A Possibilistic Analysis Approach for Assessing Environmental Risks from Drinking Groundwater at Petroleum-Contaminated Sites

L. Liu1*, R. X. Hao2 and S. Y. Cheng3

1Department of Civil Engineering, Dalhousie University, Halifax, NS B3J 1Z1, Canada
2College of Architecture and Civil Engineering, Beijing University of Technology, Beijing 100022, China
3College of Environmental and Energy Engineering, Beijing University of Technology, Beijing 100022, China

ABSTRACT. Subsurface contamination at petroleum-contaminated sites is posing a serious threat to the environment and the public and is acquiring more and more attention by governments and industries. This paper proposed a fuzzy risk assessment method and its application to a petroleum-contaminated site. The method consists of three parts: (i) calculation of fuzzy steady-state contaminant concentrations in the aquifer based on an analytical solute transport model, (ii) possibilistic analysis of fuzzy criteria for different risk levels, and (iii) environmental risk assessment based on the Euclidean method. The method can effectively reflect fuzzy natures of environmental quality at a site and evaluation criteria for different risk levels. Results of an illustrative case study indicate that environmental risks at a petroleum-contaminated site can be effectively evaluated through the developed methodology. The assessment results are useful for the related site remediation and management decision.

Keywords: Environmental risk assessment, fuzzy number, fuzzy set, petroleum contaminant, possibility, probability, uncertainty

1. Introduction

Development of the petroleum industry is currently associated with a number of environmental concerns. Among them, soil and groundwater contamination is acquiring more and more attention by the public, governments, and petroleum industries themselves (Cheremisinoff, 1991). This situation is especially true in western Canada where petroleum production, processing, upgrading, and utilization are active. Generally, the major sources of soil and groundwater contamination are leaking storage tanks used by commercial, industrial and residential sectors. Similar contamination can also occur during landfarming, petroleum sludge disposal, heavy oil upgrading, and others (Livingston & Islam, 1999).

Soil and groundwater contamination has lead to a variety of impacts, risks, and liabilities to the communities and for the industries themselves. For example, it is proved that one gallon of gasoline can render one million gallons of water unsuitable for drinking. In Canada, about 10% of the 200,000 underground storage tanks are leaking and contaminating the surrounding environment, causing losses of thousands of dollars yearly to petroleum industries. In Alberta, there are over 130,000 abandoned drilling sumps that are associated with a number of subsurface contamination problems. It is estimated that rendering all these abandoned drilling sumps will require a minimum expenditure of 10 billion dollars. Therefore, in-depth and effective environmental risk assessment of groundwater contamination due to leaking petroleum contaminants is important and desired for evaluating the need for site remediation actions and providing support for decisions related to prevention, detection, and correction of the leakage and contamination problems (Huang et al., 1999).

In recent years, risk assessment techniques have become widely used as aids in the decision-making process related to contaminated soil and groundwater. Risk assessment could give managers a more rational basis on which to make decisions (Guyonnet et al., 1999). The general formulation of the environmental risk problem captures the entire process of identifying the source term of the risk agent, its fate and transport through porous media, estimation of human exposure, and conversion of such exposure into the risk level. This process involves a number of chemical, physical, biological, geological, and socioeconomic factors due to their direct or indirect relations to the environmental impacts/risks. The related parameters generally show high degrees of intrinsic variability and substantial levels of uncertainty since many system components in real-world problems may not be known with certainty (Robin et al., 1991; Woodbury & Dudicky, 1991). This makes the study systems more complicated and harder to quantify. Thus, effective reflection of uncertainties, which is essential for generating reliable and realistic outcomes, has been a major concern for risk assessment (Lein, 1992).

