Station B is further upstream than Station C and should result in a higher probability of having at least two hours of warning time.

Upon examining the results, the inferiority of Station B was attributed to the fact that the wastewater spill occurs between Stations B and C. As explained in Chapter 3, the wastewater spill results had to be analyzed differently because background concentrations of wastewater constituents already exist in the river. The concentration that triggered a detection (1.62 mg/L of TKN as identified in Section 4.2.2) was much higher than the MDL used for the other six spill scenarios and resulted in greater amounts of warning time. A lower MDL results in an earlier time of detection as expected. However, since the concentration of interest at the intake is assumed to be equal to the MDL, the time of arrival at the intake is also earlier. Since concentration profiles tend to spread and decrease in steepness as a spill disperses (Chapra, 1997), the arrival time at the intake is even earlier than the arrival time at the monitoring station, resulting in less warning time. Concentration profiles at Station A, Station C, and the intake are shown in Figure 5-11 for an arbitrary spill simulation to illustrate this effect.

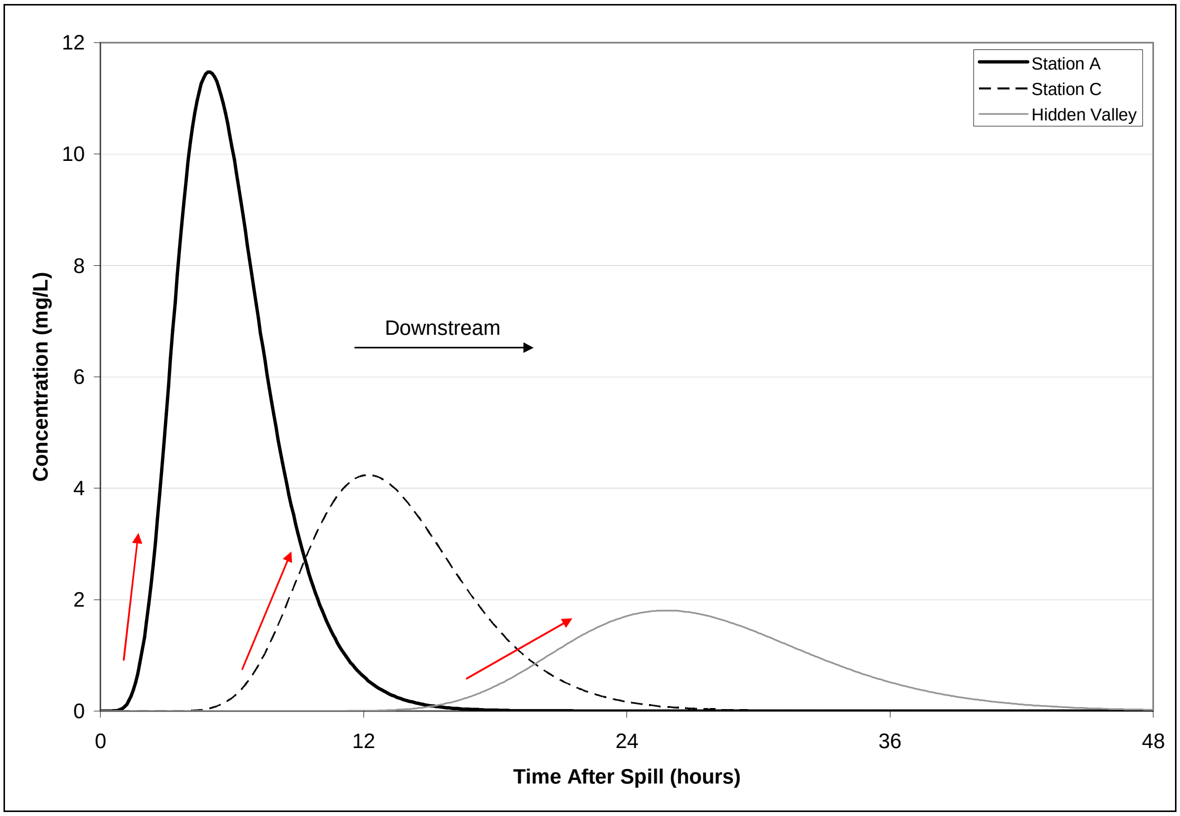

Figure 5-11. Change in Concentration Profiles Moving Downstream

Average concentration profiles based on all 1000 wastewater spill simulations are plotted for Station C and the Hidden Valley Intake in Figure 5-12 to further illustrate how warning time increases when the concentration at which a detection is triggered increases.

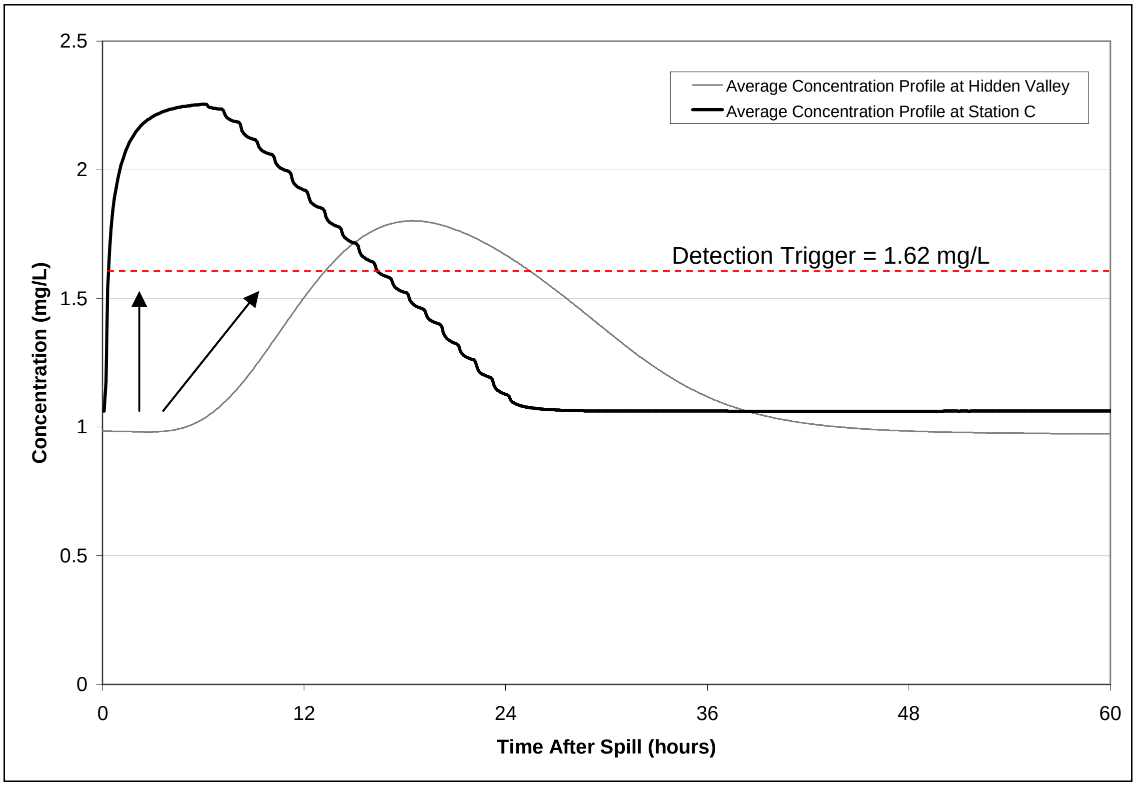

Figure 5-12. Average Concentration Profiles after a Wastewater Spill

As can be seen in Figure 5-12, the effect of increased warning time is even more amplified because Station C's concentration profile peaks very rapidly since it is located immediately downstream of the WWTP.

When the results of all seven spill scenarios were combined, the probability of having two hours of warning time at the stations located downstream of the WWTP improved relative to Stations A and B. This resulted in Station B being dominated by Station C. Strictly speaking, Station B should be removed from further analyses. However, Station B was left as an option because Station C is not located at a bridge. In the event that Station C is not technically feasible, Station B is still a viable alternative.

5.4.2 Effect of Different Risk Weightings

The probability of detection is largely based upon the cumulative weightings of the spills located upstream of a given station. Recall that the weightings represent the relative risk levels of each spill scenario. These cumulative weightings are shown in Table 5-7 for both weighting schemes.

Table 5-7. Cumulative Weightings Upstream of Each Monitoring Station

	A	B	C	D	E
Weighting Scheme #1	0.284	0.426	0.568	0.71	1
Weighting Scheme #2	0.15	0.2	0.45	0.55	1

The cumulative weightings upstream of Stations A through D are lower for the second weighting scheme, which resulted in decreased probabilities of detection at these stations. For example, when all spills are assumed to be of equal risk there is a 28.4% probability of detection at Station A for sample intervals of 0.1 to 6 hours, because the threats upstream have a total weighing of 0.284 (and all upstream events are detected at low sample intervals). Using the second set of weightings, there is a 15% probability of detecting events at Station A for the same sample intervals, once again because the threats upstream have a total weighting of 15%. Therefore, the risk weightings directly impact the probability of detection.

The different weighting schemes also lead to different amounts of warning time. The second weighting set improves the probability of having two hours of warning time for all of the stations at most sample intervals. This improvement is most obvious for Stations C and D, once again due to the increased weighting applied to the wastewater spill scenario results which provide greater amounts of warning time (as discussed in Section 5.4.1).

Since weighting scheme 1 tends to improve the probability of detection at most stations and weighting scheme 2 tends to improve the probability of having at least two hours of warning time at most stations, accurately defining the relative levels of risk associated with each scenario is essential prior to making a final decision on station location. If more emphasis is given to upstream threats, for example, Stations A and B will become more attractive because their probabilities of detection will increase while maintaining large warning times. When a full risk assessment has been completed as part of the legislated source water protection planning, the weightings can be updated, if necessary, and the objective function values recalculated for each station.

5.4.3 Effect of Selected MDL

In order to determine if the chosen MDL impacts the design decisions, the results were reanalyzed using an MDL of 0.1 mg/L (for all but the wastewater scenario). A comparison of results for both MDLs using weighting scheme 2 is presented in Figure 5-13 for a sample interval of 1 hour.

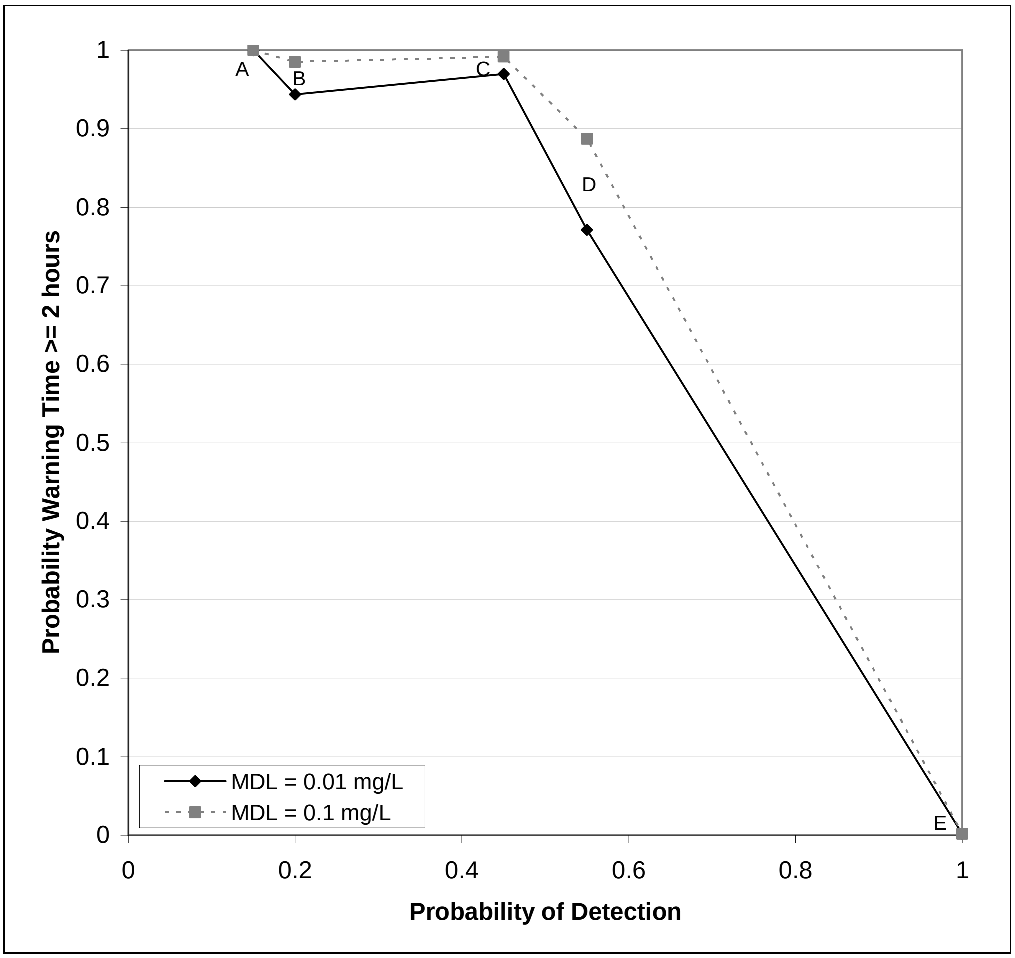

Figure 5-13. Comparison of Different MDLs at a Sample Interval of 1 hour

As can be seen in Figure 5-13, the MDL does have an impact on the probability of having at least two hours warning time, particularly at Stations B, C, and D. For reasons described in Section 5.4.1, as the MDL increases the amount of warning time also increases.

Since a higher MDL results in a contaminant being detectable for a shorter duration, the MDL can also impact the probability of detection if the sample interval is greater than the duration a contaminant is detectable. This is only the case for large sample intervals, such as 12 hours or more.

Since MDLs of 0.01 mg/L and lower are used by many common analyzers (Grayman et al., 2001), the analysis continued with the original MDL.

5.4.4 Monitoring Station Location and Sample Interval

For the purpose of this thesis, the following assumptions were made:

- Weighting scheme 2, which represents best estimates of relative risks associated with the seven identified spill scenarios, is appropriate;
- It is desired to prevent contaminated water from entering the Hidden Valley Reservoir and a minimum of two hours of warning time are required;
- Maximizing the probability of detection is also important;
- An MDL of 0.01 mg/L is appropriate and no contaminants are at levels of concern less than this MDL; and
- The concentration of interest at the intake is equal to the MDL.

Using these assumptions, a preferred monitoring station location and sample interval were identified as described in the following sections.

5.4.4.1 Monitoring Station Location

The results for each monitoring station and each sample interval are presented together in Figure 5-14. Each line represents the Pareto front for a different sample interval, with a point representing each potential monitoring station. Dominated points are shown but are not part of the Pareto front.

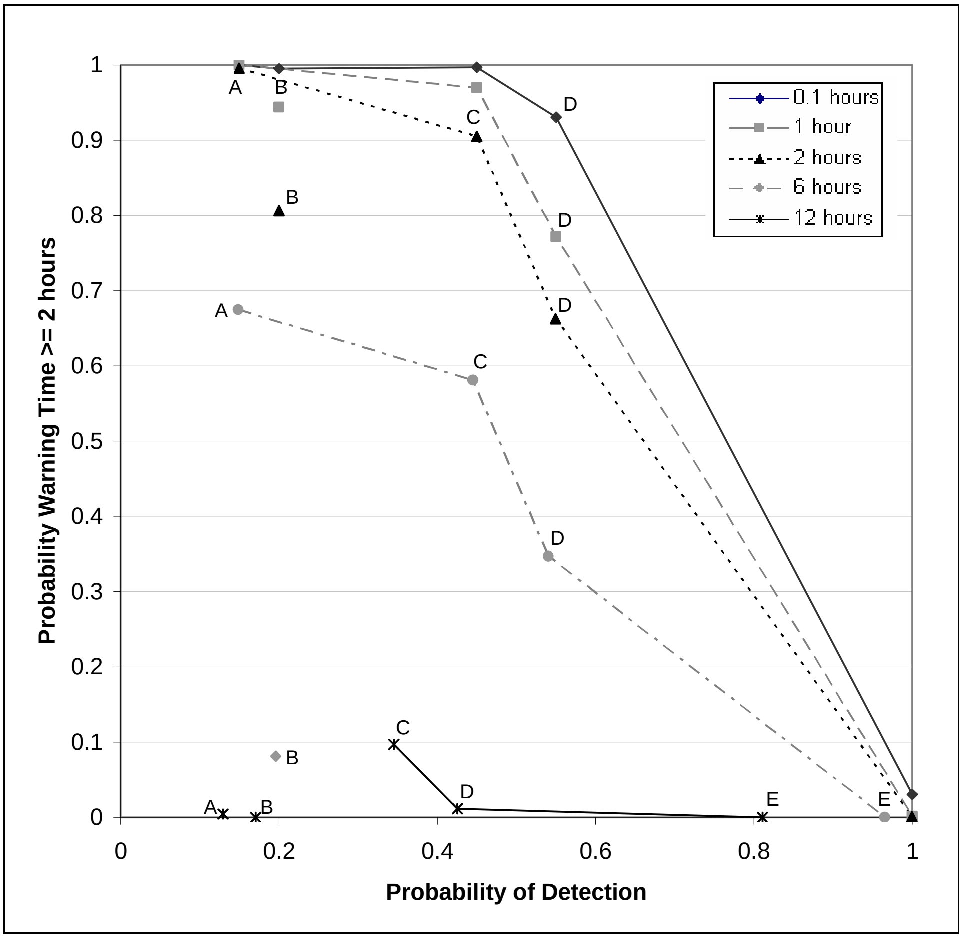

Figure 5-14. Results Using Different Sample Intervals (Weighting Scheme 2)

For most sample frequencies and both sets of risk weightings, Stations C and D appear to be a good compromise in objectives. This was expected as these stations are located downstream of many of the major identified threats, but are located sufficiently far upstream of the intake to provide more warning time than Station E.

If one of the two objectives is deemed to be of much greater importance, then stations at either extreme may be preferable. For example, if decision makers want to maximize threats coverage and are comfortable with contaminated water entering the Hidden Valley Reservoir, Station E may be preferred. A late detection from Station E is still of value because contaminated water can be diverted from the Hidden Valley Reservoir and prevented from entering the Mannheim Water Treatment Plant. Although it provides almost no warning time, Station E may be selected by decision makers depending on the tradeoff they choose between the two objectives. Alternatively, if operations staff want as much warning time as possible, Station A may be selected even though it provides less coverage of all possible threats.

In some cases there is a large difference between Station D and Station E with respect to the probability of achieving at least two hours of warning time. This was expected as there is a 10 km gap between these stations. A station located between them may be a good compromise in the objectives; however, no bridges are located along this stretch of the river. If a suitable location between D and E can be found, this may be an attractive option to decision makers.