To deal with uncertainties within complicated environmental systems, previously, there have been a few studies of risk assessment for petroleum waste management through employing probability theory (e.g. Monte Carlo simulation). For example, Lo proposed an oil spill risk simulation model based on a probabilistic approach (Lo, 1991). Hallenbeck &
Flowers undertook a study of risk assessment for worker exposure to benzene (Hallenbeck & Flowers, 1992). Rubin et al. used numerical simulations and presented the results in the form of a concentration cumulative distribution function as a possible probabilistic form to be used in risk assessment studies (Rubin et al., 1994). However, problems with data availability and solution algorithm may exist and affect their practical applicability (Lein, 1992; Bardossy et al., 1991).

Another major approach for identifying uncertainties in risk assessment process is through fuzzy set theory in situations when probabilistic information is unavailable (uncertainties present as fuzzy membership functions rather than probability distribution functions) (Bardossy et al., 1991). There have been few studies of fuzzy risk assessment for petroleum waste management. In comparison, many applications to other areas have been reported. For example, Ganoulis et al. (1995) proposed a fuzzy arithmetic for ecological risk management under uncertainty. Donald & Ross (1996) used fuzzy logic and similarity measures for risk management of hazardous wastes. Dahab et al. (1994) developed a rule-based fuzzy set approach for risk analysis of nitrate-contaminated groundwater. Dou et al. (1995) used fuzzy numbers to represent the uncertainty in simulations of groundwater flow. Bardossy et al. (1995) simulated unsaturated flow using a fuzzy approach. Vollmer et al. (1997) suggested representing cleanup guidelines with the help of fuzzy number.

A possibilistic approach for assessing risks associated with petroleum-contaminated sites is provided in this study, as an extension of the previous studies. The approach will then be applied to an illustrative case study for demonstrating its effectiveness in reflecting uncertainties and its applicability and implication for providing decision support for effective site remediation and management.

2. Methods for Uncertainty and Risk Analysis

Uncertainties inherently exist in many environmental processes due to sparse and imprecise natures of the available information. Probabilistic and possibilistic methods are major approaches for dealing with them. Generally, most of the previous risk analysts argued that risk should be measured through considering the probability of a damage that may occur following exposure of a target to contaminants. These studies generally dealt with uncertainties through the Monte Carlo method (Bobba et al., 1995; Haas et al., 1996; Labieniec et al., 1997). Thus, risk was expressed as a probability distribution over a number of adverse consequences. However, in practical applications, probability theory often assumes that there exists a historical run for the observation of events. The assumed probability distributions will then be assigned to the related parameters. In fact, in a risk assessment context, analysts often suffer from lack of data or imperfect knowledge about the processes. Therefore, it is always questionable whether statistically meaningful probabilities can be derived. This may frustrate rigorous probabilistic studies. Another problem with the probability theory is its law of excluded middle and contradiction. Intuitively, we know this is not true in many problems (Lai & Hwang, 1992).

From risk assessment and decision-making standpoints, if information collected during field investigations reveals the existence of probability distribution functions, the probability theory should certainly be the preferred approach. But this is unlikely to be the case in many field situations, where parameter distributions are often selected based on the best estimates and/or literature. Another important question is whether the probabilistic approach is always appropriate, whether true probabilities can be assigned to the parameters, or whether these values should simply be considered as possible. The fact is that, when using the probabilistic approach, scenarios that combine low probability parameter values have very little chance of being randomly selected. However, in an environmental risk assessment context where human health is often at stake, the mere possibility that a scenario might occur can be an important element in the decision-making process. For example, if probabilities are arbitrarily assumed, some possible combinations of parameter values will disappear from the analysis as a result of Monte Carlo averaging. This might have detrimental consequences in terms of health impacts because unlikely parameter combinations may generate the most damage.

In recent years, application of fuzzy set approach to environmental studies has received increasing attention, as environmental hazards are often perceived by the public in terms of possibilities rather than probabilities. Also, it is typically more difficult for planners and engineers to specify probability distribution than to define membership functions. For risk assessment, the fuzzy set approach has advantages in its effectiveness in reflecting uncertainties and its applicability to practical problems, especially in situations when probabilistic information is unavailable. Another advantage lies in that the fuzzy set approach can appropriately handle the previous dilemma faced by probabilistic approach, through considering all possible combinations of uncertain parameters. Thus, extension of this method to the petroleum waste management area would be desired for generating effective evaluations and decisions.