Given the assumptions stated above, Stations A and E were not selected as they do not represent a compromise in objectives. Station B was not selected because it is dominated, so the choice was narrowed down to Station C or D. Station D is recommended as the preferred location because it was assumed that a slightly higher risk of having less than two hours of warning time was worth the increased probability of detection provided by Station D. As a result of the conservative approach used for calculating the time of detection (see Section 3.6.1.2), the actual amount of warning time provided at Station D is expected to be even better than the conservative results presented here.

5.4.4.2 Sample Frequency

Since the monitoring station location is fixed, and the sample interval can easily be adjusted, it was decided that the required sample interval would be selected after the location was chosen. As discussed, more frequent sampling results in improved objective values. This is shown in Figure 5-14 for weighting scheme 2.

For Station A, it can also be seen that there is almost no difference in either objective between sampling every 6 minutes and every 2 hours. This suggests that if Station A were to be selected, there is minimal benefit in sampling more often than every 2 hours.

Further downstream, the tradeoff between sampling every 6 minutes to improve the warning time and sampling less often with shorter warning times needs to be evaluated. It may not be worth sampling ten or twenty times as often if sample frequencies of every hour or two produce acceptable probabilities of detection and sufficient warning time.

The probability of having two hours of warning time decreases substantially at a sample interval of 12 hours, particularly at Stations A through D. It can also be seen that a sample interval of 12 hours never results in a probability of having two hours of warning time of more than 10%. The probability of detection also decreases for most stations at a sample interval of 12 hours. This is due to the fact that not all concentration profiles exist above the MDL for more than 12 hours and some events will pass by the monitoring station without being detected. Therefore, a sample interval of 12 hours is not appropriate at any location.

Sampling every 6 minutes at Station D results in a probability of having two hours of warning time of 93%. Sampling every hour results in a probability of 77%, and sampling every two hours results in a probability of only 66%. Therefore, sampling at least every hour, if not continuously, is recommended in order to ensure a high probability of having two hours of warning time at Station D.

5.5 Sources of Uncertainty Impacting Results

As identified in Chapter 3, each simulation involved the following sources of uncertainty:

- Time of the spill (flow);
- Mass of the spill;
- Duration of the spill; and
- Water quality parameters associated with the spill.

As discussed in Section 5.1.1, the uncertainty associated with the water quality parameters had minimal impact on the distribution of the relevant results. It was further hypothesized that neither the spill mass nor the spill duration contributes significantly to the distribution of the results.

In order to determine the sources of uncertainty impacting the results, a second set of 1000 Monte Carlo simulations were performed for Spill Scenario 1 with the flow at the time of the spill as the only uncertainty. Scenario 1 was chosen for this analysis because it is the furthest spill upstream so results at both a near and far monitoring station could be assessed. The mass of the spill was set to 1500 kg for each of the simulations and the spill duration was set to one hour. The resulting average concentration profiles at Station C and the intake are shown in Figure 5-15. This figure also shows the corresponding concentration profiles for the original results that considered mass and duration as uncertain.

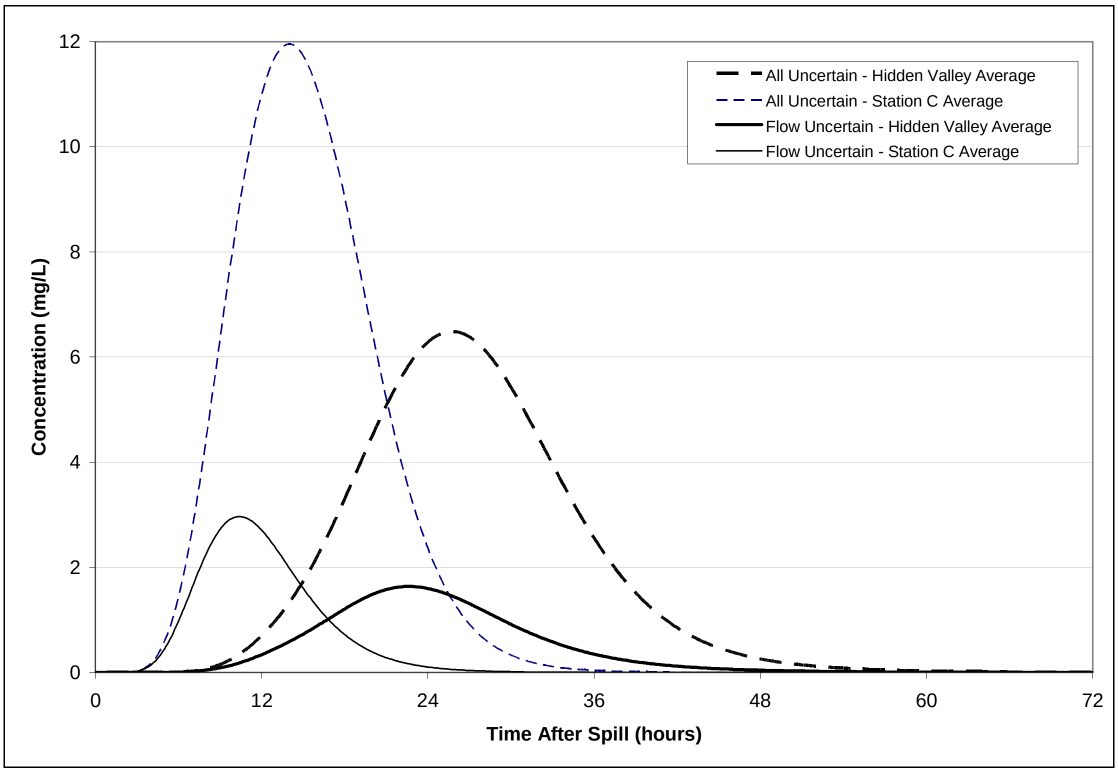

Figure 5-15. Average Concentration Profiles With and Without Uncertain Inputs

As can be seen in Figure 5-15, the shape of the concentration profile for each set of simulations is quite different. For the case when all inputs are uncertain, the peak concentration is much higher, the time to peak later, and the profile wider. This was expected due to the wide range of spill masses and durations that were modelled. Longer durations result in later peaks and wider concentration profiles. Despite these differences, one similarity is that the time of arrival of both profiles is very similar at Station C and the intake. At an MDL of 0.01 mg/L, the time of arrival at Station C for both average profiles is exactly the same (2.7 hours) and the times of arrival at the intake are within 1.5% of each other.

Since warning time is a function of the time of arrival at the monitoring station, time of arrival at the intake, and the sample frequency, the distribution of warning time for both sets of simulations should also be similar. This is clearly shown in Figure 5-16 and Figure 5-17, which illustrate CDFs of warning time for the case that considered all inputs uncertain and the case that considered only flow uncertain. Both plots were created using a sample frequency of 0.1 hours and were based on the results from Spill Scenario 1. Stations A and E were chosen as examples to represent the distributions at a station immediately downstream of the spill and a station some distance downstream, respectively.

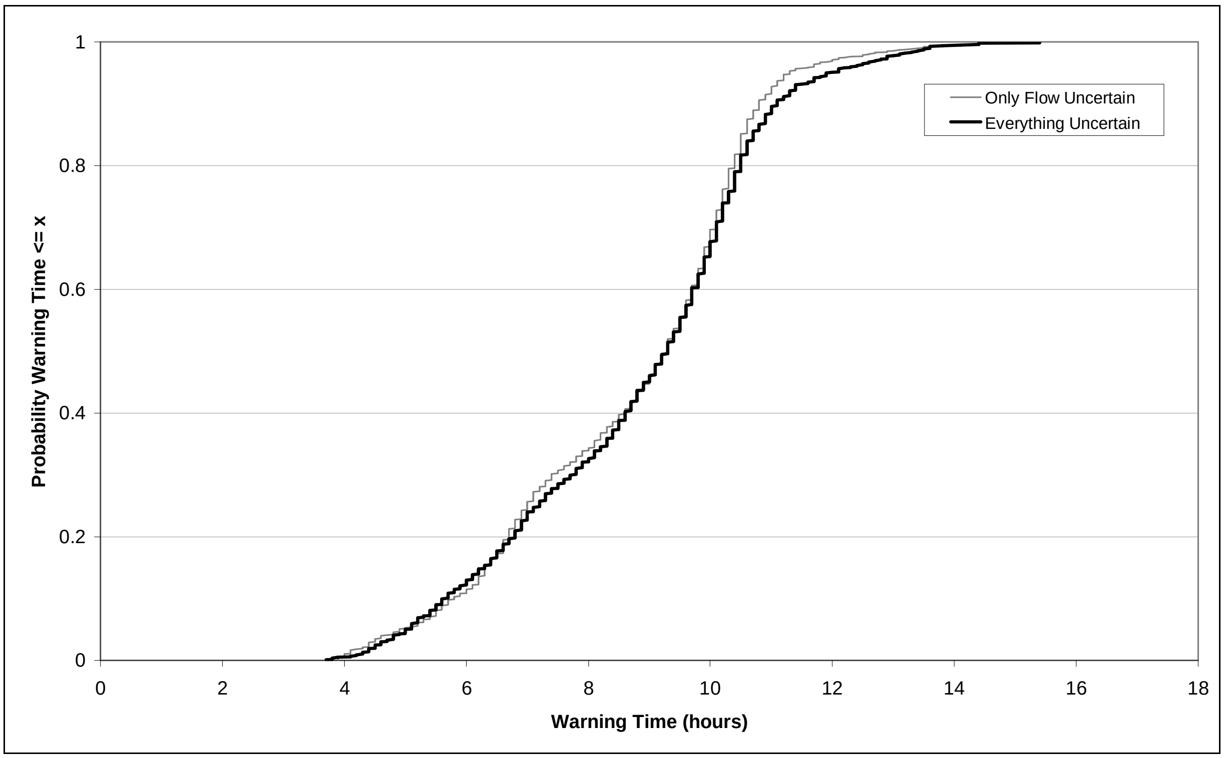

Figure 5-16. CDF of Warning Time at Station A (Sample Interval = 0.1 hours)

As can be seen in Figures 5-16 and 5-17, almost no difference in warning time exists between each set of simulations. Therefore, the only source of uncertainty significantly affecting the probability of having two hours of warning time *for a given spill scenario* is the flow at the time of the spill. Advection controls the time of arrival, which is ultimately controlled by the flow at the time of the spill.