3. Methodology

3.1. Possibilistic Analysis

(1) Fuzzy numbers

The concept of fuzzy set theory was introduced by Zadeh (1965). It is a generalization of ordinary set theory for solving real world problems which are often obscure or indistinct. Fuzzy numbers are used for representing the uncertainties. A fuzzy number X can be defined as a set of ordered pairs:

\[ X = \{(x, \mu_X(x)) : x \in R; \mu_X(x) \in [0,1]\} \quad (1) \]

where \( x \) is a particular value of \( X \), and \( \mu_X(x) \) represents a membership function of \( X \). Values of a membership function
are located in a closed interval \([0, 1]\). The closer \(\mu_X(x)\) is to 1, the more “certain” one is about the value of \(x\).

Figure 1 presents two types of fuzzy membership functions (trapezoidal and triangular) for illustrating uncertainties associated with an aquifer parameter \(X\). The triangular membership function means that (i) it is most likely that \(X\) is equal to \(c\), and (ii) values lower than \(a\) or greater than \(b\) are considered impossible for \(X\); the trapezoidal one represents that (i) it is most likely that \(X\) lies between \(c_1\) and \(c_2\) without preference in this range, and (ii) \(X\) will never be lower than \(a\) or greater than \(b\). A membership function is normally defined based on characteristics of the uncertain information.

(2) Arithmetic operations on fuzzy numbers

Fuzzy arithmetic has the capability of performing point-to-point operations on fuzzy sets. It defines the fuzzy addition, multiplication, and division as follows, which is the base of fuzzy modeling:

\[
A \oplus B = \left[ (x + y) \right] \min \{ \mu_a(x), \mu_b(y) \} \\
A \otimes B = \left[ (x \times y) \right] \min \{ \mu_a(x), \mu_b(y) \} \\
A \Theta B = \left[ (x + y) \right] \min \{ \mu_a(x), \mu_b(y) \}
\]

where \(A (A = [x, \mu_X(x)])\) and \(B (B = [y, \mu_Y(y)])\) are two fuzzy sets; symbols \(\oplus\), \(\otimes\), and \(\Theta\) represent fuzzy arithmetic operations of addition, multiplication, and division; and \(+, \times\) are normal arithmetic operations. When the result of calculation leads to more than one membership value for a given fuzzy set, the highest membership value is selected. In most cases, \(A \Theta B\) must be approximated. One approach is to reduce this set by disregarding any number from the division operation that is not integer (Dubois & Prade, 1980).

3.2. Solute Transport Model

The solute transport model is used for estimating pollutant concentrations in the aquifer immediately down-gradient from the contaminated site, given a constant source concentration in a multilayered system (Domenico, 1987). The solution is obtained by coupling the equations for solute migration through soil with the mass balances in the aquifer, via boundary conditions at the soil-aquifer interface. The main underlying assumptions include: (i) soil infiltration is controlled by its permeability, and is sufficient to keep the soil saturated; (ii) the source concentration equals to the pollutant’s aqueous solubility; (iii) the pollutant migrates vertically through the soil as a dissolved phase; (iv) the mass flux at soil-aquifer interface mixes homogeneously over a certain aquifer thickness; and (v) there is enough mass at the source to attain a steady-state concentration in the aquifer below the site. For a multi-layered system in subsurface, the solution is as follows:

\[
c = \frac{C_0}{1 + \frac{qH}{IL} \left[ 1 - \exp \left( \frac{\int_a^b \frac{I}{\theta(z)D(z)} \, dz}{q} \right) \right]}
\]

where \(c\) = average pollutant concentration in the aquifer immediately downgradient from the contaminated area (after mixing); \(C_0\) = source concentration; \(L\) = length of the site in the direction of groundwater flow; \(I\) = infiltration rate through the layers overlying the aquifer; \(E\) = thickness of the aquifer; \(\theta(z)\) and \(D(z)\) = volumetric water content and diffusion-dispersion coefficient at depth \(z\), respectively; \(q\) = Darcy flux in the aquifer domain; and \(H\) = thickness of the mixing layer.