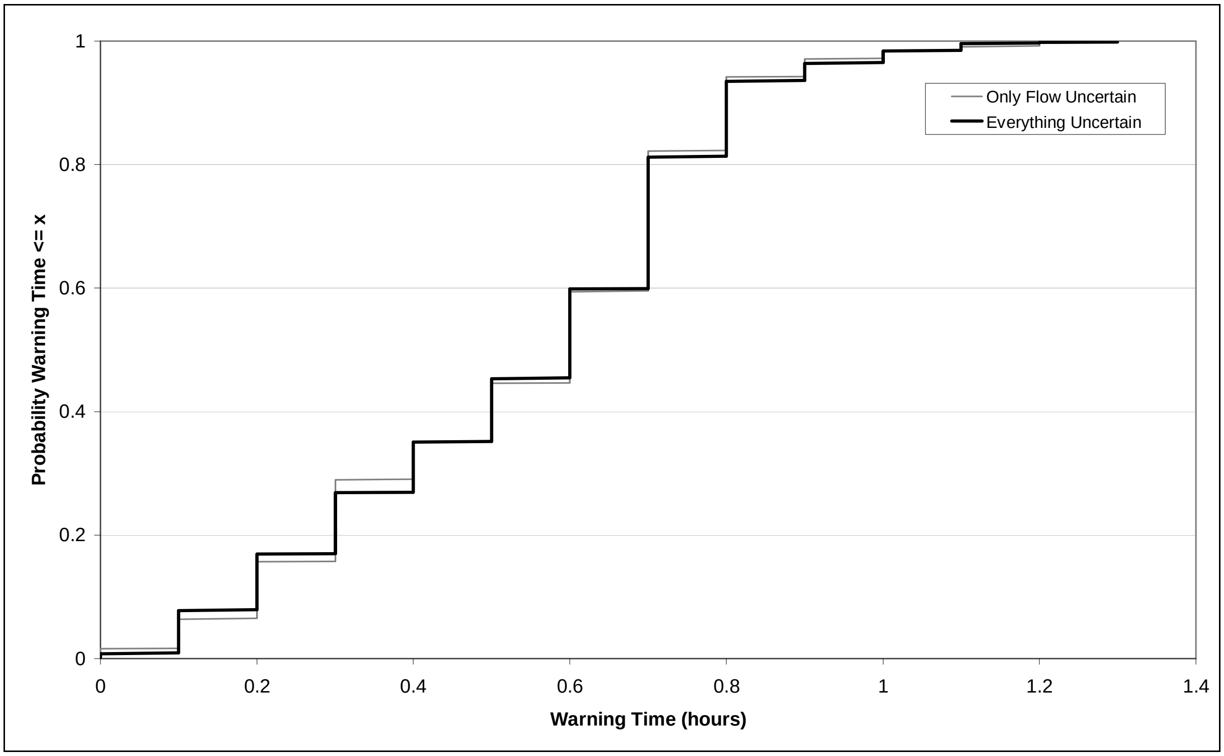

Figure 5-17. CDF of Warning at Station E (Sample Interval = 0.1 hours)

The probability of detection is also a function of arrival time at a given station as well as the duration a contaminant is above the MDL. The detection duration is clearly longer when the spill duration is uncertain (see Figure 5-15); however, this does not significantly impact the results because this is only an issue for cases when the sample interval is greater than the detection duration. Since near-constant sampling has been recommended, the effect of uncertain spill durations on the distribution of the duration contaminants are detected at a given station does not impact the final decision.

In addition to flow, the distribution of spill locations also impacted the distribution of the results because different spill scenario weightings were shown to produce different probabilities of detection and warning time. Therefore, the flow at the time of the spill and the spill location are the most important sources of uncertainty affecting the design of an early warning source water monitoring station.

6 Conclusions and Recommendations

6.1 Conclusions

The value of using a probabilistic approach for designing an early warning source water monitoring station was demonstrated in this research. A new approach for sampling the flow at the time of the spill was developed so that non-constant flows and tributary effects could be simulated. The use of multi-objective optimization techniques was shown to be a useful way in which to assess the probabilistic modelling results.

The impact of various sources of uncertainty was examined in this thesis. Although the spill mass, spill duration, and model parameters impacted the peak and duration of the resulting concentration profiles, the relevant model outputs were only impacted by the uncertainty associated with the river flow at the time of the spill and the location of the spill.

Since model parameter uncertainty did not have an impact on the design of a source water monitoring station, there was no need to devote resources towards a more refined description of model parameter uncertainty as it had no impact on the decision making process. As a result of not refining the parameter ranges through calibration or conditioning experiments, the model in its current state should not be used as a real-time predictive tool.

Results showed that the risk weightings applied to the modelled spill scenarios impact the results, particularly with respect to the probability of detection. Therefore, if the risk weightings are modified the recommended station design may also change, particularly if scenarios further upstream are determined to be of higher risk.

The choice of method detection limit (MDL) also impacts the results. Higher MDLs were shown to produce larger amounts of warning time (assuming that the concentration of interest at the intake is equal to the MDL).

For the case study application, Station A, located furthest upstream from the Hidden Valley Intake, almost always provides at least two hours of warning time at the expense of a low probability of detection. Station E, located closest to the intake, provides excellent threats coverage but almost no chance of having two hours of warning time for any events at any sample interval. Since both objectives were assumed to be of importance, a location that represents a better tradeoff in the objectives is required. Station D, located at the Victoria Street Bridge, has a much higher probability of having two hours of warning time than Station E and a 55% probability of detection at low sample intervals. Providing less than two hours of warning time was assumed to be preferable to decreasing the probability of detection any further. Therefore, it was concluded that a monitoring station near the Victoria Street Bridge represents the best tradeoff solution, assuming the risk weightings applied are accurate.

The selection of sample frequency was found to be dependent upon the location of the monitoring station. At Station A, a sample interval of two hours or more is most appropriate because there is minimal benefit in sampling more often. However, at stations closer to the intake more frequent sampling is required to achieve sufficient warning time. At Station D, near constant sampling is required in order to ensure a high probability of having at least two hours of warning time.

6.2 General Recommendations

The probabilistic method presented in this thesis should be used for future monitoring station designs. The use of multi-objective methods for assessing potential station designs, as presented in this thesis, is recommended in order to assess the tradeoff between warning time and threats coverage.

Future early warning monitoring station designs should focus upon defining accurate probability distributions for flow and spill location as these were the only sources of uncertainty impacting the relevant results for this case study. Since this result may be different for other applications, initial experiments are recommended to confirm the findings of this thesis. For example, if flow is relatively constant, the uncertainty associated with water quality model parameters may become more significant.

Future research could include the probability of incomplete mixing to account for the fact that not all spills located upstream of a station are detected. The use of a two or three dimensional model could also be considered.

6.3 Recommendations Specific to the Case Study

Prior to implementing an early warning source water monitoring station upstream of the Hidden Valley Intake, it is recommended that the following points be confirmed:

- The risk weightings that should be applied to each of the spill scenarios modelled (which should be based upon the results of the risk assessment that will be completed as part of source water protection planning efforts);
- The importance of having at least two hours of warning time;
- The comfort level operations staff have with contaminated water entering the Hidden Valley Reservoir;
- The concentration of interest at the intake;
- The tradeoff between maximizing threats coverage and maximizing the probability of having sufficient warning time;
- The types of monitoring technologies that may be implemented and their respective MDLs; and
- The feasibility of implementing a station where no bridge exists (e.g. Station C or somewhere between Stations D and E).

Once these issues have been resolved, the objective values should be recalculated, if necessary, in order to generate final results upon which to base a decision.

If the assumptions made as part of this thesis are found to be acceptable, it is recommended that the Regional Municipality of Waterloo proceed with constructing an early warning source water monitoring station near the Victoria Street Bridge. It is further recommended that the station should sample at least every hour to increase its likelihood of providing two hours of warning time. The parameters for which the station will monitor and the technology it will use should be determined based on the contaminants identified as part of the risk assessment and any technological or cost constraints.

It is also recommended that the Regional Municipality of Waterloo considers undertaking a study to determine the optimal lateral positioning of the monitoring station. Upon completing the risk assessment, the positioning of various threats may indicate that the station should be located closer to one bank or another, though it will likely be most ideally situated on the river's center line.

Since the recommended monitoring station will fail to detect spills from both King Street and Highway 8, the implementation of a second monitoring station immediately downstream of the Highway 8 Bridge should be considered. Although this station will likely not provide enough warning time to prevent contaminated water from entering the Hidden Valley Reservoir, it can provide notification of events that the conventional intake analyzers may fail to detect.

It is also recommended that the Regional Municipality of Waterloo consider developing a predictive model for use in the event of a spill. This can be used in conjunction with the early warning monitoring station to predict the time of arrival and duration of real-time spill events. The model used for this thesis may be adapted for use as a real-time prediction tool if additional data are collected and calibration or conditioning experiments are conducted.

Early warning source water monitoring stations act as an additional barrier in the production of safe drinking water and are an important part of source water protection. It is recommended that the implementation of such a station should complement traditional source water protection efforts, not replace them. Identifying and mitigating risks to water quality and quantity should remain the focus of source water protection.

7 References

Ambrose, R. B., Wool, T. A. and Martin, J. L. (1993). *The Water Quality Analysis Simulation Program, WASP5, Part A: Model Documentation.* Environmental Research Laboratory. Athens, GA: USEPA.

Andrews, D., Manager, Wastewater Operations, Regional Municipality of Waterloo. (2007). "Typical Effluent and Raw Wastewater Quality at the Waterloo WWTP." Conversation of May 29, 2007.

Beven, K. and Binley, A. (1992). "The future of distributed models - model calibration and uncertainty prediction." *Hydrological Processes*, 6(3), 279-298.

Boano, F., Revelli, R. and Ridolfi, L. (2006). "Stochastic modelling of DO and BOD components in a stream with random inputs." *Advances in Water Resources*, 29(9), 1341-1350.

Brooks, S. (1998). "Markov chain Monte Carlo method and its application." *The Statistician*, 47(1), 69-100.

Chapra, S. C. (1997). *Surface Water-Quality Modeling.* B. J. Clark, D. A. Damstra, and J. W. Bradley. Boston: McGraw-Hill.

Cooke, S. (2006). "Water Quality in the Grand River: A Summary of Current Conditions (2000-2004) and Longer Term Trends." Grand River Conservation Authority, Cambridge, Ontario. Available at: http://www.grandriver.ca/water/2006_WaterQuality_complete.pdf.

Cox, B. A. and Whitehead, P. G. (2005). "Parameter sensitivity and predictive uncertainty in a new water quality model, Q^2." *Journal of Environmental Engineering*, 131(1), 147-157.

Cozzolino, L., Mucherino, C., Pianese, D. and Pirozzi, F. (2006). "Positioning, within water distribution networks, of monitoring stations aiming at an early detection of intentional contamination." *Civil Engineering and Environmental Systems*, 23(3), 161-174.

Dames and Moore. (1994). *User's Manual - Linked Watershed/Waterbody Model.* Tampa, FL: Southwest Florida Water Management District.

Davies, J.-M. and Mazumder, A. (2003). "Health and environmental policy issues in Canada: the role of watershed management in sustaining clean drinking water quality at surface sources." *Journal of Environmental Management*, 68(3), 273-286.

Deng, Z.-Q., Bengtsson, L., Singh, V. P. and Adrian, D. D. (2002). "Longitudinal dispersion coefficient in single-channel streams." *Journal of Hydraulic Engineering*, 128(10), 901-916.

Dixon, W. and Chiswell, B. (1996). "Review of aquatic monitoring program design." *Water Research*, 30(9), 1935-1948.

Environment Canada. (2004). "Threats to Canada's Water Supply." Available at: http://www.nwri.ca/envirozine/issue42-e.html.

Environment Canada. (2005). "National Spill Statistics and Trends, 1984-1995." Available at: http://www.ec.gc.ca/ee-ue/default.asp?lang=en&n=95DE537D.

Fenicia, F., Savenije, H. H. G., Matgen, P. and Pfister, L. (2007). "A comparison of alternative multiobjective calibration strategies for hydrological modeling." *Water Resources Research*, 43(W03434).

Fischer, H. B. (1975). "Discussion of 'Simple method for predicting dispersion in streams' by McQuifer R.S. and Keefer T.N." *Journal of Environmental Engineering Division, American Society of Civil Engineers*, 101(EE3), 453-455.

Gostin, L. O. (2000). "Water quality laws and waterborne diseases: *Cryptosporidium* and other emerging pathogens." *American Journal of Public Health*, 90(6), 847-853.

Grand River Conservation Authority. (2006a). "The Grand River Watershed." Cambridge, Ontario. Available at: http://www.sourcewater.ca/index/document.cfm?Sec=7&Sub1=0&sub2=0.

Grand River Conservation Authority. (2006b). "Hidden Valley Weir Provisional Rating Curve."

Grand River Conservation Authority. (2006c). "Hourly Flow Data for Grand River, Canagagigue Creek, Conestogo River, and Laurel Creek".

Grand River Conservation Authority. (2006d). "Surveyed Cross Sections for the Grand River".

Grayman, W. M., Deininger, R. A. and Males, R. M. (2001). *Design of Early Warning and Predictive Source-Water Monitoring Systems.* Denver: AWWA Research Foundation and American Water Works Association.

Grayman, W. M. and Males, R. M. (2002). "Risk-based modeling of early warning systems for pollution accidents." *Water Science & Technology*, 46(3), 41-49.

Gullick, R. W., Grayman, W. M., Deininger, R. A. and Males, R. M. (2003). "Design of early warning monitoring systems for source waters." *Journal AWWA*, 95(11), 58-72.

Harmancioglu, N. B. and Alpaslan, N. (1992). "Water quality monitoring network design: A problem of multi-objective decision making." *Water Resources Bulletin (American Water Resources Association)*, 28(1), 179-192.

Harmancioglu, N. B. and Alpaslan, N. (1994). "Basic approaches in design of water quality monitoring networks." *Water Science & Technology*, 30(10), 49-56.

Hibbs, D. E. and Gulliver, J. S. (1999). "Processes controlling aqueous concentrations for riverine spills." *Journal of Hazardous Materials*, 64(1), 57-73.

International Life Sciences Institute. (1999). "Early warning monitoring to detect hazardous events in water supplies." ILSI, Washington. Available at: http://www.ilsi.org/file/EWM.pdf.

Iwasa, Y. and Aya, S. (1991). "Predicting longitudinal dispersion coefficient in open channel flows." *Proceedings, International Symposium on Environmental Hydraulics*, Balkema, Rotterdam, 505-511.

Jobson, H. E. (1996). "Prediction of Traveltime and Longitudinal Dispersion in Rivers and Streams." USGS, Available at: http://water.usgs.gov/osw/pubs/disp/dispersion.pdf.

Jobson, H. E. (1997). "Predicting travel time and dispersion in rivers and streams." *Journal of Hydraulic Engineering*, 123(11), 971-978.

Jobson, H. E. and Schoellhamer, D. H. (1987). *User's Manual for a Branched Lagrangian Transport Model.* Reston, Virginia: USGS.

Kashefipour, S. M. and Falconer, R. A. (2002). "Longitudinal dispersion coefficients in natural channels." *Water Research*, 36(6), 1596-1608.

Kilpatrick, F. A. (1993). "Techniques of Water-Resources Investigations of the United States Geological Survey." Simulation of Soluble Waste Transport and Buildup in Surface Waters Using Tracers. USGS, Denver, CO.

Klepper, O. and Hendrix, E. M. T. (1994). "A method for robust calibration of models under different types of uncertainty." *Ecological Modelling*, 74(1994), 161-182.

Lindenschmidt, K.-E. (2006). "The effect of complexity on parameter sensitivity and model uncertainty in river water quality modeling." *Ecological Modeling*, 190(1), 72-86.

Lindenschmidt, K.-E., Rauberg, J. and Hesser, F. B. (2005). "Extending uncertainty analysis of a hydrodynamic-water quality modelling system using high level architecture (HLA)." *Water Quality Research*, 40(1), 59-70.

Liu, H. (1977). "Predicting dispersion coefficient of streams." *Journal of the Environmental Engineering Division, American Society of Civil Engineers*, 103(EE1), 59-69.

MacKenzie, M. C., Palmer, R. N. and Millard, S. P. (1987). "Analysis of statistical monitoring network design." *Journal of Water Resources Planning and Management*, 113(5), 599-615.

Martin, J. L. and McCutcheon, S. C. (1999). *Hydrodynamics and Transport for Water Quality Modeling.* Florida: CRC Press.

Martin, J. L. and Wool, T. (2002). *A Dynamic One-Dimensional Model for Hydrodynamics and Water Quality - EPDRiv1 Version 1.0 Users Manual.* Athens, Georgia: AScI Corporation.

McCuen, R., H. (1998). *Hydrologic Analysis and Design.* Upper Saddle River, New Jersey: Prentice Hall.

McIntyre, N. R., Wagener, T., Wheater, H. S. and Chapra, S. C. (2003a). "Risk-based modelling of surface water quality: a case study of the Charles River, Massachusetts." *Journal of Hydrology*, 274(1-4), 225-247.

McIntyre, N. R., Wagener, T., Wheater, H. S. and Zeng, S. Y. (2003b). "Uncertainty and risk in water quality modelling and management." *Journal of Hydroinformatics*, 5(2003), 259-274.

McIntyre, N. R. and Wheater, H. S. (2004). "A tool for risk-based management of surface water quality." *Environmental Modelling & Software*, 19(12), 1131-1140.

Mikol, Y. B., Richardson, W. R., Van Der Schalie, W. H., Shedd, T. R. and Widder, M. W. (2007). "An online real-time biomonitor for contaminant surveillance in water supplies." *Journal AWWA*, 99(2), 107-115.

Morgan, M. G. and Henrion, M. (1990). *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis* Cambridge: Cambridge University Press.

Muleta, M. K. and Nicklow, J. W. (2005). "Decision support of watershed management using evolutionary algorithms." *Journal of Water Resources Planning and Management*, 131(1), 35-44.