When there is only one layer overlying the aquifer, Eqs. (5) can be simplified as follows, through the introduction of diffusion-dispersion coefficient \(D\) defined by Freeze and Cherry (1979):

\[
c = \frac{C_0}{1 + \frac{qH}{IL} \left[ 1 - \exp \left( \frac{\int_a^b \frac{I}{\theta D} \, dz}{q} \right) \right]}
\]
where $e$ = thickness of the soil layer. Equation (6) is the pollutant source term of solute transport model for calculating pollutant migration in the aquifer to a distance located downgradient from the source.

Due to sparse and imprecise natures of the available information, many parameters of the analytical solute transport model for calculating steady-state pollutant concentration would be represented as fuzzy numbers associated with membership functions. This is a model fuzzification process. With fuzzy arithmetic operations, when some parameters are presented as fuzzy numbers, the resulting modeling outputs (contaminant concentrations) would also be fuzzy. For a number of fuzzy variables $X_1, X_2, \ldots, X_k$ and a fuzzy function $f(X_1, X_2, \ldots, X_k)$, the following operation procedure can be used (Dubois & Prade 1988):

1. Select a level, $\alpha$, of membership grade (a level of possibility);
2. Select values $a_i$ and $b_i$ corresponding to the $\alpha$-cut for each fuzzy number $X_1, X_2, \ldots, X_k$ (Figure 1);
3. Calculate the minimum and maximum values of $f$, considering all of the values located within the $\alpha$-cut for each fuzzy member;
4. Use the minimum and maximum values as lower and upper limits of the $\alpha$-cut of $f$;
5. Repeat the operation for another $\alpha$-cut;
6. Build the fuzzy outputs of $f$ using the minimum and maximum values of $f$ for each $\alpha$-cut.

After this procedure, the membership function of steady-state contaminant concentration in the aquifer can be obtained.

### 3.3. Possibilistic Analysis of Risk Levels

Groundwater standards and related regulations are generally established for safeguarding water quality and preventing adverse effects on public health. The first step in the development of regulations and related guidance is to identify all contaminants that are known to or likely to occur in public water systems. Then, through undertaking epidemiological investigation and laboratory analysis, the relationship between contaminant concentration and human health impacts can be identified. This relationship is useful for establishing the maximum contaminant level (MCL with a meaning of the maximum allowable contaminant concentration), which can be used as a threshold for evaluating potential risks associated with the site contamination.

In the process for developing the MCL, conditions for the toxicological studies were considered deterministic. However, in the real world, many related parameters are uncertain. Due to the sparse of available data related to the MCL, fuzzy numbers would be used for reflecting the uncertainties.

Figure 2 shows a schematic framework for developing fuzzy criteria of different risk levels. It represents the projections of MCL value under each risk level corresponding to

![Figure 2](Image URL)
different membership grades. Here, “risk” is a linguistic variable; “very low”, “low”, “medium”, “fairly high”, and “high” are the fuzzy restrictions. The combination of a fuzzy restriction and a linguistic variable leads to a fuzzy criteria for one risk-level. Totally, five fuzzy criteria for different risk-level can be obtained through the combinations, including “very low risk”, “low risk”, “medium risk”, “fairly high risk”, and “high risk” criteria. To develop each fuzzy criteria, the corresponding fuzzy restriction is linked to a number of numerical values on the scale which represents the MCL value. Then, a membership grade between zero and one is assigned to each branch which connects the fuzzy restriction with a MCL value. The membership grade shows the belief degree in which that MCL value belongs to the fuzzy criteria. The belief degree is based on the information from epidemiological investigation, toxicological study, laboratory analysis, and related expertise and experience.