National Utility Service. (2006). "2005-2006 International Water Report & Cost Survey." Available at: http://www.nusinc.com/downloads/2006WaterSurvey.pdf.

Ning, S.-K. and Chang, N.-B. (2004). "Optimal expansion of water quality monitoring network by fuzzy optimization approach." *Environmental Monitoring and Assessment*, 91(1-3), 145-170.

Ning, S.-K. and Chang, N.-B. (2005). "Screening the relocation strategies of water quality monitoring stations by compromise programming." *Journal of the American Water Resources Association*, 41(5), 1039-1052.

Ontario Ministry of the Environment. (1990). "Environmental Protection Act". Regulation 360.

Ontario Ministry of the Environment. (2006a). "Assessment Report: Guidance Modules (DRAFT)".

Ontario Ministry of the Environment. (2006b). "Clean Water Act". Bill 43.

Ontario Ministry of the Environment. (2006c). "Legislature Passes Clean Water Act: Protecting Drinking Water Sources Will Benefit All Ontarians ". Available at: http://www.ene.gov.on.ca/envision/news/2006/101801.htm.

Ontario Ministry of the Environment. (2006d). "Provincial Water Quality Monitoring Network Data for the Grand River Watershed."

Ontario Spills Action Centre. (2006a). "Database of Spills Reported in the Grand River Watershed, 2003-2005."

Ontario Spills Action Centre. (2006b). "Spills Action Centre: 2006 Summary Report." Toronto, Ontario. Available at: http://www.ene.gov.on.ca/programs/6240e.php.

Organisation for Economic Co-operation and Development. (2007). "OECD Factbook 2007: Economic, Environmental and Social Statistics - Water Consumption." Available at: http://lysander.sourceoecd.org/vl=2758430/cl=51/nw=1/rpsv/factbook/08-01-01.htm.

Osidele, O. O., Zeng, W. and Beck, M. B. (2006). "A random search methodology for examining parametric uncertainty in water quality models." *Water Science & Technology*, 53(1), 33-40.

Ostfeld, A. and Salomons, E. (2004). "Optimal layout of early warning detection stations for water distribution systems security." *Journal of Water Resources Planning and Management*, 130(5), 377-385.

Ostfeld, A. and Salomons, E. (2005). "Securing water distribution systems using online contamination monitoring." *Journal of Water Resources Planning and Management*, 131(5), 402-405.

Palmer, R. N. and MacKenzie, M. C. (1985). "Optimization of water quality monitoring networks." *Journal of Water Resources Planning and Management*, 114(3), 478-493.

Perera, B. J. C. and Ng, A. W. M. (2001). "River Water Quality Modelling - Parameter Uncertainty, Sensitivity and Estimation." Water Pollution. C. A. Brebbia, ed. Wessex Institute of Technology, United Kingdom.

Portielje, R., Hvitved-Jacobsen, T. and Schaarup-Jensen, K. (2000). "Risk analysis using stochastic reliability methods applied to two cases of deterministic water quality models." *Water Research*, 34(1), 153-170.

Reckhow, K. H. (1994). "Water quality simulation modeling and uncertainty analysis for risk assessment and decision making." *Ecological Modeling*, 72(1004), 1-20.

Reddy, M. J. and Kumar, D. N. (2007). "Multiobjective differential evolution with application to reservoir system optimization." *Journal of Computing In Civil Engineering*, 21(2), 136-146.

Regional Municipality of Waterloo. (2006). "Mannheim Water Treatment Plant and Waterloo Wastewater Treatment Plant Flow Data".

Reinelt, L. E., Horner, R. R. and Mar, B. W. (1988). "Nonpoint source pollution monitoring program design." *Journal of Water Resources Planning and Management*, 114(3), 335-352.

Rutherford, J. C. (1994). *River Mixing.* West Sussex: John Wiley and Sons Ltd.

Sanders, T. G., Ward, R. C., Loftis, J. C., Steele, T. D., Adrian, D. D. and Yevjevich, V. (1983). *Design of Networks for Monitoring Water Quality.* Chelsea, Michigan: Water Resources Publications.

Schaffranek, R. W. (1987). *Flow Model for Open-Channel Reach or Network.* P. P. 1384. Reston, VA: U.S. Geological Survey.

Seo, I. W. and Baek, K. O. (2005). "Discussion of "Estimation of the longitudinal dispersion coefficient using the velocity profile in natural streams" " *Journal of Hydraulic Engineering*, 131(10), 927-928.

Seo, I. W. and Cheong, T. S. (1998). "Predicting longitudinal dispersion coefficient in natural streams." *Journal of Hydraulic Engineering, American Society of Civil Engineers*, 124(1), 25-32.

Singh, A., Minsker, B. and Goldberg, D. E. (2003). "Combining reliability and Pareto optimality - an approach using stochastic multi-objective genetic algorithms." *World Water Congress 2003*, Philadelphia, Pennsylvania.

Singh, S. K. and Beck, M. B. (2003). "Dispersion coefficient of streams from tracer experiment data." *Journal of Environmental Engineering*, 129(6), 539-546.

Spear, R. C. and Hornberger, G. M. (1980). "Eutrophication in Peel Inlet – II. Identification of critical uncertainties via generalized sensitivity analysis." *Water Research*, 14(1), 43-49.

Stantec Consulting Ltd. (2006). *Intake Protection Zones Delineation Study - Grand River Hidden Valley Intake, City of Kitchener.* Kitchener, Ontario: Stantec Consulting Ltd.

Technical Experts Committee. (2004). "Watershed-Based Source Protection Planning: Science-based Decision-making for Protecting Ontario's Drinking Water Resources".

U.S. Army Corps of Engineers-WES. (1990). *CE-QUAL-RIV1: A Dynamic, One-Dimensional (Longitudinal) Water Quality Model for Streams: User's Manual, Instruction Report.* Vicksburg, MS: U.S. Army Corps of Engineers Waterway Experiment Station.

United States Environmental Protection Agency. (2007). "An Introduction to Water Quality Monitoring." Available at: http://www.epa.gov/owow/monitoring/monintr.html.

Waldon, M. G., O'Hara, K. K. and Everett, D. E. (1989). "Detection Probabilities and System Analysis of the Mississippi River early Warning Organic Compound Detection System." Baton Rouge, Louisiana: Louisiana Department of Environmental Quality.

Wallis, S. G. and Manson, J. R. (2004). "Methods for predicting dispersion coefficients in rivers." *Proceedings of the Institution of Civil Engineers, Water Management*, 157(WM3), 131-141.

Walton, T., Water Quality Specialist, Regional Municipality of Waterloo. (2006). "Information about Mannheim WTP." Conversation of June 30, 2006.

Yapa, P. D. and Shen, H. T. (1994). "Modelling River Oil Spills - A Review." *Journal of Hydraulic Research*, 32(5), 765-782.

Zheng, Y. and Keller, A. A. (2006). "Understanding parameter sensitivity and its management implications in watershed-scale water quality modeling." *Water Resources Research*, 42(5), W0542.

Appendix A: Water Quality Model Parameter Ranges

Note: Parameter ranges listed below are based on (Perera and Ng, 2001; Martin and Wool, 2002; McIntyre et al., 2003a; Cox and Whitehead, 2005; Lindenschmidt, 2006; Osidele et al., 2006).

Parameter	Description	Units	Min	Max
TH_K1	Theta for CBODu1 Decay	--	1.024	1.15
TH_KDN	Theta for Denitrification for CBODu1	--	1.024	1.15
TH_K1N	Theta for Organic Nitrogen to NH_3	--	1.024	1.15
TH_KNH3	Theta for Ammonia to NO_3 Transformation	--	1.024	1.15
TH_KDNO2	Theta for Sediment Denitrification	--	1.024	1.15
TH_BENP	Theta for Benthic Ortho Phosphate Release Rate	--	1.024	1.15
TH_BENN	Theta for Benthic Ammonia Release Rate	--	1.024	1.15
TH_SOD	Theta for Sediment Oxygen Demand	--	1.024	1.15
TH_ARB1	Theta for Arbitrary 1 decay	--	0	1
TH_AGRO	Theta for Phytoplankton/Algae Growth	--	1.024	1.15
TH_ADIE	Theta for Phytoplankton/Algae Death	--	1.024	1.15
TH_MGRO	Theta for Macrophyte Growth	--	1.024	1.15
TH_MDIE	Theta for Macrophyte Death	--	1.024	1.15
TH_SORP	Theta for Ortho Phosphate Loss/Adsorption	--	1.024	1.15
APCONT	Phytoplankton Phosphorus Content	mg P/mg B	0.009	0.011
ANCONT	Phytoplankton Nitrogen Content	mg N/mg B	0.07	0.1
MPCONT	Macrophyte Phosphorus Content	mg P/mg B	0	0
MNCONT	Macrophyte Nitrogen Content	mg N/mgB	0.02	0.4
ONEQUI	Oxygen/Nitrogen Ratio for Denitrification	mg O_2/mg N	0.34	0.36
ONITRI	Oxygen/Nitrogen Ratio for Nitrification	mg O_2/mg N	4.5	4.64
OPDECY	Oxygen Consumption by Plant Decay	mg O_2/mg B	1	2
ADN	CBODu1 Denitrification Rate	1/day at 20C	0	1
AKN	Ammonia to NO_3 Transform Rate	1/day at 20C	0.025	2
ATB	Bottom Heat Exchange Rate	1/day at 20C	0	0
ATS	DO Concentration at which Algal death is half max rate	Watts/m^2/C	0	0
APO4	Ortho Phosphate Loss Rate	1/day at 20C	0.62	0.76
KALGDK	Phytoplankton Decay Rate	1/day at 20C	0.003	2
KNCBDN	NO_3 Concentration at which Denitrification is 1/2 Rate	mg/L	0	1
KOCB1	DO Concentration at which CBODu1 Decay is 1/2 Max Rate	mg/L	0	1

Parameter	Description	Units	Min	Max
KNPOOL	NH_3 and NO_3 Concentration at which Algal growth rate is 1/2 max	mg/L	0.01	0.3
KP04X	Total phosphorus Concentration at which Algal growth rate is 1/2 max	mg/L	0	0.1
KDNO2	Sediment Denitrification Rate	1/day at 20C	0	0.1
ACK	Organic Nitrogen decay to NH_3 Transform Rate	1/day at 20C	0.02	0.4
LAMBDA0	Non-algal Light Extinction Coefficient	1/m	0.1	0.3
LAMBDA1	Linear Algal Self Shading Coefficient	(1/m)/ (ug Chl-a/L)	0	0.01
LAMBDA2	Non-linear Algal Self Shading Coefficient	(1/m)/ (ug Chl-a/L)$^{(2/3)}$	0.04	0.06
ALPHAO	Algae to Chlorophyll Conversion Factor	ug Chl-a/mg B	10	100
XONS	Organic Nitrogen Settling Rate	m/day	0.001	0.1
CBODSR	CBODu1 Settling Rate	m/day	-0.36	0.5
FCBOD	Fraction of Algal/Macrophyte Death which goes to CBODu1	fraction	0	1
KPDK	Organic Phosphorus to Ortho Phosphate Transform Rate	1/day at 20C	0.01	0.7
KPSET	Organic Phosphorus Settling Rate	m/day	0.001	0.1
AKARB1	Arbitrary Contaminant Decay Rate	1/day at 20C	0	1
SOD	Sediment Oxygen Demand		0	4.4
AK1	CBODu1 Decay Rate	1/day at 20C	0.004	4