3.4. Environmental Risk Assessment

Based on the obtained fuzzy membership functions for contaminant concentration and risk-level criteria, the Euclidean method is used for further risk assessment (Hipel, 1982). In the method, an Euclidean distance between fuzzy contaminant concentration and each fuzzy risk-level criteria is calculated:

\[ D(C_i, R_i) = \left( \sum_{i} (C_i - R_i)^2 \right)^{1/2} \quad \forall i, i = 1 \text{ to } n \]  

(7)

where \( D(C_i, R_i) \) denotes Euclidean distance between fuzzy contaminant concentration \( C_i \) and fuzzy risk-level criteria \( R_i \); \( n \) is the total number of certain values in \( C_i \) and \( R_i \) used for distance calculation. Thus, when \( C_i \) and \( R_i \) are known, \( D(C_i, R_i) \) can be calculated through Equation (7). The obtained \( D(C_i, R_i) \) values can then be directly used for environmental risk assessment, through identifying the closest natural expression by minimizing the Euclidean distance.

4. Application

The developed methodology is applied to a site with an underground storage tank that is leaking gasoline into the subsurface. The leaking gasoline consists mainly of aliphatic hydrocarbons (e.g. pentane and butane) and aromatic hydrocarbons (e.g. benzene and toluene). Most of these hydrocarbons are classified as either priority pollutants or hazardous substances. Among them, benzene is classified as a human carcinogen by the International Agency for Research on Cancer and as a hazardous waste and a priority pollutant by the U.S. Environmental Protection Agency (USEPA, 1984). In this study, we focus on the health impacts and risks of benzene.

The type of soil underneath the tank is clay with low hydraulic conductivity (Table 1). However, there has been no building up of the hydraulic head around the tank, and infiltration through the clay layer is gravity-driven. It is assumed that infiltrating water mobilizes dissolved benzene at a concentration equal to its aqueous solubility (1750 mg/L). The aquifer underlying the clay layer is a layer of homogeneous sand. The contaminated zone has a length of 50 m in the direction of groundwater flow. Other parameters, including clay porosity, clay dispersivity, aquifer hydraulic conductivity, aquifer hydraulic gradient, and thickness of mixing layer, are presented in Table 1. All these parameters are fuzzy in nature. They are graphically represented as fuzzy numbers with triangular membership functions (Figure 3).

Table 1. Fuzzy Numbers of the Parameters used in the Case Study

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Minimum Value</th>
<th>Maximum Value</th>
<th>Most Likely Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clay hydraulic conductivity (m/s)</td>
<td>5\times10^{-11}</td>
<td>2\times10^{10}</td>
<td>10^{-10}</td>
</tr>
<tr>
<td>Clay porosity</td>
<td>0.4</td>
<td>0.55</td>
<td>0.5</td>
</tr>
<tr>
<td>Clay dispersivity (m)</td>
<td>0.03</td>
<td>0.10</td>
<td>0.05</td>
</tr>
<tr>
<td>Aquifer hydraulic conductivity (m/s)</td>
<td>8\times10^{4}</td>
<td>2\times10^{-3}</td>
<td>10^{-3}</td>
</tr>
<tr>
<td>Aquifer hydraulic gradient</td>
<td>0.005</td>
<td>0.015</td>
<td>0.010</td>
</tr>
<tr>
<td>Thickness of mixing layer (m)</td>
<td>10</td>
<td>40</td>
<td>30</td>
</tr>
</tbody>
</table>

The results of fuzzy modeling calculation are shown in Figure 4. It is indicated that the steady-state benzene concentrations in the aquifer range from 12 to 638 μg/L. The most likely concentration (65 μg/L with possibility = 1) is the value obtained using the most likely input values in Table 1. Thus, the fuzzy benzene concentration can take on membership grades as follows:

\[ C(\text{benzene}) = [50|0.0, 10|0.0, 15|0.1, 30|0.3, 45|0.6, 60|0.9, 80|0.9, 100|0.9, 200|0.7, 300|0.6] \]  