Appendix B: Complete Results with an MDL of 0.01 mg/L

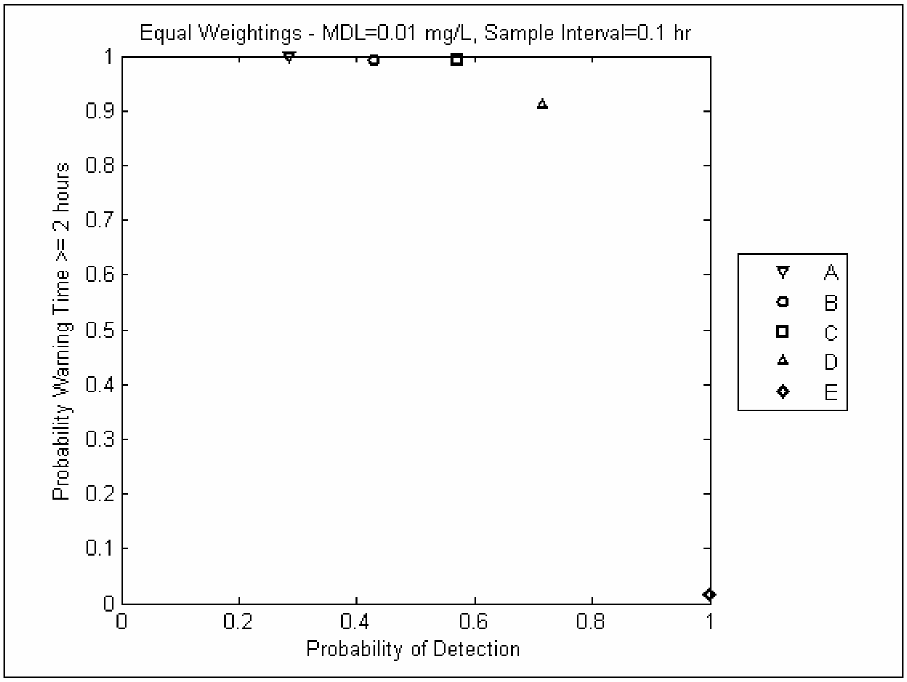

Figure B - 1. Results Using Weighting Scheme 1 and Sample Interval of 0.1 hours

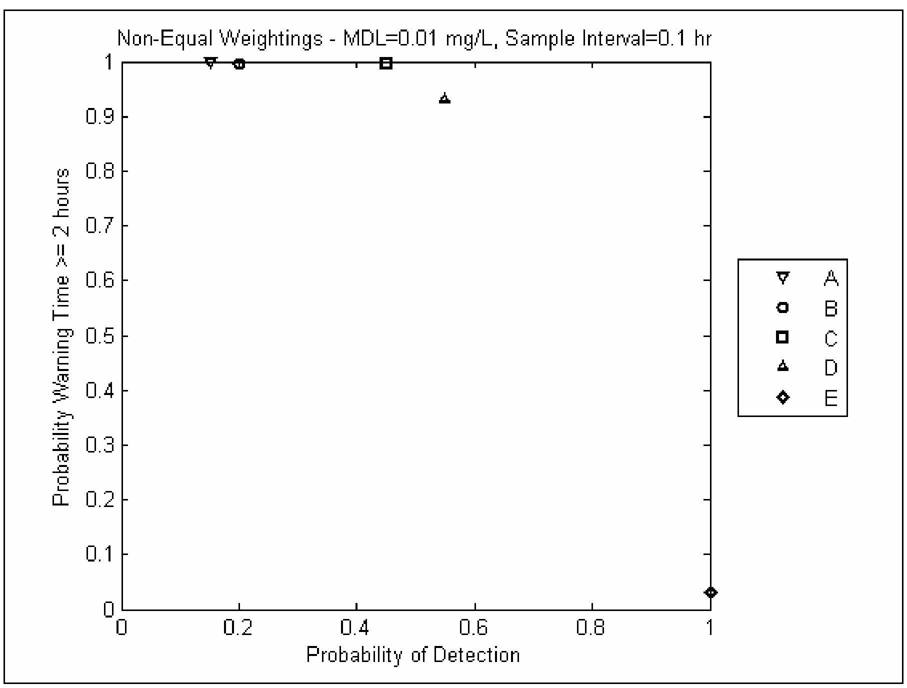

Figure B - 2. Results Using Weighting Scheme 2 and Sample Interval of 0.1 hours

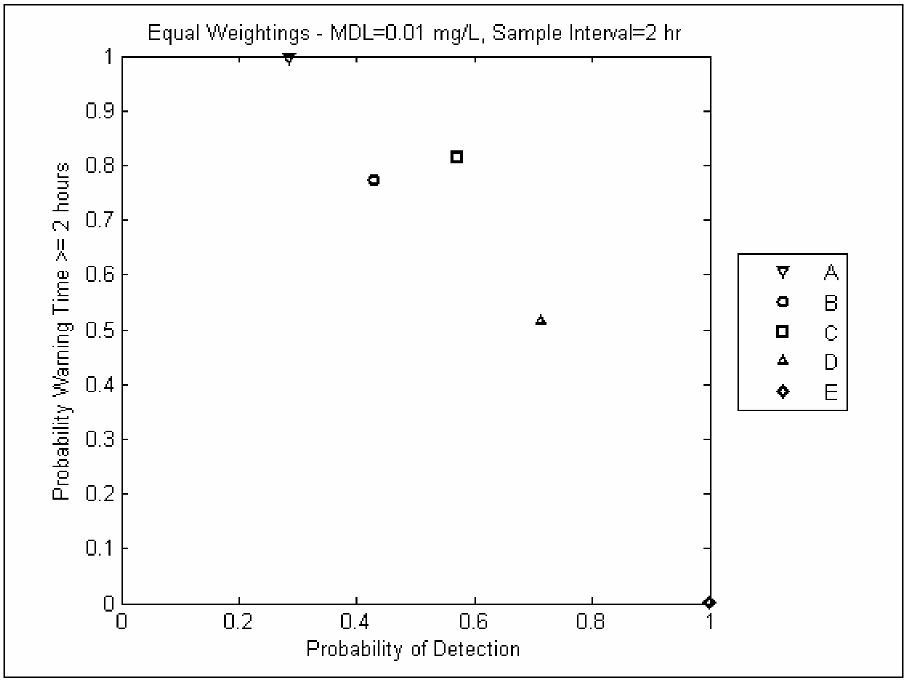

Figure B - 3. Results Using Weighting Scheme 1 and Sample Interval of 2 hours

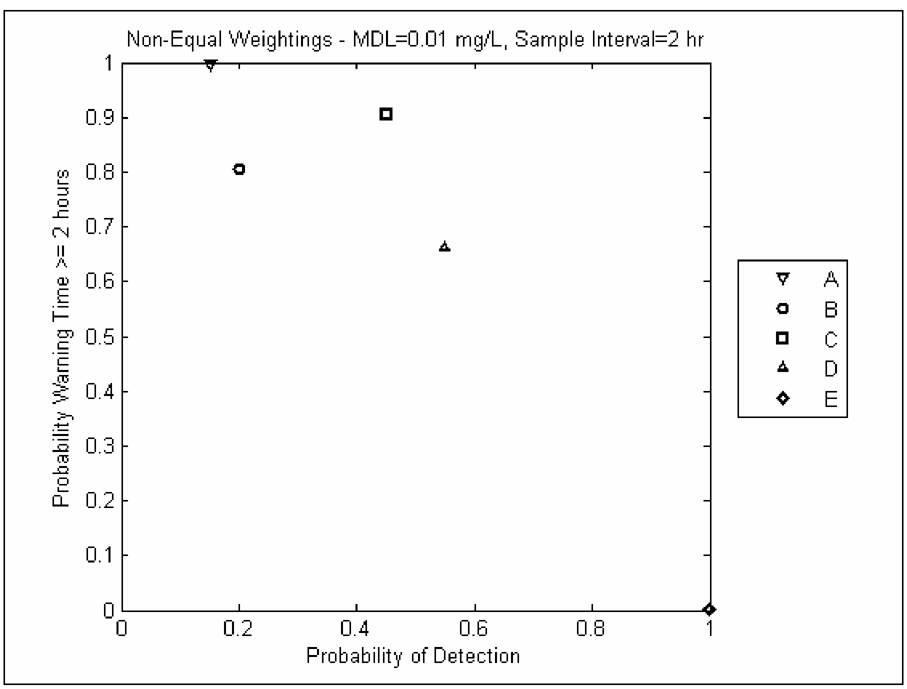

Figure B - 4. Results Using Weighting Scheme 2 and Sample Interval of 2 hours

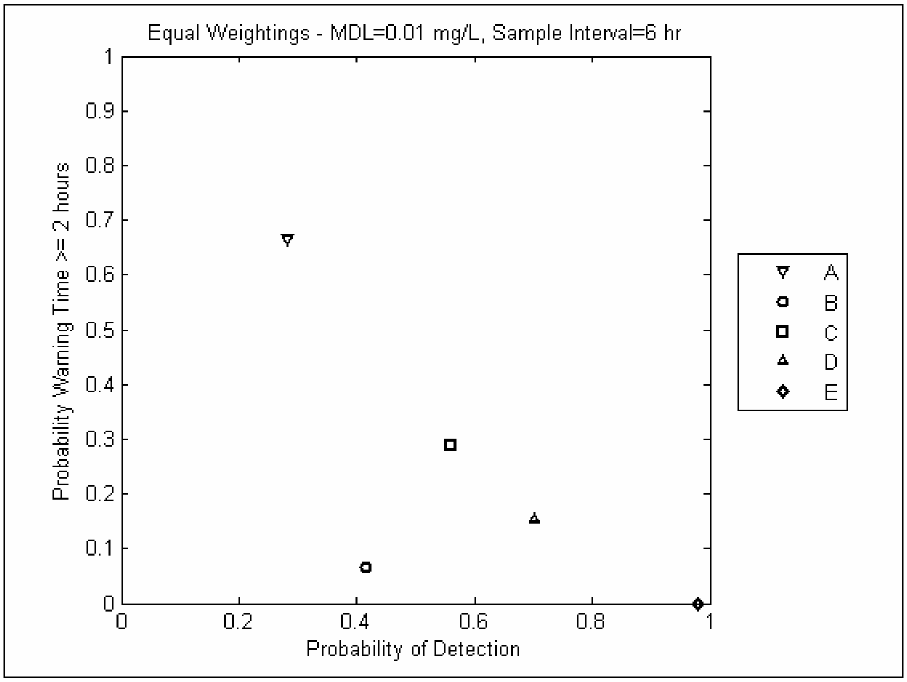

Figure B - 5. Results Using Weighting Scheme 1 and Sample Interval of 6 hours

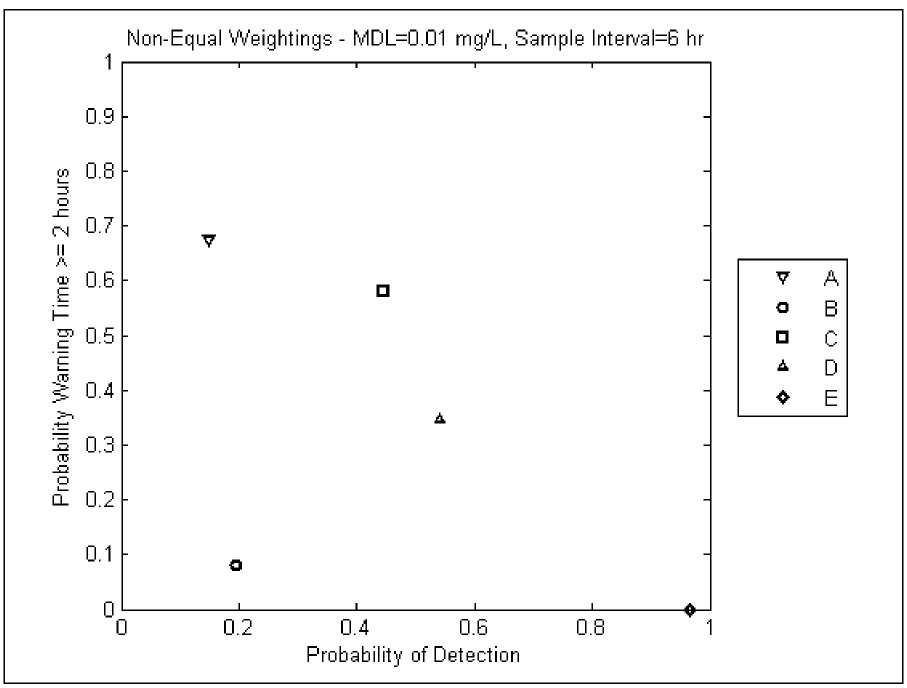

Figure B - 6. Results Using Weighting Scheme 2 and Sample Interval of 6 hours

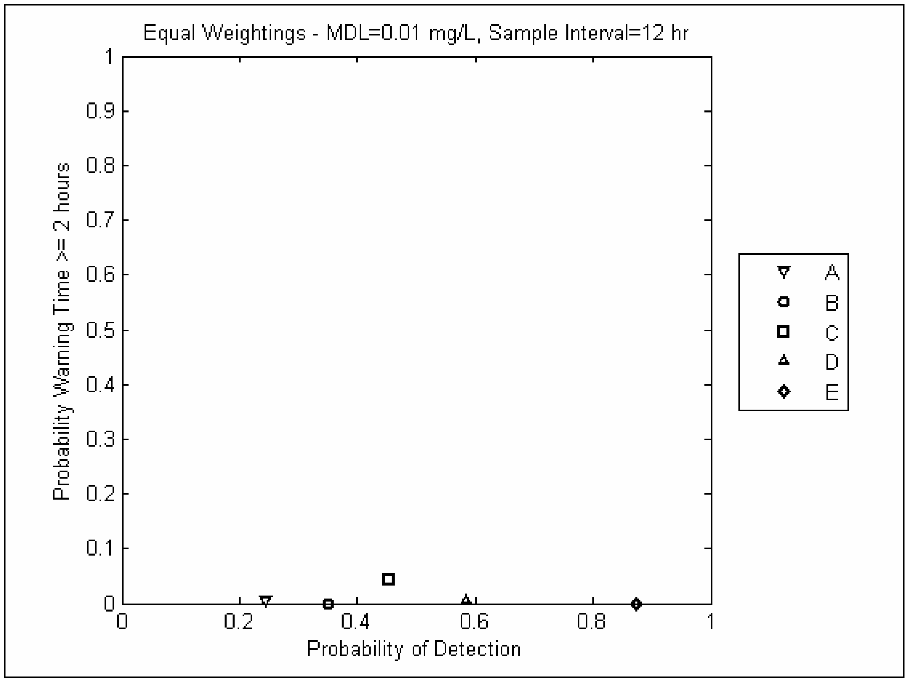

Figure B - 7. Results Using Weighting Scheme 1 and Sample Interval of 12 hours

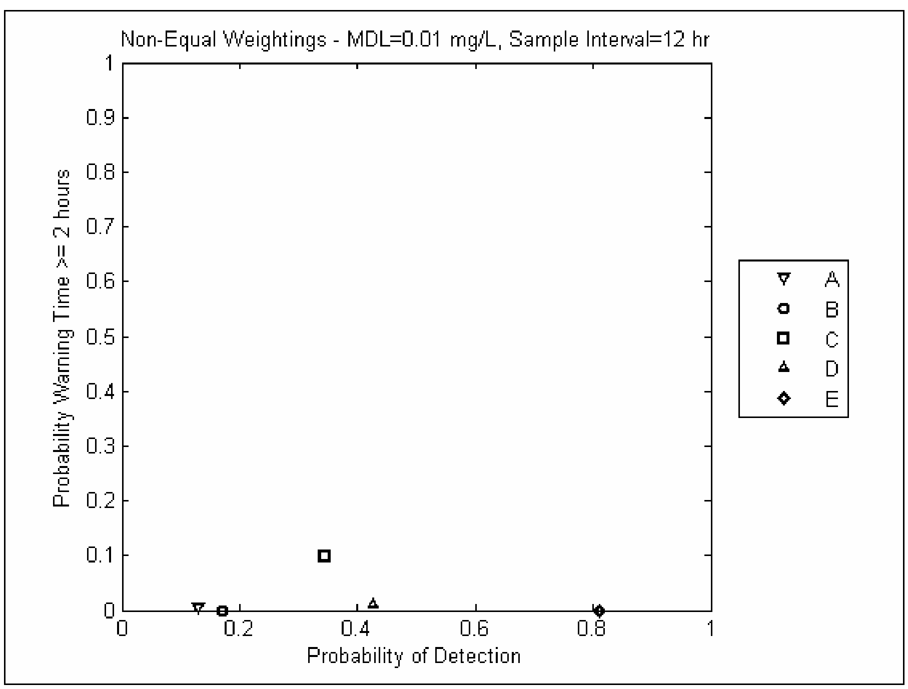

Figure B - 8. Results Using Weighting Scheme 2 and Sample Interval of 12 hours