(8a)

According to the existing water quality standards (CEPA, 1988; USEPA, 1989; Health Canada, 1988), we define fuzzy criteria for five risk-levels as follows:

\[ R_1 = [5|1.0, 10|0.6, 15|0.3, 30|0.0, 45|0.0, 60|0.0, 80|0.0, 100|0.0, 200|0.0, 300|0.0] \]  

(8b)

\[ R_2 = [5|0.2, 10|1.0, 15|0.4, 30|0.1, 45|0.0, 60|0.0, 80|0.0, 100|0.0, 200|0.0, 300|0.0] \]  

(8c)
$R_3 = [5|0.0, 10|0.1, 15|0.3, 30|1.0, 45|1.0, 60|0.5, 80|0.0, 100|0.0, 200|0.0, 300|0.0]$  \quad (8d)

$R_4 = [5|0.0, 10|0.0, 15|0.0, 30|0.0, 45|0.0, 60|0.3, 80|1.0, 100|0.3, 200|0.0, 300|0.0]$  \quad (8e)

$R_5 = [5|0.0, 10|0.0, 15|0.0, 30|0.0, 45|0.0, 60|0.0, 80|1.0, 100|1.0, 200|1.0, 300|1.0]$  \quad (8f)

where $R_1$, $R_2$, $R_3$, $R_4$, and $R_5$ are the criteria for “very low”, “low”, “medium”, “fairly high”, and “high” risk levels, respectively.

Based on Figure 4, and Equations (7) and (8), the Euclidean distance between the fuzzy benzene concentration and the fuzzy risk-levels can be estimated as follows:

$D(C, R_1) = 2.26$  \quad (9a)

$D(C, R_2) = 2.19$  \quad (9b)

$D(C, R_3) = 1.82$  \quad (9c)

$D(C, R_4) = 1.43$  \quad (9d)

$D(C, R_5) = 1.48$  \quad (9e)

The results indicate that, among the five fuzzy risk-levels, the obtained benzene concentration has the closest Euclidean distance to $R_4$ and $R_5$. Therefore, the environmental impacts from the contaminated site has “fairly high” or “high” risk level. Remediation should then be carried out to reduce the risk level and ensure low benzene concentrations in the groundwater that satisfy the regulated water quality standards.

The advantages of this environmental risk assessment method include: (1) it is an integrated approach that incorporates the fuzzy solution of contaminant concentrations and the fuzzy criteria of environmental risk levels within a general assessment framework, and (2) it can effectively reflect uncertainties presented as fuzzy numbers which is especially meaningful for practical problems with sparse information available.

Figure 3. Fuzzy parameters.

Figure 4. Modeling solution of fuzzy benzene concentration.
for the assessment task.

5. Conclusions

This paper presented an integrated environmental risk assessment framework and its application to a petroleum-contaminated site. It consists mainly of three parts: (i) calculation of fuzzy steady-state contaminant concentrations in the aquifer based on an analytical solute transport model, (ii) possibilistic analysis of fuzzy criteria for different risk levels, and (iii) environmental risk assessment based on the Euclidean method. Results from an illustrative case study indicated that environmental risks at a petroleum-contaminated site can be effectively evaluated through the developed methodology. The risk assessment framework proposed in this study can effectively handle uncertainties presented as fuzzy numbers. Fuzzy nature of water quality and risk-level criteria were reflected in the related simulation and evaluation models. This framework is especially useful for situations when probabilistic information is unavailable. Application of the proposed approach to risk assessment of groundwater contamination represents a new attempt to the area of petroleum waste management under uncertainty. The results are useful for the related site remediation and management decisions.

Acknowledgments. This research has been supported by the Natural Sciences and Engineering Research Council of Canada. The authors thank the anonymous reviewers for their comments and suggestions that helped in improving the manuscript.

References


